SPRINGER NATURE
Reference

Bharath Sriraman
Editor

Handbook of the History and Philosophy of Mathematical Practice

Springer

Handbook of the History and Philosophy of Mathematical Practice

Bharath Sriraman
Editor

Handbook of the History and Philosophy of Mathematical Practice

With 514 Figures and 75 Tables

Springer

Editor
Bharath Sriraman
Department of Mathematical Sciences
The University of Montana
Missoula, MT, USA

ISBN 978-3-031-40845-8 ISBN 978-3-031-40846-5 (eBook)
https://doi.org/10.1007/978-3-031-40846-5

© Springer Nature Switzerland AG 2024
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

In Memoriam
Reuben Hersh (1927–2020)
Sometimes being kind is more important than being right
Yuri Manin (1937–2023)
Good proofs are proofs that make us wiser
Chandler Davis (1926–2022)
The mathematics tells truth about the world
We are its ventriloquist, yet
some words it won't let us put in its mouth
&
C and K
for their unwavering support through the vicissitudes of the years. . .

Preface

This book is the outcome of co-operation between the communities of mathematicians, philosophers, and historians of mathematics. I would like to thank the editorial advisory board consisting of eminent scholars working in the aforementioned communities for being a part of this project. Next, I am grateful to the section editors for shaping their respective sections and the authors who contributed to these sections. Finally, I would like to mention specific individuals without whom this book would have been difficult to accomplish logistically. These individuals are Thomas Hempfling (Springer/Birkhaüser), Sarah Annette Goob, and Ruth LeFevre (Palgrave).

विद्या ददाति विनयं विनयात् याति पात्रताम्।
पात्रत्वाद्धनमाप्नोति धनाद्धर्मं ततः सुखम्॥

Missoula, USA
April 2024

Bharath Sriraman
Editor

Contents

Volume 1

Section I From Origins to the Natural Historians/Philosophers . . . 1
Bharath Sriraman

Introduction to the Handbook of the History and Philosophy of Mathematical Practice 3
Bharath Sriraman

The History and Philosophy of Mathematical Practice: From Origins to Natural Historians/Philosophers – A Conversation 7
Bharath Sriraman and Reuben Hersh

Yuri Ivanovich Manin (1937–2023): Mathematics, Philosophy, and Poetry 13
Athanase Papadopoulos

The Nature of Mathematical Objects 35
Carlo Cellucci

The Ghost and the Spirit of Pythagoras 63
Arkady Plotnitsky

The Metaphysics of Platonism 91
Ulf Persson

Mathematical Practices Can Be Metaphysically Laden 109
Colin Jakob Rittberg

Definitions (and Concepts) in Mathematical Practice 135
V. J. W. Coumans

Mathematics and the Method of Abstraction 159
Kit Fine

Leibniz on Number Systems 167
Lloyd Strickland

Fermat and Descartes in Light of Premodern Algebra and Viète 199
Jeffrey A. Oaks

Descartes' Transformation of Greek Notion of Proportionality 235
Piotr Błaszczyk

Leonhard Euler Between Mathematics and Natural Philosophy: An Introduction to Natural Science Anleitung zur Naturlehre 269
Danilo Capecchi

Algebraic Versus Geometric Thought and Expression in the Early Calculus 323
Viktor Blåsjö

Evolution of Leibniz's Thought in the Matter of Fictions and Infinitesimals 341
Monica Ugaglia and Mikhail Katz

René Thom: From Mathematics to Philosophy 385
Athanase Papadopoulos

Heuristics and Mathematical Practice 431
Otávio Bueno

Finding a Fit Among Philosophical Finitisms 443
Eamon Darnell and Aaron Thomas-Bolduc

Section II Mathematician's Reflections on Practice and Working Philosophy of Mathematics 463
Athanase Papadopoulos

Introduction to *Mathematicians' Reflections on Practice and Working Philosophy of Mathematics* 465
Athanase Papadopoulos

The Anthyphairetic Interpretation of Knowledge as Recollection in Plato's *Meno* 80d-86e and 97a-98b 471
Stelios Negrepontis

The Mystery of Plato's Receptacle in the *Timaeus* Resolved 521
Stelios Negrepontis and Demetra Kalisperi

The Indivisible Line Coincides with the Platonic Intelligible Being . . . 599
Stelios Negrepontis

Brouwer's Intuitionism 645
Victor Pambuccian

Mathematical Practice as Philosophy, with Galois, Riemann, and Grothendieck 701
Arkady Plotnitsky

"That Which a Minority Construct": Abelian Mathematics in Abel, Galois, Noether, and Grothendieck 749
Arkady Plotnitsky

Pavel Florensky and His World 801
Athanase Papadopoulos

The Practice of Mathematics in a Totalitarian Society 869
A. B. Sossinsky

Philosophy Behind Mathematics 881
Ken'ichi Ohshika

Some Disillusioned Thoughts About School and Teaching 897
Valentin Poénaru

Volume 2

Section III Influences in Twentieth- and Twenty-First-Century Mathematical Practice **907**
Bharath Sriraman and Kyeonghwa Lee

Influences in Twentieth- and Twenty-First-Century Mathematical Practice 909
Bharath Sriraman and Kyeonghwa Lee

The Pell Equation in the Pythagoreans, Theaetetus, and Hindu Mathematics 915
Stelios Negrepontis, Vassiliki Farmaki, and Marina Brokou

Ancient Indian Mathematics: Sulbasutras – A Mathematical Review 1023
S. G. Dani

An Ethnoarithmetic Excursion into the Javanese Calendar 1097
Natanael Karjanto and François Beauducel

Historiography of Mathematics from the Mathematician's Point of View 1127
Viktor Blåsjö

A Tale of Three Cities: Thebes, Babylon, and Alexandria 1149
Maurice Burke

Sir Isaac Newton and *Opticks* 1179
Jonah Britton and Bharath Sriraman

Olry Terquem's Forgotten Problem and an Enumerative-Combinatorial Perspective on the Euclidean Algorithm 1189
Robert G. Donnelly, Molly W. Dunkum, and Rachel McCoy

The Ethics of Mathematical Practice 1219
Paul Ernest

Grothendieck: A Short Guide to His Mathematical and Philosophical Work (1949–1991) 1257
Fernando Zalamea

An Uncertain Travel 1297
Elijah Liflyand

The Unreasonable Effectiveness of Mathematics to Generate New Spatial Imaginaries 1307
Nathalie Sinclair

High Mathematics Meets High Finance 1329
Tom P. Davis

Section IV Practical, Historical, and Philosophical Instances of Probability **1361**
Egan J. Chernoff

Introduction to Practical, Historical, and Philosophical Instances of Probability 1363
Egan J. Chernoff

How Markov's Little Idea Transformed Statistics 1371
Jeffrey S. Rosenthal

The Use and Abuse of Probability in Evolutionary Biology 1383
Jason Rosenhouse

Probability Models in Statistical Data Analysis: Uses, Interpretations, Frequentism-as-Model 1411
Christian Hennig

Section V Constructing the Reals: Paradoxes and Foundations **1459**
Paul M. Livingston and Bharath Sriraman

Introduction to the History and Philosophy of Mathematical Practice in Constructing the Reals 1461
Paul M. Livingston

The Continuum and the Evolution of the Concept of Real Number . . . 1473
John L. Bell

On Bolzano and Greek Concepts of Continuity 1563
Piotr Błaszczyk and Marlena Fila

Bolzano's Theory of *meßbare Zahlen*: Insights and Uncertainties Regarding the Number Continuum 1595
Elías Fuentes Guillén

Abel's Approach to Elliptic Integrals 1633
John K. Dagsvik

The Absolute Arithmetic Continuum and Its Geometric Counterpart 1677
Philip Ehrlich

Volume 3

Section VI Practices in Euclidean and Non-Euclidean Geometry **1719**
Bharath Sriraman

Euclidean and Non-Euclidean Geometry in the History and Philosophy of Mathematical Practice 1721
Bharath Sriraman

Cultures of Mathematical Practice in Alexandria in Egypt: Claudius Ptolemy and His Commentators (Second–Fourth Century CE) 1727
Alberto Bardi

Menelaus' Spherics in Greek and Arabic Mathematics 1745
Athanase Papadopoulos

Copernicus and Axiomatics 1789
Alberto Bardi

Visual Representations of Euclidean Geometry: Diagrammatic Reasoning in Oliver Byrne's Work 1807
Andrea Pedeferri

From Friezes to Quasicrystals: A History of Symmetry Groups 1823
Franka Miriam Brückler and Vladimir Stilinović

The Notion of Space in Grothendieck 1865
John Alexander Cruz Morales

Geometry in the Twentieth Century: A Return to Euclid – The Work of Herbert Busemann 1885
Athanase Papadopoulos

Section VII Logic **1903**
Bharath Sriraman

Logic in the History and Philosophy of Mathematical Practice 1905
Valentina Harizanov

Model Theory of Nonstandard Structures with Applications 1921
Roman Kossak

Computability Theory 1933
Valentina Harizanov, Keshav Srinivasan, and Dario Verta

Reverse Mathematics 1963
John Stillwell

Countable Nonstandard Models: Following Skolem's Approach 1989
Rumen D. Dimitrov and Valentina Harizanov

Fermat's Last Theorem 2011
Colin McLarty

Section VIII Proof **2035**
Bart Van Kerkhove and Joachim Frans

Proof in the History and Philosophy of Mathematical Practice: An Introduction 2037
Joachim Frans and Bart Van Kerkhove

Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research 2045
Valeria Giardino

The Social Epistemology of Mathematical Proof 2069
Line Edslev Andersen

The Values of Mathematical Proofs 2081
Rebecca Lea Morris

Formal Proofs in Mathematical Practice 2113
Danielle Macbeth

The Dark Side of the History of Proof 2137
Karine Chemla

Section IX Ontology **2163**
Bharath Sriraman and Michael N. Fried

Ontology in the History and Philosophy of Mathematical Practice: An Introduction 2165
Michael N. Fried

The Algorithmic-Device View of Informal Rigorous Mathematical Proof 2179
Jody Azzouni

Counterpossibles in Mathematical Practice: The Case of Spoof Perfect Numbers 2261
Alan Baker

What Mathematicians Do: Mathematics as Process and Creative Rationality 2289
William Byers

Platonism, *De Re*, and (Philosophy of) Mathematical Practice 2307
Marco Panza

Section X Pluralism **2337**
Michele Friend

One Mathematic(s) or Many? Foundations of Mathematics in Twentieth-Century Mathematical Practice 2339
Andrei Rodin

Multiverse and the Society 2365
Mirna Džamonja

Implicitly Defining Mathematical Terms 2383
Demetra Christopoulou

Non-deductive Justification in Mathematics 2401
A. C. Paseau

Volume 4

Section XI "Experimental" Mathematics **2417**
Jessica Carter

Experimental Mathematics in Mathematical Practice 2419
Jessica Carter

The Role of Experiments in *Experimental Mathematics* 2431
Henrik Kragh Sørensen

Revisiting "The New 4CT Problem" 2459
Mark McEvoy

Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice 2481
Giovanni Ferraro

Mathematical Experiments on Paper and Computer 2503
Dirk Schlimm and Juan Fernández González

Thought Experiments in Mathematics: From Fiction to Facts 2523
Irina Starikova

The Intimate Interplay Between Experimentation and Deduction: Some Classroom Implications 2551
Michael de Villiers and Hans Niels Jahnke

Section XII Views from Other Domains **2587**
Andrew Aberdein

Introduction to Views from Other Domains 2589
Andrew Aberdein

Inverting Hierarchies: The Sociology of Mathematical Practice 2597
Michael J. Barany and Milena I. Kremakova

Practice Makes Perfectoid 2619
Michael J. Barany

Mathematics Education Research on Mathematical Practice 2637
Keith Weber and Matthew Inglis

Argumentation in Mathematical Practice 2665
Andrew Aberdein and Zoe Ashton

Rhetorical Approaches to the Study of Mathematical Practice 2689
G. Mitchell Reyes

Bayesian Perspectives on Mathematical Practice 2711
James Franklin

Polymath as an Epistemic Community 2727
Patrick Allo, Jean Paul Van Bendegem, and Bart Van Kerkhove

Is Mathematics Connected to Religion? 2757
Stanisław Krajewski

Section XIII New Perspectives **2783**
Valeria Giardino

New Perspectives: An Introduction 2785
Valeria Giardino

What Are Mathematical Practices? The Web-of-Practices Approach 2793
José Ferreirós

Experiments in Mathematics: Fact, Fiction, or the Future? 2821
Jean Paul Van Bendegem

The Social Constitution of Mathematical Knowledge: Objectivity, Semantics, and Axiomatics 2847
Paola Cantù

The Epistemological Subject(s) of Mathematics 2879
Silvia De Toffoli

Agency in Mathematical Practice 2905
Yacin Hamami

The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts 2925
Fenner Stanley Tanswell and Matthew Inglis

Section XIV Semiology of Mathematical Practice 2953
Juan Luis Gastaldi

Introduction to the Semiology of Mathematical Practice 2955
Juan Luis Gastaldi

Signs as a Theme in the Philosophy of Mathematical Practice 2971
David Waszek

Structural Semiotics as an Ontology of Mathematics 3003
Roi Wagner

Shaping, Revisited 3033
Reviel Netz

Mathematical Practices and Written Evidence: General Reflections Based on a Historian's Experience 3047
Karine Chemla

What Happens, from a Historical Point of View, When We Read a Mathematical Text? 3073
Lucien Vinciguerra

Symbolic Algebra as a Semiotic System 3101
Ladislav Kvasz

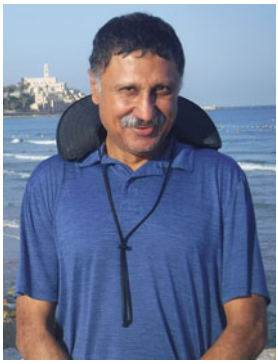
Reading Mathematical Texts with Structuralist Semiotics 3135
Anna Kiel Steensen

The Design of Mathematical Language 3151
Jeremy Avigad

How to Do Maths with Words: Neural Machine Learning Applications to Mathematics and Their Philosophical Significance 3191
Juan Luis Gastaldi

Index 3227

About the Editor



Bharath Sriraman is a Professor of Mathematics at the University of Montana-Missoula working in the areas of cognition, creativity, mathematics, and philosophy. In 2016, he was named the University of Montana Distinguished Scholar. Prof. Sriraman has held 40+ visiting professorships at institutions in Norway, Iceland, Sweden, Germany, Turkey, Iran, Malaysia, South Korea, Canada, South Africa, Colombia, and Argentina, which include two US Fulbright awards. To date, Prof. Sriraman has published more than 330 journal articles, book chapters, proceedings papers, and reference work entries, and served as the Editor of 33 volumes, in his areas of interest. He is the founder and Editor-in-Chief of *The Mathematics Enthusiast*, an independent, peer-reviewed, open-access international journal. In addition to editing and curating the *Handbook of the History and Philosophy of Mathematical Practice*, he previously served as the Editor-in-Chief of the *Handbook of the Mathematics of the Arts and Sciences* (Springer, 2021). In keeping with the adage *Omne Trium Perfectum*, his third Major Springer Reference Works Project is the Handbook of Visual, Experimental and Computational Mathematics: Bridges Through Data in preparation for release in 2027. In his spare time, he is an amateur arborist.

Editorial Advisory Board

Andrew Aberdein Florida Institute of Technology, USA

Jody Azzouni Tufts University, USA

Otávio Bueno University of Miami, USA

William Byers Concordia University, Canada

Carlo Cellucci Sapienza University of Rome, Italy

Chandler Davis University of Toronto, Canada (1926–2022)

Paul Ernest University of Exeter, UK

Michele Friend George Washington University, USA

Reuben Hersh University of New Mexico, USA (1927–2020)

Kyeong-Hwa Lee Seoul National University, South Korea

Yuri Manin Max Planck Institute for Mathematics, Germany (1937–2023)

Athanase Papadopoulos University of Strasbourg, France

Ulf Persson Chalmers University of Technology, Sweden

John Stillwell University of San Francisco, USA

David Tall University of Warwick, UK

Section Editors

Andrew Aberdein Florida Institute of Technology, Melbourne, FL, USA

Jessica Carter Department of Mathematics, Centre for Science Studies, Aarhus University, Ny Munkegade Aarhus, Denmark

Egan J. Chernoff University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Joachim Frans Vrije Universiteit Brussel, Centre for Logic and Philosophy of Science, Brussels, Belgium

Michael N. Fried Graduate Program for Science and Technology Education, School of Education, Ben-Gurion University of the Negev, Beer Sheva, Israel

Michèle Indira Friend Department of Philosophy, George Washington University, Washington, DC, USA

Univ. Lille, CNRS, Centrale Lille, Univ. Artois, UMR 8181 –UCCS – Unité de Catalyse et Chimie du Solide, Lille, France

Juan Luis Gastaldi ETH Zürich, Zürich, Switzerland

Valeria Giardino Institut Jean Nicod, Chargée de Recherche/Research Fellow (CRCN) Centre, National de la Recherche Scientifique, Paris, France

Bart Van Kerkhove Vrije Universiteit Brussel Centre for Logic and Philosophy of Science, Brussels, Belgium

Kyeonghwa Lee Seoul National University, Seoul, South Korea

Paul M. Livingston Philosophy Department, University of New Mexico, Albuquerque, NM, USA

Athanase Papadopoulos Institut de Recherche Mathématique Avancée and Centre de Recherche et d'Expérimentation sur l'Acte Artistique of the University of Strasbourg and the Centre National de la Recherche Scientifique, Strasbourg, France

Bharath Sriraman University of Montana, Missoula, MT, USA

Contributors

Andrew Aberdein School of Arts and Communication, Florida Institute of Technology, Melbourne, FL, USA

Patrick Allo Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

Line Edslev Andersen Aarhus University, Aarhus, Denmark

Zoe Ashton Department of Philosophy, Ohio State University, Columbus, OH, USA

Jeremy Avigad Carnegie Mellon University, Pittsburgh, PA, USA

Jody Azzouni Department of Philosophy, Tufts University, Medford, MA, USA

Alan Baker Department of Philosophy, Swarthmore College, Swarthmore, PA, USA

Michael J. Barany Science, Technology & Innovation Studies, The University of Edinburgh, Edinburgh, UK

Alberto Bardi Department of the History of Science, Tsinghua University, Beijing, China

François Beauducel Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France

Institut de recherche pour le développement, Research and Development Technology Center for Geological Disaster, Balai Penyelidikan dan Pengembangan Teknologi Kebencanaan Geologi (BPPTKG), Yogyakarta, Indonesia

John L. Bell Department of Philosophy, The University of Western Ontario, London, ON, Canada

Viktor Blåsjö Universiteit Utrecht, Utrecht, The Netherlands

Piotr Błaszczyk Institute of Mathematics, Pedagogical University of Cracow, Cracow, Poland

Jonah Britton University of Connecticut Storrs, Storrs, CT, USA

Marina Brokou Department of Mathematics, Athens University, Athens, Greece

Franka Miriam Brückler Department of Mathematics, Faculty of Science, University of Zagreb, Zagreb, Croatia

Otávio Bueno Department of Philosophy, University of Miami, Coral Gables, FL, USA

Maurice Burke Montana State University, Bozeman, Montana, USA

William Byers Concordia University, Montreal, QC, Canada

Paola Cantù Centre Gilles Gaston Granger, Aix-Marseille University, CNRS, Aix-en-Provence, France

Danilo Capecchi Sapienza University of Rome, Rome, Italy

Jessica Carter Department of Mathematics, Aarhus University, Aarhus C, Denmark

Carlo Cellucci Sapienza University of Rome, Rome, Italy

Karine Chemla SPHERE (UMR7219, CNRS & Université Paris Cité), Paris, France

Egan J Chernoff College of Education/Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK, Canada

Demetra Christopoulou Department of Mathematics, National and Kapodistrian University of Athens, Athens, Greece

V. J. W. Coumans Institute for Science in Society, Radboud University, Nijmegen, The Netherlands

John Alexander Cruz Morales Universidad Nacional de Colombia, Bogotá, Colombia

John K. Dagsvik Research Department, Statistics Norway, Oslo, Norway

S. G. Dani UM-DAE Centre for Excellence in Basic Sciences, Vidyanagari Campus of the University of Mumbai, Mumbai, India

Eamon Darnell University of Toronto, Toronto, ON, Canada

Tom P. Davis FactSet Research Systems Ltd., London, UK

Silvia De Toffoli Scuola Universitaria Superiore IUSS Pavia, Pavia, Italy

Rumen D. Dimitrov Department of Mathematics & Philosophy, Western Illinois University, Macomb, IL, USA

Robert G. Donnelly Department of Mathematics and Statistics, Murray State University, Murray, KY, USA

Molly W. Dunkum Department of Mathematics, Western Kentucky University, Bowling Green, KY, USA

Mirna Džamonja IRIF(CNRS & Université de Paris-Cité), Paris, France

Philip Ehrlich Ohio University, Athens, OH, USA

Paul Ernest University of Exeter, Exeter, UK

Vassiliki Farmaki Department of Mathematics, Athens University, Athens, Greece

Juan Fernández González Graduate School of Design, Harvard University, Cambridge, MA, USA

Giovanni Ferraro University of Molise, Campobasso, Italy

José Ferreirós Universidad de Sevilla, Sevilla, Spain

Marlena Fila Institute of Mathematics, Pedagogical University of Cracow, Cracow, Poland

Kit Fine New York University, New York, NY, USA

James Franklin School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, Australia

Joachim Frans Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

Michael N. Fried Program for Science and Technology Education, School of Education, Ben-Gurion University of the Negev, Beer Sheva, Israel

Elías Fuentes Guillén Institute of Philosophy of the Czech Academy of Sciences, Prague, Czech Republic

Department of Mathematics, Faculty of Sciences, UNAM, Mexico City, Mexico

Juan Luis Gastaldi ETH Zurich, Zurich, Switzerland

Valeria Giardino Chargée de Recherche/Research Fellow (CRCN) Centre National de la Recherche Scientifique, Institut Jean Nicod, Paris, France

Yacin Hamami Postdoctoral Researcher FNRS, Philosophy Department, University of Liège, Liège, Belgium

Valentina Harizanov Department of Mathematics, The George Washington University, Washington, DC, USA

Christian Hennig Dipartimento di Scienze Statistiche "Paolo Fortunati", Università di Bologna, Bologna, Italy

Reuben Hersh University of New Mexico, Albuquerque, NM, USA

Matthew Inglis Department of Mathematics Education, Loughborough University, Loughborough, UK

Hans Niels Jahnke Faculty of Mathematics, University of Duisberg-Essen, Essen, FR, Germany

Demetra Kalisperi Department of Mathematics, Athens University, Athens, Greece

Natanael Karjanto Department of Mathematics, University College, Sungkyunkwan University, Suwon, Republic of Korea

Mikhail Katz Department of Mathematics, Bar Ilan University, Ramat Gan, Israel

Anna Kiel Steensen ETH, Zurich, Switzerland

Roman Kossak The Graduate Center, CUNY, New York, NY, USA

Stanisław Krajewski Faculty of Philosophy, University of Warsaw, Warsaw, Poland

Milena I. Kremakova The Sociological Review Magazine, The Sociological Review Foundation, London, UK

Ladislav Kvasz Institute of Philosophy, Czech Academy of Sciences, Prague, Czech Republic

Kyeonghwa Lee Seoul National University, Seoul, South Korea

Elijah Liflyand Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel

Paul M. Livingston Philosophy Department, University of New Mexico, Albuquerque, NM, USA

Danielle Macbeth Haverford College, Haverford, PA, USA

Rachel McCoy Department of Mathematics, Western Kentucky University, Bowling Green, KY, USA

Mark McEvoy Hofstra University, Hempstead, NY, USA

Colin McLarty Case Western Reserve University, Cleveland, OH, USA

Rebecca Lea Morris Independent Scholar, Minneapolis, MN, USA

Stelios Negrepontis Department of Mathematics, Athens University, Athens, Greece

Reviel Netz Stanford University, Stanford, CA, USA

Jeffrey A. Oaks Department of Mathematical Sciences, University of Indianapolis, Indianapolis, IN, USA

Ken'ichi Ohshika Department of Mathematics, Gakushuin University, Tokyo, Japan

Victor Pambuccian School of Mathematical and Natural Sciences, Arizona State University – West Campus, Phoenix, AZ, USA

Marco Panza CNRS, IHPST (CNRS and University of Paris 1, Panthéon Sorbonne), Paris, France

Chapman University, Orange, CA, USA

Athanase Papadopoulos Institut de Recherche Mathématique Avancée (IRMA) and Centre de Recherche et d'Expérimentation sur l'Acte Artistique (ITI CREAA), Université de Strasbourg and CNRS, Strasbourg Cedex, France

A. C. Paseau Oxford University, Oxford, UK

Andrea Pedeferri Union College, Schenectady, NY, USA

Ulf Persson Chalmers Institute of Technology, Gothenburg, Sweden

Arkady Plotnitsky Literature, Theory, and Cultural Studies Program, Philosophy and Literature Program, Purdue University, West Lafayette, IN, USA

Valentin Poénaru IHES, Bures-sur-Yvette, France

G. Mitchell Reyes Lewis & Clark College, Portland, OR, USA

Colin Jakob Rittberg Centre for Mathematical Cognition, Loughborough University, Loughborough, UK

Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussel, UK

Andrei Rodin Institute of Philosophy, Russian Academy of Sciences, Moscow, Russia

Saint-Petersburg State University, St. Petersburg, Russia

Jason Rosenhouse James Madison University, Harrisonburg, VA, USA

Jeffrey S. Rosenthal University of Toronto, Toronto, ON, Canada

Dirk Schlimm Department of Philosophy, McGill University, Montreal, QC, Canada

Nathalie Sinclair Simon Fraser University, Burnaby, BC, Canada

Henrik Kragh Sørensen Section for History and Philosophy of Science, Department of Science Education, University of Copenhagen, København Ø, Denmark

A. B. Sossinsky Strasbourg, France

Keshav Srinivasan Department of Mathematics, George Washington University, Washington, DC, USA

Bharath Sriraman University of Montana, Missoula, MT, USA

Irina Starikova UFMG, Belo Horizonte, Brazil

Vladimir Stilinović Department of Chemistry, Faculty of Science, University of Zagreb, Zagreb, Croatia

John Stillwell University of San Francisco, San Francisco, CA, USA

Lloyd Strickland History, Politics, and Philosophy, Manchester Metropolitan University, Manchester, UK

Fenner Stanley Tanswell Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

Department of Mathematics Education, Loughborough University, Loughborough, UK

Aaron Thomas-Bolduc Independent Researcher, Port Alberni, BC, Canada

Monica Ugaglia Montecassiano, Italy

Jean Paul Van Bendegem Center for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

Bart Van Kerkhove Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

Dario Verta Department of Mathematics, George Washington University, Washington, DC, USA

Michael de Villiers Stellenbosch University, Stellenbosch, South Africa

Lucien Vinciguerra Université de Lille UMR 8163 – STL – Savoirs Textes Langage, Lille, France

Roi Wagner ETH, Zurich, Switzerland

David Waszek ITEM, Ecole Normale Supérieure, Paris, France

Keith Weber Rutgers University, New Brunswick, NJ, USA

Fernando Zalamea Universidad Nacional de Colombia, Bogotá, Colombia

Section I

From Origins to the Natural Historians/Philosophers

Bharath Sriraman

Introduction to the Handbook of the History and Philosophy of Mathematical Practice

Bharath Sriraman

Abstract

This Handbook offers a comprehensive survey of the story of mathematics and its practices. It is neither her story, nor his story, but story. In order to do so, the 100+ chapters spanning 3200+ pages in four volumes examine our collective mathematical practices over time and ways in which this shaped its history and philosophy. "Mathematical practice" is left as an undefined term in order to allow the reader to conceptualize it in whatever way they choose to. The book was inspired by conversations with three eminent souls Reuben Hersh, Yuri Manin, and Chandler Davis and is dedicated to them in memoriam.

Keywords

Origins · Natural history · Natural philosophy · Mathematicians reflections · Influences in twentieth- and twenty-first-century mathematical practice · Probability · Constructing the reals · Paradoxes and foundations · Euclidean and non-Euclidean geometry · Logic · Proof · Ontology · Pluralism · Experimental mathematics · Views from other domains · New perspectives · Semiology of mathematical practice · Reuben Hersh · Yuri Manin · Chandler Davis

B. Sriraman (✉)
University of Montana, Missoula, MT, USA
e-mail: sriramanb@mso.umt.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_148

1 Introduction

It is apropos that the introduction to this Handbook and its completion occurs in Old Jaffa. Reuben Hersh's progenitor landed on these shores to escape the clutches of Pilsudski's army in Poland more than a century ago, to emigrate yet again to America in 1920. This Handbook is dedicated in memoriam to three individuals, namely, Reuben Hersh, Yuri Manin, and Chandler Davis. Reuben played an important role when I was conceptualizing it and offered his generous help. I was asked "Why are you doing this?" My reply was a cheeky "Deeds not words" (Emmeline Pankhurst's words), to elicit loud laughter from Vera John-Steiner who remarked ". . .which will require many, many words!" Keeping with another thing, I learned from Reuben and Vera, "Be pithy but draw in the reader," I will be as succinct as possible. Yuri Manin and Chandler Davis also offered valuable advice in crafting this book. Yuri's advice was to draw in the collective knowledge of the community that engages in mathematical practice to have something that makes us wiser, like a good proof. Chandler's tremendous gift with words (prose and poetry) in addition to his convictions helped me to get to the finish line. Alas, all the individuals mentioned here are no longer with us. It is my hope that this book channels their spirit and courage.

This Handbook offers a comprehensive survey of the story of mathematics. It is neither her story, nor his story, but story. In order to do so, it seemed wiser to examine our collective practices over time and ways in which this shaped its history and philosophy. What is "mathematical practice" then? It is an undefined term as far as I am concerned. The contents of the book will perhaps aid the reader in conceptualizing it in whatever way they choose to.

The book has been organized in four volumes consisting of 14 sections in all, and 100+ chapters. Each volume is self-contained and offers a different starting and ending point in the telling of the story of mathematical practice(s). Volume 1 consists of sections: (a) From Origins to the Natural Historians/Philosophers; and (b) Mathematician's Reflections on Practice and Working Philosophy of Mathematics. Volume 2 contains (c) Influences in Twentieth and Twenty-First-Century Mathematical Practice; (d) Practical, Historical and Philosophical Instances of Probability, and (e) Constructing the Reals: Paradoxes and Foundations. Taken as a whole, the first two volumes offer more or less a complete story of mathematical practice including reflections from those who had to practice mathematics in highly unusual sociopolitical conditions in the twentieth century. Volumes 3 and 4 become more technical in their rendition and cover the following topics (f) Practices in Euclidean and Non-Euclidean Geometry; (g) Logic; (h) Proof; (i) Ontology; (j) Pluralism; (k) Experimental Mathematics (not in a strictly Borweinian sense); (l) Views from Other Domains; (m) New Perspectives; and finally (n) Semiology of Mathematical Practice.

Taken as a whole, this Handbook has been a labor of love, and I hope it is of some value to all those who engage in mathematical practice.

Old Jaffa, July 27, 2023

The History and Philosophy of Mathematical Practice: From Origins to Natural Historians/Philosophers – A Conversation

Bharath Sriraman and Reuben Hersh

Contents

1 A Conversation 8
2 Coda 10

Abstract

In this chapter, the authors engage in a dialogue on what constitutes mathematical practice from pre-Socratic Greeks onto the advent of natural philosophy. The conversation occurred through dialogues between the authors over the period 2016–2019. A coda with relevant content from the Origins section is provided.

Keywords

Origins of mathematics · Mathematical practices · History of mathematics · Platonism · Pythagoreans

Reuben Hersh: deceased.

B. Sriraman (✉)
University of Montana, Missoula, MT, USA
e-mail: sriramanb@mso.umt.edu

R. Hersh
University of New Mexico, Albuquerque, NM, USA

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_117

1 A Conversation

H = Hersh; S = Sriraman

S: What would be a good starting point for the opening section on Origins? We could start with the Pythagoreans?

H: Pre-Socratic might be a better way to characterize it. Poor Pythagoras has been turned into an academic industry. Have you seen the book on Ancient Loons?

S: The one by Davis? I was supposed to do a review of it, but it seemed too esoteric in my opinion. The book was very anecdotal since it reported on email conversations between Pingree and Davis, a bit like what we have been doing for over a decade. There is a chapter on Pythagoras in it and parts of the book are euphemistic in my opinion.

H: Is Carlo (Cellucci) doing a chapter in the opening section?

S: We agreed on a topic related to what mathematical objects really are. It seemed related to the book he is currently writing and fitting to be in the opening section of the book.

H: Do you know at which point he starts tracing the history of mathematical objects?

S: Apparently Herodotus writes of an Egyptian King asking the nature of mathematical objects. I don't think there is too much emphasis on ontology per se. There will also be a chapter from Ulf Persson on Platonism.

H: He (Ulf) visited me many years ago. He has had an interesting career. Is he going to get into metaphysics? We had some conversations about Platonism as well if I remember correctly.

S: I have given free rein in the sense that there is a lot of misappropriation of the expression "Platonism" – it seems it is used differently by philosophers as opposed to things written about the psychology of mathematics and those doing mathematics education. One could also say Poor Plato. Ulf wants to get into the physics or science of things. From what I understand in our email exchanges, he wants to compare mathematical objects to atoms – if one believes in the existence of the latter, does that have any implications for believing in the former?

H: My use of Platonism is quite clear – you can call it mathematical Platonism as it pertains to the practice of mathematics. People get too hung up with the ontology of things which I have come to now think of as nonsensical. I found a thesis from someone called Rittberg that was recent. I haven't had the energy to read it yet. You may want to look at it and invite him.

S: I will read it. Have you changed your view about pluralism as you wrote in the Festschrift? I recall the paragraph about the Bourbaki where you wrote that formalism was a kind of a front to hide behind when they were confronted with paradoxes.

H: . . .and Realism is what they really believed in. One would think they were schizophrenic. I talked about this as being a sort of duality. Accepting our mental models of objects as real and getting agreement with other mathematicians suggest a sort of sociocultural realism. This is not Platonic realism per se,

although one can argue for consistency of these two views or compatibility. Pluralism is not really a philosophical position. Tell me about the full handbook.

S: My goal is to split it up into sections starting with Origins. Believe it or not, I promised Springer about 100 chapters! I am getting mathematicians, historians, and philosophers involved, so it is supposed to be a handbook on the history and philosophy of mathematical practice. You did read the proposal I sent to Springer; there are sections on ontology, logic, pluralism, semiology, experimental mathematics, proof, and new perspectives in addition to sections on constructing the reals, reflections from mathematicians, views from other domains, and maybe probability.

H: It was very well written. I am glad you are focusing on practice; people still do not understand or know what mathematicians do. You have to bear in mind that philosophers typically theorize about mathematics; they are not really interested in the activity of doing mathematics or the mathematics, although some have started to theorize about mathematical practices. The former results in models which as we know people argue about back and forth – take a position, the latter in more nuanced – mathematical practice is about the activity of doing mathematics. Maybe this will renew an interest in the messy part of doing mathematics. I am unsure whether philosophers will delve into this.

S: There might be a schism in what historians will write as opposed to what philosophers will write. I am not too worried. Creating the links will be for the astute reader. I also asked Yuri and Chandler for help with finding interested authors. Do you have any suggestions in terms of significant shifts in practice? I am thinking of breaking it down into historical periods, but the danger in doing this is that it would seem the development of practice as being linear and not indicating jumps when something significant comes along like a change in notation.

H: Such as?

S: If you consider the Arithmetica in its original notational form which was rhetoric, syncopated, and symbolic with use of the Greek alphabet and special symbols, and see the modern treatment of arithmetic of algebraic curves, one sees a significant shift in the practice of Diophantine arithmetic. Now the study of elliptical curves is a specialty on to itself. I am also hoping to clear up the misappropriation of both Plato and Pythagoras. I don't think mathematics had the same status as philosophy for Plato and like anything else people write all sorts of thing without reading the original texts. I know people from Stelios' circle – real Greeks who might step up to the plate.

H: Negrepontis? Good – the more diverse the people you find in terms of expertise, the better the book will be. Did you know there was a beautiful book that was published in the Dolciani series by Isabella Bashmakova that goes along what you were saying about practices in Diophantine arithmetic. I don't often use the adjective "beautiful," but in this case, it applies. Have you thought of writing a chapter on Descartes, or perhaps Leibniz?

S: I already have an author writing a chapter on Leibniz' binary arithmetic. My main idea is to indicate shifts in practice by tracing it forward to the period of

natural philosophy and perhaps even further. I mean a lot of so-called mathematical practices were inspired by problems in astronomy and later physics. There are also other ways to unfold practice as in how Euclidean geometry inspired a lot of new practices involving proportionality such as the logarithm. Napier's original axioms built off of Euclid in the notion of progressions or comparing arithmetic and geometric progressions with notation and arithmetic that is completely different from what modern books suggest. Or even the development of algebra, although this has been done a lot in books that give a historical approach to the study of abstract algebra.

H: You may want a separate section just on geometry? There are significant shifts in practice in its history. There are plenty of books on the topic. How will you develop this topic?

S: Yes, I am teasing out the practices in Euclidean geometry in a different section focused on practices in Euclidean/non-Euclidean geometry. My intention is to get someone who knows the work of Menelaus to write about practices in spherical geometry that predated the fashionable Gauss/Lobachevsky/Bolyai triage, which overshadows it. It also offers a chance for people to learn about the Arab contributions to mathematical practices, since they had translated Menelaus' Spherica.

H: If you are going the natural philosophy route, the choices are limitless in terms of cherry picking which or what practices changed. You should make it clear that these were your choices. You know I probably won't be around to see the book in print.

S: Please don't say that. Regardless, the book will be dedicated to you posthumously. However, if you believed in Gilgul like your distant ancestors, you might see it in a different form. . .

2 Coda

Reuben Hersh passed away on January 3, 2020. We planned on writing this introductory chapter together once all the chapters for the Origins section had materialized. In the 3 years since his passing, the Origins section developed into the *largest* section of the handbook with mathematical practices described over two millennia, from the Pythagoreans onto the work of René Thom in the twentieth century.

Cellucci's chapter on the nature of mathematical objects opens the section on Origins. It considers mathematical objects as

> "independently existing entities, or abstractions, or logical objects, or simplifications, or mental constructions, or structures, or fictions, or idealizations of sensible things, or idealizations of operations."

Persson's chapter examines misconceptions in the use of Platonism. In particular, it takes "exception to the post-Plato development of Plato's Platonism, the so-called Neo-platonism, which is an exalted version of Plato's thought but shorn off its irony." Persson writes,

> ". . .it constitutes a literal reading of Plato's metaphor positing a separate abode for forms and interpreting the immortality of the "soul" in personal terms. Those religious ideas are but a distraction irrelevant to the issue of mathematical Platonism, but unfortunately providing easy fodder for contemporary attacks and dismissals of it."

While Cellucci argues for mathematical objects as hypotheses introduced to solve mathematical problems by the analytic method, Persson uses Physics as an example of a human activity and compares mathematical practice in many ways analogous as a collective feat of human imagination. In both of these chapters, mathematical objects are not conferred any special ontological status.

Plotnitsky's chapter clears up the misappropriation of the terms "Pythagorean" and "Platonism" by explaining clearly what Pythagorean thinking entailed in terms of the logos and the alogon. By delving into original sources to distinguish between Pythagorean thinking and Platonist thinking, the intended meaning of the status of mathematics for Plato becomes clear. Plotnisky writes:

> While Plato conceived of the ultimate architecture of the ideal realm on the *mathematical* model, he distinguished the philosophical thinking he aimed at from *mathematics*, at least the mathematics as practiced by mathematicians of his time. Plato thought that only philosophy, and specifically dialectic, could achieve, for humans, the maximally close proximity to the ultimate architecture of (ideal) reality. He expressed doubts that mathematics as a technical discipline, or at least mathematicians could do so, given their (Pythagorean) attitude, too tolerant of the presence of the alogon with the logos of mathematics, however, circumvented this presence may be.

The different chapters of the Origins section examine in depth:

(a) The spirit of the Pythagoreans that lingers in algebra and geometry
(b) Different meanings and the metaphysical aspects of Platonism
(c) How definitions, concepts, and heuristics develop in mathematical practice
(d) The influence of natural philosophy in the contributions of Descartes, Fermat, Leibniz, and Euler; and last but not least
(e) Evolving notions of measurement, abstraction, and consistency in mathematics

Taken as a whole, the Origins section provides a substantial glimpse into the changing nature of mathematical practices over the course of 2500 years. It culminates with a chapter on René Thom, in which one learns of his connection with the work of Aristotle and to those of Leibniz, Riemann, and others.

Yuri Ivanovich Manin (1937–2023): Mathematics, Philosophy, and Poetry

Athanase Papadopoulos

Contents

1 Introduction 14
2 Tarkovsky 14
3 Mathematics and Language 18
4 Analogies 1: Cognitive Development, Internet, and Human Condition 19
5 Brain, Computers, and the Internet 21
6 History 21
7 Analogies 2: Man and Cosmos 23
8 Proof and Truth, Mathematics and Tarkovsky Again 24
9 Painting, the Brain, and Reverse Perspective 26
10 A Poem 28
11 Platonism 29
12 In Guise of a Conclusion 31
References 32

Abstract

This is a personal view of and commentary on Yuri Manin's relationship with mathematics, philosophy, literature, poetry, art, and culture. His life epitomized those of the great natural philosophers of yore.

Keywords

Yuri Manin · Biography · Linguistics · Language · Philosophy of mathematics · Proof · History of computer science · Evolution of mathematical ideas · Andrei Tarkovsky · Cosmism · Mathematics and poetry

AMS Codes

01A70 · 91F20 · 97M80 · 68-03

A. Papadopoulos (✉)
Université de Strasbourg and CNRS, Strasbourg, France
e-mail: athanase.papadopoulos@math.unistra.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_143

1 Introduction

Yuri Manin passed away a few months ago. Such a loss, even if in principle is to be expected, always comes as a shock, and in Manin's case, it is a collective shock. He was an important editorial advisor of the present Handbook. Bharath Sriraman, the chief editor, asked me to write an article on Manin, and at the same time, he told me that it was Manin who had recommended me as coeditor. I am happy to render him this small tribute. It goes without saying that I would have preferred not to have to write it, and that, instead, Manin would have been able to see this Handbook.

I will not make the praise of Manin as a mathematician, because, first, it is well known that he was a first-order mathematician whose works on algebraic geometry, number theory, mathematical physics, complexity theory, computing, and probably other fields have had a crucial impact on the development of all these domains, and secondly, and most of all, it is difficult for me not to follow Jules Tannery's aphorism: "Il est des hommes auxquels on ne doit pas adresser d'éloges, si l'on ne suppose pas qu'ils ont le goût assez peu délicat pour goûter les louanges qui viennent d'en bas." (There are men whom we should not praise, unless we assume they have the rather indelicate taste for praises that come from below.) (Tannery 1912, p. 102). Instead, I would like to talk about some aspects of Manin's thinking and life that are less known to those who have only heard of his purely mathematical work.

Manin was thoroughly involved, besides mathematics, in philosophy, linguistics, psychology, sociology, history, literature, and poetry. In fact, mathematics, for him, was part of the vast field we call "culture," in which his thought was evolving in several directions at the same time. Even in his mathematical writings, he often addresses social, cultural, and philosophical issues. This chapter will be a quick review of some of Manin's relationships with all these domains, relationships that he lived on a daily basis. The word mathematics will only appear because it is part of the expression "philosophy of mathematics."

2 Tarkovsky

I will begin with a personal recollection: a conversation I had with Manin on July 21, 1994.[1] I was visiting the Max-Planck-Institut in Bonn for a few months. Manin had already joined this institute as a permanent professor,[2] and it was the first time I talked with him. A few days before, I was in Paris, and I had purchased there a book by Andrei Tarkovsky, his *Diary* of the years 1970–1986, in French translation, published by the *Cahiers du Cinéma* (Tarkovsky 1994). On that day of July 21, I had just finished reading it. Tarkovsky recounts in his diary his everyday life in

[1] I can tell the exact date because I keep a diary.

[2] In fact, Manin was hired as one of the directors. The Max Plank Institutes have no "permanent professors" but only directors. (I thank J.-P. Bourguignon for this clarification.)

Moscow, his troubles with censorship,[3] his exile in Italy, his illness, and the last days of his life, in Paris. One may also read there, among many other things, Tarkovsky's reflections on art and literature. Figures like Bergman, Bresson, Dostoevsky, Bach, Pushkin, Shakespeare, Solzhenitsyn, Rostropovich, etc. are present throughout the book. In the first pages, Tarkovsky mentions Manin.

On that day, at the 4 o'clock tea of the MPI, I met Manin. I asked him: "Do you know that Tarkovsky talks about you in his book?" He responded with a question: "Which Tarkovsky?" For a fraction of a second, I was puzzled by his question, but then I realized that Manin, like all educated Russians, knows that there are two Tarkovskys: Andrei, the filmmaker, and his father, the poet Arseni Tarkovsky, whose poems resonate throughout the film *The mirror*. I told him "the filmmaker," with a hand gesture, that of someone who is shooting a film. He was not aware of this. I had the book in my office. I fetched it and I showed him the spot. He was visibly moved. He leafed through the book, photocopied a few pages and asked me to buy one copy for him when I go back to France. I told him I would.

On the next day, I saw Manin anew, at tea time, and we talked again about Andrei Tarkovsky. He told me that he had met him twice: the first time, at Moscow University, where the latter had come to present his film *Andrei Rublev*, and the second time, at the home of Otar Iosseliani, who (I learned this from Manin) had studied mathematics before starting his career as a filmmaker. We also talked about Krzysztof Zanussi, who (I also learned this from Manin) prior to embarking on a career in cinema, studied physics. I told him that Tarkovsky mentions Zanussi in his memoirs, toward the end, when the latter lent him his apartment in Paris.

Here is what Tarkovsky writes in his diary on September 1, 1970:[4]

> Going through old papers I came across the transcript of a university debate on Rublev. God, what a level. Abysmal, pathetic. But there is one remarkable contribution by a maths professor called Manin, Lenin Prize winner,[5] who can hardly be more than thirty. I share his views. Not that one should say that about oneself. But it's exactly what I felt when I was making Andrei. And I'm grateful to Manin for that.

Here are now, according to the same diary, an excerpt of what Manin said on that occasion:

> Almost every speaker has asked why they have to be made to suffer all through the three hours of the film. I'll try to reply to that question. It is because the twentieth century has seen the rise of a kind of emotional inflation. When we read in a newspaper that two million

[3] From Tarkovsky's diary, one gets the impression that these troubles arose not because of an opposition to the authorities, but because of the spiritual themes that are present in his movies.

[4] For all these quotes from Tarkovsky's book, I am using the English translation which appeared in 1994 (Tarkovsky 1994).

[5] [My note] The Lenin prize was one of the most prestigious prizes (maybe the most prestigious one) of the Soviet Union. Manin received it in 1967. Among the other recipients, let me mention Andrei Sakharov, Sergei Prokofiev, Dmitry Shostakovich, Sviatoslav Richter, and, among the mathematicians, Vladimir Arnol'd and Andrei Kolmogorov.

> people have been butchered in Indonesia, it makes as much impression on us as an account of our hockey team winning a match. The same degree of impression! We fail to notice the monstrous discrepancy between these two events. The channels of our perception have been smoothed out to the point where we are no longer aware. However, I don't want to preach about this. It may be that without it life would be impossible. Only the point is that there are some artists who do make us feel the true measure of things. It is a burden which they carry throughout their lives, and we must be thankful to them.

Tarkovsky comments: "It was worth sitting through two hours of rubbish for the sake of that last sentence."

A few days later, I talked again with Manin about Tarkovsky's memoirs, and this time I wanted to ask him about Djouna, an astrologer-medium-paratherapist who practiced in Moscow, of whom Tarkovsky speaks all throughout his diary. She used to treat cancer and all other kinds of diseases by paramagnetism. Tarkovsky, who speaks of her as "an uneducated and capricious woman, a mountain woman," was being treated by her, and he used to advise his friends to do the same. At some point, she became ill, and Tarkovsky cared for her, using her own methods. Incidentally, when Tarkovsky arrived in Italy, he met another healer – "of Djouna's type," he writes – who predicted his future.

I was puzzled by all this and I asked Manin if he had heard of Djouna. He responded: "Of course! everyone in Moscow has heard of her." I told him that I was surprised by the fact that Tarkovsky was so close to her, and he said: "If you had known Moscow in the 70s and 80s, you would not be surprised." At that moment, I thought of Mikhail Bulgakov's *Master and Margarita*, with its supernatural, nonconformist, spiritistic side, where you meet witches, black magic, etc. All the Russian mathematicians I know have read it.[6] In fact, I understood later that I was wrong, that paramagnetism and hypnosis have nothing to do with magic, that they are part of science, of medicine.[7]

Manin talked, at several occasions, about Moscow in the 1960s–1980s, which hosted a cultural community consisting of small or larger circles, including mathematicians, artists and poets, a community one could find in Moscow and Leningrad and probably nowhere else in the world in terms of vigor and enthusiasm. In the speech he made at his reception ceremony as a foreign associate at the Paris Academy of Sciences (Manin 2006),[8] he recalls:

> Moscow of the 1960s–1980s, where I spent the most creative time of my life, was also boiling with humanities studies and quests for meaning of human culture, history, and psychology. I was happily participating in seminars and conferences organized by my

[6] I learned later on that Bulgakov was a medical doctor, just as Chekhov was. Incidentally, Bulgakov wrote a book titled *A country doctor's notebook*, in a very classical style (unlike *The master and Margarita*), which he thought of as a parallel to Chekhov's medical tales.

[7] I also know that the oncologist Léon Schwartzenberg, who was treating Manin and who, together with his partner Marina Vlady had become close to him during the last moments of his life, declared once that he could not do his treatment on Tarkovsky in an efficient manner because the latter refused to take the classical drugs.

[8] The passage is also included in the introduction of (Manin 2007).

> friends linguists and philologists, delivered talks and published amateurish papers on subjects ranging from the trickster figure in mythology to the epistemology of Lévi-Strauss.

Manin talks here about the period before he left to the West. From what I know, this form of culture is still bubbling today, in Moscow and Saint Petersburg, in the form he describes. As a matter of fact, Manin was among the Russian mathematicians living in the West who returned regularly to their home country.

Several years later, I thought about my conversation with him about Tarkovsky and magnetism, especially after I watched Tarkovsky's *Mirror*, a film that opens with a scene where a young stuttering child recovers his speech thanks to a therapy that mixes magnetism and hypnosis. I thought about Manin because I knew that language and linguistics were among his favorite research subjects all throughout his life, and that in his theory, hypnosis plays a fundamental role. For example, in his essay *The early development of speech and consciousness (Phylogeny)*, reprinted in (Manin 2007), he discusses works of several scientists on hypnosis and he describes a number of hypnosis experiments in which the researcher makes, as in Tarkovsky's *Mirror*, verbal suggestions and studies their effects on the patient's behavior. In the same essay, Manin reports on instances of auditory hallucinations, discussing the part played by consciousness in the behavior of an individual. He recalls that "hypnosis became possible only after development of speech" (p. 176 of the English version of (Manin 2007)). In the same essay, he writes:

> For many years I led a seminar in Moscow on psycholinguistics and evolution of mind and consciousness. Among its participants and contributors were linguists, ethnologists, neurobiologists, psychologists and psychiatrists. All of us had different backgrounds and different interests, and we tried to find common viewpoints and problems that might possibly be clarified by throwing together our diverging experiences.

I will say more about Manin's ideas on language and linguistics below.

Almost three decades after my first conversation with him, we talked again about Tarkovsky. It was during covid period. I am extracting from an email Yuri sent me on March 20, 2020, with the subject "Tarkovsky":

> [. . .] We have also stopped listening or reading news, although Xenia once a day browses some *Zeitung* in the Web. I am trying to be quiet and keep my feelings. My ultimate joke about all of it: after all, we have read Gogol, Kafka, etc. [. . .] Let's hope that this crazy time will soon pass. Yesterday my son from Palo Alto sent me an email with subject "Let us watch Tarkovsky", with the address [. . .]

Manin sent me the address of a website where one can watch free of charge all of Tarkovsky's movies and which I am happy to use until now. I then reminded him of our 1994 conversation about Tarkovsky's diary. In his answer, he gave me a few more details about his intervention in the debate that took place at the university, after the film projection:

> It was the first ever public screening of *Rublev*; it was at the Moscow University Club (a former Orthodox Church, now again). I was then a young math professor at the university, and I was

invited to attend it. After the screening, several persons from the public (I do not know who, probably, a party committee required it from them) delivered very critical speeches. I was very angry, ran to the podium and defended Tarkovsky and his creativity. I have long forgotten exactly what I said, but Andrei kept a rendering of my words in his diary.

3 Mathematics and Language

In his Kyoto ICM talk (1990) (Manin 1990), Manin presents mathematics as "a specialized dialect of natural language, and its functioning as a special case of speech." Language is complex, a complexity which is that of man. In several works, Manin considers mathematics, like language, a *metaphor*. He explains the meaning he gives to this crucial word by referring to a sentence in the book by J. P. Carse, *Finite and infinite games* (Carse 2013), according to which a metaphor is the "joining of like to unlike such that one can never become the other [. . .] At its root all language has the character of metaphor, because no matter what it intends to be about it remains language, and remains absolutely unlike whatever it is about."[9] Language is also an essential component of the talk Manin gave at his reception to the Paris Academy of Sciences (Manin 2006). He declares there: "Language and its functioning, both in contemporary mathematics and mathematical logic, and at its very early pre-historical stages about which we can only speculate, permanently fascinated me."[10] From this talk, we learn that it is in his effort to understand the linguistic dimension of mathematics that Manin was led to an extension of the notion of Kolmogorov complexity, to the initiation of the project of quantum computation, and to his contribution to the theory of error-correcting codes.

Since I mentioned the work of Kolmogorov, and since my article opens with references to Tarkovsky's *Diary*, let me report on the following. By the beginning of 1972, Tarkovsky's third feature-length film, *Solaris*, was finished, but it had still to pass several censorship committees. Tarkovsky gives a list of the observations, questions, corrections, and recommendations that he received from various bodies: the cultural department of the Central Committee, the ideology department of the same committee, etc. For instance, recommendation Nb. 5 says: "Cut out the concept of God"; Nb. 7 says: "Cut out the concept of Christianity"; and Nb. 8 says: "The conference. Cut out the foreign executives." The motive for Gibarian's suicide (not in Lem) ought to be that he is sacrificing himself for his friends and colleagues. Tarkovsky found most of the recommendations absurd. He writes: "I simply cannot work out what they meant, or rather, what was in their minds, when they gave me those corrections." Nb. 21 says: "Put in some quotations from Kolmogorov (about the finite nature of man)." It probably shows that either there was a mathematician in

[9] The joining of like to unlike was at the basis of the philosophies of Heraclitus, Anaximander, and the Pythagoreans.

[10] The passage is also included in the Introduction to his book *Mathematics as metaphor* (Manin 2007).

one of these committees, or that Kolmogorov's theories went deep in the collective body of the society.

Let us return to language.

The question of the relationship between mathematics and language is also addressed in the interview *Good proofs are proofs that make us wiser* (Aign and Schmidt 1998) in which Manin, answering a question about the cultural role of mathematics, sketches a wide historical perspective on language, showing how natural language, used for the purposes of science and the scientific description of nature, became progressively specialized, ending up with the language of mathematics:

> In my opinion, the basis of all human culture is language, and mathematics is a special kind of linguistic activity. Natural language is an extremely flexible tool of communicating essentials required for survival, of expressing one's emotions and enforcing one's will, of creating virtual worlds of poetry and religion, and of seduction and conviction. However, natural language is not very well fit for acquiring, organizing, and keeping our growing understanding of nature, which is the most characteristic trait of the modern civilization. Aristotle was arguably the last great mind that stretched this capability of language to its limits. With the advent of Galileo, Kepler, and Newton, the natural language in sciences was relegated to the role of a high-level mediator between the actual scientific knowledge encoded in astronomical tables, chemical formulas, equations of quantum field theory, databases of human genome on the one hand, and our brains on the other hand. Using the natural language in studying and teaching sciences, we bring with it our values and prejudices, poetical imagery, passion for power, and Trickster's skills, but nothing really essential for the content of the scientific discourse. Everything that is essential is carried out either by long lists of more or less well-structured data, or by mathematics. For this reason, I believe that mathematics is one of the most remarkable achievements of culture, and my lifelong preoccupation with mathematics in the capacity of researcher and teacher still leaves me with awe and admiration by the end of every working day. However, I do not believe that I can convincingly defend this conviction in the context of contemporary public debate on science and human values.

Manin's approach to language and linguistics involves ideas on psycholinguistics, social interactions, self-consciousness, the theory of development of the central nervous system, animal behavior, the human psychopathology of language, and others. I would like to mention here the relationship he makes between, on the one hand, the history of language and, on the other hand, that of the known cultures, from the prehistoric man who painted on cave walls, passing through the Greeks of the time of Homer, mentioning the Winebago Indians, the biblical world, Sanskrit culture, etc. until modern man. This historical parallel, which is part of Manin's aim to understand the process of language, its structural properties, and its evolution, is also an example of the notable analogies that he envisioned, and of which I will say a few more words in what follows.

4 Analogies 1: Cognitive Development, Internet, and Human Condition

An important analogy Manin makes in his theory of language is the comparison between the linguistic development of the child and the history of speech and language. Among the questions he addresses is the following: Since a child can

only learn to speak to the extent that he or she is immersed in a language environment, that is, being in contact with adults who speak with him and with each other, who was then the "adult" for the first men? Among the other problems he points out is the one of "gifted individuals" in the face of language skills: (Manin 2007, p. 171): "[Why,] despite the democratization of education, such national poets as Pushkin and Dante so far exceeded the average level of language abilities as to become the nation's language teachers for centuries? ".

Manin also talked about analogies between the history of languages and that of computers. In the article *Semantic spaces* (Manin and Marcolli 2016) which he wrote with Matilde Marcolli, the two authors discuss mathematical, and in particular geometric, constructions which help bridging, on the one hand, database, usually associated with computer processing and computer memory, and, on the other hand, natural language, which resides in the human brain and functions in human communication. Manin develops an analogous theory in an article titled *Cognitive Networks: Brains, Internet, and Civilizations* he wrote in collaboration with his son Dmitry (Manin and Manin 2017). In the same paper, the cognitive activity of the nervous system is discussed in analogy with the Internet: neurones and neural networks are compared with computers. The two Manins, with the aim of understanding the functioning of the cognitive activities of the brain, explore a comparison between certain neural networks and computers. The whole nervous system is compared with the Internet. Dissociative phenomena like multiple personalities, automatisms, fugues, and hypnosis are also mentioned in this "brain/Internet metaphor." The Saussurean trichotomy langue/langage/parole is involved in this study.[11]

Another analogy of the same kind, between something a priori special and something universal, appears right in the Introduction to his book *Mathematics as metaphor*, and it is about mathematics, presented as "a specimen and a metaphor of the human condition." Here, Manin makes the following comment: "This attitude is not as esoteric as it might sound, if only because through mathematics, as through poetry and philosophy, dozens of generations communicate to each other, often over the heads of their contemporaries, because in each generation the community of mathematicians is small and scattered all over the world."

I will say more on such analogies after a few remarks on computers and on Manin's ideas on history.

[11] The famous linguist Ferdinand de Saussure (1857–1913), considered as the founder of modern linguistics, in his *Cours de linguistique générale*, makes a distinction between the following three notions: *langue*, the result of a social convention transmitted by society to man and for which the latter does not play any important role, *langage*, as the capacity that allows human beings to communicate and interact with each other, and *parole*, as the personal use of language. Note that English makes no difference between "langue" and "langage."

5 Brain, Computers, and the Internet

Manin, in several works, mentions the brain/Internet metaphor regarding information transmission, and he introduces Kolmogorov complexity in this discussion. In his lecture *Kolmogorov complexity as a hidden factor of scientific discourse* (Manin 2012a), he explains the fact that the characteristic features of scientific theories, from epicycles to the standard model of particles, split into two distinct parts: the first one with relatively small Kolmogorov complexity (laws, basic equations, periodic tables, natural selection, etc.) and the second one with indefinitely large Kolmogorov complexity (examples include the initial and boundary conditions, phenotypes, population, etc.). More generally, in his works on this topic, he divided scientific theories into two parts, "Kolmogorov simple" and "Kolmogorov indefinitely complicated."

Manin has always highlighted the fact that computers have a non-negligible impact on pure mathematics, in that they allow us to do experiments. In the interview (Gelfand 2009), he declares that they gave to the ordinary mathematician a unique possibility of "doing large-scale experiments in mental reality." He says that thanks to computers, we are capable of doing things that only mathematicians of the stature of Euler and Gauss were capable of doing without computers. He declares: "By now, what Euler and Gauss could do, any mathematician can do, sitting at his desk. So if he doesn't have the imagination to distinguish some features of this Platonic reality, he can experiment."

In his paper *The early development of speech and consciousness (Phylogeny)*, which we already quoted, Manin states a metaphor which expands the beautiful expression "the smoke of centuries"[12]: "The sharpness of historical vision drops quickly with distance, and the milestones of subjective past conveniently fall on a logarithmic scale." I will talk about some of Manin's ideas on history in the next section.

6 History

Manin's vision on history is exemplified in his article *Time and periodicity from Ptolemy to Shrödinger: Paradigm shifts* vs. *continuity in history of mathematics* (Manin 2019), in which he discusses Kuhn's notion of *paradigm shift* applied to the history of mathematics. His argument involves examples from the ancient laws of planetary motion until those of elementary particle physics. He starts by quoting Thomas Kuhn who, in his influential book on philosophy of science, *The structure of scientific revolutions* (Kuhn 1962), defended the thesis that history of natural sciences can be divided into periods such that the passage between a period and

[12] Manin says that this expression is in Pushkin. Actually, I found it in Walter Scott's *Rob Roy* (1817), and also in a poem by Dovid Knut (1900–1955), a Russian-Jewish emigre of the first generation who published his poems in Paris in the 1920s.

the next one occurs through a revolution, that is, a radical change in the basic assumptions. Manin maintains that Kuhn's theory does not apply to pure mathematics, a domain, he says, which is different from the other sciences. He recalls that mathematics "has a position mediating, or bridging, daily life, common sense, philosophy, and physics." He asserts that only one part of mathematics is sensitive to the so-called revolutions or paradigm shifts, namely the part that is subject to the "reality tests,"[13] whereas the rest of mathematics, which we call "pure mathematics," experiences only a continuous development. He considers that the intergenerational continuity of that field has been neglected by historians. Placing himself from a point of view which is above the vision field of most professional historians of mathematics and science general, he focuses on the example of the mathematical theory of time and of periodicity. To explain what he means, he starts with Ptolemy's dynamical model of the solar system representing the visible motions, a system which involves complex combinations of circles, with epicycles and secondary epicycles. He then continues with the history of the numbers π and e, the appearance of Fourier sums and Fourier integrals, all the way to the attempts to explain the randomness that appears at the electron level by methods of probability, leading directly to the works of Heisenberg, Pauli, and Schrödinger on quantum mechanics, to Feynman integrals, and to the contemporary cosmological models. He declares that at the end, quantum interaction in quantum mechanics is essentially described by Fourier sums, series, and integrals in complex spaces equipped with Hermitian metrics replacing real Euclidean spaces with real metrics; thus, we are back to Euclid and Ptolemy,[14] and Manin's conclusion is that, in some sense, "quantum mechanics is a complexification of Ptolemy's epicycles."

Regarding continuity versus revolution in the history of mathematics, let me also mention a passage from an interview with M. Gelfand which carries the significant title *We do not choose mathematics as our profession, it chooses us* (Gelfand 2009). The passage comes as a response to the question "What will happen [in mathematics] in the next twenty years?". Manin declares:

> I don't foresee any revolutionary changes, because in my view there have been none in the last 300 years. Every time new and powerful intuitions arose, mathematics retained its character, in some strange way. This is also a theme of a lecture, one I've not given. I would like to show the development of the idea of the integers from the most remote times to Kolmogorov complexity, and all this can be done almost without appealing to new mathematics. One and the same idea persists. It changes a bit in one era or another, its verbal casing changes. But all the same it stays completely invariant and so lives on. Nothing is forgotten. And so I don't foresee anything extraordinary.

[13] Manin puts this expression in quotation marks, and it is not clear whether he intends the literal or the Freudian meaning of this expression; both may fit here (although, from what I know, Manin was not a strong supporter of Freud, maybe because of Freud's materialistic side).

[14] The first terms $\sum_{n=1}^{N} a_n e^{inx}$ of the Fourier series is a trigonometric polynomial which can be put in parallel with Ptolemy's superposition of periodic motions.

7 Analogies 2: Man and Cosmos

A brilliant element of Manin's argument in the paper *Time and periodicity from Ptolemy to Shrödinger: Paradigm shifts* vs. *continuity in history of mathematics* I just discussed involves a comparison between the evolution of knowledge on the scale of civilizations and on that of the development of cognition in the individual brain of a growing human being. This is part of the general pattern he expanded in several works, some of which I alluded to in Sect. 4 making parallels between historical developments, on the one hand, at the civilizational level, and, on the other hand, at the human level.

All these analogies remind me of a general metaphor relating the "cosmos" (or "macrocosmos") and the "microcosmos," which goes back to Ancient Greece (Anaximander, Plato, the Stoics, etc.), which was taken up by early Christian writers, according to which man is presented as a "microcosm" recapitulating in him the whole of "macrocosm."[15] Manin strongly felt this relation, as an analogy between the microcosmos and the macrocosmos. I think that this is an aspect of his thought which has never been emphasized, if ever it was noted. This reminds me of cosmism, a philosophical, scientific, literary, spiritual, and intellectual movement which was born in Russia in the nineteenth century and which attained its apogee at the beginning of the twentieth century. This movement claims an alliance, solidarity, and harmony between man and the whole creation.[16] Manin's ideas have a taste of cosmism even if they stay at the level of *metaphor*, or *theory*.[17] And if one does not like the term cosmism in Manin's work, one can talk at least about cosmology, as in the paper *Time and Periodicity from Ptolemy to Schrödinger* (Manin 2019) which I already mentioned. I also would like to mention here his talk at the 2014 Plenary Session of the Pontifical Academy of Sciences, titled *Big Bang, Blow Up, and Modular Curves: Algebraic Geometry in Cosmology* (Manin 2014),[18] in which he starts by quoting T. S. Eliot, "In my beginning is my end," and where he sets his goal as follows: "We are trying to explain how it could be that in my end (that is, in the universe's end) is my (that is, the universe's) beginning." Again, the analogy – here an identification – between man and the universe. Let me also quote another Russian

[15] To quote again Russian authors, I mention Sergei Bulgakov, who writes that man is a "microcosm, and his imprint therefore lies upon the entire world, the macrocosm," and that "by his body man belongs to the creaturely world, which, he, as microcosm, has in himself and unifies and generalizes," (Bulgakov 2008, pp. 173, 279), etc.

[16] The founder of cosmism is considered to be Nikolaï Fiodorov (1829–1903). The movement was developed especially in the first three decades of the twentieth century, before it was squashed brutally by Stalin. This movement inspired writers, philosophers, composers, theater directors, and artists, like Leo Tolstoi, Alexander Scriabin, Vassily Kandinsky, Casimir Malevic, Vladimir Soloviev, and Sergei Bulgakov.

[17] The word *theory* comes from the Greek word "theōria," which means "contemplation."

[18] The article is a joint work with Matilde Marcolli.

author, on the same topic, Pavel Florensky,[19] from his booklet *Hamlet* (Florensky 2006, pp. 24–25), talking about "development" as a result of either universal and supraindividual human reason, or of the individual human reason, which he calls "microcosm":

> [. . .] Motion and development come either from the universal, supra-individual human reason, or from the common, individual human reason. The latter is more easily observable given its reduced dimensions, and its more or less conscious construction, it is a microcosm. Nevertheless, the microcosm corresponds to the macrocosm, in this case the natural, spontaneous and irresistible dialectics of the universal human consciousness; there is a reciprocal correspondence between phylogeny and ontology, between the philosophical structure of the individual and the historical destinies of the world consciousness.

To return to Tarkovsky's universe, I would like to mention this scene from the *Sacrifice* in which the character, in the middle of a nuclear accident, knowing that he is ready to perish, instead of making, as one would expect, an evaluation of his life, makes an evaluation of existence in general. He strives for salvation in many forms, frightened by an eschatological feeling. He seeks help, but finds none, and is finally seized by this destructive eschatological feeling, not only for himself after his death, but also for the universe after its disappearance. I can also quote Andrei Tarkovsky again, from his book *Captured time* (Tarkovsky 2004): "There must be other possibilities of working with sound, which would allow us to be more precise, more faithful to the inner world that we try to recreate on the screen, and not only to the inner world of the artist but to what is inside the world itself, to what is essential to it and independent of us." Again a relation between the inner world of man and the world itself. Finally, I can quote a line from Arseni Tarkovsky's poem *Life, life* (1965): ". . . I tailored the age to fit me . . .."

In think that this feeling of the cosmos and of a strong bond between man and the universe, exemplified in Manin's writings, is a common feature of all Russian thought.

8 Proof and Truth, Mathematics and Tarkovsky Again

This section consists essentially of five quotes: two by Manin and three by Tarkovsky.

All of us have read (or heard of) Bourbaki's famous sentence, from the introduction to *Set Theory*, the opening treatise of the *Elements*: "Depuis les Grecs, qui dit

[19] The mathematician, theologian, physicist, philosopher, and historian of art Pavel Florensky (1882–1937) was dubbed the "Russian Leonardo" and the "Pascal of the Gulag." Solzhenitsyn said about him that he was "perhaps the most remarkable person devoured by the Gulag." The translation of his books and articles in French, English, Romanian, and other languages have taken an accelerated pace in the last two decades. I wrote an article on him for the present Handbook (Papadopoulos 2023).

mathématiques dit démonstration."[20] Manin belongs to those who go against the grain. In his Kyoto ICM talk, *Mathematics as metaphor* (Manin 1990), he writes:

> [. . .] What is relevant is the imbalance between various basic values which is produced by the emphasis on proof. Proof itself is a deviate of the notion of 'truth'. There are a lot of values beside truth, among them 'activities', 'beauty' and 'understanding', which are essential in high school teaching and later. Neglecting precisely these values, a teacher (or a university professor) tragically fails.

Let me complement this discussion about "truth" by some words of Tarkovsky, whose name I associate with that of Manin for reasons I explained at the beginning of my article. This is a quote, again from Tarkovsky's diary, in which he talks at the same time of truth, the cosmos, Euclid, and the infinite, topics which were all dear to Manin. He writes, on September 5, 1970:

> [. . .] Since [truth] is human, it is limited, rigidly confined, in human terms, within the framework of the human milieu. There is no conceivable link between what is human and the cosmos. And the same applies to truth. To attain greatness within our own limits—which are Euclidean and insignificant in juxtaposition with infinite—is to illustrate that we are merely human."

In the next section, I will say a few words about Manin as an artist, and in a later section I will quote one of his poems. Let me, before that, quote a passage, written in Manin's poetical style, about proof, as a conclusion of his paper *Foundations as superstructure* (Manin 2012b):

> One metaphor of proof is a route, which might be a desert track boring and unimpressive until one finally reaches the oasis of one's destination, or a foot path in green hills, exciting and energizing, opening great vistas of unexplored lands and seductive offshoots, leading far away even after the initial destination point has been reached.

Finally, let me conclude this section with two quotes on mathematics from Tarkovsky's *Diary.* On March 24, 1982, he writes:

> What is expressed in mathematics is not so much objective laws of the world as laws of the human psyche, laws of logic, of the human mind. It's really a kind of mental game, and yet the exact sciences, as they are called, command extraordinary respect—physics, astronomy, and so on. It's very odd, in fact I would say an amazing, paradoxical misconception.

On April 2 of the same year, he writes:

> Our reason cannot reach the other parameters or other dimensions which exist, and so it fantasizes about them in the form of mathematics and physics. Science involves the discovery not so much of objective laws of nature, as of the laws which govern the

[20] Since the Greeks, talking about mathematics means proving.

> functioning of our consciousness. A kind of music; image; symbol; sign. A mathematical symbol of truth, corresponding to our capacity for knowing it with our brains.

I am always surprised to see how much the filmmaker understood what mathematics is about, even though he never studied mathematics.

9 Painting, the Brain, and Reverse Perspective

I would like to mention an episode on painting recounted in the French version of Manin's book, published in 2021 (Manin 2007, pp. 426–427).[21] He talks there about a man whose full name he does not give, referring to him only by his initials, YM. In fact, before I read this text of Manin, I had already heard this story from him, as something that happened to him personally. In the book, the episode is part of a discussion regarding the asymmetry of the human brain, the tensions between its two hemispheres, and the main features of the division of its cognitive functions. The story is the following.

In 1973, Manin, 36 years old, who was already a professor at the Lomonosov State University in Moscow, was suddenly affected by a brain disease called *arachnoïdite*, an inflammation of the arachnoid – one of three membranes that surround and protect the spinal cord. This disease causes neurological disorders, and sometimes it attacks the back and the limbs. In Manin's case, the sickness was probably a result of a poorly treated flu. As a consequence, he underwent a treatment with antibiotics that lasted several months. He was not able to do mathematics anymore. One day, in the fourth month of his illness, while riding the subway home on the outskirts of Moscow, he felt a sudden urge to paint, although he had never painted before. He got off at the train stop, took the metro in the backwards direction, that is, toward the city center, and there he stopped at an artists' materials store. He bought brushes, tubes of oil paint, and blank canvases. Back home, he started to paint. For several months, painting became an important activity in his life, perhaps the only original activity. In the book (Manin 2007) (p. 427 of the French version), Manin lists a number of canvases he painted during that period.

In retrospect, after he had returned to normal activity, and after he had stopped drawing and painting, Manin examined his paintings, trying to understand the states through which he passed. Some of his drawings seemed strange to him, until he understood that they satisfy a theory developed by a specialist in functional cerebral asymmetry who works at the Institute of Biochemistry and Evolutionary Physiology of the Russian Academy of Sciences. This theory is reported on in an article on artistic activity and cerebral asymmetry (Nikolaenko 2003). The research is based on clinical observations of drawings made by patients with brain damage. The conclusion of this study is that someone who makes drawings and who is in the state in

[21] The episode is told in a chapter titled *Spontaneous Artistic Activity, Origin of logograms, and Mathematical Intuition*. This chapter is not contained in the English version of the book.

which Manin was during his illness – a state where the right hemisphere of his brain (or his "right brain") is predominant – tends to produce images that seem too close to the person looking at them, and this is why they are drawn in reverse perspective. In contrast, the left hemisphere, which (for the majority of human beings) is the verbal and logical hemisphere, tends to produce representations of space with the standard proportion scheme and with an ordinary perspective. The above story, which describes Manin's personal experience, is an illustration of this fact.

Manin explains at the beginning of his article that the asymmetry between the left and right hemispheres of the brain is not only a functional division, but the source of a dynamic competition between the two hemispheres which, when it occurs, leads to a back and forth motion around a certain balance of the brain. The point of his article is to say (again) that such an equilibrium can also be observed in the study of civilizations and their cultures, especially in periods of transition. He discusses the balance of the brain in the context of the origin of writing, its development and its transmission at the level of the cultural human heritage, and also in connection with the drawings that survive on the prehistoric caves in the South of France, representing animals and hunting scenes, compared to other forms of realistic representations, talking also about poetry, etc. At the same time, he addresses the case of history of mathematics, noting in particular the tension, in Euclid's *Elements*, between the use of the natural language to express the *logic* of the proofs, and that of the figures, which express the underlying *imagination*.

Returning to Manin's drawing experience, let me say that personally, I sometimes had to deal with colleagues who, while we were discussing in front of the blackboard, used to draw awkward pictures; for instance I noticed that some of them, when they draw a cube the perspective is always reversed, and I always wondered how that could be. I understood much later, when I read Manin's explanations on the brain and reverse perspective, that what I experienced was something that could be explained, and the drawing, here, is a tool that helps us exploring the neural mechanisms that are at the base of the aesthetic experience of the person who performed it.

There are other explanations. The manner we draw, in reverse or in Renaissance (call it "usual") perspective, is also a question of how we perceive the world, and I cannot help but mention here Byzantine iconography, of which Russian iconography is the full heir, whereby the drawing is usually done in reverse perspective. In this connection, I would like to quote again Pavel Florensky, from his essay *Reverse Perspective* (1920) (Florensky 2002, p. 219):

> The drawings of children, in their lack of perspective and especially their use of reverse perspective, vividly recall mediaeval drawings, despite the efforts of educators to instill in children the laws of linear perspective. It is only when they lose their spontaneous relationship to the world that children lose reverse perspective and submit to the schema with which they have been indoctrinated. This is how all children behave, independent of each other. This means that it is not mere chance, or a willful invention by one of them putting on Byzantine airs, but a representational method that derives from a characteristic perceptual synthesis of the world. Since the way children think is not weak thinking but a particular type of thinking which, moreover, is capable of unlimited degrees of perfection, including genius, and indeed is primarily akin to genius, it must be admitted that the use of reverse perspective

> to depict the world is also far from being an unsuccessful, ill-understood, imperfectly learned linear perspective, and is rather a distinctive grasp of the world that should be reckoned with as a mature and independent representational method.

After he recovered his original functioning of the brain, Manin stopped painting, and he came back to mathematics. He did not stop being an artist; mathematics was also, for him, an art. In his reception speech as foreign associate to the Paris Academy of Sciences (Manin 2006), he tells how he came back later to mathematical research, and in particular to his favorite field, algebraic geometry. About this field, he writes that he sees the set of solutions of systems of polynomial equations with several variables – an object that is at the heart of algebraic geometry – as "a geometric entity, a shape situated in a multidimensional space, stretching to infinity along some directions and capriciously closing upon itself at others. Variety and complexity of such forms is infinitely richer than anything one could glimpse at modern exhibitions of abstract art, but algebraic geometers managed to find patterns, discover connections and laws of this immense world."

Manin also talks about mathematics as an art while he gives his point of view on the heritage of Georg Cantor (another Russian-born mathematician). He writes, at the beginning of the paper (Manin 2004): "Georg Cantor's grand meta-narrative, Set Theory, created by him almost singlehandedly in the span of about fifteen years, resembles a piece of high art more than a scientific theory."

Let me also mention a passage in the interview (Gelfand 2009): "Once I was asked to deliver a lecture for the general public in an art museum, and I decided that the Banach–Tarski paradox is a great subject for the presentation *The Abstract Art of Mathematics*."

Let me mention, to close this subject, that on the dust cover of Manin's *Mathematics as metaphor*, published by the AMS, there is a picture of a house, with the inscription *Geomêtre* over the entrance door; the house is drawn in reverse perspective. I never asked him an explanation.

From visual art and abstract art, let us pass to poetry.

10 A Poem

Manin, in an article he wrote on Georg Cantor (Manin 2004), has, as an epigraph, the following citation from V. Tasić's (2001): "God is no geometer, rather an unpredictable poet."

Like God, Manin was a poet. In the French version of *Mathematics as metaphor* (Manin 2007, p. 536),[22] he wrote a poem related to his painting period. The title of the poem is *Je peins à l'huile*. I reproduce it here, in French:

[22] Manin was very happy of the fact that the French version of the book, which appeared in 2022, included the poems. The French translation of his book was done by Claire Vajou, but it is certain that the translation of his poems was done with his close collaboration, maybe by him alone (he was completely fluent in French).

La touche vient toucher la touche
comme la phalène — la page.
La touche vient toucher la touche
comme le canon glissant — la tempe.
La touche vient toucher la touche.
Le cerveau se languit sans langue.
L'hémisphère taiseux travaille
comme fait une avalanche —
le bruit arrive en retard . . .

It would be a redundancy to say that Manin was permanently reading poetry. We once talked about the Greek poet Constantine Cavafy, and he told me he has his works on his shelf, in several editions.

11 Platonism

We talked about mathematics as a language, but mathematics is also ideas, and Manin had his own idea about ideas. Let us hear again what he says, in his Reception speech at the Paris Academy of Sciences (Manin 2006):[23]

> [. . .] This Platonic feeling that even most abstruse mathematical ideas are somehow predestined to be in harmony with the physical world, always constituted for me one of the most irresistible attractions of our trade. Stéphane Mallarmé wanted to make us aware that poetry is made of words rather than ideas. To a certain degree, this is true about mathematics as well, but in a more profound sense, this is fundamentally wrong. (I suspect that this is wrong for poetry as well).

At several occasions, Manin expressed his "Platonic feeling" of preexisting mathematical notions. In the interview *Good proofs are proofs that make us wiser* (Aign and Schmidt 1998), he declares:

> I see the process of mathematical creation as a kind of recognizing a preexisting pattern. When you study something—topology, probability, number theory, whatever—first you acquire a general vision of the vast territory, then you focus on a part of it. Later you try to recognize "what is there?" and "what has already been seen by other people?". So you can read other papers and finally start discerning something nobody has seen before you.

In his interview with Gelfand (2009), he expresses the same idea, slightly differently:

> I must explain to you how I imagine mathematics. I am an emotional Platonist (not a rational one: there are no rational arguments in favor of Platonism). Somehow or other, for me mathematical research is a discovery, not an invention. I imagine for myself a great castle, or something like that, and you gradually start seeing its contours through the deep mist, and begin to investigate something. How you formulate what it is you've seen depends on your type of thinking and on the scale of what you have seen, and on the social circumstances around you, and so on.

[23] The passage is also reproduced in the Introduction of *Mathematics as metaphor* (Manin 2007).

Staying with Plato, let me mention Manin's book with I. Yu. Kobzarev, *Elementary Particles: Mathematics, Physics and Philosophy* (Kobzarev and Manin 1989), a philosophical dialogue, formally written in the style of Plato's dialogues. There are three characters: a theoretical physicist, an experimental physicist, and a philosopher, and the aim of the dialogue is to present the development of elementary particle physics, including its present state. However, unlike in Plato's dialogues, there is no "wise man" involved in this dialogue; such a figure is replaced by that of a scientist. Let me take this opportunity to quote Manin again, from *Mathematics as metaphor* (p. 197):

> The central figure of a philosophic dialogue is a wise man, whereas modernity generally and systematically replaces wisdom by training. Wisdom seems to be an inborn faculty slowly ripened by life experience; as such it is rarely met and even more rarely put to any use. Training is a democratic surrogate for wisdom which, in spite of all of its (mainly aesthetic) drawbacks, is superior in one respect: it produces professionals.

Let me also mention, since we are talking about Platonism, that the adjective "Platonic," which is systematically used in such a context (the mathematician discovering preexistent objects) indeed is part of Plato's teaching.[24] The term "Platonic" is nevertheless abusive, because the fact referred to is not sufficient to characterize Plato's view on mathematical objects.[25] In fact, Plato was absolutely against the idea that mathematical objects may be used to describe the real world: They can only describe immutable things, whereas the real world is permanently subject to movement and change. In contrast, Manin has always been interested in mathematics as a means of describing reality. Even in his work on algebraic geometry or number theory, he was resolute in thinking in terms of applications to theoretical thinking. Besides that, he taught us that mathematics is not only the language of nature, but it is also an adequate language for talking about art, literature, and the deep human psyche. More lyrically, he writes that mathematics is a "novel about Nature and Humankind" and he adds that his wish is "to restore a certain balance between the technological and the humanitarian sides of mathematics" (Manin 2007, p. 28).

[24] The idea that abstract mathematical objects exist independently of us and our knowledge of them is one of the basic teachings of Plato, who also considered that knowledge (which, for him, is, to a large extent, mathematical), is simply recollection (*anamnesis*) of innate knowledge, about preexisting things. The standard reference for this theory of *anamnesis* is Plato's *Meno* 81e1–82b8. It is worth noting that in this passage, Plato, to illustrate his idea, brings in Socrates helping a slave boy recollecting a specific knowledge, which, in fact, is a mathematical knowledge: how to construct a square whose area is twice that of another given square. S. Negrepontis explained to me that the method Socrates uses for this recollection turns out to be anthyphaeretic division. Finally, let me recall that "Platonism," as a belief in preexisiting objects (outside our thought), is contrary to the intuitionistic trend in mathematics (mostly represented, in the modern period, by Brouwer), which holds that mathematical objects exist only when we can construct them.

[25] As Plotnitsky rightly points out in his essay *Continuity and discontinuity in, and between, mathematics and physics* (Plotnitsky 2023): "mathematical Platonism (which is a twentieth-century conception) should not be confused with Plato's own philosophy." In fact, the term "Platonic" is often used and misused. A Japanese friend told me that in Japan, this word is used exclusively (as a Japanese word) as an adjective for a love relation, to say that it is purely mental.

It has often been said (by mathematicians) that there are two kinds of mathematicians: those who build theories and those who solve problems. If this is true, then what we recalled in this article says that Manin belongs definitely to the first category. But I think that this distinction is rather secondary, and even irrelevant, in Manin's case. For him, mathematics was a language, and he was interested in language much more than in mathematics. According to Aristotle, language is what distinguishes man from the rest of the animals, as we have been reminded above. I would like to quote, closer to us, Emmanuel Levinas: "[. . .] l'Infini se passe dans le Dire [. . .]."[26]

I started this chapter with a personal recollection, and I will also finish it with a personal recollection.

12 In Guise of a Conclusion

While I was writing this chapter, I often regretted not having had more conversations with Manin about all of this. In one of our last contacts, an email dated September 14, 2022, answering an invitation I made for him and Xenia to visit us in Strasbourg, he wrote me that he was tired and that he did not travel anymore. I regret not having had the reflex to tell him: "In this case I will come to see you in Bonn."

Athanase Papadopoulos, Xenia Semenova, Yuri Manin and Norbert A'Campo. Strasbourg, September 13, 2019

[26] "[. . .] The Infinite occurs in the Saying [. . .]" [My bad translation], from his book *Autrement qu'être ou au-delà de l'essence* (Levinas 1990, p. 230).

Acknowledgments I would like to thank Arkady Plotnitsky and Susumu Tanabé who kindly made several interesting comments on a previous version of this chapter. This work is supported by the lnterdisciplinary Thematic lnstitute CREAA, as part of the ITl 2021–2028 program of the University of Strasbourg, CNRS and Inserm (ldEx Unistra ANR-10-IDEX-0002), and the French lnvestments for the Future Program.

References

Aign M, Schmidt VA (1998) Good proofs are proofs that make us wiser, Interview with Yuri I. Manin, The Berlin Intelligencer. Springer, Berlin, pp 16–19. Also published in [18], p. 207–213

Bulgakov S (2008) The Bride of the lamb, transl. B. Jakim. Eerdmans Publishing Company, Grand Rapids

Carse JP (2013) Finite and infinite games. Free Press, New York

Florensky PA (2002) Reverse perspective (1920). In: Beyond vision, essays on the perception of art, compiled and edited by Nicoletta Misler, transl. Wendy Salmond. Raktion Books, London

Florensky PA (2006) Hamlet, translated from the Russian by Evdokiya Sichov, Ed. Allia, Paris

Gelfand M (2009) We do not choose mathematics as our profession, it chooses us: interview with Yuri Manin. Notices AMS 56(10):1268–1274

Kobzarev IY, Manin YI (1989) Elementary particles. In: Mathematics, physics and philosophy. Ser: Fundamental theories of physics, vol 34. Kluwer, Dordrecht

Kuhn TS (1962) The structure of scientific revolutions. University of Chicago Press, Chicago, several later editions

Levinas E (1990) Autrement qu'être ou Au-delà de l'essence. Librairie générale française, Paris

Manin YI (1990) Mathematics as metaphor. In: Proceedings of the International Congress of Mathematicians, Kyoto, Japan, vol II, pp 1665–1671. Also published in [Manin 2007], p. 27–32

Manin YI (2004) Georg Cantor and his heritage. Proc Steklov Inst Math 246:195–203. Also in [Manin 2007], p. 45–54

Manin YI (2006) À la recherche du temps perdu (In search of lost time), Académie des Sciences, Réception des Associés étrangers élus en 2005, 12 décembre 2006

Manin YI (2007) Mathematics as metaphor. Selected essays of Yuri I. Manin. With foreword by Freeman J. Dyson. American Mathematical Society, Providence. French version: Les Mathématiques comme Métaphore (Essais choisis), Translated by Claire Vajou, Les Belles lettres, Paris, 2021

Manin YI (2012a) Kolmogorov complexity as a hidden factor of scientific discourse, lecture given at the Pontifical Academy of Sciences, Plenary Session on complexity and analogy in science: theoretical, methodological and epistemological aspects, Vatican City, 5–7 November 2012

Manin YI (2014) Big bang, blow up, and modular curves: algebraic geometry in cosmology, lecture given at the Pontifical Academy of Sciences, Plenary Session on evolving concepts of nature, Vatican City, 24–28 October 2014

Manin YI (2019) Time and periodicity from Ptolemy to Schrödinger: paradigm shifts vs. continuity in history of mathematics. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Cham, pp 129–138

Manin YI (2012b) Foundations as superstructure. In: Prosorov O (ed) Proceedings of the international conference on philosophy, mathematics, linguistics: aspects of interaction (PhML-2012), Euler International Mathematical Institute St. Petersburg, St. Petersburg, Russia, May 22–25, 2012. College Publications, London. Stud. Log. (Lond.) 70, 101–115 (2017)

Manin DY, Manin YI (2017) Cognitive networks: brains, Internet, and civilizations. In: Sriraman B (ed) Humanizing mathematics and its philosophy. Springer International Publishing AG, Berlin, pp 85–96

Manin YI, Marcolli M (2016) Semantic spaces. Math Comput Sci 10(4):459–477

Nikolaenko NN (2003) Artistic thinking and cerebral asymmetry. Acta Neuropsychol 1(2):48–67

Papadopoulos A (2023) Pavel Florensky and his world, this Handbook
Plotnitsky A (2023) Continuity and discontinuity in, and between, mathematics and physics. In: Essays in geometry – dedicated to Norbert A'Campo (ed: Papadopoulos A). European Mathematical Society, EMS Press, Berlin
Rota G-C (2008) Indiscrete thoughts. With forewords by R. Hersh and R. Sokolowski. Edited, with notes and an epilogue by F. Palombi. First ed. 1997, Modern Birkhäuser Classics. Basel, new ed
Tannery J (1912) Pensées. In: En souvenir de Jules Tannery (ed: Boutroux É, Borel É). Imprimerie Crété, Corbeil
Tarkovsky A (1994) Journal, 1970–1986, French translation from the Russian by Anne Chilikov, Cahiers du Cinéma, Atelier 3, 1993. Revised and augmented edition, December 2004, definitive edition, 2017. English edition: Time Within Time – The Diaries 1970–1986, translated by Kitty Hunter-Blair. Faber and Faber, London/Boston
Tarkovsky A (2004) Le temps scellé: De l'Enfance d'Ivan au Sacrifice, transl. from the Russian by A. Kichoilov and Ch. H. de Brantes. Philippe Rey, Paris
Tasić V (2001) Mathematics and the roots of postmodern thought. Oxford University Press, Oxford

The Nature of Mathematical Objects

Carlo Cellucci

Contents

1 Introduction 36
2 Mathematical Objects As Independently Existing Entities 37
3 Mathematical Objects as Abstractions 42
4 Mathematical Objects as Logical Objects 44
5 Mathematical Objects as Simplifications 45
6 Mathematical Objects as Mental Constructions 46
7 Mathematical Objects as Structures 48
8 Mathematical Objects as Fictions 50
9 Mathematical Objects as Idealizations of Sensible Things 52
10 Mathematical Objects as Idealizations of Operations 53
11 The Basis for an Alternative Conception 55
12 The Heuristic Conception of Mathematical Objects 56
13 Some Remarks about the Heuristic Conception 57
14 Conclusion 58
15 Cross-References 59
References 59

Abstract

A traditional question in the philosophy of mathematics is to give an answer to the question: What is the nature of mathematical objects? This chapter considers the main answers that have been given to this question, specifically those according to which mathematical objects are independently existing entities, or abstractions, or logical objects, or simplifications, or mental constructions, or structures, or fictions, or idealizations of sensible things, or idealizations of operations. The

C. Cellucci (✉)
Sapienza University of Rome, Rome, Italy
e-mail: carlo.cellucci@uniroma1.it

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_20

chapter also shows the shortcomings of these answers, and considers an alternative answer, according to which mathematical objects are hypotheses introduced to solve mathematical problems by the analytic method.

Keywords

Objects in Platonism · Logicism · Formalism · Intuitionism · Structuralism · Fictionalism · Heuristic view

1 Introduction

A traditional question in the philosophy of mathematics is to give an answer to the question: What is the nature of mathematical objects?

This question originated in antiquity. Herodotus says that Egypt's king Sesostris (twentieth century BCE) "made a division of the soil of Egypt among the inhabitants, assigning square parcels of land of equal size to all, and obtaining his chief revenue from the rent which the holders were required to pay him year by year" (Herodotus, *Historiae*, II, 109.7–10). If any man was robbed of a part of his land by the inundations of the Nile, "he would come to the king and tell him what had happened; then the king would send inspectors to measure how much land he had lost, so that thereafter it should pay in proportion to the tax originally imposed. Thus discovered" in Egypt, "geometry then passed into Greece" (ibid., II, 109.11–14).

Passing into Greece, however, mathematics changed nature, it became an abstract discipline, detached from its empirical origin. Specifically, "Pythagoras transformed the study of geometry into a liberal discipline, tracing its first principles and studying its problems from a purely abstract and theoretical viewpoint" (Proclus, *In primum Euclidis Elementorum librum commentarii*, 65.15–19).

The change in nature of mathematics, in particular, concerned mathematical objects. Because of its concrete nature, Egyptian mathematics was about such things as parcels of land, thus about physical things, knowable by the senses. Conversely, because of its abstract nature, Greek mathematics was about non-physical things, not knowable by the senses.

Plato assumed that these non-physical things were things existing independently of the mind, which could be known only by intuition.

Indeed, Plato says that, although mathematicians "use visible figures and make their claims about them," they "are not thinking of them but of those ideas of which they are likeness. They pursue their inquiry for the sake of the square itself and the diagonal itself, and not for the sake of the diagonal they draw, and similarly with the others" (Plato, *Respublica*, VI, 510 d 5–e 1). Mathematicians use the visible figures they draw "only as images, but what they really seek is to get sight of those ideas, which can be seen only by the mind" (ibid., VI, 510 e 3–511 a 1). In fact, "there

really exist these ideas in themselves, that we cannot perceive with the senses but only with intuition" (Plato, *Timaeus*, 51 d 4–5). They have a kind of reality that "is always the same, ungenerated and imperishable," it "is not visible nor perceptible by any sense," and is such that "only intuition has been granted to contemplate it" (ibid., 52 a 1–4).

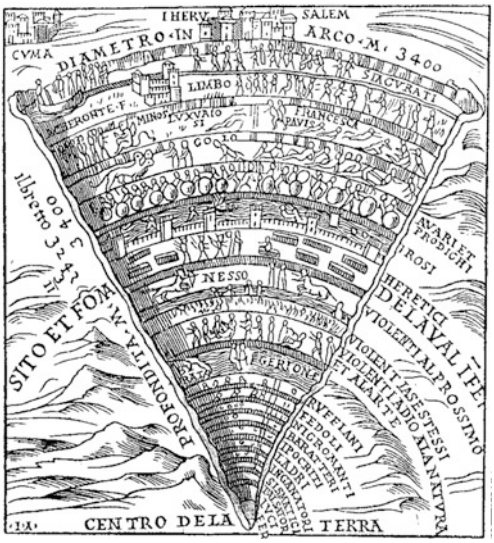
Since antiquity, the question "What is the nature of mathematical objects?" has received several answers. This chapter discusses the main ones and shows their shortcomings. Then, it considers an alternative answer, according to which mathematical objects are hypotheses introduced to solve mathematical problems by the analytic method.

2 Mathematical Objects As Independently Existing Entities

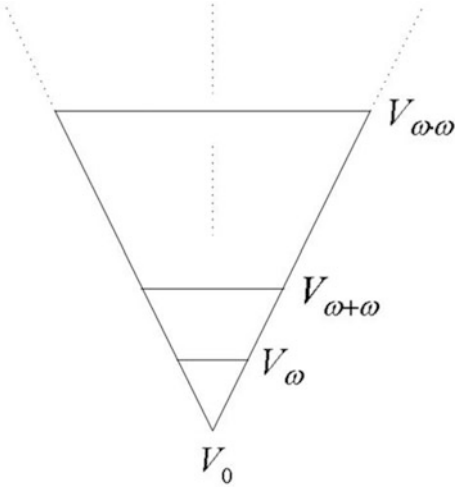
An answer to the question "What is the nature of mathematical objects?" is that they are independently existing entities. As we have just seen, this conception of mathematical objects has been first formulated by Plato, therefore it goes under the name "mathematical platonism," which was coined by Bernays (1935). Such conception has a large following between mathematicians because, as Dieudonné says, "each mathematician" has "the feeling" that "he is working with something real," namely, with independently existing entities, and this feeling "is very convenient" (Dieudonné 1970, 145).

Although the conception that mathematical objects are independently existing entities has been first formulated by Plato, it, however, is best known in Gödel's formulation.

Gödel says that mathematics "describes a non-sensual reality, which exists independently both of the acts and of the dispositions of the human mind and is only perceived, and probably perceived very incompletely, by the human mind" (Gödel 1986–2002, III, 323). Such non-sensual reality consists of sets, and precisely of "the objects of transfinite set theory" (ibid., II, 267). These objects do "not belong to the physical world," and "even their indirect connection with physical experience is very loose" (ibid.). They "exist objectively and independently of our mental acts and decisions" (ibid., III, 311). A set can be conceived as "something obtainable from" the empty set $\varnothing$ "by iterated application of the operation 'set of'," where iterated application "is meant to include transfinite iteration" (ibid., II, 259 and footnote 13). If by the operation 'set of' it is meant the power set operation $\mathrm{P}(X)$ that collects all the subsets of a set X into one set, then a set can be conceived as a member of the hierarchy defined by: $V_0 = \varnothing$; (ii) $V_{\alpha+1} = \mathrm{P}(V_\alpha)$ for any ordinal α; (iii) $V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha$, for any limit ordinal λ. Such hierarchy is commonly represented in a way that is reminiscent of the current representations of Dante's *Inferno*.



For, it is represented as follows.



But, "despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory," namely, "mathematical intuition" (ibid., II, 268). This is a superior faculty, which is not "derived from subconscious induction or Darwinian adaptation" (ibid., III, 354).

Mathematical intuition "is sufficiently clear to produce the axioms of set theory and an open series of extensions of them" (ibid., II, 268). It gives insights both "for obtaining unambiguous answers to the questions of transfinite set theory," such as

Cantor's continuum hypothesis, and "for the solution of the problems of finitary number theory (of the type of Godlbach's conjecture)" (ibid., II, 269).

Through mathematical intuition, sets "are known precisely, and general laws" about them "can be recognized with certainty," because they can be derived from axioms "by deductive, not inductive, inference" (ibid., III, 312, footnote 18). So, one can recognize "the truth of the axioms concerning them" (ibid., II, 258). In fact, for the "axioms there exists no other" rational "foundation except that they" can "directly be perceived to be true" by "an intuition of the objects falling under them" (ibid., III, 346–347).

We can arrive at directly perceiving the axioms to be true by the procedure which "consists in focusing more sharply on the concepts concerned" (ibid., III, 383). Indeed, "by directing our attention in a certain way," we will get an intuition of those concepts, and hence a "comprehensive and secure insight" into "the axioms that hold for them" (ibid.). In fact, "cultivating (deepening) knowledge of the abstract concepts" will "produce in us a new state of consciousness in which we describe in detail the basic concepts we use in our thought" (ibid.).

However, besides mathematical intuition, there is "another (though only probable) criterion of truth of mathematical axioms, namely their fruitfulness in mathematics and, one may add, possibly also in physics" (ibid., II, 269). "Fruitfulness" means "fruitfulness in consequences, in particular in 'verifiable' consequences, i.e., consequences demonstrable without the new axiom" (ibid., II, 261). The "simplest case of an application of the criterion under discussion arises when some set-theoretical axiom has number-theoretical consequences verifiable by computation up to any given integer" (ibid.).

The conception that mathematical objects are independently existing entities, however, has the following shortcomings.

1. Gödel claims that mathematics describes a non-sensual reality, which exists independently both of the acts and of the dispositions of the human mind. This raises the question: How, then, can mathematics be useful in physics? Gödel answers this question saying that mathematics "might help us considerably in knowing" the physical reality "if in some respect, or as to certain parts, it happened to be similar or isomorphic to it. In fact this would correspond closely to the manner in which mathematics is applied in theoretical physics" (Gödel 1986–2002, III, 353–354, footnote 44). Thus, according to Gödel, mathematics can be useful in physics if it is isomorphic to the physical reality. But this conflicts with the fact that different mathematical theories are commonly used to deal with the same physical phenomena. For example, on the one hand, Weyl says that group theory is exactly what is "necessary for an adequate description of the quantum mechanical relations" (Weyl 1950, xxi). On the other hand, von Neumann says that Hilbert space theory is exactly what is necessary for an adequate description of the quantum mechanical relations, because it permits to present "quantum mechanics in a unified representation" which "is mathematically rigorous" (von Neumann 1955, vii). Now, group theory and Hilbert space theory cannot be both isomorphic to the physical reality.

2. Gödel claims that, despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, namely, mathematical intuition. This raises the question: How can we have this kind of perception? Gödel assumes that, in order "to make the handling of abstract impressions (as opposed to sense impressions) possible," there must be "some physical organ," and that "such a physical organ must be closely related to the neural center for language" (Wang 1996, 233). But there is no evidence that such a physical organ exists, and indeed its existence makes no sense in evolutionary terms, because the highly abstract objects of set theory were introduced only in the nineteenth century. As to the claim that such a physical organ must be closely related to the neural center for language, experiments have shown that "mathematics and language are functionally and neuroanatomically independent" (Brannon 2005, 3178).
3. Gödel claims that mathematical intuition is sufficiently clear to produce the axioms of set theory and an open series of extensions of them. But this conflicts with the fact that, at least since Kant, it has been held that "an intuition is a singular representation" (Kant 1992, 589). How, then, can mathematical intuition produce the axioms of set theory and an open series of extensions of them, since the axioms in question are not singular propositions but universal propositions? Moreover, mathematical intuition has given no insight for obtaining unambiguous answers to the questions of transfinite set theory, such as Cantor's continuum hypothesis, and for the solution of the problems of finitary number theory, of the type of Godlbach's conjecture. These problems remain unsolved.
4. Gödel claims that, through mathematical intuition, sets are known precisely, and general laws about them can be recognized with certainty, because they can be derived from axioms by deductive inference. But this conflicts with Gödel's second incompleteness theorem, by which the consistency of no axiom system for set theory can be recognized with certainty. Therefore, one cannot say that, by mathematical intuition, sets are known precisely, and general laws about them can be recognized with certainty because they can be derived from axioms by deductive inference. For, intuition is unreliable, so it is unrealistic to use it as a foundation for the axioms of sets theory and the whole of mathematics.
5. Gödel claims that we can arrive at directly perceiving the axioms to be true by the procedure which consists in focusing more sharply on the concepts concerned. But this is untenable. For, suppose that, by focusing more sharply on the concept of set Σ, we get an intuition of that concept. Let S be a formal system for set theory whose axioms this intuition ensures us to be true of Σ. Thus, Σ is a model of S, hence S is consistent. Then, however, by Gödel's first incompleteness theorem, there is a sentence A of S which is true of Σ but is unprovable in S. Since A is unprovable in S, the formal system $S' = S \cup \{\neg A\}$ is consistent, and hence has a model, say Σ'. Then $\neg A$ is true of Σ' and hence A is false of Σ'. Thus, Σ and $\Sigma'$are both models of S, but A is true of Σ and false of Σ', so Σ and Σ' are not equivalent.

Suppose now that, once again, as Gödel says, by focusing more sharply on the concept of set Σ', we get an intuition of that concept. Then we have two different intuitions, one ensuring us that Σ is the concept of set, and the other one ensuring us that Σ' is the concept of set, where the sentence A is true of Σ and false of Σ'. Which of Σ and Σ' is the genuine concept of set? Gödel's procedure provides no answer, so through it we cannot arrive at directly perceiving which axioms for sets are true.

6. Gödel claims that, besides mathematical intuition, there is another (though only probable) criterion of truth of mathematical axioms, namely, their fruitfulness in consequences, and that the simplest case of application of this criterion of truth is when some set-theoretical axiom has number-theoretical consequences verifiable by computation up to any given integer. But this is untenable. As Gödel himself admits, this criterion of truth "cannot yet be applied to the specifically set-theoretic axioms (such as those referring to great cardinal numbers), because very little is known about their consequences in other fields" (Gödel 1986–2002, II, 269). What is more important, the criterion of truth in question fails because from false axioms it is possible to deduce true consequences. In order to assert that the axioms are true, it should be possible to demonstrate that all the consequences deducible from them are true, but this is generally impossible. This is already made clear by Kant, who says that "inferring the truth of a cognition from the truth of its consequences, would be allowed only if all of the possible consequences are true," but "this procedure is unusable, because to have insight into all possible consequences of any proposition that is assumed exceeds our powers" (Kant 1998, A790/B818). Evidence for Kant's statement is provided by the fact that the set of all consequences of the axioms of second-order Peano arithmetic PA^2 is not algorithmically enumerable. For, assume P: The set of all consequences of the axioms of PA^2 is algorithmically enumerable. Now, since the axioms of PA^2 are categorical, a second-order sentence is a consequence of the axioms of PA^2 if and only if it is true of the natural numbers. From this, by the assumption P, it follows Q: The set of all second-order sentences true of the natural numbers is algorithmically enumerable. But, by Tarski's theorem for second-order sentences, the set of all second-order sentences true of the natural numbers is not definable in the set of all-natural numbers by any second-order formula, and hence *a fortiori* is not algorithmically enumerable. Thus not-Q: The set of all second-order sentences true of the natural numbers is not algorithmically enumerable. So, from the assumption P, we have deduced a contradiction, Q and not-Q. Therefore, by *reductio ad absurdum*, we conclude not-P: The set of all consequences of the axioms of PA^2 is not algorithmically enumerable. This means that there exists no algorithm, *a fortiori* no feasible procedure, for enumerating all consequences of the axioms of PA^2. So, as Kant says, inferring the truth of a cognition from the truth of its consequences is unfeasible, because it exceeds our powers.

3 Mathematical Objects as Abstractions

Another answer to the question "What is the nature of mathematical objects?" is that they are objects obtained by abstraction from sensible things, where abstraction is the process by which one considers certain attributes of sensible things and leaves out the other ones. Therefore, objects obtained by abstraction from sensible things have attributes which also occur in sensible things. This conception of mathematical objects has been put forward by Aristotle.

Indeed, Aristotle says that "the mathematician carries out his investigation about things reached by abstraction. For, he investigates things leaving out all sensible qualities, such as weight, lightness, hardness," and "considering only quantity and continuity, in one, two or three dimensions, and the properties of these as quantitative and continuous; and he does not consider them in any other respect" (Aristotle, *Metaphysica*, K 3, 1061 a 28–35).

Thus, the mathematician investigates the same things as the physicist, because "the physicist concerns himself with all the properties, active and passive, of physical bodies" (Aristotle, *De Anima*, A 1, 403 b 11–12). And "physical bodies contain surfaces and volumes, lines and points, and these are the subject-matter of mathematics" (Aristotele, *Physica*, B 2, 193 b 24–26).

However "the mathematician, though he too treats of these things, nevertheless does not treat them as the limits of a physical body; nor does he consider the attributes indicated as the attributes of such bodies," but he "separates them" from physical bodies; for, "in thought they are separable from physical motion" (ibid., B 2, 193 b 33–34). That is, the mathematician "thinks of mathematical objects, which are not separate, as being separate, when he thinks of them" (Aristotle, *De Anima*, Γ 7, 431 b 15–16). Thus, the mathematician deals with attributes which "are not considered as properties of any particular physical body and are obtained by abstraction" (Aristotle, *De Anima*, A 1, 403 b 14–15).

On the other hand, "since, as it appears, there is nothing that exists as separate from the sensible magnitudes, it is in the sensible forms that one finds" both the things "that are spoken of in abstraction," that is, the mathematical objects, and "the properties and affections of sensible objects. For this reason, one would not be able to learn or understand anything if he had not the faculty of perception" (ibid., Γ 8, 432 a 3–8). So, although "one thinks of mathematical objects as separate, when one thinks of them," nevertheless they "are not separate" (ibid., Γ 7, 431 b 15–17).

The conception that mathematical objects are abstractions from sensible things, however, has the following shortcomings.

1. By abstraction from sensible things we cannot obtain things which do not exist in sensible things, for example, actual infinite sets. For, as Kant says, "abstraction is only the negative condition under which universal representations can be generated" (Kant 1992, 593). Through it, "nothing is produced, but rather left out" (ibid., 487). Therefore, "abstraction does not add anything," but "rather cuts off everything that does not belong to the concept" (ibid., 351).

2. By abstraction from sensible things we cannot obtain objects having that exactness which is proper to mathematical objects. For example, from an edge of a table, we cannot obtain a straight line, but only a line which is slightly zigzag and with some, however small, breadth.
3. By abstraction from sensible things we cannot obtain objects having that determinateness which is proper to mathematical objects. As Frege says, "suppose, e.g., that there are a black and a white cat sitting side by side before us. We do not attend to their colour, and they become colourless – but they still sit side by side. We do not attend to their posture, and they cease to sit without, however, assuming a different posture – but each of them is still in its place. We no longer attend to the place, and they cease to occupy one – but they continue presumably to be separate" (Frege 1984, 197–198). By these successive abstractions, from the black and the white cat we have perhaps obtained "a general concept of cat" (ibid., 198). But, "by continued application of this procedure, each object is transformed into a more and more bloodless phantom. As the end result, we obtain from each object a something emptied of all content; but the something obtained from one object differs nevertheless from the something obtained from another object, even though it is not easy to say how" (ibid.).
4. By abstraction from sensible things we cannot obtain mathematical objects whenever the attributes which are considered are subordinate to those which are left out. For example, suppose that there are a black and a white cat sitting side by side before us, and that we do not attend to what distinguishes them from all the other things. We thus obtain the concept of two "something." But, from the concept of two "something," if we do not attend to "something," we do not obtain the concept of two, since then there would not be left any more things of which one could say that there are two of them. Therefore, Frege says that if, "in considering a white cat and a black cat, I disregard the properties which serve to distinguish them," I do not obtain the concept "two" but rather "the concept 'cat'" (Frege 1960, 45). The "concept 'cat', no doubt, which we have arrived at by abstraction, no longer includes the special characteristics of either, but of it, for just that reason, there is only one" (ibid., 45–46).
5. By abstraction from sensible things, we may arrive at extravagant results. As Frege ironically says, the supporters of abstraction consider directing attention away from something an operation that "is especially effective. We attend less to a property, and it disappears. By thus making one characteristic mark after another disappear, we obtain more and more abstract concepts," therefore "inattention is a most effective logical power; this is presumably why professors are absent-minded" (ibid., 197). But directing attention away from something leads to absurdities. For example, "if, abstracting from the distinction between my house and that of my neighbour, I were to regard them as the same, and consequently act in the other house as in my own, then the erroneousness of my abstraction would soon be made clear to me" (Frege 2013, I, 107).

4 Mathematical Objects as Logical Objects

Another answer to the question "What is the nature of mathematical objects?" is that they are logical objects, because they are definable in terms of logical concepts, and all their properties are deducible from logical principles. This conception of mathematical objects has been put forward by Leibniz, Frege, and Russell.

For example, Russell says that all pure mathematics deals with objects "definable in terms of a very small number of fundamental logical concepts" and "all its propositions are deducible from a very small number of fundamental logical principles" (Russell 2010, xliii). In particular, it is possible to "reduce arithmetic, and above all the definition of numbers, to logic" (Russell 1992, 33). This view "was strongly advocated by Leibniz" but, owing to the limitations of his logic, Leibniz failed "in the endeavour to carry out in detail a view which, in its general outline, is now known to be correct" (ibid., 5).

Indeed today, "thanks to the progress of symbolic logic," all "mathematics can be strictly and formally deduced" from logical premisses, "and all the entities that occur in mathematics can be defined in terms of" the logical concepts "that occur in" those logical "premisses" (Russell 2010, 4–5). Therefore, "all mathematics is symbolic logic," and "when this fact has been established, the remainder of the principles of mathematics consists in the analysis of symbolic logic itself" (ibid., 5). Not only all mathematics is symbolic logic, but "it has now become wholly impossible to draw a line between the two; in fact, the two are one" (Russell 1993, 194).

The conception that mathematical objects are logical objects, however, has the following shortcomings.

1. Russell claims that all mathematical objects are definable in terms of logical concepts. But, while single natural numbers are definable in terms of logical concepts, not all their properties are deducible from logical principles, in particular, it is impossible to deduce from logical principles that there exist infinitely many natural numbers. Therefore, Russell is forced to introduce and axiom, the axiom of infinity, which postulates that there exist infinitely many natural numbers. But this is incompatible with the claim that all the properties of mathematical objects are deducible from logical principles. Even Russell admits that "the axiom of infinity" is "a proposition which, though it can be enunciated in logical terms, cannot be asserted by logic to be true" (Russell 1993, 202–203). For, "it seems plain that there is nothing in logic to necessitate its truth or falsehood, and that it can only be legitimately believed or disbelieved on empirical grounds" (Whitehead and Russell 1925–1927, II, 183).
2. Russell claims that all properties of mathematical objects are deducible from logical principles. Then, in particular, all arithmetic truths are deducible from logical principles L. But this is refuted by Gödel's first incompleteness theorem, by which there are arithmetic truths which are not deducible from L. Therefore, there are arithmetic truths which are not logical truths. Against this conclusion, Hale and Wright object that Gödel's first incompleteness theorem "might just as well" be "taken as revealing, not that there are arithmetic truths which are not"

logical truths, "but that logical truth likewise defies complete deductive characterization" (Hale and Wright 2001, 4, footnote 5). For, by the strong incompleteness theorem for second-order logic, not all logically valid second-order sentences are deducible from L. So, the fact that there are arithmetic truths which are not deducible from L, does not mean that there are arithmetic truths which are not logical truths. Such arithmetic truths might as well be logically valid sentences which are not deducible from L, and hence might be logical truths. Therefore, Hale and Wright conclude, "if the logicist is entitled to take logic as including second- and perhaps higher-order logic, then Gödel's incompleteness result has no specific bearing on the logicist project" (ibid., 5, footnote 5). But this objection is invalid because, by Gödel's first incompleteness theorem, the arithmetic truths which are not deducible from L are first-order sentences, and, by the completeness theorem for first-order logic, all logically valid first-order sentences are deducible from L. So, if the arithmetic truths in question were logically valid, being first-order sentences they would be deducible from L. Since they are not deducible from L, it follows that they cannot be logical truths. Therefore, contrary to Hale and Wright's objection, the claim that all arithmetic truths are deducible from certain logical principles L is refuted by Gödel's first incompleteness theorem.

5 Mathematical Objects as Simplifications

Another answer to the question "What is the nature of mathematical objects?" is that mathematics is ultimately about concrete spatiotemporal objects but, in order to demonstrate properties of such objects, abstract objects are added to the concrete ones, notably actual infinities. Their use can be justified by demonstrating, by absolutely reliable means, that it will lead to no contradiction. This conception of mathematical objects has been put forward by Hilbert.

Indeed, Hilbert says that mathematics is ultimately about "concrete" spatiotemporal "objects that are intuitively present as immediate experience prior to all thought" (Hilbert 1967b, 464). But, in order to demonstrate properties of such objects, abstract objects are added to the concrete ones, notably "the actual infinite" (Hilbert 1967a, 373). These abstract objects do not permit to demonstrate properties of concrete objects that could not be demonstrated without them, in particular "the modes of inference employing the infinite" can "be replaced generally by finite processes that have precisely the same results" (ibid., 370). But they are useful, because they significantly simplify the demonstration of such properties, which "results in the simplification and completion of the theory" (Hilbert 1996b, 1144).

However, since abstract objects are not immediately given in sensible intuition, operation with them "has proved to be inadequate and uncertain" (Hilbert 1996a, 1121). Now, the "extension by the addition" of abstract objects to concrete ones "is legitimate only if no contradiction is thereby brought about in the old, narrower domain" (Hilbert 1967b, 471). In order to ensure this, the theory must be "rigorously

formalized, so that mathematics" proper "becomes a stock of formulae" (Hilbert 1996c, 1152). Then one must demonstrate that, in the theory so formalized, "it is impossible for us to obtain two logically contradictory propositions" (Hilbert 1967b, 471). Moreover, one must demonstrate this by absolutely reliable means, namely, without using "any dubious or problematical mode of inference" (Hilbert 1996b, 1139). This will guarantee that operation with abstract objects "can never lead to a demonstrably false result" (Hilbert and Bernays 1968–1970, I, 44). To do so is always possible, because "there are absolutely no unsolvable problems" (Hilbert 1996d, 1165).

The conception that mathematical objects are simplifications, however, has the following shortcomings.

1. Hilbert assumes that a theory can always be formalized. But, by Gödel's first incompleteness theorem, this is generally impossible, because there will always be some sentence of the theory which is true but cannot be demonstrated in the theory.
2. Hilbert assumes that it is always possible to demonstrate, by absolutely reliable means, that in the theory so formalized no contradiction will arise. But, by Gödel's second incompleteness theorem this is generally impossible.
3. Hilbert assumes that to demonstrate, by absolutely reliable means, that in the theory so formalized no contradiction will arise, will guarantee that no falsity will be demonstrable in the theory. But it is not so because, by a corollary of Gödel's first incompleteness theorem, any consistent sufficiently strong formal system has a consistent extension in which some false sentence is demonstrable. In fact, Hilbert need not have waited for Gödel to realize that noncontradictoriness is no guarantee against falsity. Already Kant had made it clear that it is, "to be sure, a necessary logical condition" that in "a concept no contradiction must be contained," but this is "far from sufficient for the objective reality of the concept, i.e., for the possibility of such an object as is thought through the concept" (Kant 1998, A220/B268). It is not enough to assume, as "condition of all our judgments whatsoever," that "they do not contradict themselves," because "for all that a judgment may be free of any internal contradiction, it can still be either false or groundless" (ibid., A150/B190).

6 Mathematical Objects as Mental Constructions

Another answer to the question "What is the nature of mathematical objects?" is that they are mental constructions based on intuition. This means that they are not merely ideas conceived by the mind but are entities whose construction can be wholly accounted for in terms of acts of the mind based on intuition. This conception of mathematical objects has been put forward by Brouwer.

Indeed, Brouwer criticizes contemporary mathematics because it holds that mathematical objects "have mathematical existence" when "they have been represented in spoken or written language together with" the "laws upon which

their development depends" (Brouwer 1975, 83). So, contemporary mathematics "completely abandons any geometric or arithmetical intuition and confines its subject matter exclusively to the domain of mathematical language" (Brouwer 1998, 55). On the contrary, a mathematical object "is a mental construction, essentially independent of language" (Brouwer 1975, 477). It is not merely a linguistic object but a mental construction based on intuition, because "no mathematics can exist which has not been intuitively built up" (ibid., 52). For example, a natural number is a mental construction based on the "intuition of time" (ibid., 116). Mental constructions are the same in all individuals, because "it is an essential hypothesis for human understanding that the structure of" construction "is the same for all individuals" (ibid., 599). Therefore, the mental construction of the natural numbers is the same for all individuals. The "only possible foundation of mathematics must be sought in this construction under the obligation carefully to watch which constructions intuition allows and which not" (ibid.). Since contemporary mathematics abandons any geometric or arithmetical intuition and consider mathematical objects which cannot be intuitively built up, this requires "a new practice of creative mathematical work" (ibid., 477). Indeed, "several of the theories of mathematics proper" must be "built anew" (ibid., 412).

The conception that mathematical objects are mental constructions based on intuition, however, has the following shortcomings.

1. Brouwer criticizes contemporary mathematics for completely abandoning any geometric or arithmetical intuition and confining its subject matter exclusively to the domain of mathematical language. But his criticism is not motivated by any necessity of mathematical practice, it is only due to his purely ideological assumption that to be is to be constructed in intuition. This entails that important parts and results of modern mathematics must be rejected. As Weyl says, "with Brouwer, mathematics gains the highest intuitive clarity; his doctrine is idealism in mathematics thought to and end. But, full of pain, the mathematician sees the greatest part of his towering theories dissolve in fog" (Weyl 1998, 136). And the mathematician sees such theories dissolve in fog only to be replaced with surrogate theories which are weaker and quite awkward to use in practice.
2. Brouwer claims that a mathematical object is a mental construction based on intuition, for example, a natural number is a mental construction based on the intuition of time. The mental construction of the natural numbers is the same for all individuals. But, as Gillies points out, this contrasts with the fact that, "if we look at number systems historically, we find" that "different, and generally more primitive, systems existed for centuries in many parts of the world. Often there was a crucial innovation in some relatively small group, and this change gradually spread elsewhere," so "the system of natural numbers" was "constructed slowly and often painfully over long periods of time by human beings" (Gillies 2014, 26). Generally, "some mathematical concepts are counter-intuitive rather than natural, and are the result of centuries of trial and error rather than easy inferences from an *a priori* intuition" (ibid.).

3. Brouwer claims that no mathematics can exist which has not been intuitively built up. But this entails that there cannot exist mathematical objects which are important to physics. In particular, it entails there cannot exist everywhere defined discontinuous functions. This is not an oversight, but deliberate. For, Brouwer explicitly declares that he does not care a bit for the use of mathematics in physics, since mathematics is not aimed at "expansion of human domination over nature" but at beauty, and "things as such are not beautiful, nor is their domination by shrewdness" (Brouwer 1975, 483). Mathematics "is an art, its application to the world an evil parasite" (Borwein et al. 2004, 295).

7 Mathematical Objects as Structures

Another answer to the question "What is the nature of mathematical objects?" is that, rather than about objects, mathematics is about structures, where a structure is a collection of objects with certain relations among them, which however highlights the relationships between the objects, rather than the objects themselves. Objects are simply places in structures. This conception of mathematical objects has been put forward by Bourbaki, Shapiro, and Resnik.

For example, Shapiro says that "mathematics is the science of structure" (Shapiro 2000, 257). A structure is "a collection of objects with certain relations among them," highlighting "the interrelationships among the objects, and ignoring any features of them that do not affect how relate to other objects in the system" (ibid., 259). For example, "the subject-matter of arithmetic is a single abstract structure," the structure of natural numbers, that is "the pattern common to any infinite collection of objects that has a successor relation, a unique initial object, and satisfies the induction principle" (ibid., 258). Mathematical objects are merely positions in a structure, for example, "the number 2 is no more and no less than the second position in the natural number structure; and 6 is the sixth position. Neither of them has any independence from the structure in which they are positions, and as positions in this structure, neither number is independent of each other" (ibid.).

A "structure can be characterized by an axiomatization, which is an implicit definition" (Shapiro 2004, 32). Not every implicit definition, however, "successfully characterizes a structure," in order to characterize it an implicit definition must fulfill two conditions, the "existence condition" that "at least one structure" satisfies "the axioms," and the "uniqueness condition" that "at most one structure is described" (Shapiro 2000, 285). The uniqueness condition "depends on the underlying logic of the axiomatization," indeed, if the latter is taken to be second-order logic, then "there are unique characterizations of rich mathematical structures" (ibid.). In particular, there is a unique characterization of the structure of natural numbers. As to the existence condition, "in mathematics as practised, set theory" is "taken to be the ultimate court of appeal for existence questions," so "a certain type of mathematical objects" can be said to exist if "objects of this type can be found or modelled in the set-theoretic hierarchy" (ibid., 288). In fact, "set theory is the appropriate court of appeal because it is comprehensive. The set-theoretic hierarchy is so big that just

about any structure can be modelled or exemplified there" (ibid.). Thus, ultimately, "a structure is an equivalence class in the set-theoretic hierarchy" (Shapiro 1997, 92).

Since there are axiomatizations which give unique characterizations of rich mathematical structures, all the models of such characterizations are isomorphic. But, "since isomorphic models are equivalent, the relevant properties of any model of the axiomatization are the same, and so, in a sense, any model is as good as any other," therefore "we can learn about the structure by studying an exemplification of it" (Shapiro 2004, 32). For example, "the natural number structure is exemplified by various systems of natural numbers" (Shapiro 2008, 174). Therefore, we can learn about the structure of natural numbers by studying any of the systems of natural numbers by which the natural number structure is exemplified.

The conception that mathematical objects are structures, however, has the following shortcomings.

1. Shapiro claims that mathematics is the science of structure, but thus he mistakes for mathematics what is merely the mathematics produced by a particular school, the Göttingen's school, which, through Van der Waerden's *Moderne Algebra*, influenced Bourbaki. (On the Göttingen's school, see Ferreiros 1999; Ferreiros and Gray 2006). Contrary to his claim, there are many mathematical questions, such as the distribution of prime numbers or the transcendence of π, which are not questions about a structure.
2. Shapiro claims that a certain type of mathematical objects can be said to exist if objects of this type can be modelled in the set-theoretic hierarchy. But, as Shapiro himself acknowledges, "for the structuralist, the set-theoretic hierarchy is just another structure. What makes us think that set theory itself is coherent?" (Shapiro 2000, 288). Surely, "we cannot justify the coherence of set theory itself by modelling it in the set-theoretic hierarchy" (ibid.). The circle would be too blatant. Despite this, Shapiro claims that "the coherence of set theory is presupposed by much of the foundational activity in contemporary mathematics. Rightly or wrongly, mathematics presupposes that satisfiability (in the set-theoretic hierarchy) is sufficient for existence" (ibid., 288–289). Therefore, we must "accept this presupposition and make use of it like everyone else, and are in no better (and no worse) of a position to justify it" (ibid., 289). Such "presupposition is not vicious, even if it lacks external justification" (Shapiro 1997, 136). Contrary to Shapiro's claim, however, the presupposition in question is indeed vicious. Shapiro's argument is a *petitio principii* because it is similar to the argument that we must accept the existence of God because such existence is presupposed by the existence of the world.
3. Shapiro claims that ultimately a structure is an equivalence class in the set-theoretic hierarchy. But then the primitive concept of mathematics is the concept of set, not the concept of structure. To avoid this problem, Shapiro formulates an axiomatic theory of structures. However, this theory is entirely similar to set theory, and neither clarifies nor solves any problem that set theory does not solve. Shapiro himself admits that "the set-theoretic hierarchy and the realm of

structures are little more than notational variants of each other" (Shapiro 2004, 20). Then the concept of structure is just the concept of set in disguise.

4. Shapiro claims that we can learn about the structure of natural numbers by studying any of the systems of natural numbers by which the natural number structure is exemplified. But Zermelo's and von Neumann's provide two different exemplifications of the structure of natural numbers. In Zermelo's exemplification 0, 1, 2, 3, . . .are $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}$, . . ., while in von Neumann's exemplification they are $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, ... According to Shapiro, the only thing that matters about the natural number structure is the relation between natural numbers. Now, in both such exemplifications, the relation between the natural numbers is the membership relation, and yet the membership relation has different properties in each of them, even properties which are inconsistent with each other. Thus, in Zermelo's exemplification, $1 \notin 3$ because 1 is $\{\emptyset\}$ and 3 is $\{\{\{\emptyset\}\}\}$, while in von Neumann's exemplification, $1 \in 3$ because 1 is $\{\emptyset\}$ and 3 is $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. Which is the correct exemplification? Structuralism can give no answer.

8 Mathematical Objects as Fictions

Another answer to the question "What is the nature of mathematical objects?" is that they are fictions. This conception of mathematical objects has been put forward by Field, Balaguer, and Yablo.

Thus, Field says that mathematical objects are fictions as the characters of a story, because all there is to mathematics is that "we have a good story about natural numbers, another good story about sets, and so forth" (Field 1989, 22). A mathematical proposition is true only in the same sense in which the proposition "'Oliver Twist lived in London' is true: the latter is true only in the sense that it is true according to a certain well-known story, and the former is true only in the sense that it is true according to standard mathematics" (ibid., 3).

Nevertheless, even though mathematics is about fictional objects, mathematics is useful because we can use it "as a device for drawing conclusions" about the physical world "much more easily than we could draw them by a direct proof" (Field 1980, 28). On the other hand, since mathematics is about fictional objects, "it would be extremely surprising if it were to be discovered that standard mathematics implied that there are at least 10^6 non-mathematical objects in the universe, or that the Paris Commune was defeated" (ibid., 13).

To make sure that this will not occur, it is necessary to prove that mathematics is conservative over physics. Namely, it is necessary to prove that the conclusions we arrive at by adding a mathematical theory to a physical theory which makes no reference to mathematical entities "are not genuinely new, they are already derivable in a more long-winded fashion from the premises" of the physical theory, "without recourse to the mathematical entities" (ibid., 10–11). If this could be proved, it would be made sure that any deduction "that can be made with the help of mathematics could be made (usually more long-windedly) without it" (ibid., x). One could then assert that "that mathematics is not really indispensable after all" (Field 1989, 26). And one

would be "free to use any mathematics he likes for deducing consequences, as long as" one can always dispense with "the mathematics he uses" (Field 1980, 14).

Now, it can be proved that mathematics is conservative over physics. For, "there are some fairly general strategies that can be employed to purge" physical "theories of all reference to mathematical entities" (Field 1989, 18). The paradigmatic example of this is "a rewritten version of Newtonian physics" in which individual variables range over concrete objects, which amounts to "formulating it without mathematical entities" (ibid., 18, footnote 11).

The conception that mathematical objects are fictions, however, has the following shortcomings.

1. Field claims that all there is to mathematics is that we have a good story about natural numbers, another good story about sets, and so forth, and that a mathematical proposition is true only in the same sense in which the proposition "Oliver Twist lived in London" is true. But Dickens did not first introduce the character Oliver Twist and then investigated him, discovering that Oliver was a small boy whose mother dies the moment he is born, who lives in a poorhouse and ends up getting involved with a gang of thieves, but eventually is adopted by a kind-hearted and rich man. Dickens made up the character Oliver Twist with all his features through his story. On the contrary, a mathematician first introduces certain mathematical objects and then investigates them, discovering some of their properties, afterwards some other mathematician discovers other properties, and so on. Therefore, mathematical objects are essentially different from characters in a story.
2. Field claims that it can be proved that mathematics is conservative over physics, but this claim is unwarranted. For, let M be a mathematical theory that is added to a physical theory N which makes no reference to mathematical entities. In order to prove that mathematics is conservative over physics, one should prove that, if a sentence A of N is provable in $N + M$, then A is provable already in N. But, as Shapiro points out, since there is a one-to-one correspondence between space-time points and quadruples of real numbers, "it is possible to model the natural numbers in space-time and, in effect, to do arithmetic in N. To put it differently, the natural number structure is exemplified in the universe of space-time" (Shapiro 1983, 526). Then "the Gödel's incompleteness theorems apply" (ibid., 527). Therefore, there is a sentence A of N which is provable in $N + M$ but not in N. Thus, the mathematical theory M is not conservative over the physical theory N, which refutes the fictionalist "argument for the conservativeness of mathematics over physics" (ibid., 528).
3. Field claims that the paradigmatic example of the fact that there are some fairly general strategies that can be employed to purge physical theories of all reference to mathematical entities is a rewritten version of Newtonian physics in which individual variables range over concrete objects. But this claim is unwarranted. For, in order to make a rewritten version of Newtonian physics in which individual variables range over concrete objects, one is bound to introduce entities having properties of mathematical entities, rather than properties of concrete

objects. For example, the postulates of Field's theory of space-time entail that there is a one-to-one correspondence between space-time points, conceived of as physical entities, concrete and not abstract, and quadruples of real numbers. Since real numbers are uncountable, this involves postulating uncountably many physical entities. According to Field, this is not objectionable, nor "does it become any more objectionable when one postulates that these physical entities obey structural assumptions analogous to the ones that platonists postulate for the real numbers" (Field 1980, 31). But, then, what Field calls "physical entities" are really mathematical entities, and hence the rewritten version of Newtonian physics does not really purge Newtonian physics of all reference to mathematical entities.

9 Mathematical Objects as Idealizations of Sensible Things

Another answer to the question "What is the nature of mathematical objects?" is that they are ideal objects obtained from sensible things not by abstraction but by idealization. The latter is the process by which one pushes certain attributes of sensible things to the limit, for example, the concept of sphere is obtained by pushing certain attributes of a ball to the limit. Thus, while objects obtained by abstraction from sensible things have attributes which also occur in sensible things, objects obtained by idealization have attributes which do not occur in sensible things. This conception of mathematical objects has been put forward by Husserl.

Indeed, Husserl says that "the mathematical realm" is "a world of ideal objects" (Husserl 2008, 49). These ideal objects, the mathematical objects, are the result of "idealizing the world of bodies in respect to what has spatiotemporal shape in this world," so such idealization "created ideal objects" (Husserl 1970, 32). Being the result of idealizing the world of bodies, mathematical objects express "something which cannot be 'seen'," they are "essences drawn immediately from intuition simpliciter" as "ideal essences" (Husserl 1983, 166). The ideal essences are "ideal 'limits' which it is essentially impossible to find in any sensuous intuition" (ibid., 167). They are the product of an idealization, which "is fundamentally different in its essence from the seizing upon an essence by simple 'abstraction'" (ibid., 167).

The conception that mathematical objects are idealizations of sensible things, however, has the following shortcomings.

1. Husserl claims that the mathematical realm is a world of ideal objects. This raises the question: If the mathematical realm is a world of ideal objects, why is mathematics applicable to the real world? Husserl's answer is that it is applicable to it because, "through Galileo's mathematization of nature, nature itself is idealized" and "becomes" a "mathematical manifold" (Husserl 1970, 23). So, Galileo substitutes a "mathematically substructed world of idealities for the only real world, the one that is actually given through perception, that is ever experienced and experienceable" (ibid., 49). Thus, Galileo conceives the "idea of a rational infinite totality of being with a rational science systematically mastering

it" (ibid., 22). Then, however, Husserl does not really account for the applicability of mathematics to the real world, but only for the applicability of mathematics to a mathematically substructed world of idealities. So, the problem is not solved but only transformed in the problem of why, if mathematics is applicable to a mathematically substructed world of idealities, it is also applicable to the world.

2. Husserl claims that mathematical objects are the product of an idealization of sensible things, which is fundamentally different in its essence from the seizing upon an essence by simple abstraction. Being the product of an idealization, mathematical objects have attributes which do not occur in sensible things. Then, however, Husserl not only does not really account for the applicability of mathematics to the real world, but even trivializes the question of the applicability. For, if mathematical objects have attributes which do not occur in sensible things, then all mathematical propositions are vacuously true in the real world. For example, such is the proposition "In any triangle, the interior angles are equal to two right angles." For, the proposition is of the form $P(x) \rightarrow Q(x)$, where $P(x)$ expresses "x is a triangle" and $Q(x)$ expresses "The interior angles of x are equal to two right angles." Since, according to Husserl, triangles have attributes which do not occur in sensible thing, $P(x)$ is false and hence $P(x) \rightarrow Q(x)$ is vacuously true in the real world. For the same reason, even the opposite proposition, "In any triangle, the interior angles are not equal to two right angles," being of the form $P(x) \rightarrow \neg Q(x)$, is vacuously true in the real world.

10 Mathematical Objects as Idealizations of Operations

Another answer to the question "What is the nature of mathematical objects?" is that they are idealizations of operations of collecting and segregating performed by an ideal subject. This conception of mathematical objects has been put forward by Kitcher.

Indeed, Kitcher says that mathematical objects are idealizations "of operations of collecting" and segregating "which we are able to perform with respect to any objects" (Kitcher 1983, 12). Or rather, since, "given our biological limitations, the operations in which we actually engage are limited," mathematical objects are idealizations of operations of collecting and segregating performed by an "ideal subject, whose status as an ideal subject resides in her freedom from certain accidental limitations imposed on us" (ibid., 109). The ideal subject is not a mysterious being, but only "an idealization of ourselves" (ibid., 111). The relation between the operations of the ideal subject and "the actual operations of human agents parallels that between the laws of ideal gases and the actual gases which exist in our world" (ibid., 109). Ideal gases do not exist in our world, and yet "actual gases approximately satisfy the condition set down in the Boyle-Charles Law" of ideal gases; moreover, "if the molecules of gases were of negligible size and if there were no intermolecular forces then gases would obey the Boyle-Charles Law" (ibid., 117). This "prompts us to abstract from certain features of the actual situation, introducing the notion of an ideal gas to describe how actual gases would behave if complicating

factors were removed" (ibid.). Much in the same way, the ideal subject does not exist in our world, and yet the actual operations of collecting and segregating of human agents approximate the operations of the ideal subject, and moreover, if human agents were free from their biological limitations, their operations would coincide with those of the ideal subject. This prompts us to "specify the capacities of the ideal agent by abstracting from the incidental limitations on our own collective practice" (ibid.).

That mathematical objects are idealizations of operations of collecting and segregating also includes sets. According to Kitcher, sets are members of the hierarchy which "arises from the iterated performance of one fundamental operation, that of collecting," and its description "may be viewed as a literal account of the iterated constructive activity of the ideal mathematical subject" (ibid., 133). This amounts "to ascribe to the ideal subject an ability to perform iterative collective activity through an infinite sequence of stages" (ibid., 147). Since "the sequence of stages at which sets are formed is highly superdenumerable," it follows that, if "each of the stages corresponds to an instant in the life of the constructive subject," then "the subject's activity is carried out in a medium analogous to time, but far richer than time. (Call it 'supertime')" (ibid., 146). Admittedly, this means "to idealize still further from our own thoroughly finite performances" (ibid., 147). But such idealization "is no different in principle from earlier idealizations in which we abstract from our own mortality or from our inability to survey infinite domains. The conception of the ideal subject as an idealization of ourselves does not lapse when we release the subject from the constraints of our time" (ibid.). Thus, to talk of such an ideal subject "is not to suppose that there is a mysterious being with superhuman powers" (ibid., 110).

The conception that mathematical objects are idealizations of operations of collecting and segregating that an ideal subject performs, however, has the following shortcomings.

1. Kitcher claims that, as ideal gases do not exist in our world and yet actual gases approximately satisfy the laws of an ideal gas, so the ideal subject does not exist in our world, nevertheless the actual operations of collecting and segregating of human agents approximate the operations of the ideal subject. But this overlooks that actual gases very closely approximate ideal gas behavior at high temperatures and low pressures, and the degree of approximation can be calculated. On the contrary, the ideal subject is supposed to be able to perform certain operations, such as all possible collecting of the members of an uncountable totality, which are far beyond the possibilities of any human agent.
2. Kitcher claims that the description of the hierarchy of sets may be viewed as a literal account of the iterated constructive activity of the ideal subject. Now since, as Kitcher admits, the sequence of stages at which sets are formed is highly superdenumerable, this entails that the ideal subject should be able to perform superoperations in a super-time. But, while actual gases approximately satisfy the laws

of an ideal gas, the structure of our time by no means approximates the structure of such super-time since the latter is highly super-denumerable.

Actually, an ideal subject capable of performing super-operations in a super-time would be not only super-human but also super-divine. Indeed, as Parsons points out, the mind of such an ideal subject "would differ not only from finite minds but also from the divine mind as conceived in philosophical theology" (Parsons 2005, 273). For, in the latter, either the divine mind "is thought of as in time, and therefore as doing things in an order with the same structure as that in which finite beings operate, or its eternity is interpreted as complete liberation from succession" (ibid.).

3. Kitcher claims that to view the description of the hierarchy of sets as a literal account of the iterated constructive activity of the ideal mathematical subject means to make an idealization that is no different in principle from earlier idealizations in which we abstract from our own mortality or from our inability to survey infinite domains. Hence, Kitcher concludes that it is unjustified to say that to talk of such an ideal subject is not to suppose that there is a mysterious being with superhuman powers. But this overlooks that an ideal subject who constructed the hierarchy of sets would be a being whose mind would differ not only from human minds but, by what has been said above, also from the divine mind. Therefore, to talk of such an ideal subject would be just to suppose that there is a mysterious being with superhuman powers.

11 The Basis for an Alternative Conception

Since all conceptions of mathematical objects discussed above have substantial shortcomings, one may wonder whether there is an alternative conception not subject to them. However strange this may sound, the basis for such an alternative conception of mathematical objects can be found in Plato.

Admittedly, as we have seen, Plato says that mathematical objects are ideas in themselves, which have a kind of reality not perceptible by any sense and such that only intuition can contemplate it.

But Plato also says that mathematicians, as long as they are alive, cannot contemplate this kind of reality by intuition. For, as long as they have the body and their soul is contaminated by such an evil, they will never adequately gain the possession of what they desire, that is, truth. Only after death, mathematicians can behold "the happy vision and contemplation" of the reality in question, having "calm and happy visions, in pure light" (Plato, *Phaedrus*, 250 b 8–c 4).

Since mathematicians, as long as they are alive, cannot contemplate the kind of reality of which mathematical objects consist by intuition, in order to solve mathematical problems, they introduce certain mathematical objects as hypotheses.

Indeed, Plato says that, in order to solve a mathematical problem, mathematicians must "make use of a hypothesis" (Plato, *Meno*, 86 e 3). In fact, "those who deal with

geometry, arithmetic, and the like, hypothesize the odd and the even, geometrical figures, the three kinds of angles, and other things akin to these, depending on the subject of their investigation. They take these as known and assume them as hypotheses" (Plato, *Republic*, VI 510 c 2–6). So, they arrive at knowing the properties of these objects, not by intuition, but by making hypotheses, hence through discursive thought. Thus, it is "discursive thought, not intuition, the state of geometry and the like, namely, thought intermediate between opinion and intuition" (ibid., VI 511 d 2–5).

According to Plato, this is the only way in which mathematicians, as long as they are alive, can solve mathematical problems.

12 The Heuristic Conception of Mathematical Objects

Plato's statements suggest that mathematical objects are hypotheses human beings make to solve mathematical problems. Specifically, they are hypotheses human beings make to solve mathematical problems by the analytic method.

The analytic method is the method according to which, to solve a problem, one looks for some hypothesis that is a sufficient condition for the solution of the problem. The hypothesis is obtained from the problem, and possibly other data, by some nondeductive rule and must be plausible, that is, the arguments for it must be stronger than the arguments against it, on the basis of experience. But the hypothesis is, in turn, a problem that must be solved and is solved in the same way. That is, one looks for another hypothesis that is a sufficient condition for the solution of the problem posed by the previous hypothesis. The new hypothesis is obtained from the previous hypothesis, and possibly other data, by some nondeductive rule, and must be plausible. And so on, ad infinitum. Therefore, the solution to a problem is a potentially infinite process (For details on the analytic method, see Cellucci 2013).

The status of mathematical objects is similar to that of the theoretical terms of physical theories, which are hypotheses human beings make to solve physical problems by the analytic method.

This is the heuristic conception of mathematical objects.

Of course, in the heuristic conception, since mathematical objects are hypotheses human beings make to solve mathematical problems by the analytic method, the hypotheses in question must be plausible, that is, such that the arguments for them are stronger than the arguments against them, on the basis of experience. Only mathematical objects given by hypotheses which are plausible can be said to "exist." But this does not mean that they exist literally, in the physical or in a superphysical world. For, hypotheses are only ideas conceived by the mind. On the other hand, not all ideas conceived by the mind can be accepted as mathematical objects, but only those which are plausible. Therefore, that mathematical objects exist means that they are given by hypotheses which are plausible.

13 Some Remarks about the Heuristic Conception

Some remarks about the heuristic conception of mathematical objects may be useful.

1. That mathematical objects are hypotheses human beings make to solve mathematical problems by the analytic method, does not mean that they are mental constructions in Brouwer's sense. For, hypotheses are simply ideas conceived by the mind. Conversely, mental constructions in Brouwer's sense are entities whose construction is accounted for in terms of acts of the mind based on intuition, which are "acts of mathematical construction" (Brouwer 1975, 88). This makes a great difference in practice. For example, actual infinities can be introduced as hypotheses to solve mathematical problems, but they are not mental constructions in Brouwer's sense, because they are "certainly not built up mathematically" (ibid., 80).
2. That mathematical objects are hypotheses human beings make to solve mathematical problems by the analytic method, does not mean that they are fictions. For, hypotheses are introduced to deal with the world, so they must agree with experience and are accepted until strong arguments against them emerge. Conversely, fictions are not introduced to deal with the world, so they are not supposed to agree with experience, and are accepted despite our being aware that they have no basis in experience. As Vaihinger, the founder of philosophical fictionalism, says, "whereas every hypothesis seeks to be an adequate expression of some reality still unknown and to mirror this objective reality correctly, the fiction is advanced with the consciousness that it is an inadequate" manner "of conception, whose coincidence with reality is, from the start, excluded" (Vaihinger 2008, 268).
3. That mathematical objects are hypotheses human beings make to solve mathematical problems by the analytic method, means that they are mental-cultural human products. To solve some mathematical problem, one or more mathematicians may conceive the idea of certain new mathematical objects, and use such idea as hypothesis to solve the problem by the analytic method. Afterwards, other mathematicians may find the hypothesis useful to solve other problems by the analytic method. And so on. Thus, in addition to being conceived by the mind of one or more mathematicians, mathematical objects become part of the cultural process as cultural products. Therefore, as White says, the introduction of new mathematical objects is "an event that takes place" both "in one or more nervous systems" and "in the cultural tradition," where "the human nervous system" is "the catalyst which makes the cultural process possible" (White 1947, 303). Of course, like all other cultural products. Newly introduced mathematical objects may encounter resistance. For example, such has been the case of the number zero, negative numbers, infinitesimals, complex numbers, actual infinities.
4. The hypotheses by which mathematical objects are introduced characterize their identity, but do not entail that those mathematical objects exist – just as saying that hippogriffs are animals having the foreparts of a griffin and the body of a

horse characterizes their identity, but does not entail that hippogriffs exist. Identity does not entail existence. As already said above, in order to say that mathematical objects exist, one must show that the hypotheses by which they are given are plausible.

5. That mathematical objects are introduced as hypotheses by some mathematician does not mean that such mathematician knows all their properties. Gödel says that "the creator necessarily knows all properties of his creatures, because they can't have any others except those he has given to them" (Gödel 1986–2002, III, 311). But it is not so. A painter does not necessarily know all properties of his creatures, otherwise there would be no need for art critics. It is one thing to introduce mathematical objects as hypotheses, and another thing to discover all their properties. For example, "when Cardano introduced his complex numbers, he could have had no inkling of the many magical properties which were to follow – properties which go under various names, such as the Cauchy integral formula, the Riemann mapping theorem, the Lewy extension property" (Penrose 1991, 96). These properties, and many others, do not pertain to later modifications of the concept of complex number, but are "properties of the very numbers" that "Cardano had first encountered in about 1539" (ibid.). Cardano could have had no inkling of such properties because mathematical objects, entering into relation with other mathematical objects, acquire new properties, other than those the creator has given to them, and to discover them requires an investigation.

14 Conclusion

It has been said above that the hypotheses by which mathematical objects are introduced characterize their identity. This, however, does not mean that the hypotheses characterize their identity completely and conclusively. The identity may undergo changes, because mathematical objects are not closed in themselves, they are open to interactions with other mathematical objects, from which they gain new determinations.

Moreover, mathematical objects introduced to solve a particular problem may turn out to be useful to solve other problems. Such objects, then, consolidate, crystallize, and acquire a stability which makes them amply independent of the problem for which they were originally introduced, and become subjects of study themselves.

For example, Menaechmus introduced conic sections to solve the problem of finding two mean proportional lines between two given lines (see Cellucci 2013, Sect. 4.9). But, afterwards, conic sections turned out to be useful to solve other problems, and eventually became subjects of study themselves as in Apollonius of Perga's treatise *Conics*, one of the greatest scientific works from the ancient world.

When mathematical objects become subjects of study themselves, they may appear so familiar to us that they may seem to be independently existing entities. We may have the impression that, as Hermite says, they are not "arbitrary product of our minds," but entities which "exist outside of us with the same character of

necessity as the objects of objective reality; and we find or discover them and study them as do the physicists, chemists and zoologists, etc." (Hermite 1905, 398). But this is an illusion, which arises from our great familiarity with them. Mathematical objects are not independently existing entities, because the conception that mathematical objects are independently existing entities has the severe shortcomings already spelled out in detail earlier in this chapter.

On the other hand, other mathematical objects, also introduced to solve a certain mathematical problem, do not consolidate nor acquire stability. They hang fluttering between being and non-being, and eventually disappear from mathematics.

For example, Newton and Leibniz introduced infinitesimals to solve problems concerning continuous change. To that aim, Newton said it was useful to imagine "infinitely small additions" by which "quantities increase during each infinitely small interval of time" (Newton 1967–1981, III, 81). Leibniz said it was "useful to imagine infinitely small quantities, or, as Nicolaus Mercator called them, infinitesimals, which may not be taken to be zero when the assignable ratio between them is sought," but "are rejected whenever they are added to incomparably greater quantities" (Leibniz 1971, IV, 63). But infinitesimals never acquired stability, no precise definition of them was ever given, and they were even assigned contradictory determinations. Therefore, they were the object of severe criticism, such as that of Berkeley who said that infinitesimals "are neither finite quantities, nor quantities infinitely small, nor yet nothing," and hence may be called "the ghosts of departed quantities" (Berkeley 1992, 199). As a result, infinitesimals never crystallized and became subjects of study themselves, but hung fluttering between being and non-being, and in the nineteenth century disappeared from mathematics, only to be revived in the twentieth century by Robinson, in a completely different form, as "the infinitely small" numbers "of a non-standard model of Analysis" (Robinson 2016, 282).

That mathematical objects can gain new determinations from interactions with other mathematical objects, and may acquire stability becoming subjects of study themselves or may not crystallize eventually disappearing from mathematics, shows that mathematical objects have an open-ended character because such is the character of the hypotheses by which they are introduced.

15 Cross-References

- ▶ Heuristics and Mathematical Practice
- ▶ Platonism, *De Re*, and (Philosophy of) Mathematical Practice
- ▶ What Mathematicians Do: Mathematics as Process and Creative Rationality

References

Berkeley G (1992) *De motu* and *The analyst*. Springer, Dordrecht

Bernays P (1935) Sur le platonisme dans les mathématiques. L'Enseignement Mathematique 34:52–69

Borwein J, Bailey D, Girgensohn R (2004) Experimentation in mathematics: computation paths to discovery. A K Peters, Natick
Brannon E (2005) The independence of language and mathematical reasoning. Proc Natl Acad Sci 102:3177–3178
Brouwer LEJ (1975) Collected works, vol 1. North-Holland, Amsterdam
Brouwer LEJ (1998) The structure of the continuum. In: Mancosu P (ed) The debate on the foundations of mathematics in the 1920s. Oxford University Press, Oxford, pp 54–63
Cellucci C (2013) Rethinking logic: logic in relation to mathematics, evolution, and method. Springer, Berlin
Dieudonné J (1970) The work of Nicholas Boubaki. Am Math Mon 77:134–145
Ferreiros J (1999) Labyrinth of thought. Birkhäuser, Cham
Ferreiros J, Gray JJ (eds) (2006) The architecture of modern mathematics: essays in history and philosophy. Oxford University Press, Oxford
Field H (1980) Science without numbers. Blackwell, Oxford
Field H (1989) Realism, mathematics and modality. Blackwell, Oxford
Frege G (1960) The foundations of arithmetic. Harper, New York
Frege G (1984) Collected papers on mathematics, logic, and philosophy. Blackwell, Oxford
Frege G (2013) Basic laws of arithmetic. Oxford University Press, Oxford
Gillies D (2014) Should philosophers of mathematics make use of sociology? Philos Math 22:12–34
Gödel K (1986–2002) Collected works. Oxford University Press, Oxford
Hale B, Wright C (2001) The reason's proper study: essays towards a neo-Fregean philosophy of mathematics. Oxford University Press, Oxford
Hermite C (1905) Letter to Stieltjes, 13 May 1894. In: Hermite C, Stieltjes TJ (eds) Correspondance, vol 2. Gauthier-Villars, Paris, pp 398–400
Hilbert D (1967a) On the infinite. In: van Heijenoort J (ed) From Frege to Gödel: a source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, pp 367–392
Hilbert D (1967b) The foundations of mathematics. In: van Heijenoort J (ed) From Frege to Gödel: a source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, pp 464–479
Hilbert D (1996a) The new grounding of mathematics: first report. In: Ewald W (ed) From Kant to Hilbert: a source book in the foundations of mathematics, vol 2. Oxford University Press, Oxford, pp 1117–1134
Hilbert D (1996b) The logical foundations of mathematics. In: Ewald W (ed) From Kant to Hilbert: a source book in the foundations of mathematics, vol 2. Oxford University Press, Oxford, pp 1134–1148
Hilbert D (1996c) The grounding of elementary number theory. In: William EW (ed) From Kant to Hilbert: a source book in the foundations of mathematics, vol 2. Oxford University Press, Oxford, pp 1149–1157
Hilbert D (1996d) Logic and the knowledge of nature. In: Ewald W (ed) From Kant to Hilbert: a source book in the foundations of mathematics, vol 2. Oxford University Press, Oxford, pp 1157–1165
Hilbert D, Bernays P (1968–1970) Grundlagen der Mathematik. Springer, Berlin
Husserl E (1970) The crisis of European sciences and transcendental phenomenology. Northwestern University Press, Evanston
Husserl E (1983) Ideas pertaining to a pure phenomenology and to a phenomenological philosophy I. Nijhoff, The Hague
Husserl E (2008) Introduction to logic and theory of knowledge: lectures 1906/07. Springer, Cham
Kant I (1992) Lectures on logic. Cambridge University Press, Cambridge
Kant I (1998) Critique of pure reason. Cambridge University Press, Cambridge
Kitcher P (1983) The nature of mathematical knowledge. Oxford University Press, Oxford
Leibniz GW (1971) Mathematische Schriften. Olms, Hildesheim
Newton I (1967–1981) Mathematical papers. Cambridge University Press, Cambridge

Parsons C (2005) Mathematics in philosophy: selected essays. Cornell University Press, Ithaca
Penrose R (1991) The emperor's new mind. Concerning computers, minds, and the laws of physics. Penguin, London
Robinson A (2016) Non-standard analysis. Princeton University Press, Princeton
Russell B (1992) The philosophical importance of mathematical logic. In: Russell B (ed) Collected Papers, Volume 6: logical and philosophical papers, 1909–1913. Routledge, London, pp 32–40
Russell B (1993) Introduction to mathematical philosophy. Dover, Mineola
Russell B (2010) Principles of mathematics. Routledge, London
Shapiro S (1983) Conservativeness and incompleteness. J Philos 80:521–531
Shapiro S (1997) Philosophy of mathematics: structure and ontology. Oxford University Press, Oxford
Shapiro S (2000) Thinking about mathematics: the philosophy of mathematics. Oxford University Press, Oxford
Shapiro S (2004) Foundations of mathematics: metaphysics, epistemology, structure. Philos Q 54:16–37
Shapiro S (2008) Mathematical objects. In: Gold B, Simons RA (eds) Proof and other dilemmas: mathematics and philosophy. Mathematical Association of America, Washington, pp 157–177
Vaihinger H (2008) The philosophy of 'as if'. Routledge, London
von Neumann J (1955) Mathematical foundations of quantum mechanics. Princeton University Press, Princeton
Wang H (1996) A logical journey: from Gödel to philosophy. The MIT Press, Cambridge
Weyl H (1950) The theory of groups and quantum mechanics. Dover, Mineola
Weyl H (1998) The current epistemological situation in mathematics. In: Mancosu P (ed) From Brouwer to Hilbert: the debate on the foundations of mathematics in the 1920s. Oxford University Press, Oxford, pp 123–142
White LA (1947) The locus of mathematical reality: an anthropological footnote. Philos Sci 14:289–303
Whitehead AN, Russell B (1925–1927) Principia mathematica. Cambridge University Press, Cambridge

The Ghost and the Spirit of Pythagoras

The Twentieth- and Twenty-First-Century Mathematics Between and Beyond Geometry and Algebra

Arkady Plotnitsky

Contents

1 Introduction: The Ghost of Pythagoras 64
2 The Spirit of Pythagoras: Geometry and Algebra – From Modernity to Modernism 69
3 Radical Pythagorean Mathematics and Mathematical Modernism 74
4 The Thinkable, the Unthinkable, and the Real in Mathematics and Radical Pythagorean Mathematics 83
5 Summary/Conclusion 87
6 Cross-References 88
References 88

Abstract

This chapter introduces the concept of *radical* Pythagorean mathematics, as part of mathematical modernism, which emerged around 1900 and extends to our own time. Arguably, the greatest discovery of Pythagorean mathematics of (sixth-fifth century BC) was that of the incommensurability of the side and the diagonal of the square. This discovery established the complex relationships between geometry and arithmetic and, correlatively, the possibility of a type of logos, a logos containing the alogon within it – a logical thought containing the unthinkable within it. I shall call the mathematics defined by this combination "Pythagorean mathematics." I shall argue that, as part of mathematical modernism, there emerged a new attitude toward the unthinkable in thought, an attitude that accepts this situation as a positive, enabling condition of thinking and knowledge. I shall call the mathematical thinking defined by this attitude, while keeping the irreducible relationship between geometry and algebra (which subsumes arithmetic within it), *radical* Pythagorean mathematics. The original Pythagorean attitude, at

A. Plotnitsky (✉)
Literature, Theory, and Cultural Studies Program, Philosophy and Literature Program, Purdue University, West Lafayette, IN, USA
e-mail: plotnits@purdue.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_7

least in one reconstitution of Pythagorean thought, was, as against the Platonist attitude, tolerant of this condition, but still sought to overcome it by finding a logos without the alogon within it. This attitude, often sliding into a Platonist one, has remained dominant throughout the history of mathematics, physics, and philosophy, including during the modernist period, notable as much for the presence of the radical Pythagorean thinking as for a resistance to it.

Keywords

Algebra · Geometry · Mathematical modernism · Pythagorean mathematics · Radical Pythagorean mathematics · Quantum theory

1 Introduction: The Ghost of Pythagoras

Pythagorean mathematics (the sixth–fifth century BC) had a profound influence on the subsequent history of mathematics and physics, from ancient Greece to our own time. The Pythagorean theorem alone would be sufficient to testify to this influence. To name only a few important examples of its role beyond Euclidean geometry, it figured in calculus (a tangent to a curve), complex numbers (the modulus of a complex number), differential and then Riemannian geometry (a tangent space to a manifold and a tangent bundle), and vector spaces, in mathematics. In physics, it was indispensable to all three main areas of modern (mathematical-experimental) physics: classical physics, special and general relativity theory, and finally, and most pertinently to my argument here, quantum mechanics, where the analog of the Pythagorean theorem in Hilbert spaces (vector spaces, possibly of infinite dimensions) over **C**, which define the mathematical formalism of the theory, plays a key role.

Arguably, the greatest, or at least, conceptually most remarkable, discovery of Pythagorean mathematics was, however, that of the incommensurability (alogon) of the side and the diagonal of the square, actually a consequence of the Pythagorean theorem as well. This discovery led to a great crisis, responsible for the shift from arithmetic to geometry as the *dominant* way of mathematical thinking in ancient Greek mathematics, eventually making possible a theory of proportion applicable to incommensurable as well as commensurable magnitudes, and with it, establishing the more complex relationships between geometry and arithmetic. Obviously (hence, my emphasis), arithmetic never left the stage of and remained indispensable to ancient Greek mathematics. This complexity has accompanied the relationships between geometry and arithmetic, and eventually algebra, throughout the subsequent history of mathematics until our own time, sometimes returning algebra to a more dominant role in mathematics. It had, however, taken more than two millennia before a proper theory of irrational numbers was developed by R. Dedekind and others, although algebra had of course worked with irrational numbers long before then.

Conceptually and philosophically, this discovery established, against the Pythagoreans' own grain, at least a *possibility* of new type of logos, a logos containing the alogon within it – mathematical or any logical thought containing within it that which is beyond it, or in our Latinized terms, the rational containing the irrational within itself. My double qualification, "against the Pythagoreans' own grain" and a "possibility," is prompted by the complexities, to be considered presently, of strictly attributing this logos to the Pythagoreans or, assuming the Pythagoreans entertained this possibility (as they, as I shall suggest, might have), of aiming to establish the scope of such a logos in their thinking. However, even establishing a possibility of such a logos was a momentous event. I shall call the mathematics defined by this possibility, along with the complex interplay of geometry and arithmetic, and eventually algebra, "Pythagorean mathematics" (to be distinguished from mathematical Pythagoreanism (e.g., Kahn 2001)). I shall also argue that, as part of mathematical modernism, which took the center stage of mathematics around 1900 and has continued to occupy it until our own time, there emerged a *new attitude* toward the unthinkable in thought (the alogon in a logos), in mathematics itself and in physics, as, from Galileo on, a mathematical science of nature. This attitude accepts the presence of the alogon within a logos as an enabling condition of thinking and knowledge. I shall call the mathematical thinking, in mathematics and physics, accompanied by this attitude, while keeping the irreducible relationship between geometry and algebra (which subsumes arithmetic within it), "radical Pythagorean mathematics."

The "original" Pythagorean attitude *appears* to have been, as against the Platonist attitude, *tolerant* of this condition, but saw it as an obstacle to thought and sought to overcome it in favor of a logos without the alogon within it. The actual thinking of the Pythagoreans themselves concerning the subject is, again, difficult to ascertain, and it might be debated (hence, my quotation marks around "original" and my emphasis on "appears"). This makes it more prudent to speak of the *Pythagorean* attitude rather than that of any particular Pythagorean figure and similarly of the *Platonist* attitude (which I shall contrast to the Pythagorean one) rather than Plato's own, incomparably greater textual evidence for assessing it notwithstanding, given that Plato's views on this (or most other subjects) are a matter of interpretation and debate as well. Still, while each attitude qua attitude may be ascertained with more (even if not complete) certainty and evidence in the work of many specific figures during the modern (from roughly sixteenth century on) and then the modernist period, when referring to actual ancient Greek thinking, one deals with a reconstitution of this thinking, a reconstitution that, while it may be plausible, may be questioned in one degree or another. Given this qualification it would be more accurate to speak of a *conservative* rather than, as I did above, *original* Pythagorean attitude, but for simplicity, I shall from now on use the Pythagorean attitude, thinking, or mathematics, only qualifying when necessary, while always using the *radical* Pythagorean attitude, thinking, or mathematics. The (conservative) Pythagorean attitude, often sliding into the Platonist one, has remained dominant throughout the history of mathematics, including in twentieth- and twenty-first-century mathematics, the epoch of mathematical modernism, in which, amidst a more

overt dominance of algebra, geometry, in a new synergy with algebra, came to play a new role and gave rise to new forms of geometry as well. The same attitude has been no less prevalent throughout the history of physics, from Descartes and Galileo (or indeed Aristotle) on, or the history of philosophy, from and before Plato on, until our own time. While it emerged along with modernism in mathematics and physics, radical Pythagorean attitude, advocated as it is here, has always been and remains a minority position, and the modernist period was notable as much for the presence of radical Pythagorean thinking as for a resistance to it.

The discovery of the incommensurables undermined the Pythagorean belief that the harmony of the ultimate nature of reality could be expressed in terms of numbers and their commensurable ratios and led to a shift from the dominance of arithmetical to that of geometric thinking. The discovery and the narrative of this discovery were accompanied by a narrative, a legend, of the fate of the one who made this discovery public, sometimes identified with the discoverer. Hippasus is often mentioned in either or both roles, but others are mentioned as well. The account is given in the first scholium of Book X of *Euclid's Elements*. According to T. Heath:

> The scholium quotes . . . the legend according to which "the first of the Pythagoreans who made public the investigation of these matters perished in a shipwreck," conjecturing that the authors of this story "perhaps spoke allegorically, hinting that everything irrational and formless is properly concealed, and, if any soul should rashly invade this region of life and lay it open, it would be carried away into the sea of becoming and be overwhelmed by its unresting current." There would be a reason also for keeping the discovery of irrationals secret for the time in the fact that it rendered unstable so much of the groundwork of geometry as the Pythagoreans had based it upon the imperfect theory of proportions which applied only to numbers. . .. [T]he discovery of incommensurability must have necessitated a great recasting of the whole fabric of elementary geometry, pending the discovery of the general theory of proportion applicable to incommensurable as well as to commensurable magnitudes. (Euclid 1989, vol. 3, p. 1)

This recasting becomes especially apparent when one compares Book V and Books VII–IX, which bring to the *Elements* theories of proportion, one for numbers and one for all magnitudes. As a result, there also emerged the separation between continuous (geometrical) and discontinuous (numerical) entities and thus geometry and arithmetic. (I am grateful to an anonymous reader of this chapter for suggesting this formulation and for directing my attention to the books of the *Elements* mentioned here.) Neither separation, I argue, has ever become complete but instead has continued to be underlain by a complex interplay of continuity and discontinuity, and geometry and arithmetic and then algebra. I further argue that this interplay has never been brought to a complete synthesis either in the history of mathematics from that point on, a history that, on the other hand, has multiplied this interplay and the forms of geometry and arithmetic and then algebra (itself a product of this unfolding multiplicity) themselves. The same, in part correlatively, might be said for the history of mathematical (or philosophical) concepts of continuity and discontinuity, and their interplay, which took an entirely unexpected form with modernism, especially in set theory in mathematics and quantum theory in physics, which, as explained

below, related continuous mathematics to discrete physical phenomena in a way never found or even contemplated in physics previously.

A remarkable aspect of the discovery of the incommensurables was the capacity of mathematics to discover the alogon within it by a rigorous mathematical procedure, by the logos of mathematics itself. In this respect, it was akin to K. Gödel's incompleteness theorems. These theorems proved mathematically the incapacity of a mathematical system, if it is rich enough to contain arithmetic (as is most of mathematics), to mathematically account for itself, by demonstrating all of its propositions to be provable as either true or false, which is Gödel's first incompleteness theorem, and, as a consequence, to be able to mathematically establish its consistency, which is Gödel's second incompleteness theorem. The second theorem still allows that such a system could be proven to be inconsistent. I shall discuss the significance of Gödel's findings in the present context in Sect. 4.

Once, however, the irrational within the rational, the alogon within a logos, is established, the question arises concerning the status of and the attitude toward this irrationality. Plato is the case in point in ancient Greek thinking, in his concern with the logos of philosophy as a logos without the alogon within it, against the mathematical, Pythagorean, logos of his time, that emerged, even if, again, only as a possibility, with the discovery of the incommensurables. The subject is discussed in several dialogues, such as *Meno*, *Theaetetus*, *Statesman*, and the *Republic*. While Plato conceived of the ultimate architecture of the ideal realm on the *mathematical* model, he distinguished the philosophical thinking he aimed at from *mathematics*, at least the mathematics as practiced by mathematicians of his time. Plato thought that only philosophy, and specifically dialectic, could achieve, for humans, the maximally close proximity to the ultimate architecture of (ideal) reality. He expressed doubts that mathematics as a technical discipline, or at least mathematicians could do so, given their (Pythagorean) attitude, too tolerant of the presence of the alogon with the logos of mathematics, however, circumvented this presence may be. According to S. Negropontis, citing the *Republic*:

> "But, surely," said I, "if you should ever nurture in fact your children whom you are now nurturing and educating in logos ("toi logoi"), you would not suffer them, I presume, to hold rule in the state, and determine the greatest matters, being themselves as irrational ("alogous") as the lines ("grammas"), so called in geometry." "Why, no," he said. ([*Republic*] 534d3-7)

> This is, not just a silly linguistic quip, as some Platonists think but a quite revealing passage that contrasts the philosophic "logos" to the geometric "alogoi" lines. According to the definition, which is due to Theaetetus and appears in Book X of the Elements, a line a is "alogos" (literally, "ratio–less") with respect to a given ("protetheisa") line b if a^2 and b^2 are incommensurable. (Negropontis 2019, p. 23)

Negropontis then argues that, while Plato adopted Theaetetus' mathematical *theory*, he also wanted to transfer it into dialectic so as to avoid the alogon (Negropontis 2019, p. 24). I leave aside technical details of this problematic, extensively

considered by Negropontis and other commentators. My point here is that ancient Greek thought appears to have contemplated or at least introduced the possibility of two forms of logos, Pythagorean and Platonist. The first is that of a logos inhabited by the alogon, and the second is that of a pure logos without the alogon in it and, in considering the first type, the aspiration of overcoming the alogon within a logos, an aspiration equally dominant in mathematics, physics, and philosophy throughout their history. I shall, accordingly, define Platonism as the aspiration toward a logos without the alogon. This definition allows me to bring most known forms of Platonisms together with that of Plato himself. They may be different on other grounds, including when, in contrast to Plato, it applies to matter, to the material constitution of physical reality, which may be seen as a form of idealism, an idealism of matter, insofar as it allows thought to fully capture this constitution. On the other hand, it is the possibility of a logos with the alogon within it that makes this chapter invoke Pythagoras' ghost of mathematical modernism, as against its "Plato's ghost," of Gray's title of his book on mathematical modernism (Gray 2007).

It is, again, a separate question whether any of the Pythagoreans ever embraced the first logos without aiming to overcome the alogon in it, which, as I said, appears unlikely. There is also a question of the scope of Pythagorean thinking concerning the relationships between the commensurable and the incommensurable even in mathematics itself and, especially, whether they extended the concept of a logos containing the alogon within it beyond mathematics, for example, to cosmology. I make no definitive claims as regard either case. My main concern here is the *spirit* of Pythagoras, defined by posing the problem of a logos containing the alogon within, in the subsequent history of mathematics and, then, beginning with Galileo and R. Descartes, mathematical physics, moreover, primarily during the modernist period, leading to our own time. On the other hand, it would, I think, be difficult to entirely rule out that some among the Pythagoreans or, philosophically, other pre-Socratics have contemplated this problem in this form, possibly on a cosmological scale, and even adopted a more positive attitude toward such as a logos. F. Nietzsche argued that such an attitude could be found, even apart from the discovery of the incommensurables, among the pre-Socratics during, in his terms (linked to the ancient Greek tragedy), "the tragic age of the Greeks," juxtaposed by Nietzsche to the Socratic age, defined by the aspiration to discover a logos without the alogon (Nietzsche 1996). One should still keep in mind that such claims reflect a modern reconstitution of ancient Greek thinking, especially in the case of the pre-Socratics, where we deal with fragmented texts, often via citations by later authors. It is also a separate question whether Plato had ever achieved his goal of establishing a philosophical logos free from the alogon in his texts, for they, too, sometimes work against the grain of this aspiration, as in his discussion of *khōra* in *Timaeus*, which could be read as more open to the alogon within the logos of philosophy (51a) (Derrida 1995).

2 The Spirit of Pythagoras: Geometry and Algebra – From Modernity to Modernism

While unavoidably invoking earlier developments, in particular, ancient Greek mathematics and the mathematics that emerged along with *modernity* (such as algebra, analytic geometry, and calculus), this chapter is primarily concerned with mathematical *modernism*, which took the center stage around 1900, building on several developments of nineteenth-century mathematics, and which extends to our own time. In this section I focus on the relationships between geometry and algebra in modernism and in radical Pythagorean mathematics. Their other, specifically epistemological, features will be discussed in the next two sections.

First, I shall define two general categories, modernity and modernism, in relation to which mathematical modernism will be positioned. Modernity is a broad cultural category. It refers to the period of Western culture extending from about the sixteenth century to our own time: we are still modern, although, during the last 50 years or so, modernity entered a new stage, postmodernity, shaped by the rise of digital information technology. Postmodernity, however, also amplified modernist epistemological trends, such as the role of the unthinkable in thinking, in art, philosophy, and mathematics and science (e.g., Lyotard 1984). Modernity is defined by several interrelated transformations, sometimes known as revolutions, although each took a while. Among them are scientific (defined by the new cosmological thinking, beginning with the Copernican heliocentric view of the Solar System, and the introduction of physics as the mathematical-experimental science by Descartes, Galileo, and Newton); industrial or, more broadly, technological (defined by the transition to the primary role of machines in industrial production and beyond); philosophical-psychological (defined by the rise of the concept of the individual human self, beginning with Descartes' concept of the Cogito); economic (defined by the rise of capitalism); and political (defined by the rise of Western democracies).

One might add to this standard list the mathematical revolution, rarely discussed as such, although sometimes considered as part of the scientific revolution. Although algebra was crucial to modern mathematics and physics, geometry remained dominant for a long time and has never lost its significance. Thus, while the laws of classical mechanics, embodied in its equations, are algebraic, they are grounded in geometrical picture of the world, including the motion of bodies, such as, paradigmatically, planets moving around the Sun. *Modern* geometrical thinking had continued to define physics, including relativity (although it has modernist aspects), until quantum mechanics (QM), which is more algebraic in character. As discussed below, QM has its geometry, too, which makes it Pythagorean on this account, alongside the radical Pythagorean character of its epistemology, defined by placing the ultimate constitution of nature beyond knowledge or even conception and thus by bringing the alogon into the logos of QM.

In contrast to modernity, "modernism" is a well-established denomination only when applied to literature and art. There, modernism refers to certain developments in the first half of the twentieth century, from roughly the 1900s on, represented by such figures as W. B. Yeats, J. Joyce, F. Kafka, R. M. Rilke, and V. Woolf, in

literature; P. Picasso, W. Kandinsky, and P. Klee, in art; and A. Schoenberg and I. Stravinsky in music. On occasion, it has been applied to the philosophy of the same period, such as that of Nietzsche, E. Husserl, and M. Heidegger. The denomination has been rarely used in considering mathematics and physics, or science, as opposed to "modern," used frequently, but with different periodizations. In mathematics, "modern" tends to refer to the mathematics that had emerged in the nineteenth century, with such figures as K. F. Gauss, H. Abel, A-L. Cauchy, and E. Galois, while in physics it refers to all mathematical-experimental physics, from Galileo and Descartes on, which uses the mathematics (algebra, analytic geometry, and calculus) developed around the same time. This is fitting because this physics emerged along with and shaped the rise of modernity making it scientific. After the discovery of relativity and quantum theory, the term "classical physics" was adopted for the preceding physics, still considered modern by virtue of its mathematical-experimental character. As Heidegger argued in commenting on Galileo and Descartes, "modern science is experimental because of its mathematical project" (Heidegger 1967, p. 93). I also adopt the designation "modern" for the mathematics emerging at the same time. If modernity is scientific, it is because it is also mathematical.

Arguably the most prominent recent example of using the term modernism in referring to mathematics is Gray's *Plato's Ghost: The Modernist Transformation of Mathematics* (2007), in part following (Mehrtens 1990). The book covers developments in *modern* mathematics (such as topology, set theory, abstract algebra, and rethinking the foundations of geometry) that had reached their *modernist* stage around 1900. It characterizes modernism arguably most essentially by the independence from relation to the object of nature, *physis*, and thus from physics, a feature difficult to deny to mathematical modernism (e.g., Gray 2007, p. 20). The book focuses, in considering this separation and beyond, primarily on geometry, with Hilbert's *Foundations of Geometry* (Hilbert 1999) as the conceptual center, analysis, and on post-Cantorian logical foundations of mathematics, where Hilbert, again, figures prominently. On the other hand, it somewhat marginalizes algebra, where this separation is more pronounced even during the modern period, and, with it, the Pythagorean character of modernism, as concerns both its mathematical architecture (defined by the interplay of geometry and algebra, also in analysis) and its epistemology, defined the alogon within its logos. As a result, Gray misses the ghost and, more importantly, the spirit of Pythagoras in mathematical modernism (the spirit that, as we have seen, has also inhabited the geometry of continuity from the Pythagoreans on), ultimately surrendering it to Plato's ghost, the ghost of a (geometrical) logos without the alogon.

Historically, just as modernist literature and art, modernist mathematics and science are characterized by major transformations. However, the complex and multifaceted nature of these transformations and their causes makes it difficult to uncontestably define their nature. It is hardly surprising, then, that, *conceptually*, the thinking concerning modernism in any field is diverse and, in each case, only partially reflects the nature of modernism in a given field or in the relationships among fields. This limitation cannot be avoided by this chapter either. Accordingly, I only argue that the form of modernist mathematics especially considered here,

radical Pythagorean mathematics, emerged coextensively with mathematical modernism but not that it exhausts it, let alone the development of mathematics during that period, not all of which is modernist by any definition of modernism. Radical Pythagorean mathematics has its history, too. For one thing, this mathematics is shadowed by the ghost of Pythagoras or, taking advantage of both the proximity of words "ghost" and "spirit" and a more creative meaning of "spirit," animated by the spirit of Pythagoras. The radical Pythagorean attitude, however, emerged or in any event became manifested during mathematical modernism, including in the way modernist mathematics functioned in physics. While, mathematical modernism is in part defined by its divorce from physics, modernist physics has never divorced itself from mathematics. Quite the contrary; it gave mathematics a greater and even unprecedented significance in physics. Indeed, ironically, modernist physics depends on mathematics that emerged as a result of this divorce.

Now, I understand algebra as the mathematical formalization of the relationships between symbols, arithmetic as the mathematical practice dealing specifically with numbers, geometry as the mathematical formalization of spatiality, especially in terms of measurement, and topology as the mathematization of the structure of spatial or spatial-like objects apart from measurements, through their continuity or discontinuity. The corresponding mathematical fields are algebra, number theory, geometry, and topology. Analysis deals with questions of limit, and related concepts, such as continuity, differentiation, integration, and so forth. There are multiple intersections between these fields, and fields that branch off these basic fields.

Defining algebra as the mathematical formalization of the relationships between symbols makes it part of all modern mathematics. The ancient Greek geometry was more self-contained, although one might detect elements of proto-algebraic symbolism there as well, especially at later stages of its development, as in Diophantus (third century AD), who shares the honor of "the father of algebra" with Muhammad ibn Mûsâ al-Khwârizmî (eighth to ninth century). Geometrical and topological mathematical objects always have algebraic components in their structure, while algebraic objects may, but need not, have geometrical or topological components. Two other meanings of algebra are important as well. The first, standing at the origins of algebra as a mathematical discipline, is that of algebra as the study of algebraic equations. The second is that referring to algebraic structures such as groups, which are crucial to geometry and topology. All these aspects of algebra are important to mathematical modernism, even in fields, like geometry and topology, that are disciplinarily juxtaposed to algebra. It is not only a matter of having an algebraic component as part of the mathematical structure of geometrical or topological objects but also of defining these objects algebraically.

An emblematic case of the role of algebra in modernist mathematics is algebraic topology. As part of topology, algebraic topology does have an earlier history, extending from G. Leibniz, with L. Euler as the main precursor before the rise of the discipline as such with B. Riemann, H. Poincaré, and others. General or point-set topology has a much longer history, extending to Plato and Aristotle, and possibly the Pythagoreans (e.g., Thom 1988; Papadopoulos 2019). What makes algebraic topology a mathematical discipline is the fact that one can associate algebraic

structures (initially numbers, eventually groups and other abstract algebraic structures) to the architecture of spatial objects that are invariant under continuous transformations, independently of their geometrical properties, associated with measurements. This makes topology topo-*logy* vs. geo-*metry*. In retrospect, topology is inherently categorical. It relates, functorially, the objects of topological and algebraic categories, a form of algebraic thinking that is one of the culminating concepts of modernism, for example, in A. Grothendieck's algebraic geometry. While keeping in mind this defining difference between topology and geometry, I shall for the purposes of this article group them together under the rubric of geometry as the mathematical science of spatiality, in juxtaposition to algebra as the science of mathematical symbolization.

Is, then, modernist geometry, a form of algebra? It is tempting to argue such a case, and the present author has done so previously, even if in a qualified way (Plotnitsky 2018, 2019b). Here, however, I would like to take a more nuanced view and consider, on Pythagorean lines, the *irreducible* interplay of geometry and algebra in modernist mathematics, an interplay that is underlain by that which is neither, any more than anything else we can conceive of, by the alogon within the logos of mathematics.

Even at the most immediate level, geometry has retained its significance in modernism on several counts, two of them in particular. First, the algebra defining modernist geometrical and topological objects has a special form that may be called "spatial algebra." Spatial algebra arises from algebraic structures that *mathematically* define geometrical or topological objects and reflects their proximity to $\mathbf{R}^3$ and mathematical spatial objects there that are close to our phenomenal intuition. This proximity may be, and commonly is, left behind in rigorous mathematical definitions and treatments of such objects, beginning with $\mathbf{R}^3$ itself. The same type of algebra may also be used to define mathematical objects that are no longer available to our phenomenal intuition. Among examples of such objects are a projective line (a set of lines through the origin of a vector space, such as $\mathbf{R}^2$, with projective curves defined algebraically, as algebraic varieties) and an infinite-dimensional Hilbert space (the points of which are typically square-integrable functions or infinite series, although a Euclidean space of any dimension is technically a Hilbert space, too). Spatial algebra is an algebraization of spatiality that makes it rigorously mathematical, topologically or geometrically, as opposed to something that is phenomenally intuitive spatially, even in the case of more conventional spatial objects in $\mathbf{R}^3$.

In addition, which is the second count on which spatial algebra retains its connection to geometrical thinking, analogies with $\mathbf{R}^3$ continue to remain useful and even indispensable. Thus, the analogues of the Pythagorean theorem or parallelogram law in Euclidean geometry, which holds in infinite-dimensional Hilbert spaces, are important, including in applications to physics, especially QM, the formalism of which is based in Hilbert spaces over $\mathbf{C}$. Indeed, our *thinking* concerning geometrical and topological objects is not entirely translatable into algebra. This was well understood by Hilbert in his axiomatization of Euclidean geometry, although this axiomatization clearly had a spatial-algebraic character (Hilbert 1999). According to D. Reed:

> [A]fter a chapter in which [Hilbert] provides himself with more tools to link geometry and algebra [following Descartes' analytic geometry], he goes on to demonstrate in a truly spectacular way:
>
> (*) a "theory of plane area can be derived from the axioms" (but not a theory of volume);
> (*) Desargues' theorem, which states that if two triangles are situated in a plane so that pairs of corresponding sides are parallel then the lines joining the corresponding sides pass through one and the same point or are parallel . . ., expresses a criterion for a "plane" geometry to form part of "space" geometry; and
> (*) Pascal's theorem [technically the affine version of Pascal's theorem], which states that if A, B, C, and A^I, B^I, C^I are two sets of points on two intersecting lines and of AB^I is parallel to BC^I and AA^I is parallel to CC^I then BA^I is parallel to CB^I . . . is dependent in a very specific way on the so-called Axiom of Archimedes.
>
> None of these statements can be given a simple unequivocal expression in the realm of algebra even though models from 'analytic geometry' are used in the demonstrations. In other words, while algebra is useful as a tool in the demonstration of geometrical statements it is not useful in formulating the statements themselves. (Reed 1995, pp. 33–34)

While Reed is right to stress the significance of geometrical thinking and expression in mathematics, his claim that algebra is not useful in formulating geometrical statements is a simplification even in Euclidean geometry, where our geometrical intuition is more applicable and where certain proofs could be purely geometrical. Besides, a more natural setting for Desargues' and Pascal's theorems is projective geometry, which these theorems helped to usher in, a setting that we cannot visualize and that is spatial-algebraic.

One can gain further insight into this situation by considering the principle due to J. Silverman and J. Tate: "Think Geometrically, Prove Algebraically!" It was introduced in their book on "the rational points of elliptic curves" (a modernist subject):

> It is also possible to look at polynomial equations and their solutions in rings and fields other than **Z** or **Q** or **R** or **C**. For example, one might look at polynomials with coefficients in the finite field $\mathbf{F}_p$ with p elements and ask for solutions whose coordinates are also in the field $\mathbf{F}_p$. You may worry about your geometric intuitions in situations like this. How can one visualize points and curves and directions in $\mathbf{A}^2$ when the points of $\mathbf{A}^2$ are pairs (x, y) with $x, y \in \mathbf{F}_p$? There are two answers to this question. The first and most reassuring is that you can continue to think of the usual Euclidean plane, i.e., $\mathbf{R}^2$, and most of your geometric intuitions concerning points and curves will still be true when you switch to coordinates in $\mathbf{F}_p$. The second and more practical answer is that the affine and projective planes and affine and projective curves are defined *algebraically* in terms of ordered pairs (r, s) or homogeneous triples $[a, b, c]$ without any reference to geometry. So in proving things one can work algebraically using coordinates, without worrying at all about geometrical intuitions. We might summarize this general philosophy as: *Think Geometrically*, *Prove Algebraically.*" (Silverman and Tate 2015, p. 277)

Affine and projective planes and curves can in principle be defined without any reference to our phenomenal intuition, which grounds the geometrical thinking Silverman and Tate refer to, rather than spatial algebra, which defines them rigorously.

Even in these more intuitively accessible cases, we think algebraically, too, by using spatial algebra, except possibly in dealing with low-dimensional topological and geometrical objects, where topological and geometrical arguments could be used more rigorously. I would argue, however, that spatial algebra is still irreducible there because one commonly converts topological operations into algebraic ones. This conversion, in low dimensions, was essential to the origin of algebraic topology. On the other hand, the more recent development of low-dimensional topology, following, among others, W. Thurston's pioneering work from the 1970s on, is a geometrically oriented trend that counters modernist algebraization but only to a degree, because the algebraic structures associated with these objects remain crucial. Some of the most powerful algebraic tools of algebraic topology have been used and even developed in this field. It is also true that a mathematician can develop and use intuition in dealing with discrete geometries, say, that of the Fano plane of order 2, which has the smallest number of points and lines (seven each). However, beyond the fact that they occur in the two-dimensional regular plane, the diagrammatic representations of the Fano plane are still difficult to think of as other than spatial algebra, in this case combinatorial. Euclidean intuitions are limited even when we deal with algebraic curves in the Euclidean plane, let alone in considering curves and other objects of finite or projective geometries, abstract algebraic varieties, Hilbert spaces, the spaces of noncommutative geometry, or geometric groups, a great example of the extension, by a reversal, of spatial algebra to algebraic objects as such.

And yet, these geometrical intuitions and, thus, geometry are indispensable. Silverman and Tate's "*think* geometrically" may even be read as giving geometry a greater creative significance, although proving theorems requires thinking, too. Our geometrical thinking shapes our practice of spatial algebra and other areas of abstract algebra, for example, when dealing with such objects as ideals of rings, even apart from algebraic geometry where manifolds or schemes are constructed of ideals (maximal or prime, respectively) of commutative rings. For the moment, however, I would like to emphasize the creative role of intuitive geometrical *thinking*, which may need algebra to achieve the necessary mathematical rigor – Prove algebraically! – but is crucial to mathematical thought as geometry – Think geometrically!

3 Radical Pythagorean Mathematics and Mathematical Modernism

Both defining features of radical Pythagorean mathematics – the reciprocal interplay of geometry and algebra (and the modernist rethinking of each) and mathematical thinking inhabited by the unthinkable, a mathematical logos inhabited by the alogon – are, I argue, found, often working together, in mathematical modernism and modernist physics, especially quantum theory. These features are, however, rarely, if at all, addressed in discussions of mathematical modernism (whether the term is used or not), perhaps because the idea of the role of the unthinkable as an enabling condition of thinking is as resisted by commentators on mathematical modernism as by most mathematicians and physicists. As noted, arguably the most prominent aspect of

mathematical modernism in these discussions is mathematics' capacity for an independence from its relation to nature and thus from physics and, in part correlatively, from our daily phenomenal intuition and concepts, in sum, a fundamentally abstract character of modernist mathematics. While these features are difficult to deny to mathematical modernism, the latter, I argue, has accomplished quite a bit more, including in physics, not the least by its Pythagorean and ultimately radical Pythagorean nature.

According to Gray, referring what he calls "the crisis of continuity": "This is the widespread feeling among mathematicians around 1900, documented in many sources, that the basic topic of analysis, continuity, was profoundly counterintuitive. This realization marks a break with all philosophy of mathematics that present mathematical objects as idealizations from natural ones: it is characteristic of modernism" (Gray 2007, p. 20). Given Gray's overall argument, "natural" here refers to both objects of nature and common objects of our daily phenomenal intuition, which can also be related to each other, as they are in classical physics or in daily life. While Gray's claims here are valid, they require qualifications. First of all, mathematics separated, abstracted, itself from representing natural objects and thus from physics, and even from our daily phenomenal world much earlier, already with the Pythagoreans, or during modernity, with the rise of algebra, especially the study of algebraic equations, pursued largely apart from physics. It is true, however, that both geometry and analysis had kept close relationships with representing nature and thus physics and also with our daily thinking and its representation of nature, a representation that classical physics refines and mathematically idealizes, including by using geometry and calculus. It is also easier to argue this type of case in its disciplinary and institutional sense, insofar as the possibility of this separation was used to ground the disciplinary field of mathematics from modernity to modernism. The drive toward this independence had been emerging in *modern* mathematics from the 1800s on, but with the rise of mathematical modernism, it reached the stage of breaking with representing (mathematically idealizing) natural objects in most areas of mathematics, including geometry and analysis, especially following the rise of set theory and G. Cantor's investigations of the nature of mathematical infinity and his continuum hypothesis. Cantor's set theory was a major and, in some respects, the most significant event in the history of "the crisis of continuity" invoked by Gray, and it separated the nature of continuity in mathematics from our phenomenal intuition of continuity especially radically and irrevocably. Quantum theory, beginning with M. Planck's discovery of it in 1900, was about to contribute to this crisis as well, in QM by way of complex relationships between mathematical continuity and physical discontinuity, discussed below.

This double divorce of modernist mathematics from both the representation of natural objects and our general phenomenal intuition was stressed by H. Weyl, himself a major figure of mathematical modernism, half a century earlier. Weyl made this point in his 1918 *The Continuum* (Weyl 1928), followed closely by his 1918 classic *Time Space Matter* (Weyl 1952), both shaped by modernist mathematics, by which point quantum theory (not yet QM) enters his thinking concerning continuity and its role in physics (Weyl 1952, p. 23). According to Weyl: "the

conceptual world of mathematics is so foreign to what the intuitive continuum presents to us that the demand for coincidence between the two must be dismissed as absurd" (Weyl 1928, p. 108). "Coincidence" is, however, not the same as "relations," which are unavoidable, at least insofar as it is difficult to think of continuity spatially apart from our phenomenal intuition. On the other hand, it is possible to define the continuum algebraically. This situation, as just noted, emerged with Cantor's introduction of set theory and, with it, the multiplicity of infinities, the infinity of infinities, and his continuum hypothesis. It was made more dramatic by Gödel's incompleteness theorems of 1931 and then P. Cohen's proof of the undecidability of the continuum hypothesis in the 1960s. These findings made us realize not only that we do not know but also and more radically that we *cannot know* how a continuous line, straight or curved (which does not matter topologically), is constituted by its points.

Weyl's overall point pertains to most of modernist mathematics. Physics, such as relativity and quantum theory, was able to take advantage of this divorce between mathematics and our general phenomenal intuition. Weyl added to his statement cited above: "Nevertheless, those abstract schemata supplied us by mathematics must underlie the exact sciences of domains of objects in which continua play a role" (Weyl 1928, p. 108). This comment was made primarily with Einstein's relativity in mind, as Weyl's next book, *Time Space Matter*, was already in the works. This view retains a modernist twist because the *abstract* schemata used in relativity and then QM are divorced from our phenomenal intuition. Both used modernist mathematics, respectively, that of Riemannian geometry and that of infinite-dimensional Hilbert spaces, but with a crucial difference, in part due to the fundamental role of discontinuity in QM, even though both cases deal with continuous mathematics and in fact geometry.

Relativity still did this in a realist way. This is possible, because the break from our phenomenal intuition need not entail a divorce from realism or ontology, even a physical one, let alone mathematical, which is possible in the absence of a physical ontology, including in QM. Relativity was a radical departure from classical physics because the relativistic law of addition of velocities (defined by the Lorentz transformation) in special relativity, $s = \frac{v+u}{1+(vu/c)^2}$, for collinear motion ($c$ is the speed of light in a vacuum), runs contrary to any possible intuitive representation. This concept of motion is, thus, no longer a mathematical refinement of a daily concept of motion as the classical concept of motion is. Relativity was the first physical theory that defeated our ability to form a phenomenal conception of an elementary individual physical process. It was a form of incommensurability, in this case between nature and our phenomenal intuition, rather than in any mathematical sense. The theory, however, still allowed for a mathematized representation of physical reality.

By contrast, QM, at least in certain nonrealist interpretations (defined below as reality-without-realism, RWR, interpretations), only used mathematics for providing probabilistic or statistical predictions concerning the outcomes of quantum experiments, without providing a representation of the ultimate constitution of what is

responsible for these events, even a purely mathematical one in the absence of a physical one. (The latter view has been entertained, including by Heisenberg in his later work.) These probabilistic predictions are, moreover, only possible by using ad hoc rules *added* to the formalism rather than being part of it, such as Born's rule, which relates, by using complex conjugation, complex quantities of the QM formalism (the so-called amplitudes) to real numbers corresponding to the probabilities of quantum events. One takes the square moduli of the eigenvalues of the operators associated with quantum variables, such as position, momentum, or energy (or equivalently, multiply these eigenvalues by their complex conjugates), which gives one real numbers, corresponding, once suitably normalized, to the probability densities of observed events. In departure from the classical probability theory, the rule for adding the probabilities of alternative outcomes is changed to first adding the corresponding amplitudes and deriving the final probability by squaring the modulus of the sum by Born's rule. Accordingly, in these interpretations, QM used "abstract schemata" of modernist mathematics in a radically new way and established a new type of relationships between the mathematics of a physical theory and the experimental data considered.

Representations and the corresponding physical ontologies are possible and necessary in dealing with some strata of reality considered in QM, specifically, that observed in measuring instruments, defining, in Kant's language, *quantum phenomena*. Let's call this reality Reality A. On the other hand, a knowledge or even conception of the ultimate nature of reality responsible for quantum phenomena corresponding to or idealized as, in Kant's language, quantum *objects*, and thus the ultimate-level ontology are no longer offered. (As will be seen presently, this view of quantum objects goes beyond Kant.) Let's call this reality Reality X. Importantly, the existence of Reality X is, rather than merely assumed, *inferred* from the effects it has on Reality A. To use a mathematical metaphor, X is defined by the conceptual equation $f(\mathrm{X}, \mathrm{A}) = 0$ (with f symbolically standing the sum of the conditions of quantum experiments), which has a "solution" that cannot, however, be represented in terms of A and is irrational relative to A, even though effects of Reality X on Reality A can be probabilistically predicted by QM, which makes the relationships between A and X strictly mathematical. We neither know nor can even conceive how these relationships are physically possible. Reality X is thus *constructed by thought* as *unconstructible by thought*, at least as things stand now, analogously to the way the incommensurables (irrationals) were constructed by Pythagorean arithmetic, in this case literally as (in our terms) a solution of the equation $X^2 - 2 = 0$. If one adopts the radical Pythagorean view, this situation is assumed to be an enabling condition of thought and knowledge, rather than inhibiting both, with a hope to be eventually overcome, as it happened in Pythagorean geometry.

This view may also be designated "the reality-without-reality (RWR) view," the language introduced by the present author previously (e.g., Plotnitsky 2016, 2018, 2019a, b). Reality X exists, is *real*, and has effects on Reality A but allows for no representation or even conception, as would be required by realism. Although the RWR epistemology is the same as that of the radical Pythagorean view, the RWR view is more general: it does not depend on the relationships between algebra and

geometry that co-define the radical Pythagorean view, and it can apply beyond mathematics and physics, for example, in other sciences or in philosophy.

While one can, thus, trace the history of the RWR view to the Pythagoreans and their discovery of the alogon within the logos of arithmetic, its modern history may be argued to have originated with Kant's philosophy, which distinguished noumena as objects or things-in-themselves and phenomena or representations in our thought (Kant 1997). It was, however, expressly introduced in physics with QM, beginning with Heisenberg's discovery of the theory (Heisenberg 1925) and Bohr's interpretation of it, developed, with some significant revisions over two decades and presented in several articles. According to Bohr, in QM we are no "longer in a position to speak of the autonomous behavior of a physical object, due to the unavoidable interaction between the object and the measuring instruments" (Bohr 1937, p. 79). Hence, as he said elsewhere, "in quantum mechanics, we are not dealing with an arbitrary renunciation of a more detailed analysis of atomic phenomena, but with a recognition that such an analysis is *in principle* excluded" (Bohr 1987, vol. 2, p. 62). The RWR view, in its most radical form, is, however, different from that of Kant. While Kant places things-in-themselves (objects) beyond representation or knowledge (phenomena), he still allows that a *conception* of them could be formed and even be true, although this truth cannot be guaranteed (Kant 1997, p. 115).

What is, however, the difference between placing the ultimate nature of the reality, physical or mathematical, beyond representation or knowledge and placing it beyond conception? For, if, as Bohr says, we are "not . . . any longer in a position to speak of the autonomous behavior of a physical object," would not this already imply that this behavior, or again, the RWR-type of reality so idealized, is beyond conception? For, if we had such a conception, would we not be able to says something about it or represent it? The situation is not so straightforward. First, there is difference between *some* conception of this reality and a theory that would enable us to provide a representation, especially a mathematized one. Bohr, however, makes a stronger claim: we are no longer in a position to *speak* of the autonomous behavior of quantum objects or the reality thus idealized *at all* and, thus, it follows, to have a conception of this reality. The question then becomes that of whether our inability to do so only characterizes the quantum-mechanical situation as things stand now (A) or reflects that this reality is beyond the reach of our thought altogether, ever (B), a view adopted by the present author (Plotnitsky 2016, 2018, 2019a, b). While Bohr, thus, at least assumes (A) and while there are signs that he might have entertained (B), he never stated so, which leaves whether he assumed (B) open to interpretation. Logically, once (A) is the case, then (B) is possible. I am not saying that (B) is necessary but only that it is *possible*. There does not appear to be any experimental data so far compelling one to prefer either (A) or (B). Both views are thus far equivalent as far as physics is concerned. But they are different philosophically in defining how far our thought could, in principle, reach in investigating the ultimate constitution of nature.

The qualification "as things stand now," however, applies to (B) as well, even though it might appear otherwise given that this view precludes any conception of

the ultimate reality not only now but also ever. It still applies because a return to realism, to a logos without the alogon, is possible on experimental or theoretical grounds. This return may take place either because quantum theory, as currently constituted, is replaced by an alternative theory that allows for a realist interpretation, or because (B), or (A), becomes obsolete even for those who hold it and is replaced by a realist view with QM or QFT in its present form. It is also possible, however, that either (A) or (B) will remain viable in interpreting of QM or quantum field theory (QFT), which handles high-energy quantum physics, where this view equally applies, and thus all quantum physics as currently constituted and experimentally confirmed (Plotnitsky 2019b).

As noted from the outset, the epistemological cost of the radical Pythagorean or RWR view is not easily absorbed and to some, beginning again with Plato, or the Pythagorean themselves, and then Einstein, is unacceptable. This is not surprising, because this view defies many basic assumptions concerning the workings of nature and thought. These assumptions, in part arising due to the neurological constitution of our brain, have served us for as long as human life itself, and within certain limits are unavoidable, including in physics, although, while respected by classical physics, they were, as noted, already challenged by relativity. QM made this challenge much greater, without entirely renouncing these features, which are applicable at levels other than that of the ultimate constitution of nature, specifically in describing measuring instruments.

In QM, the RWR view arises by virtue of the algebraic rather than, as in classical physics or relativity, geometrical, relationship between the mathematics of a physical theory and physical reality in its ultimate constitution, Reality X. This shift compelled Einstein to speak of "the Heisenberg method" as fundamentally algebraic (Einstein 1936, p. 378). Indeed, as just explained, "the Heisenberg method" established new, mathematically algebraic and epistemologically irreducibly probabilistic, relationships between a (mathematized) physical theory and nature. It is worth nothing in this connection that probability theory is, too, essentially algebraic. The origin of probability theory, in the work of G. Cardano, B. Pascal, and P. Fermat coincides with the emergences of algebra. As I. Hacking argued in explaining why the theory emerged in the seventeenth century rather than earlier, some form of algebra was necessary for probability theory (Hacking 2006). QM, however, reshaped, on RWR lines, how probability works, as against previous uses of probability, for example, in classical statistical physics. There the relationships between them is underlain by a geometrical picture of the behavior of the individual constituents of the systems considered, assumed to follow the laws of classical mechanics. By contrast, in quantum physics, even the behavior of the elemental individual quantum objects too had to be treated probabilistically.

In confronting this situation, Heisenberg was compelled to conceive of QM, on RWR lines, as theory designed only for predicting the probabilities of what was observed in measuring instruments, as quantum phenomena (Reality A), without representing the behavior of quantum objects, even the elementary ones (Heisenberg 1925). Quantum objects were or, more precisely, idealized Reality X. Even before his paper was published he said: "What I really like in this scheme is that one can really

reduce *all interactions* between atoms and the external world … *to transition probabilities*" between quantum measurements (Heisenberg, Letter to Kronig, 5 June 1925; cited in Mehra and Rechenberg 2001, vol. 2, p. 242; emphasis added). The "scheme" worked! It was Heisenberg's renunciation of a geometrical representation of quantum objects and behavior by replacing the geometry of trajectories with the algebra of probabilities that led him to the discovery of QM. (QM was co-discovered by E. Schrödinger by using a realist approach, while, however, still at least allowing for an RWR-type interpretation.)

Heisenberg used and in fact reinvented mathematics never previously used in physics, in essence (although Heisenberg did not present it in these terms) that of infinite-dimensional Hilbert spaces over **C**. A radically new, for physics, feature of his algebra was the noncommutative nature, $PQ - QP \neq 0$, for the corresponding physical variables, Hermitian operators in Hilbert spaces over **C** (replacing the functions of real variables of classical mechanics), for example, and in particular, those associated, respectively, with the momentum, P, and the coordinate, Q. This feature proved to be momentous. Most famously, it came to represent in the formalism of QM, Heisenberg's uncertainty relations making impossible simultaneous measurements or predictions for the momentum (P) and the coordinate (Q) *associated* with a quantum object (but observed only in measuring instruments), the impossibility defined by Bohr as "complementarity." While unbeknownst to Heisenberg, who reinvented it, noncommutativity was, however, a familiar feature of mathematics by then and may indeed be seen as part of mathematical modernism.

Heisenberg's approach or "method" was, thus, grounded in the primacy of algebra, as against classical physics and relativity or the preceding quantum theory, specifically, Bohr's 1913 atomic theory, which retained a geometrical, orbital, representation of electrons' motion in so-called stationary states, although it renounced any conception of transitions between such states. The algebraic nature of the theory was, again, emphasized by (and worried) Einstein:

> [P]erhaps the success of the Heisenberg method points to a purely algebraic method of description of nature, that is, to the elimination of continuous functions from physics. Then, however, we must give up, in principle, the space–time continuum [at the ultimate level of reality]. It is not unimaginable that human ingenuity will some day find methods which will make it possible to proceed along such a path. At present however, such a program looks like an attempt to breathe in empty space. (Einstein 1936, p. 378)

The theory has, however, proven to be extraordinarily successful, and it remains our standard theory of quantum phenomena in both low- and high-energy quantum regimes, governed by QM and QFT, respectively. Quantum electrodynamics (QED), which is part of QFT, is the best experimentally confirmed physical theory ever. Quantum theory is, however, "not a *description* of nature," because, in the RWR view, it offers no representation of the ultimate constitution of nature, which was Einstein's imperative for a fundamental theory and to which he referred here.

A few qualifications are in order, which will also bring geometry back into quantum theory. First, in saying that "we must give up, in principle, the space–

time continuum," Einstein must have had in mind giving up the spacetime continuum in *representing* the ultimate reality considered. This view defined his philosophy of physics, embodied in general relativity, which was based in Riemannian geometry and, thus, the concept of continuous, actually differential, manifolds. The idea that this reality may ultimately be discrete had been around for quite a while by then. It was proposed as a possibility by Riemann as early as 1854 (Riemann 1854, p. 33). The idea has acquired new currency in view of QM and then QFT and was advocated by, among others, Heisenberg in the 1930s and is still around. In the RWR view, the ultimate nature of physical reality is beyond representation and even conception and, as such, may not be seen as either continuous or discrete. Discreteness only pertains to quantum phenomena, observed in measuring instruments, while continuity has no physical significance at all. It is, however, a feature of the formalism of QM, which thus relates to discrete phenomena by predicting the probabilities or statistics of their occurrence by means of continuous mathematics – a radically new form of the relationships between the continuous and the discrete in physics.

This situation also gives a greater complexity and geometry to the algebraic method of QM. It is true that the continuous functions (over **R)** used in classical physics or relativity (for variables such as position or momentum) are replaced by operators and their algebra in Hilbert spaces over **C**. Hilbert spaces open the possibility, amply used by physicists, to think geometrically. Besides, some continuous functions are retained, because these Hilbert spaces are those of continuous functions, considered as infinite-dimensional vectors in dealing with continuous variables such as position and momentum, although these variables themselves are, again, represented by operators. These functions, vectors, are those of complex (rather than, as in classical physics, real) variables, and the vector spaces that they comprise or associated objects, such as operator algebras, have special properties, such as noncommutativity. As noted, these vectors, of which Schrödinger's wave function, ψ, is the most famous example, play an essential role in calculating (via Born's rule) the probabilities of the outcomes of quantum experiments.

The ultimate nature of reality in QM, in RWR-type interpretations, Reality X, idealized by quantum objects may thus be seen as a form of "irrationality," the term invoked by Bohr, in the sense of its incommensurability with our concepts, physical, mathematical, and philosophical, which we use in quantum theory in considering Reality A, that of observed quantum phenomena. The situation exhibits a parallel with the case of the diagonal of the square in Pythagorean mathematics, especially given that our relationships to Reality X in QM are strictly mathematical. The parallel merits a further elaboration as an instructive case of the Pythagoras's ghost in, and Pythagoras's spirit of, modernist mathematics and physics. Bohr's invocation of "irrationality" in this connection is often misunderstood and, as a result, sometimes held against him. This irrationality not any irrationality of QM, of which Bohr expressly speaks as a "rational" theory (Bohr 1987, vol. 1, p. 48). Bohr's use of "irrationality" is based on an analogy with irrational numbers, reinforced by the role of complex numbers and specifically the square root of -1, i (an irrational magnitude in the literal sense because it cannot be presented as a ratio of two integers) in

the formalism of QM (Bohr 1987, vol. 1, p. 76), while everything that we can actually observed is represented in terms of real, indeed, technically, *rational*, numbers. The deliberate nature of this analogy is confirmed by the fact that Bohr expressly reflected on this connections between the mathematical and quantum-theoretical concepts of the irrational, via the idea of a Riemann surface (which was part of the genealogy of his concept of complementarity) in his final interview (Plotnitsky 2017, p. 1130). On the other hand, the ultimate nature of the reality considered cannot be represented at all. The Pythagoreans' incommensurables must have been on Bohr's mind, not the least because in both cases one deals with the problem of measurement, with the mathematics and physics of *measurement*. QM offers no longer either algebraic (in parallel with the arithmetic of the Pythagoreans) or (in contrast to the post-Pythagorean Greek geometry) geometrical means of representing this irrationality. One can only, by using the geometry of Hilbert spaces (over **C**!), relate, probabilistically, to its effects upon what is observed in measuring instruments, as Reality A impacted by Reality X. The irrationality is that of the unrepresentable or inconceivable nature of the ultimate constitution of reality (idealized as quantum objects), while the "geometry" is that of a Hilbert space. Although this geometry no longer represents the incommensurable (the alogon), it still relates to it by means of an algebra of probability, which is possible only because of the relationships between the geometry and algebra of a Hilbert space. This is the logos of QM, which, in the spirit of the Pythagoreans, elevated to the radical Pythagorean thought, contains the alogon within it.

In ancient Greek mathematics, the problem was resolved geometrically by a new theory of proportion, accommodating both commensurable and incommensurable magnitudes. The arithmetical problem was only resolved, by essentially modernist means (algebra played a major role), in the nineteenth century, after more than 2000 years of effort, with Dedekind and others, albeit in view of the undecidability of Cantor's continuum hypothesis, resolved as ultimately unresolvable. It remains to be seen whether quantum-mechanical "irrationality" will ever be resolved by discovering a way to represent quantum objects and behavior, or the reality they idealize, Reality X, in which case this idealization will need to be replaced by a new one.

As things stand now, QM is a *rational* theory of something that is *irrational* in the sense of being inaccessible to a representation or knowledge, or even to thought. QM was a replacement of a *rational representational* theory, classical mechanics, with a *rational probabilistically or statistically predictive* theory (Bohr 1987, vol. 1, p. 48). QM is both a reversal of what happened in the crisis of the incommensurable in ancient Greek mathematics, which compelled it to move from arithmetic to geometry, in their theory of measurement, and a new form of the relationships between geometry and measurement, which relates a mathematical continuity to a physical discontinuity in terms of probabilities. "Thought . . . is the measure of the Universe," the greatest gift of Prometheus, the first Pythagorean, and the highest form of measurement, P. B. Shelley says in *Prometheus Unbound* (II, iv. 73). The phrase echoes Protagoras's "man is the measure of all things" but gives it the Pythagorean

meaning of a logos that must contain the alogon within it, because, as the poem says a few lines later, "the deep truth is imageless" (II, iv. 116).

4 The Thinkable, the Unthinkable, and the Real in Mathematics and Radical Pythagorean Mathematics

Analogous, even if not identical, situations to the one just considered in physics are found in modernist mathematics, in both Pythagorean aspects – the RWR nature of the ultimate reality, the alogon in a logos, and the Pythagorean interplay of geometry and algebra. For example, and in particular, this type of situation arises with Cantor's set theory, although it has been rarely, if ever, considered as such in the philosophy of mathematics, even following Gödel's incompleteness theorems of 1931 and then Cohen's proof of the undecidability of the continuum hypothesis in 1960s, which invite such a consideration. The nature of mathematical reality has of course been debated since the pre-Socratics, including the Pythagoreans, and it was at stake in Plato, whose ghost has overshadowed this debate in modernist mathematics, along with and against the ghost and the spirit of Pythagoras. As discussed from the outset, Plato responded to this thought by his attempt to overcome the alogon within the logos of philosophy, if not that of mathematics, leaving the task of this overcoming to mathematicians, who have attempted to do so ever since. The twentieth- and twenty-first-century philosophy of mathematics and mathematical logic have continued this project. Thus, mathematical Platonism (a twentieth-century philosophical trend, which should be distinguished from Plato's thought) assumes the existence of mathematical reality as independent of our thinking, which, however, could represent this reality, at least in principle. This view has been debated in turn, and several alternative, if sometimes intersecting, positions have emerged since and some before, such as logicism, formalism, intuitionism, conventionalism, and constructivism. I shall not enter these debates, apart from noting that the question of the domain, "location," of this reality has never been adequately answered by mathematical Platonists. It is difficult to think of that which is not material and yet is outside human thought, unless it is divine, which, however, is not a common assumption among mathematical Platonists. By contrast, I assume the reality of human thought, different for each of us, as the only domain in relation to which one can speak of the reality of mathematical objects and concepts, in juxtaposition to the material reality of nature considered in physics. It is possible to assume that the Platonist mathematical reality is a potentiality – the same, even if multiply branched, potentiality – in principle realizable by human thought. I shall comment on this possibility below.

What could be claimed to exist in our thought without much controversy are mathematical specifications, from strict definitions to partial and indirect characterizations, implying more complete future specifications. Such specifications would involve concepts, structures, logical propositions, or still others elements, which could be both geometrical and algebraic in the broader sense of these terms adopted here. The question, then, becomes whether anything else exists in our thought

beyond such specifications and the local ontologies they define, at least in our conscious thought, because we can unconsciously think of other properties of a given object or field, which may either be eventually made conscious or never become conscious and thus known. This qualification does not, however, change the nature of the situation, because one could either claim that the unconscious could only contain such specifications or that the object or broader reality in question, as different or exceeding such specifications, somehow exists there. Hence, essentially the same alternative remains in place. Some unconscious specifications may never become conscious. It is possible that some mathematics, even very great mathematics, just as poetry, never left the unconscious and there are accounts by mathematicians or poets of dreaming of mathematics or poetry which could not be remembered as what it actually was, although sometimes it can. Only our memory of dreams is conscious but never dreams themselves. Whatever is in our unconscious can, however, enter our consciousness. Our unconscious does a great deal of mathematical thinking, as it does most of our thinking in general, and some of this thinking may never be realized by our conscious thought. Sometimes, not uncommonly in the logical foundations of mathematics (Hilbert held this view), the consistency of a given definition of a mathematical entity is identified with the existence of the corresponding object, which becomes a form of mathematical Platonism, if this existence is assumed outside human thought.

My assumption here is that nothing mathematical exists beyond what can be thus specified by thought, possibly, again, in one's unconscious. This differentiates the situation from that of quantum theory in the RWR view. While quantum phenomena or quantum theories are specified by thought, one *assumes* the existence ofan ultimate physical reality, which is, in the RWR view, beyond representation or even conception and thus specification. By contrast, in the present view, even if some ultimate mathematical reality may be shared, it cannot be assumed to exist independent of thought; only physical matter in its ultimate constitution is assume to exist in this independent way. This assumption has been challenged as well, even by denying that the ultimate reality is material rather than mental, with Plato as the most famous ancient case (under the assumption, however, that this reality is unique and could be shared, just as the material reality in physics is now) and Bishop Berkeley as the most famous modern case. Still this is a common assumption in modern philosophy of physics, although assuming this reality to be beyond representation or even conception, now or especially ever, is uncommon. As just noted, it is possible that there are (mentally) real things that exist in our unconscious that will never become conscious. It is equally possible, however, that they will enter our consciousness at one point or another. The ultimate constitution of nature, in an RWR-type interpretation, is not assumed to become available, at least as things stand now, and possibly ever.

Our mathematical specifications could be made more complete or modified by new concepts, structures, and logical propositions, which would change the objects or concept in question, as say, a Riemann surface, as it developed during over, by now, a long period. However, in the present view, it is no longer possible to see such changes as referring to the same mathematical object (which can, again, be a broad

and multiple entity), approached by our evolving concepts. Instead, they create new objects or concepts. Thus, new multiplicities of Riemann surfaces are created by each modification of the concept. There are no Riemann surfaces as such, existing by themselves and in themselves; there is only what we can think about them at a given point of time. This view of mathematical reality has thus a constructivist flavor, in part following Kant's view of mathematics as the synthetic construction of mathematical concepts by thought (Kant 1997). (I use the term "constructivist" more generally, rather than referring to the trend known as "constructivism" in the foundational philosophy of mathematics, from intuitionism on, relevant as it may be, in part given Kant's influence.). Mathematics is only what can actually be thought, created by (conscious or unconscious) thought, and then expressed, communicated, and so forth, thus in accord with the Greek meaning of *máthēma* as that which can be known and learned, or taught.

It is not uncommon to encounter a situation in which a mathematical entity that cannot be given, now or possibly ever, an adequate mathematical specification, and is only specified partially mathematically or otherwise (e.g., by means of philosophical concepts), but that could only be consistently related to, indirectly, by means of a more properly specified mathematical concept or set of concepts. In the present view, however, any mathematical entity can only be seen as existing or real if it is sufficiently specified in some way, possibly in terms of ordinary or philosophical concepts, perhaps partially supplemented by mathematical concepts, which may be unconscious. It cannot be assigned existence or reality as independent of thought, somebody's thought.

In view of Gödel's incompleteness theorems, mathematics, at least if it is rich enough to contain arithmetic, cannot completely represent (and thus specify) itself: it cannot mathematically formalize all of its concepts, propositions, or structures and ultimately itself so as to guarantee its consistency. But it does not necessarily follow that the corresponding unspecifiable reality actually exists. One can make this assumption, as Gödel ultimately did on Platonist lines, claiming that there is, at least for now, no human means, mathematical or other, to specify this reality. But this assumption is not a consequence of Gödel's incompleteness theorems, which only imply that it is impossible to prove that all possible specifications, within any such system, are consistent. (As noted, Gödel's theorems do allow that a given system could be proven to be inconsistent.) What may be inconceivable is *why* this is the case and what is the nature of the reality responsible for this situation. This is, however, not a mathematical or even meta-mathematical question, any more than the question why our interaction with nature by means of quantum physics enables us to make correct probabilistic predictions. These questions concern the neurological nature of our thought, although they may not ultimately be answerable by neuroscience either. That is, the ultimate reasons why we think and for the way in which we think may also be beyond the reach of our thought, the alogon within a logos of neuroscience, which, again, need not mean that we cannot advance neuroscience under this condition.

One could speak in considering such as yet unspecified mathematical objects or concepts in terms of a hypothetical *potentiality*, defined by the assumption that a

mathematical object or concept of a certain type could or should exist. Such potentialities are, moreover, only partially, probabilistically, determined by what is sometimes called plausible reasoning, very important in mathematical thinking (Polya 1990). There are different and possibly incompatible way in which this potentially may be become reality. Consider the equation $X^2 + 1 = 0$ and complex numbers. While this equation (which may be assumed to exist in our thought) had no real-number solution, one could have and some had envisioned that it should have a solution and that a mathematical entity or a multiple of such entities, a new type of number, should exist. This hypothesis came to be realized insofar as complex numbers eventually became a mathematical reality. In the present view, however, they were not a mathematical reality before they were correspondingly specified in somebody's thought, in somebody's unconscious, by the time of Gauss, a crucial figure in giving complex numbers a mathematical reality – each time one thinks of them, consciously or unconsciously, but not otherwise.

In sum, in the present view, in mathematics, all reality is constructed, and this construction may, ontologically, involve multiple "mathematics." This multiplicity is also a consequence of Gödel's undecidability, as exemplified by Cantor's continuum hypothesis, mentioned above. This hypothesis was crucial not only for the question of continuity but also for the question of Cantor's hierarchical order of infinities (the infinity of which was one of his discoveries) and thus for the whole edifice of Cantor's set theory. The hypothesis was proven undecidable by Cohen. It follows, however, that one can extend classical arithmetic in two ways by considering Cantor's hypothesis as either true or false, that is, by assuming either that there is no such intermediate infinity or that there is. This allows one, by decisions of thought, to extend arithmetic into mutually incompatible systems, ultimately infinitely many such systems, because each on them will contain at least one undecidable proposition. I am not claiming that Gödel's incompleteness theorems form a necessary condition of this kind of multiplicity of possible mathematics, given, for example, that the existence of Cantorian and non-Cantorian set theory, or the existence of Euclidean and the various non-Euclidean geometries established earlier, does not depend on these theorems. Gödel's theorems, however, have given an inevitability of this situation. (I am grateful to an anonymous reader of this chapter for suggesting this qualification.) It is in principle possible to assume that all such possible constructions form a single, if multiply branching, potentiality, ultimately realizable in principle. I shall not assess this view except by noting that even if one adopted it, one would still only allow for a vast constructible mathematical potentiality rather than independent mathematical reality. Would this view be *in practice* equivalent to mathematical Platonism? While it may be in practice, the difference in principle remains important, both in general and because that it would be impossible to assume that, being infinite, this potentiality could ever be realized. In practice, mathematics creates new mathematical realities and, with them, new mathematical potentialities all the time, quite apart from any undecidable propositions.

This process will only end when mathematics is no longer with us, and 1 day it might not be, in contrast to physical reality, which, in either realist or RWR view, is assumed to exist apart from us, to have existed before we existed, and to continue to

exist when we will no longer exist. For now, however, as long as we exist, we can pursue our projects (logoi) of relating to this physical reality, either by continuing to assume it to be beyond construction (the alogon) or possibly by attempting to construct it as representable, by means of constructed mathematical realities that we can objectively share, rather than as mathematical objects existing independently of us.

5 Summary/Conclusion

This chapter introduced the concept of radical Pythagorean mathematics and considered its role in mathematical modernism, a trend that emerged around 1900, and has played a key role in the subsequent history of mathematics and its use in physics extending to our own time. As its name suggests, however, the history of mathematical thinking that this concept reflect is much longer and extends all the way back to ancient Greek mathematics, specifically to, arguably, the greatest discovery of the Pythagoreans, that of the incommensurability of the side and the diagonal of the square. This discovery established the complex relationships between geometry and arithmetic, and correlatively, the possibility of a new type of logos, a logos containing the alogon within it – a logical thought containing the unthinkable within it. I termed the mathematical thinking defined by these two features "Pythagorean mathematics." I have then argued that, as part of mathematical modernism, there emerged a new attitude toward the unthinkable in thought, an attitude that accepts this situation as a positive, enabling condition of thinking and knowledge. Ancient Greek mathematics was, at most, tolerant of this attitude, ultimately preferring what I termed Platonist mathematical thinking, defined by the belief in a logos without the alogon within it. I termed the mathematical thinking defined by the new attitude, which still keeps the irreducible relationship between geometry and algebra (which subsumes arithmetic within it), radical Pythagorean mathematics.

I have argued that the Pythagorean tension between geometry and algebra and yet their mutual dependence on and often symbiotic relationships with each other have never left mathematics and remain part of modernist mathematics and thus present-day mathematics. This is a good thing because they enrich and help to invent new forms of each other, sometimes to the point where one could no longer unconditionally distinguish geometry and algebra, and both topology and analysis may become part of mathematical architectures that thus emerge. That which underlies both and keeps them in this interplay – drifting them apart, bringing them close, intermixing them – may be beyond either, or beyond our thought altogether, thus forming the alogon within the logos of mathematics, even though we can still relate to this unthinkable by means of algebra or geometry, or their complex interplay. This possibility opens the way to radical Pythagorean mathematical thinking, which, against Platonism, affirms the unthinkable as a friend rather than an enemy of thought, because it does not preclude, but, on the contrary, helps the development of new concepts and theories. It only changes their nature, and in the first place the nature of thinking itself by making the unthinkable an unavoidable part of thought.

Thought remains our only measure of the Universe, even if the Universe may, as the Pythagoreans might have sensed, be beyond thought. The thought of this "beyond thought" is still only a thought. But it is also a testimony to thought's capacity to think beyond itself, without which, the Pythagoreans discovered, thought could not be the true measure of the Universe.

6 Cross-References

► Philosophy Behind Mathematics
► Platonism, *De Re*, and (Philosophy of) Mathematical Practice

Acknowledgments I would like to thank Athanase Papadopoulos for productive exchanges on the subjects considered here and the anonymous reader of the first version of this chapter for constructive criticism and helpful suggestions.

References

Bohr N (1937) Causality and complementarity. In: Faye J, Folse HJ (eds) The philosophical writings of Niels Bohr, vol 4. Ox Bow Press, Woodbridge, pp 83–91
Bohr N (1987) The philosophical writings of Niels Bohr, 3 vols. Ox Bow Press, Woodbridge
Derrida J (1995) Khōra. In: Derrida J (ed) On the name (ed: Dutois T, trans: Wood D). Stanford University Press, Stanford, pp 89–130
Einstein A (1936) Physics and reality. J Franklin Inst 221:349–382
Euclid (1989) Thirteen books of the elements, 3 vols (ed: Heath TL). Dover, Mineola
Gray J (2007) Plato's ghost: the modernist transformation of mathematics. Princeton University Press, Princeton
Hacking I (2006) The emergence of probability: a philosophical study of early ideas about probability, induction and statistical inference. Cambridge University Press, Cambridge, UK
Heidegger M (1967) What is a thing? (trans: Barton WB, Deutsch V). Gateway, South Bend
Heisenberg W (1925) Quantum-theoretical re-interpretation of kinematical and mechanical relations. In: Van der Waerden BL (ed) Sources of quantum mechanics. Dover, Mineola, pp 261–277
Heisenberg W (1930) The physical principles of the quantum theory (trans: Eckhart K, Hoyt FC). Dover, Mineola, rpt 1949
Hilbert D (1999) Foundations of geometry (trans: Unger L, Bernays P). Open Court, La Salle
Kahn C (2001) Pythagoras and the pythagoreans. Hackett, Indianapolis
Kant I (1997) Critique of pure reason (trans: Guyer P, Wood AW). Cambridge University Press, Cambridge, UK
Lyotard J-F (1984) The postmodern condition: a report on knowledge (trans: Bennington G, Massumi B). University of Minnesota Press, Minneapolis
Mehra J, Rechenberg H (2001) The historical development of quantum theory, 6 vols. Springer, Berlin
Mehrtens H (1990) Moderne Sprache, Mathematik: Eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjekts formaler Systeme. Suhrcampf, Frankfurt am Main
Negropontis S (2019) Plato on geometry and the geometers. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer/Nature, Berlin, pp 1–88
Nietzsche F (1996) Philosophy in the Tragic Age of the Greeks (trans: Cowan M). Gateway, Southlake

Papadopoulos A (2019) Topology and biology: from Aristotle to Thom. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer/Nature, Berlin, pp 89–128
Plotnitsky A (2016) The principles of quantum theory, from Planck's quanta to the Higgs boson: the nature of quantum eeality and the spirit of Copenhagen. Springer/Nature, New York
Plotnitsky A (2017) On the character of quantum law: complementarity, entanglement, and information. Found Phys 47:1115–1154
Plotnitsky A (2018) "The Heisenberg method": geometry, algebra, and probability in quantum theory. Entropy 20(9):656–702
Plotnitsky A (2019a) Transitions without connections: quantum states, from Bohr and Heisenberg to quantum information theory. Eur Phys J. https://doi.org/10.1140/epjst/e2018-800082-6
Plotnitsky A (2019b) On the concept of curve: geometry and algebra, from mathematical modernity to mathematical modernism. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer/Nature, Berlin, pp 153–212
Polya G (1990) Mathematics and plausible reasoning, volume 1: induction and analogy in mathematics. Princeton University Press, Princeton
Reed D (1995) Figures of thought: mathematics and mathematical texts. Routledge, London
Riemann B (1854) On the hypotheses that lie at the foundations of geometry. In: Pesic P (ed) Beyond geometry: classic papers from Riemann to Einstein. Dover, Mineola, pp 23–40
Silverman J, Tate J (2015) Rational points on elliptic curves. Springer, Heidelberg/New York
Thom R (1988) Les intuitions topologiques primordiales de l'aristotélisme. Revue Thomiste 88 (3):393–409
Weyl H (1928) The continuum: a critical examination of the foundation of analysis (trans: Pollard S, Bole T). Dover, Mineola, rpt 1994
Weyl H (1952) Space time matter (trans: Brose HL). Dover, Mineola

The Metaphysics of Platonism

Ulf Persson

Contents

1 Introduction 92
2 Axiomatics 97
3 Imagination Gone Astray Seduced by Formalism 105
4 Faith in External Reality, Including Mathematics 107

Abstract

This chapter constitutes an attempt to present a modern version of what Platonism in mathematics really entails. Thus one can view it as an attempt to get a closer approximation to Platonic Platonism and not get stuck in Plato's Platonism, which admittedly has great historical value, but when it comes to its specific details only such. In particular, the presentation takes exception to the post-Plato development of Plato's Platonism, the so-called Neo-platonism, which is an exalted version of Plato's thought but shorn off its irony. In particular, it constitutes a literal reading of Plato's metaphor positing a separate abode for forms and interpreting the immortality of the "soul" in personal terms. Those religious ideas are but a distraction irrelevant to the issue of mathematical Platonism, but unfortunately providing easy fodder for contemporary attacks and dismissals of it.

Physics is of course a human invention, but not a frivolous one, but created in response to a physical reality "out there." Mathematics as practiced is also a human activity on par with physics, and the content of mathematical Platonism is that it is also created in response to a "reality" which is "out there." While notions of beauty and virtue are ultimately a matter of social convention, knowledge is not, and just as there exists physical knowledge, we can also talk about mathematical knowledge not only as just being tautological. The latter view becomes seductive if you try to anchor mathematics in Popper's World One through

U. Persson (✉)
Chalmers Institute of Technology, Gothenburg, Sweden
e-mail: ulfp@chalmers.se

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_2

formalism, when it naturally resides in World Three of Popper, whose discovery incidentally Popper attributes to Plato. While beauty in mathematics is on par with beauty in the arts, it is part of a humanistic aspect of mathematics, not its Platonic. To exclusively place mathematics in the humanistic fold invites the abuses of post-modernism, whose classical antecedents are to be found in the teachings of the sophists, tellingly the mortal enemies of Plato.

Keywords

Platonism · Philosophical · Forms · Ideal · Canonical · Metaphor · Axiomatic · Deduction · Induction · Intuition · Proof · Formalism · Mechanical · Metaphysical · Cardinalities · Set · Subset · Mathematical · Meta-mathematics · Logic · Integers · Fiction · Invention · Discovery · Reality · Minds · Experiments · Consistent · Game · Intelligence · Cognitive · Computational · Programming · Physical · Empirical · Countable · Imagination

1 Introduction

Platonism as a philosophical theory is well-known, and most of us probably came across his famous simile already as children. The image of people in the cave is a metaphor, and as all metaphors it is intended to stimulate thought not to kill it, which invariably becomes the case once you take them literally. Metaphors taken literally are indeed just silly, and for that reason Platonism has acquired a reputation of being unsophisticated, indeed philosophically naive, and long since outgrown, and its last worshipers to be found among primitive tribes made up of mathematicians. The analogy with Cartesian dualism naturally comes to your mind. Most of us in our daily lives (including philosophers) make a distinction between the world of the spirit, that is, our consciousness (or "soul"), on the one hand, and the world of matter, on the other. Likewise the working mathematician takes it for granted that the concepts he struggles with are indeed "real," real in the sense of existing outside himself and beyond the reach and domination of his will. It is for this reason, Yuri Manin speaks about Platonism as psychologically inescapable and intellectually indefensible. But is the issue of Platonism really on the same level as Cartesian Dualism? It is my intention to argue that it is not. Now such an argument can never be as convincing as a mathematical proof (and of course there is a great irony here which should not escape most readers), how can you prove that there is a Platonic world, and more to the point what is really meant by that? Traditional deduction in mathematics, as initiated by the Greek, is based on unproven statements as well as means of reasoning that obviously cannot be rationally supported. Contemplating the deductive procedure, we find ourselves in a metaphysical mode in which we cannot escape resorting to faith. As I understand it, the Greek did not strive for absolute certainty, the axiomatic method had as its purpose not to achieve absolute knowledge, for that the Greeks were too sophisticated, but to achieve transparency. A proof can be criticized, which is not the same thing as rejection, but being

amenable to modification and ultimately to improvement. Incidentally this constitutes the core of democracy, but admittedly most people would balk at the suggestion that the best education in democracy for young people would be to study Geometry in the time-honored tradition of Euclid. Yet what most people who acquire something positive from a mathematical education "take home" with them is an appreciation of the clarity and solidity of mathematical reasoning and the concomitant intellectual honesty and ultimate consensus that characterize it, even if they do not necessarily have any mathematical interest nor talent themselves.

I claim that philosophy is the poetry of science. Many people, especially philosophers, do not take kindly to such a suggestion, but see it as a disparagement, indicating groundless speculation, haziness, and obscurity in thought and ultimate ineffectiveness; but this is not my intention, the view of Bertrand Russell and Karl Popper was instead that philosophy is pre-science, and that it develops by amputation, once a field of philosophy has matured enough to possess clear concepts and allow itself systematic reasoning, it becomes part of science and is no longer philosophy. Much can be said for this view. The folly of the grand metaphysical theories, especially those of Hegel, in the eighteenth and beginning of the nineteenth century was that they treated lofty subjects in a systematic and technical way, in other words taking metaphors literally. The Old Hebrews understood this through their refusal to give their God a name; God was not meant to be part of any ordinary human discourse, to be treated as one entity among others. Thus in a metaphysical mode you cannot argue by straightforward arguments but need to resort to evocations, any other approach would just be a travesty. It is in this spirit I am forced to operate in the discussion to follow.

At the time of Plato you could speak about the True, the Good, and the Beautiful, as if they were more or less on the same level; nowadays in a more modern materialistic mode, most poignantly expressed by Darwin's Evolution by Natural Selection, which philosophers can only ignore at their peril, we make a distinction between ontology, that is, what really is, existing independently of man, and what is based merely on human whims and conventions, whose ramifications nevertheless, one should never forget, can be quite far-reaching and tangible. The question of mathematical Platonism is whether it has a real ontology, that is, existing independently of man; or whether it is to be thought of as a human invention on par with Art, Political and Legal constitutions, or Financial conventions, as suggested by Reuben Hersh. This latter view does not mean, that it thereby trivializes mathematics. After all even if you think of morality as merely a human invention, that does not imply that you do not feel strongly about moral crimes and transgressions, and that you do not have a strong intuitive feeling of what is right and wrong, and maybe even ready to die for it. In fact, as I understand it, the motivation for such an attitude is to demystify mathematics and thereby enrich it as a human endeavor on par with the highest achievements in Art. And also the fact that the value of money ultimately rests on a collective convention does not enable the individual to flaunt it. For the individual lack of money could be as real as lack of food. No serious philosopher of mathematics claims that mathematics is subject to the whims of the individual, as

little as an individual can on his own reinterpret the value of money, or moral concepts, although there has been no lack of trying.

Does mathematical objects exist in the same way as atoms? Atoms exist in space-time, but mathematical objects seem to have no tangible existence outside of the minds of humans. Now if you think of it deeper, the distinction is not as clear-cut as might at first be seen as being the case. The original distinction was between objects in the real world, and objects in an ideal world. A straight line, without thickness and irregularities cannot be manifested in the real messy world. But can an ideal line exist, a line as reasoned with in mathematics? There are many "lines" in the real world, and we think of them as approximations of the real thing, which we form an idea of through our minds. One common interpretation of this is that there exists an ideal line, of which all the lines we encounter are but distortions of. This gives the misleading idea that a Platonic object is the canonical object, which when applied to everyday objects such as chairs and tables becomes rather silly. Is there a Platonic chair, and if so how does it look like, can you sit on it? From a more Aristotelian point of view, the ideal object is not given, but constructed in our minds, thus it has no tangible physical existence. (Never mind that according to a material point of view going back at least to Hobbes, even thoughts have material manifestations as arrangements, be they elusive, of atoms in the brain. But the same "thought," whatever that means, can have different manifestations not only in different brains, but in the same brain at different times.) But our life is filled with concepts which we encounter in many very different ways, but which nevertheless are related to each other. To take a trivial example "crossing a street." The concept of crossing a street is of course very different from every instance of it, it would be silly indeed to single out a specific ideal street crossing as the canonical one of which all other street crossings are distorted images of. A dog cannot learn commands without having an abstract idea of what they refer to. For it to obey the command "do not cross the street" it need to understand what this entails in situations it has never encountered before. In what sense can we physically encode a concept? It does not really work to consider the agglomerate of all "crossing the street" situations, yet we cannot deny that the concept has some real existence with tangible consequences. One may of course cheat, as one does when trying to program a concept, by single out some easily identifiable features, but this does not capture the concept of a concept, although it could have its uses.

To bring the discussion down to a more concrete level, what about the integers? The classical way of approaching integers is via cardinalities. Cardinalities are abstractions which belong to conglomerations of objects, not to the objects themselves; in fact the latter does not even make sense. But if we have an empty universe do we have the notion of cardinalities at all, and hence of integers? Some Victorian philosophers such as J.S. Mill would probably argue not. But if we have a profusion of objects, different cardinalities can be manifested in many different ways and then it can be tempted to define numbers as equivalence classes of collections with respect to the same cardinality, but this is almost as impractical as defining a concept, such as "crossing a street" with all the manifestations of "crossing a street." (Frege tried a definition, obsessed as he was of getting an objective, Platonic definition of number

and to avoid the horror of basing it on psychology, according to him the least reliable and developed of all the sciences; but what he came up with was actually this anti-climactic definition.) One may of all the possible agglomerations choose a canonical sequence, such as $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$,... as suggested by von Neumann. This is a clever idea, or rather a clever codification of the integers to be implemented in the machine of logical thinking, and as such with definite applications; but it does not catch the Platonic concept of integers, which of course von Neumann never intended, although it may have been thought so by some; once again a case of the silliness of taking metaphors literally. What we are witnessing here is a twentieth-century attempt to bring integers into the machinery of logic; but as C.S. Peirce remarked, the notion of integers is more basic than logic itself. But here we have a clear-cut example of a mathematical object, more fundamental than the notion of a straight line in geometry. We can talk about ideal lines, as we encounter approximate lines in ordinary life, but we cannot talk about ideal numbers in the same way, because numbers with which we are familiar make already up an abstraction with which we are on very concrete terms with; it is not like that we have encountered defective numbers all the time, and decide that one set is the real ideal thing.

So to return to the initial question of the previous paragraph. What is more real and tangible, the mathematical concept of numbers or atoms? Atoms you can neither see nor touch but are mental constructions, and in fact originally they started out as pure items of speculations, be they the fruit of a powerful paradigm or thought, later on their identities and putative properties were progressively pinned down, acquiring in the process more and more reality, yet as late as the end of the nineteenth century, physicists turned philosophers such as E. Mach doubted their reality, seeing them as mere convenient fictions. (Einstein, who was much influenced by Mach, tried to rebut him by his technical description of Brownian motion, as being caused by random collisions between small organisms, such as bacteria, and molecules. Thus to a bacterium at least, atoms have a tangible existence! But of course they are presumably too small and primitive to experience this tangibility.) Numbers on the other hand have to our minds, and hence to us, a much more intimate existence, although in this regard we may fool ourselves. It is one thing to have a firm feel for small numbers, such as a million, but what about numbers we can define by recursion, by far outstripping anything so called astronomical? In fact thoughts, especially our own (as others we can only know through the medium of our own), so effervescent from a material point of view, are the only entities to which we have direct access. Logically this leads to the position of solipsism, or less dogmatically to idealism. In fact the fashionable attitude of instrumentalism, in which theories are but figments of the mind, is a species of idealism, as Popper pointed out, a stand more and more seductive as modern physics becomes more and more sophisticated. In order to make the discussion of mathematical Platonism meaningful we must posit the existence of an outside world beyond the realm of individual or human thoughts. In other words that thoughts mean something being occasioned by something "out there." We know the world through our senses, but our perception of it should not be confused by our conception of it. We are not instructed by the world, it comes with

no ready-made conceptions, but those we construct ourselves in response to our interaction with it. We form provisional hypothesis which we can test and modify on the basis of the perceptions we receive as a response. We may distinguish between two processes of familiarizing ourselves with our environment, namely the individual and the collective. In the first case it is a matter of making sense of sense perceptions, to be able to "see," "hear" and less cerebrally "smell," "touch," and "taste" and also to use your muscles and move around. Learning how to do so is a matter of automation, in which our conceptual understandings necessarily adhere to certain templates developed through evolution otherwise it would not be automatic and reliably effective. We can talk about evolution through natural selection, as a learning process for the species. In the collective process we try to delve deeper and trying to find an underlying explanatory coherence, in other words to seek the Platonic reality behind the confusing sensual manifestations. We are now speaking about science, in which we try to transcend our biological constraints and free ourselves from the templates with which we are born. This process is not automated and far from having any guarantees of success. As Popper teaches us, this trial-and-error approach carries over a spectrum from the most basic individual and biological to the most sophisticated scientific, so it is indeed impossible to draw a sharp line between the scientific and the individual.

Our intuitive sense of reality derives from the availability of many corroborating senses. The starry sky above us is only accessible by sight, while the terrestrial also allows us to move around in and touch.

> The Greeks were familiar with non-Euclidean geometry, namely the spherical one, but they had no need to think of it as a separate geometry, as the sphere is embeddable in ordinary Euclidean three-dimensional space. Spherical geometry was used for the celestial sphere, only accessible by sight, and Euclidean geometry for the space around us in which we can move. Obviously the celestial sphere is not as real as the terrestrial one (making a pun on the fact that the Earth is approximately spherical). The celestial sphere is an ideal one, and also perfect as a sphere, while the terrestrial one is a messy distortion. I would like to imagine that the Greeks would have made an even stronger distinction between the celestial and the sub-lunar world then they actually did. In fact that they would have believed that the objects in the sky were infinitely far away, and thus for ever beyond the reach of man, only mysteriously known by sight. However, there is of course no historical evidence for this being the case, although the idea no doubt occurred to some without leaving any documentary trace. However closer inspection of the sky do reveal, using elementary principles of Euclidean Geometry, that the Moon is actually at a finite distance from us, although substantial by purely terrestrial scales. From this we eventually derive the idea that the Moon is but a big stone in the sky, and that all the objects in the universe are actually at a finite distance and hence in principle accessible by means other than pure sight, making them much more "real." But what would have happened had we been born on a planet circling a star in the intergalactic void, we would have never been able to reach the lowest rung on the ladder which allows us to climb astronomical distances, the galaxies visible to us, would in fact be assumed to be infinitely far away without any means of falsification. (One may argue that a radiating body infinitely far way would not have infinite energy to send rays across such distances. But of course the physics infinitely far away could have been very different, or maybe the objects on the celestial sky may only have been part of a collective hallucination.)

Fiction differs from reality as there is only one channel, the written or the spoken word. There are many questions which can be asked about real people which cannot be asked about fictional ones. A real person has, for example, parents, who in their turn have parents, and you can ask all kinds of questions about them, but which do not make sense for fictional people (was Sherlock Holmes paternal grandfather bald when he died?) meaning they are neither true nor false. Historical people tend to be perceived as fictional, in fact many fictional figures live more vividly in the minds of people than most historical personages, to that we will return.

There is a material world out there, it is described and explained by subjects such as physics, chemistry, and biology. Those subjects are pursued by human beings and they exist in the minds of humans, but physics shall not be confused with the physical world. The pursuit of physics is made possible by interacting with the physical world, and this interaction can be very direct as being subjected by mechanical forces or very indirect through sophisticated experiments. (Never mind that our friend Mach considered forces "fictive.") The reality of our conception of it depends on that there are so many channels through which it can be tested, and the ways they cohere instill in us our convictions. Convictions are of course subjective feelings belonging to the realm of psychology, but after all we are human beings; but at the heart of Poppers claim for science stands that a falsification is real and objective. (In practice an experiment is complicated and its interpretation rests on further tested assumptions hence subjective convictions enter, yet this does not really invalidate the underlying principle.) Mathematics is different as the study of mathematics seems to coincide with mathematics itself. Thus if the study of mathematics is a human activity on the par with the pursuit of physics, mathematics itself should be a human activity and nothing more. In fact it is popular to think of mathematics as a language, following the adage of Galileo, that Nature is written in the language of mathematics. In fact much of the activity of theoretical physics consists in mathematical manipulation (which can be seen as a kind of experimentation) while that of experimental physics is more of an engineering activity. (In string theory the only resource is mathematics, experimental falsification being (so far) not feasible in fact inaccessible.) The implication of mathematics being a language is that it is ultimately free of solid content, that it only transmits contents, through organizing hard synthetic facts, not that it creates any. Mathematics by itself does not say anything about the world, just as language by itself is vacuous, it just conveys. This presumed vacuity can also be expressed in other ways, as we will see.

2 Axiomatics

The classical way of grounding mathematics to a solid reality is to present it axiomatically. This means treating mathematics as a game, rules are set up, and the game consists in following the rules. Originally the axioms and postulates were intended to match reality, and hence they were picked with great care. The axioms of Euclid were not chosen randomly but felt as being intuitively true. Once you had the axioms, which were meta-logical, you could use some simple logical principles to

churn out a wealth of propositions. Mathematics became applied logic. The rules of logic are ironclad and proceed with an inexorability akin to mechanical forces in the physical world. Mathematics literally kicks back at you, and you cannot flaunt its strictures. The experience of working with mathematics becomes a struggle against a resistant medium beyond your control. It is very different from writing fiction; in fiction you can impose your will, in mathematics you cannot. If the inequalities go the wrong way, you cannot just cheat and turn them around, as would be your privilege as an author of a tale. (But even authors of tales experience that they do not have complete freedom. Part of the restrictions are of course due to restraints imposed by a desire for realism and concomitant credibility, others more interesting is that authors report that characters take on a life of their own and "do" unexpected things that the author simply has to accept. Such experiences are no doubt genuine, yet it is difficult to conceive of them to have any objective character beyond that of psychological projections.) Now likening mathematics to a game is of course a metaphor not to be taken literally. In the typical game, such as chess and go, there is a competition between two opposing minds, and a mathematician would ask very different questions of the game that an actual player would. (This is why mathematical ability and playing say chess has not very much in common pace popular mythology.) Playing chess is a human activity, the objectivity lies in the fact that you can judge whether a game is played correctly or not, and the players come up against restraints given by the rules as well as the positions on the game. You cannot play chess against yourself (although a machine can). If we include games such as solitary, the simile gets closer to mathematics, yet even in solitary games you are not supposed to redo moves, while in mathematics and science in general, redoing moves constitute the essence of the "game" in the form of "trial and error." The relevant notion of a "game" is not the one of competition, where human involvement is crucial as well as there being a definite goal namely the notion of winning, but game as something well-defined and thus apart from "real life" implying frivolity and ultimate irrelevance. They do exist, and one can easily argue for their ontology, which basically consists in that rules may be arbitrary and chosen by humans, yet the consequences are unpredictable and beyond our control, thus having a tangible existence, regardless of relevance or interest, constituting a kind of self-sufficient mini-universe.

One aspect of the "game" nature of mathematics inherent in axiomatization is formalization, championed by David Hilbert. (Hilbert is often described as a formalist when it comes to mathematics. I think this is unfair; his formalism does not necessarily reflect his personal philosophical attitude toward mathematics, to him formalism was merely instrumental to serve some specific purposes, namely to prove once and for all the solidity, that is, the consistency of mathematics.) It means that the objects a mathematician deals with have no meaning in themselves only the ways they are being manipulated have meaning. That is, they are subject to specific rules in a game, and have no existence beyond that. Related to this is the notion that mathematics is about nothing at all and we have no notion of what we are speaking, as provocatively expressed by Russell; and furthermore that it is nothing but a collection of tautologies. Closely related to this is the idea that deduction does not

create any new knowledge (as opposed to induction), that all the conclusions are already to be seen in the premises (the axioms). The underlying assumption, or should we say vision, is that deduction can be done mechanically (in principle if not in fact) and one can imagine that all strings of permissible symbols are systematically generated and in particular all valid deductions. A variant of this, inspired by Kolomogorovs quantification of randomness, argues in the same vein, that the information content of any proposition in a logical system cannot exceed that of the axioms. The point is that any game creates a mini-universe which may refer to nothing at all, but by virtue of its existence, becomes something by itself of which you can ask all kinds of concrete and meaningful questions. Of a logical system we can ask the question of whether it is consistent, and the iconic inquiry of that type was supplied by Gödel, who noted that any game (formal system) can be encoded by the integers, thus any question about it can be coded as an arithmetical question, thus almost, but if not quite, embed a meta-question into one expressed in the system itself and thus draw the famous conclusions (provided of course that it it is powerful enough to encode the integers as well). A strict formalistic approach prohibits meta-mathematics (statements about mathematics) to become part of mathematics itself, just as the Hebrews were not allowed to give their God a name, and for that you pay a price, known as Gödel's theorem, and an application of Cantor's diagonal trick.

The deductive method is of course not confined to mathematics, but only in mathematics does it manifest its pure (Platonic?) form. The prime extramathematical application is of course the legal system. The legal system is an ambitious attempt to encode what is right and what is wrong. This is made much more complicated by the fact that not only the acts themselves enter, but the intentions behind them as well. It is interesting to note that the legal systems goes much further back in history than does mathematics and science proper which we should date to the pre-Socratic philosophers. The existence of legal codes and the institutions which go with them is inseparable from civilization itself, and as indicated above, it stands to reason that the axiomatic system was inspired by the legal system. There are of course clear analogies which make comparisons meaningful and thus highlight the crucial differences. While a set of axioms is usually quite simple, in modern societies the law consists of a huge number of individual regulations compiled to a large extent on ad hoc decisions. Furthermore the legitimacy of say the axioms of Euclid rests on a shared intuition (and when axioms are no longer based on intuition they form a mini-universe whose existence cannot be denied only tolerated), those of a legal code depends on external authority. Laws restrict the actions of individuals and societies in very tangible ways, mathematics restricts as well, but not so tangibly to most people. And more importantly laws are imposed by human will affecting that of others often contradicting them. Mathematical results have to be derived; laws have to be interpreted, because of the latter there has to be lawyers. This comes to the fore during a trial. The point of a trial is to come to a verdict, preferably within a short amount of time. Guilt or non-guilt is to be established (and sometimes a middle is not to be excluded) based on testimonies and through a deductive contest between a prosecutor and a defendant the outcome to be judged by some kind of jury, at times reduced to the judge. The aim of a trial is not to achieve absolute justice, and

although thought of as an inspiration it is not thought of as a realistic ambition; but to get a formal, that is, legal, justice. This reflects human fallibility. One may prefer the judgment of a single wise mind with a direct access to truth, but the advantage of a more formal procedure is that it is open for inspection. But due to the formality of the process it can be, and often is, distorted. The mission of the two main actors is not primarily to establish the truth but look out for the interests of their clients. As those interests are opposed this is seen as a check, both against omissions of relevant issues as well as specious reasoning. Cicero bragged that his final speeches were not supposed to survive a critical scrutiny the next day, but to dazzle the jury for the moment. The procedure is thought of as being democratic and consequently encouraged rhetorics where arguments are chosen to persuade and convince, not necessarily to lead to the truth. This is an attitude embraced by the Greek sophists, the mortal enemies of Platonism, and as such ironically more human. The correctness of a proof is from a Platonic sense independent of human opinion, in practice it is up to the judgment of experts. One may, as noted above, see mathematical proofs as Platonic entities beyond the ultimate judgment of mere human mathematicians. (Is Wiles proof correct? Most people could not tell, so a committee of experts is appointed to decide the matter. Admittedly they have no private axes to grind, and they are in principle indifferent to anything but the truth, yet they are human. So in the eyes of society there is no real difference between a proof deemed correct and a culprit proved guilty. We do feel that the putative guilt of someone is a matter on which there is a clear answer, after all a specific event once took place, but any human verdict is inevitably provisional and can be upset by the events of the future. In the same way, the fact that a proof may be deemed correct by a human committee does not mean that its ultimate correctness is just one of a social convention, as opposed to its actual acceptance by society.)

We have already referred repeatedly above to deductive reasoning being mechanical and unforgiving hence ultra-human. The mind of the mathematician may be very similar to the mind of a lawyer, and in fact many great mathematicians such as Fermat and Cayley were lawyers professionally (but their passion was of course mathematics). One may point to the same desire for precision in pursuit of clarity, and the same predilection for stringing long chains of deductive reasoning. While in mathematics such an approach is natural, its legal implementation appears in comparison to be pedantic not to say a parody, the reason for that being that the world of human affairs is very complex and confusing. Utmost precision needs a very precise language in which deduction is performed, and this allows the deductive process to be literally mechanized not just in principle.

Logical reasoning has much in common with machines. To put together a machine you have to have a clear idea of your intentions, and the way of trying to implement them are subject to the nonnegotiable restraints imposed by an unforgiving external reality. Some of those restraints can be formulated but most of them are discovered by trial and error. The guiding principle is one given by cause and effect, although the existence of such notions is up to doubt by skeptical philosophers. (Most notably by Hume. Cause and effect implies a “before” and an “after” and hence some notion of time. Mathematics, as opposed to human mathematical

practice as documented by the history of mathematics (which is not so much a history of mathematics as a history of mathematical practice), is in its Platonic sense, timeless. A mathematical theorem may be proven in many very different ways, without us asking which is the true cause of its veracity. Even in physics, so heavily mathematized, one may explain a phenomenon in different ways, such as reformulation of Newtonian mechanics using Hamiltonian, without us asking what is the real cause. Such fanciful explanations of the reflection of light due to Fermat, later to be developed into principles of least action initiated by P.L. Maupertuis, are fully legitimate as long as you do not pursue the antropomorphic interpretations they cry out for. Another case of metaphors taken too literally.) A machine exists first in the imagination of its inventor, but the process of translating a vision, often rather vague, into a physical object is not straightforward. External reality stimulates the imagination by putting up obstacles to overcome, the imagination is starved as it needs resistance to become properly stimulated. The translation involves a lot of trial and error, and thus involves a learning process, which will come in good stead for future projects. The finished result has grown out of a process of communication between the mind of the inventor and a resistant external reality. Once it is there as a material object it will exhibit a wealth of consequence which were never intended nor even predicted. Thus the object goes beyond the initial vision and becomes independent of its creator, and in the process blurring the distinction between invention and discovery which is often held up as crucial in any exploration of mathematical Platonism. A human invention is characterized by being the brain-child of a mind and thus brought forth by intention. Thus it is meaningful to ask "why"-questions; in fact it is essential. If you want to assemble something you can either follow a manual step-by-step, this is considered fool-proof, but often very frustrating; or better still you want to fathom the underlying thought which explains why the pieces look as they look and how they are meant to fit together. (It seems as if people in general prefer to get instructed by a Global Positioning System (GPS) rather than to figure things out by looking at a map. The same goes for mathematics, most people seem to prefer given explicit step-wise instructions on how to do it, rather than being given a "map" and figuring things out for themselves.) The inventions of tools goes way back into human prehistory and does not seem to be confined to humans alone, although admittedly no other species has pursued it to such lengths inventing machines to construct other machines and so on. The ambition also to implement mathematics into machines, to turn the spirit into flesh, has also a long story. One of the first examples being the abacus. (There is constant amazement on how the Romans could survive with their clumsy method of representing number, how could they compute with them at all? Those numbers served other purposes I suspect, purposes to which they are still used, when it came to calculation I am convinced they used mechanical apparati such as the abacus. The human capacity to subitize is very limited and does not go beyond four (just try to imagine in your mind a regular heptagon!) and is not superior to that of many mammals and birds such as rats and crows. Many human cognitive abilities do not measure up to that of many animals, thus the need of mental inventions, such as numbering systems and counting schemes.) More sophisticated machines were

thought up by Pascal, Leibniz, and Babbage, having the basic ideas but not the technology to support them, technology being in the nature of a collected human inventions. (Crucial to technological process is also non-material inventions; it is not enough to invent machines, they also have to be integrated in society's infrastructure, the interplay being rather subtle and complicated.) Traditionally the most sophisticated machines, if not necessarily the most useful, were clock-works, exhibiting the most intricate "deductive" chains of constructions, and one may cynically suspect why this made clock-work intellectually exciting. (Accurate clocks turned out in fact to be rather crucial for the determination of longitude with undeniable far-reaching practical applications in sea-faring. Apparently if you take apart one of Harrison's masterpieces, you see in the mechanism various dead ends corresponding to experiments, but which could not be "edited away" as in modern computer programs.)

Implementing visions into bona-fide physical machines are hard because unlike in logical reasoning it is hard if not impossible to break it down into simple discrete steps. In fact even the simplest steps can run into unforeseen difficulties. One says that some people are handy, while others are fumbling and hence cannot implement the simplest steps. The truth is that a skilled craftsman is not so much a question of a steady and nimble hand as mental understanding of what has to be done, but steadiness of hand and nimbleness of fingers do not hurt. (It is popular to contribute the growth of human intelligence on the one hand to manipulation of tools, and on the other social interaction. It is not true that there is an opposition between practical intelligence and theoretical, in particular not that a lack of the former is a sign of the latter. Newton was a master of both. And even in purely cerebral activities one may make a distinction. Gauss in his computations continually came up with clever tricks both out of necessity and pleasure, while Euler seems to have slugged through becoming blind in the process. One may speculate that as Euler had a tremendous memory, knowing large parts of literature by heart, this came naturally to him, while Gauss did not have such questionable gifts even if his computational cognitive capacities were impressive enough.) However, the programmable digital computer changed all that, enabling unpractical people, who had no talent nor desire to acquire so-called tacit knowledge, to construct machines through clear and simple step-wise processes, thus enabling straightforward implementation of logic. Now every program may be seen as a machine, be it virtual; as a logical system it can be seen as successful (consistent) if it runs, and useful and interesting if it eventually comes to an intended goal. (Getting to the goal, however, can take a long and unpredictable time, the so-called halting problem whose intrinsic undecidability is closely akin to Gödels incompleteness theorem and makes nonsense of many theoretical criteria such as Kolmogorov complexity.) A program also starts as a vision, and its implementation is also a matter of trial and error as anyone who seriously has tried to do programming has programming has inevitably found out. Although one may conceive a mathematical proof as a program, in practice programming has much more in common with constructing physical machines than the kind of activity which mathematics involves. (In mathematics you often get stuck, while in programming you can have a dialog with a machine in a much more palpable sense.) A program (and its execution) can be codified in a long string of "0s" and "1s," but such a string

written down on a piece of paper is inert and harmless, but once implemented in a physical machine (the hardware, the details of which tend to be completely unknown to programmers) which furthermore is integrated in the infrastructure of Society can cause unpredictable havoc. The perceived dangers of evolving AI lies in fact in this integration and in any ambition to be able to contain it, we have our best chances. However, the idea of a chain reaction of evolving intelligence is based on the ability of a program (a seed) to produce even better programs. Or in more abstract terms, to create an intelligence superior to human in the sense of being able to do all what humans can do, including creating an intelligence superior to its own, and so on. We note here how the abstract notion of intelligence which AI people seem to take for granted is given a Platonic existence without any protests. The crucial assumption that a transhuman intelligence can create an intelligence superior to its own reminds you of the ontological proof of God's existence due to Anselm. This is no coincidence as much of ultra-logical thinking is in the spirit of medieval scholasticism.

However, the general technological development seems to feed on itself and its course, and it and its impact on society is unpredictable, as any invention carries with it unintended consequences. It is far from certain that it will have a beneficial effect on humanity, as far as the latter has any objective meaning. Also on a more fundamental level there has been a blind development with a far more important impact. Ever since Big Bang matter, as we know it, has been created, more and more chemical elements have emerged as a result of nuclear processes in stars and their sometimes dramatic finales. Furthermore more and more complicated chemical compounds have arisen, in particular those that would eventually define life, at least on Earth, putting on stage a biological evolution of DNA sequences. The point is that this evolution has not been designed, and hence it is pointless to ask "why"-questions as Aristotle was fond of. Thus one may present as a criteria of the natural as opposed to the artificial, the absence of conscious design, in particular things arise blindly devoid of any aforethought. But Aristotelian final causes (so-called teleological) die hard, and there is a general feeling that evolution inexorably tends to higher and higher complexity (as in the perceived explosion of AI (which incidentally is supposed to evolve naturally, and thus should not be termed artificial)), although this is in contradiction with the physicists dogma of the inevitable rise of entropy.

There has been attempts to axiomatize nature, in other words to lay down true and complete natural laws. Once this is done, the pursuit of science can ostensibly become independent of empiricism and done, so to speak, cerebrally and formally. The first attempt to do so was Euclid axiomatization of physical space. (I am grateful to Davis and Hersh, who pointed that out in one of their books.) The very fact that Euclid is concerned with an external reality explains the excitement one as a schoolboy encounters because it gives you a feeling of immense power over that spatial reality by incorporating it all in your mind. You would not have felt the same excitement if you had been presented with the "axioms" of groups and asked to play along. Much of the ancient Greek legacy has been lost and it is possible that it contained extensions of Euclid who also wrote on Optics presumably in the same spirit. Archimedes also pursued physics in a mathematical sense, be it with physical

intuition he also applied to purely mathematical problems. (But such physical demonstrations did not qualify as rigorous proofs in his mind, only as suggestions of what could be true.) Naively at least it would not have been so difficult to formally axiomatize, and Newton tried to write Principia in the style of Euclid. In fact Classical Mechanics became for all intents and purposes a mathematical venture yet not totally separated from empirical considerations. The mechanical world can be seen as a real manifestation of a putative axiomatic system; as it has a real presence one concludes that it is consistent (In the parlance of logicians, it has a model). Also it means that there are no "hidden elephants," striking phenomena meets the eye, so to speak, while in a purely formal axiomatic system there may be striking theorems that have not yet been discovered and the ignorance of which prevents you from viewing its universe properly. The dynamical world is thereby richer than the merely geometrical one. The visual reality that meets us is that of changing perspectives, which could have been worked out by the Greeks, but if so there are no surviving documentations of it, it was of no concern to Euclid. (All the mathematics is there of course, but maybe there was no need for it, after all the theorems in Euclid are concerned with "Die Dinge an sich" not their ephemeral appearances to the human eye. It also illustrates the principle of selection which comes into play in the human pursuit of mathematics.) The mechanical world presents us with a wealth of phenomena that demands our attention and explanation, so even if you pursue it purely mathematically, say as an Euler you are aware of the physical reality around you, both as a corroboration and as an inspiration. (Even if Newton pursued physics mathematically, he also had a hands-on approach I doubt Euler had. His empirical investigations into the nature of light still provide an example to follow when it comes to doing science with its mixture of theoretical hypotheses and empirical observation.) While classical mechanics can be axiomatized the reasoning about it tends not be as formal as the mathematical which more closely follow the ideal set up by logic. This divergence between the mathematician and the physicist has become more and more pronounced. Some people (The British mathematician Brian Davies, for example) claim that this divergence is an argument against mathematical Platonism; I would argue to the contrary, it merely shows a divergence of mathematical practice, not mathematics itself. To axiomatize science means first to find all the natural laws and to formulate them in such a way that they can be manipulated by formal logical steps. Both of those conditions, especially the first, seem doubtful ever to be accomplished. Natural laws are for one thing provisional as Popper emphasizes, or as Hertz remarked, everything you learn from experience can be corrected by it. In fact from this point of view, mathematics is not, according to Popper, a proper science, as their conclusions are beyond doubt; in other words deduction versus induction. (As the British philosopher of history R.G. Collingwood put it: Deduction compels, induction allows.) One may argue, as did Mach, that scientific theories are not part of reality but of human imagination. (In the terminology of Popper, not part of World Two but solely of World Three.) And instead of trying to encapture reality into axiomatic systems, once and for all, one speaks about (mathematical) models, where rather than of having reality trying to approximate an ideal world, tentative approximations to reality are proposed subject to unending modifications prompted by empirical encounters. As an example we have in quantum

mechanics particle models and wave models both useful, but which one is true? This is a meaningless question, because we do not know "Das Ding an sich" and we have only more or less useful conceptions serving in different circumstances. Physics is a very mathematical science, and hence it is, as indicated above, somewhat misleading to speak of physics as applied mathematics, the physicists develop their own mathematics so to speak, while in other fields, say as economics mathematics is supplied more or less wholesale by the mathematicians. (One may argue that Game theory as introduced by von Neumann and Morgenstern is indigenous economics mathematics; true its most relevant applications are in economics, but it is a piece of pure mathematics conceived by a mathematician. It is also noteworthy that one may speak about a physical intuition powerful enough to solve problems in pure mathematics, but not of an economics intuition to perform the same feat.)

3 Imagination Gone Astray Seduced by Formalism

Even in logic there is an empirical element, as suggested above, in mathematical exploration you may encounter hitherto unknown principles of reasoning which makes it hard to completely encode a deductive process. Such principles of reasoning, however persuasive, may lead you astray, because after all principles of reason may be part of human psychology rather than of an objective world. In fact the fancier flights of logical reason may land you in a Carrollean Wonderland. A strict logical and axiomatic approach to mathematics runs the risk of becoming a travesty of mathematics. The effort to ground mathematics in Poppers World Two should be likened to represent pictures pixel by pixel. It has its undeniable uses, but a pixel representation does capture the essence of mathematics as little as it does of a picture. (Given the same pixels, it may very well induce different perceptions among different people; a fortuitous face, for example, could become obviously apparent to some but remain hidden permanently to others.) Convictions in mathematics are not objective; in particular they are not induced by long chains of deductions. What makes a mathematical statement believable in a deeper sense is how it fits with everything else in mathematics. As noted initially, our sense of reality derives from a multitude of channels which all cohere. The mathematics which comes out of logic and the language of set theory stands somewhat apart from mainstream mathematics and does not really engage with anything else. I am thinking of Cantors hierarchy of cardinalities which does have something shadowy to it. For a working mathematician there is the countable infinity and then the uncountable, not much more is necessary to consider. (Dieudonné in his textbook *Foundations of Modern Analysis* says that apart from the negative result that the reals are not countable nothing else is needed. The uncountability of the reals is of course crucial, without it measure theory with countable additivity would not be possible.) One could argue that uncountability is a metaphysical concept (and should not be pursued further. Remember do not give God a name!).

> The crucial idea is the notion of a subset, which we intuitively think of as going through all the members of a set one by one and deciding at each instance whether the element should be

included or not. This is an argument done in your imagination and seems innocent enough and when it only concerns a finite subset, especially of a finite super-set, it appears almost trivial. But finite numbers can be very big indeed, in fact a number defined through your favorite recursion could be mindboggling indeed and induce a sense of vertigo far more powerful than infinity itself, which is but a negative concept. (In some sense, except the strictly logical, we can embed the Cantorian infinities in the finite ones!) Now a subset defined by an infinite number of independent choices cannot be expressed in any finite way, those which can, thus based on some rule, will obviously only be countable, but all subsets form an uncountable set, by the Cantorian diagonal trick. (The Russell paradox is widely known. but of course it is explicit in Cantor's proof that the power set has higher cardinality than the set itself. So once you have a set that can generate its own power-set you have opened the lid to Pandora's box, the underlying idea behind the intelligence explosion.) Then you get the apparent paradox that you can model an uncountable set in a countable, due to Skolem. Seen from above, i.e. with the meta perspective, its countability can be detected but not from inside, as the sought for 1-1 correspondence does not qualify as a subset with a finite rule. This is highly unsatisfactory from an ontological point of view, as we have the notion of all subsets, not just the ones given by rules, which seem to us to exhibit a mere epistemological short-coming. But is one thing to imagine the subsets of a countable infinite set and one of an uncountable one, if we try, our imagination cannot go beyond that of picking one member at a time. Of course we can define subsets of the reals easily by rules, but it is hard to conceive of any other way, in particular the totality of all. This goes of course back to the Ancient Greeks and Zeno's paradox. We have no problems in our imagination to pass along a continuum and in a sense passing through every point, but in the classical paradox, we define a countable infinite subset at each specified point we cannot but stop to pause and become aware and thus the process seems never to end. In a similar vein imaginative processes such as the axiom of choice enter, hard to imagine, but maybe even harder not to imagine. (One should be careful dismissing things as being beyond the imagination. We can easily imagine two or three balls in front of us but not a thousand, although we could easily (if tediously) draw a thousand balls on a piece of paper. So in what sense do we imagine or not imagine? Mathematics is a way of greatly extending our powers of imagination beyond the directly imaginable. Even in mainstream mathematics one encounters "questionable" constructions such as the Absolute Galoisgroup of the rationals, for which we have great difficulties finding explicit elements apart from complex conjugation. But here we have a concept which ties in with others and hence allows alternate approaches.) Thus we are forced to make more or less artificial constraints as to what should count as a subset which seems to be on the same level as human legislation. In fact the Zermelo-Fraenkel system of axioms, ostensible the axiomatic basis of all of mathematics but which most mathematicians could not care less about, seems to be a collection of traffic rules to prevent undesirable pursuits of inquiry. Furthermore the inability to settle the Continuum hypothesis goes against our naive sense of the palpability of the real line. Gödel's idea is that we have not yet found the right axioms for set theory, when we find them we will recognize them. This is a Platonic statement if any, and ties in with Plato's idea that knowledge which is revealed to us is just a matter of remembering things we have forgotten, an experience which well describes the sensation we have of suddenly understanding something pointed out to us as if we have known it all along. Also recognizing things we have forgotten is a common human experience. (Admirably reported on in St-Augustine's confessions)

The meta-mathematical ideas which Cantor had introduced caused strong resistance that the logic and imaginative feats required seemed questionable and self-referential. It caused a crisis in mathematics, or rather the philosophy of mathematics; for mainstream mathematicians they did not really matter and they went on doing business as usual, pragmatism taking precedence over purity of thought.

Intuitionism which was advocated as a remedy against embarrassing paradoxes was never taken seriously by mathematicians in general but seen as a pedantic curiosity. Hilbert's attitude was one of tolerance and the need for freedom, if Cantorian arguments would turn out to be useful in getting results why should we not embrace them? In the end we would be vindicated according to Hilbert, hence manifesting a Platonic view. Time has shown that many classical existence proofs in analysis could be given constructive proofs after all. Maybe intuitionism is a matter of mathematical epistemology rather than ontology, just as the use of non-standard analysis.

4 Faith in External Reality, Including Mathematics

We do believe in a real world out there, something that lies beyond our imagination, even if our imagination of it is the only thing we can directly experience. The reason being, as having been noted repeatedly above, that all those perceptions fit together so well. The same is the thing with mathematics. It provides a web in which everything indeed fits so well together and we never cease to wonder at how the most disparate parts of mathematics connect to each other. As a feat of collective imagination it is unsurpassed, and as the imagination can only be stimulated by external constraints, there must be something out there that corresponds to it, just as an outside reality must correspond to the coherent web of perceptions we have. There is a remarkable consensus when it comes to mathematics which has no real counterpart in other human activities, including science. Of course there will be great differences in opinion when it comes to the importance of various mathematical results and disciplines, as well as the worth and standing of fellow mathematicians, whom should be prized or not, but when it comes to the core, what persuades one persuades the other, and if there is no persuasion one way or another it is due to lack of effort and motivation. To be told in mathematics that your argument is wrong is in principle not taken personally but accepted for what it is, a fact against which it is pointless to oppose better to accept and become instructed. The error is there, not created by the opponent, but in fact existing independent of him, your "opponent" is only a messenger. There is a huge difference between reasoning in mathematics and interpreting a law; the former is inexorable while the latter a kind of negotiation with extra-legal agendas. A sly lawyer looks for loopholes due to accidental deficiencies in the law to take advantage of. What would be the mathematical equivalence of a loophole? In artistic endeavors there is inevitable subjectivity, be it at times collectively shared. Art does not penetrate deeper than the "soul" of man, although admittedly this can be deep enough. There are of course strong artistic elements in mathematics, seeing its manifestation as a feat of imagination, but this imagination does not spring entirely from man, but from an external encounter. (There is no real contradiction to view the pursuit of science in artistic terms, and thus see it as a humanistic project. As has been noted above, theories are the fruits of imagination and involve the inventions of concepts, be it in physics as well as in mathematics. Those concepts, such as holomorphic functions, K-3 surfaces, and groups, are of course human inventions, although some concepts seem more human and ad hoc

than others, but they are subjected to constraints beyond our control, and involve unintended consequences. To define the notion of a holomoŕphic function say, does not mean that we thereby know everything about it. A classical artist concerned with the problem of mimesis becomes involved with an interaction between the tricks he invents in order to transfer his vision into a resistant medium and an observable reality. Thereby he is a scientist, as artists such as Dürer and da Vinci once saw themselves as. One only needs to be reminded of the laws of perspectives, which is mathematics. The singularity with mathematics is that its reality is "invisible" and thus the metaphor should not be carried too far.) This is ultimately not a logical argument, how could such be supplied, but a protestation of faith on par with our faith in an external reality, a philosophical stand of realism, without which science would be impossible. The further we probe, the difference between mathematics and science seems to blur.

Mathematical Practices Can Be Metaphysically Laden

Colin Jakob Rittberg

Contents

1 The Set-Theoretic Pluralism Debate 111
2 Woodin's Ultimate-L Argument 115
3 Hamkins' Multiverse View 121
4 Concluding Remarks 128
References 131

Abstract

In this chapter I explore the reciprocal relationship between the metaphysical views mathematicians hold and their mathematical activity. I focus on the set-theoretic pluralism debate, in which set theorists disagree about the implications of their formal mathematical work. As a first case study, I discuss how Woodin's monist argument for an Ultimate-L feeds on and is fed by mathematical results and metaphysical beliefs. In a second case study, I present Hamkins' pluralist proposal and the mathematical research projects it endows with relevance. These case studies support three claims: (1) the metaphysical views of mathematicians can shape what counts as relevant research; (2) mathematical results can shape the metaphysical beliefs of mathematicians; (3) metaphysical thought and mathematical activity develop in tandem in mathematical practices. This makes metaphysical thought an integral part of mathematical practices.

Keywords

Set theory · Set-theoretic pluralism · Metaphysics · Ultimate-L · Mathematical practices · Set-theoretic multiverse

C. J. Rittberg (✉)
Centre for Mathematical Cognition, Loughborough University, Loughborough, UK

Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussel, UK
e-mail: Colin.jakob.rittberg@vub.be

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_22

In this chapter, I argue that the metaphysical views mathematicians hold can shape what counts as relevant research in mathematical practices. I will furthermore show how these metaphysical views are in turn shaped by the results of formal mathematical research. This makes metaphysical thought an integral part of mathematical practices: mathematical practices are metaphysically laden.

To make my argument, I present two case studies from set-theoretic practice. Set theory is arguably more heavily shaped by the metaphysical views of its practitioners than other mathematical practices. In set-theoretic practice, it is particularly visible how metaphysical beliefs and mathematical activity shape each other. This makes set theory a suitable mathematical discipline to study for the aims of this chapter. It also limits my sample space. I only show that one mathematical practice is metaphysically laden, not that all of them are. Hence the "can" in my title. I will come back to this point in the conclusion of this chapter, where I will briefly suggest why it is reasonable to expect that some other mathematical practices are metaphysically laden as well.

Set theorist Hamkins (2012, p. 449) supports the claims of this chapter when he writes that "the mathematician's measure of a philosophical position may be the value of the mathematics to which it leads. Thus, the philosophical debate here may be a proxy for: where should set theory go? Which mathematical questions should it consider?" Hamkins then introduces the *modal logic of forcing* and *set-theoretic geology* as two mathematical research projects which are inspired by his proposed philosophy, the *multiverse view*. This is one direction of the reciprocal relationship between philosophical thought and formal mathematical research I explore in this chapter: Hamkins' metaphysical position, the multiverse view, gives relevance to these two set-theoretic research projects. As for the other direction, Hamkins argues for his multiverse view on the basis of the set theorists' experience with certain formal mathematical results. The claim of this chapter is that philosophical views and mathematical projects develop in tandem; neither is prior to the other. Benedikt Löwe (personal communication, 2016) reported that when he and Hamkins started to collaborate on the modal logic of forcing during a visit of Hamkins' to Bonn in the summer of 2004, their work was neither motivated by nor based on any philosophical considerations about the nature of the multiverse. I argue in this chapter that the metaphysical views mathematicians hold and their formal mathematical activity develop together and may inform each other along the way.

I base my argument on two case studies: the interpretations of the undecidability-situation in set theory by Woodin and Hamkins. The set theorists are currently having a debate about how certain set-theoretic results should be interpreted, the so-called *set-theoretic pluralism debate*. Roughly, Woodin argues for monism and Hamkins for pluralism. Both agree that mathematical objects exist in a vaguely lower-case p platonic sense (Hacking 2014), but they disagree on the nature of this existence; their disagreement is about metaphysics not ontology. Whether and how ontological or other philosophical differences are a part of mathematical practices is left for future research. In this chapter I argue that mathematical practices can be metaphysically laden in the sense described above.

In Sect. 1 I introduce the epistemic situation the set theorists are currently in. I focus on the issue of undecidability and set-theoretic monism as the orthodox view in set-theoretic practice.

In Sect. 2 I present and discuss Woodin's Ultimate-L argument. I introduce the Ultimate-L axioms, Woodin's arguments in support of these axioms, the mathematical program they drive, and highlight the reciprocal relationship between the mathematical results and metaphysical views in Woodin's arguments.

In Sect. 3 I present and discuss Hamkins' multiverse view. I explain his pluralist proposal, introduce the modal logic of forcing and the set-theoretic geology projects, and point to the interplay of Hamkins' metaphysical views and his formal mathematical research.

In Sect. 4 I give a short summary of the preceding sections to highlight salient points and briefly suggest why it is reasonable to expect that other mathematical practices are metaphysically laden as well.

1 The Set-Theoretic Pluralism Debate

Rouse (1996, p. 27) has pointed out that "what scientists share is not a background of beliefs but a situation that they may belong to and understand in partially different ways." In this section I argue that this is also true for set theory. My focus is on the situation as given by the mathematical results set theorists have obtained.

There are limits to the formal machinery of set theory, e.g. the independence results, with which set theorists routinely come into contact (see e.g. (Džamonja and Kant 2019; Rittberg 2016b) for accessible overviews). Individual set theorists interpret this situation in partially different ways: some seek to overcome the paucity of the formal machinery of set theory by proposing new axioms, others interpret the independence results as a sign of the limitations of our epistemic grasp on set-theoretic propositions, and others still understand the situation they are in as revealing metaphysical truths. In this subsection I elaborate on these points by presenting the situation the set theorists currently face in more detail.

The currently accepted system of axioms of set theory are the Zermelo-Fraenkel axioms with Choice, ZFC for short. There are set-theoretic propositions, such as the Continuum Hypothesis CH, which can neither be proved nor disproved from the ZFC axioms. That is, even if we accept the ZFC axioms as true we still do not know whether CH is true or false. This is the well-known *incompleteness* of ZFC, witnessed here by the *independence* of CH from the ZFC axioms. Set theorists also say that CH is *undecidable in* ZFC, which, for the purposes of this chapter, is just the same as saying the CH is *independent from* ZFC. Undecidability (and hence independence) are formal notions; set theorists can prove that CH (and many other propositions) are undecidable in ZFC.

The undecidability of a proposition is relative to the axiom system considered. A proposition may be undecidable relative to ZFC, but decidable in some extension of ZFC by an axiom A, ZFC + A for short. For example, if A is the axiom stating that every set is constructible in Gödel's sense, then ZFC + A proves CH. That is, given

ZFC + A, then the Continuum Hypothesis is true. Since it is common set-theoretic practice to accept ZFC, the question is whether A should be accepted (e.g. Jech 2006, p. 157). Indeed, accepting another axiom may imply that the Continuum Hypothesis is false instead. Which axiom(s) to add to ZFC is a deeply mathematical question and highly skilled and respected mathematicians engage with it in the debate that is the focus of this chapter (Baldwin 2018; Antos et al. 2015; Steel 2014; Väänänen 2014; Hamkins 2012; Magidor 2012; Woodin 2011c; Shelah 2003). But the question cannot be answered by the formal machinery of mathematics. Nothing could prove, in any ordinary mathematical sense of the word, that A should be accepted or rejected. What is needed is thus a convincing rational argument for (or against) A, and this argument cannot be a formal proof.

Gödel put to the debate the idea that for some undecidable propositions there can be no convincing rational argument for any extension of ZFC that would decide whether the proposition is true or false. He wrote:

> The consistency of proposition A (that every set is constructible) is also of interest in its own right, especially because it is very plausible that with A one is dealing with an absolutely undecidable proposition, on which set theory bifurcates into two different systems, similar to Euclidean and non-Euclidean geometry. (Gödel 1995, p. 155)

Gödel suggests in this quotation that there may be *absolutely undecidable* set-theoretic propositions, that is, propositions which are (a) undecidable in ZFC and (b) there can be no convincing rational argument for or against the proposition. As is well-known, he later turned against this view. For the later Gödel, there are no absolutely undecidable propositions (e.g., Gödel 1947).

Notice that, unlike undecidability, absolute undecidability cannot be formally proven. It is a conviction one might have about a proposition rather than a formal result.

The question whether there are absolutely undecidable propositions lies at the heart of the set-theoretic pluralism debate. Set theorists regularly deal with the limits of their currently accepted formal machinery. They face the situation that there are set-theoretic questions which they consider relevant, such as whether the CH is true or false, which are provably unanswerable from the widely agreed upon intellectual resources of their practice. With Rouse (1996), it is unsurprising that they interpret this situation in partially different ways. For some set theorists, the limits of the available formal machinery indicate that certain set-theoretic questions may not have a "true or false"-answer, they are absolutely undecidable. These are the pluralists. Other set theorists argue that the inability of the currently accepted methods to answer "true" or "false" to certain set-theoretic questions is a paucity of the methods, not a challenge to the idea that mathematical statements are either true or false. These are the monists. The later Gödel is a monist in this sense.

The *set-theoretic pluralism debate* is the debate about how set theory as a discipline should deal with the issues of undecidability, the limits of its formal

machinery, and the questions about which axioms should be elevated to the same status the ZFC axioms currently have. This debate is also known as the *foundational debate* (Maddy 2011, 1997). Noteworthy contributions to this debate by set theorists include (Baldwin 2018; Antos et al. 2015; Shelah 2014; Steel 2014; Väänänen 2014; Magidor 2012; Shelah 2003) and the works of Woodin and Hamkins I will introduce in the following two sections. This debate has been the focus of much of the contemporary philosophizing about set theory, including Barton (2019), Centrone et al. (2019), Tanswell (2018), Rittberg (2015), Kennedy (2014), Kennedy and Kossak (2011), Maddy (2011), Koellner (2006).

It will be helpful for the purposes of this chapter to distinguish two forms of set-theoretic pluralism. *Epistemic pluralists* hold that our epistemic capacities may be too limited to know the truth-value of some set-theoretic propositions. Those propositions which are beyond our epistemic grasp are absolutely undecidable. Von Neumann's was an early suggestion of a form of epistemic pluralism in set theory:

> Das abzählbar Unendliche als solches ist unanfechtbar . . . Aber seine Grenzen scheinen sehr verschwommen und ohne anschaulich-inhaltliche Bedeutung zu sein. (von Neumann 1925) (The denumerable infinite as such is beyond dispute. . . But its boundaries seem to be quite blurred and to lack intuitive, substantive meaning; translation from (van Heijenoort 1967).)

Metaphysical pluralists hold that there are multiple universes of sets and propositions are true only relative to the universe considered. Those propositions which are true in one universe but false in another are regarded as absolutely undecidable. Hamkins is a well-known metaphysical pluralist; I discuss his views in Sect. 3.

Accordingly, *epistemic monists* hold that our epistemic grasp is in principle unlimited. *Metaphysical monists* hold that there is only one universe of sets, which is often called the "true universe of sets." Woodin, whose position I discuss in the next section, favors forms of epistemic and metaphysical monism.

My focus in this chapter is on the metaphysical variants of the set-theoretic pluralism debate. I will from now on drop the specification "metaphysical," employing it only where the argument benefits from fine-grained specification.

At this point I have introduced the undecidability-situation contemporary set theorists are in and outlined some of the different proposals endorsed by set theorists to understand this situation. Crucially for what follows, there is an orthodox interpretation of the undecidability-situation in set theory. Maddy (1988, 1997, 2011) provides a successful career-length argument that the orthodox position is monism. Criticism of Maddy's work on this issue has focused on her claim that set theorists should still be monists today (Ternullo 2019; Rittberg 2016b; Antos et al. 2015; Friend 2014). That monism is the historically orthodox position in set theory is not under dispute.

Set theorists also present monism as the orthodoxy in their publications. This orthodoxy vibrates through, for example, Woodin's (2011c) label of a monist

position as "The Set Theorist's" response to a skeptic (see also next section); Väänänen's (2014, p. 180) remark that "the phrase 'multiple realities', as well as 'parallel universes', may sound immediately self-contradictory and ill defined. We try to make sense of it anyway"; Magidor's (2012, p. 3) interpretation of his claim that "the main motivation for set theory is still its foundational goal"; Steel's (2014, p. 154) view that we need a "stronger foundation for mathematics"; Hamkins' (2012, p. 416) claim that monism is "the commonly held philosophical position" in set theory (see Sect. 3).

The empirical question whether monism is the most commonly held position among set theorists I leave unanswered in this chapter. I note, however, that numerous young set theorists (PhD students and postdocs mostly) have told me in unstructured personal discussions that they do not endorse monism.

In what follows I refer to monism as the orthodoxy in the set-theoretic pluralism debate. By this I mean that monism is part of the cultural history of set theory and still shapes the practice today. Those who challenge the position need to argue against it while those who endorse it can and often do assume it without much ado. Indeed, monists tend to focus on the problem the undecidability results pose for monism and offer proposals to engage with or even overcome these problems. This will be visible in Woodin's argument for a monist resolution of the undecidability issue in the next section. Pluralists tend to stress new insights or evidence that challenges the monist orthodoxy. I discuss Hamkins' argument for pluralism as an example in Sect. 3.

It bears stressing here that I explicitly do not claim that monism is the most commonly held position among set theorists. In fact, I believe this claim to be false. This belief I base on various personal communications I had with set theorists over the years. A satisfying answer to the empirical question which view is most common among set theorists is still missing (but see Džamonja and Kant (2019) for preliminary empirical work).

The situation the set theorists currently share is, in the parts relevant for this chapter, shaped by mathematical results about undecidability and an orthodox philosophical view about what it is the discipline studies. This situation is interpreted differently by the contemporary set theorists. Woodin, whose argument I present in the next section, interprets the undecidability results as problems to be overcome and provides an argument how this could be achieved. Hamkins, whose position I present in Sect. 3, interprets the undecidability results as ingredients to a counterargument to the monist orthodoxy. Relevant for this chapter is that neither argues purely philosophically: both mathematicians rely on mathematical results as evidence for their philosophical positions. Furthermore, both mathematicians explore mathematical programs whose relevance is grounded in their philosophical views on the undecidability issue. The argument of this chapter is that there is a reciprocal relationship between the mathematical activity and the metaphysical thoughts these two set theorists put forward. To make this visible, I discuss first Woodin's and then Hamkins' projects in detail in the following sections.

2 Woodin's Ultimate-L Argument

W. Hugh Woodin is a highly respected member of the set-theoretic community. He was affiliated with the informal but very influential group of set theorists known as the Cabal, which was particularly active in the 1980s and 1990s in Berkeley, California. Woodin has produced numerous influential publications and is widely regarded as one of the leading set theorists of today; a certain type of large cardinal (see below) carries his name: Woodin cardinals. Around the year 2000 he developed an argument for the so-called star-axiom as a contribution to the developing set-theoretic pluralism debate. In the first decade of this millennium, he gave up this argument and now proposes an argument for an Ultimate-L. In the historically minded (Rittberg 2015), I present both arguments in detail. Here I focus exclusively on the Ultimate-L argument.

Relevant contributions to Woodin's Ultimate-L argument include Woodin (2010a, b, c, 2011a, b, c), Koellner and Woodin (2009, 2010), Koellner (2011).

Woodin is a monist. As mentioned in Sect. 1, this is the orthodox position in set theory and Woodin provides little argument for it in his writings. Instead, he argues indirectly by challenging two non-monist positions.

Woodin (2011c) challenges skeptic, a literary figure who views the universe of sets as "a complete fiction" (p. 449). That is, Woodin is more concerned with an ontological rather than a metaphysical question here. Skeptic's positions do not reflect those of any set theorist contributing to the set-theoretic pluralism debate. Woodin argues that skeptic cannot account for the consistency results his Ultimate-L project will produce. Woodin sees this as evidence for the existence of "the transfinite universe" (2011c, p. 470), the definitive article indicating that there is only one such universe. Skeptic does not reappear in Woodin's later writings.

Woodin (2011d) challenges what he calls the generic-multiverse conception of truth. According to Woodin, the undecidability of CH in ZFC is "often cited" (2011d, p. 13) as evidence that the CH has no answer. Woodin formalizes this idea (2011d, p. 17), the result of which is the generic-multiverse conception of truth. Woodin does not discuss what is lost in this formalization, nor whether his formalized version successfully tracks the philosophical positions put forth in the set-theoretic pluralism debate. Woodin summarizes his argument:

> The essence of the argument against the generic-multiverse position is that assuming the Ω Conjecture is true (and that there is a proper class of Woodin cardinals) then this position is simply a brand of formalism that denies the transfinite by a reducing truth about the universe of sets to truth about a simple fragment such as the integers or, in this case, the collection of all subsets of the least Woodin cardinal. (Woodin 2011d, p. 17)

For details on the Ω conjecture and the point about Woodin cardinals, see Rittberg (2015).

Foreshadowing the next section, I note that Hamkins appropriated the multiverse terminology from an earlier version of (Woodin 2011d). Hamkins also points out in (2012, p. 445), which appeared later than (Woodin 2011d), that his multiverse view is not the generic-multiverse view.

Monism is part of what drives Woodin's work on Ultimate-L. From his monist point of view, the undecidability situation in set theory is a problem. He sees set theory haunted by a "specter of undecidability" (Woodin 2010a, p. 17) and speaks of "The ZFC Dilemma: Many of the fundamental questions of Set Theory are formally unsolvable from the ZFC axioms" (Woodin 2011c, p. 449). As becomes clear from the surrounding text, Woodin intends "dilemma" in the quotation in its informal meaning as a difficult situation or problem, rather than as a choice between undesirable alternatives. His Ultimate-L argument is a proposed (and still developing) answer to this problem. That is, Woodin holds a philosophical view, monism, for which certain formal mathematical results are problematic. The aim to resolve these problems gives relevance to Woodin's work on Ultimate-L. His metaphysical views provide the relevance for his mathematical research project here.

To understand Woodin's proposal to tame (and possibly banish, see below) the specter of undecidability, it is important to have a grip on the interplay of undecidability and large cardinal axioms. The key point is that set theorists can order undecidability along the large cardinal hierarchy. I offer a brief exploration of this issue here.

Large cardinal axioms are strong axioms of infinity: they state that certain infinite sets exist whose existence cannot be proven from ZFC. Kanamori (2009, p. 472) lists 28 different types of large cardinal axioms, but since there is no clear-cut definition of what makes a set-existence statement a large cardinal axiom, this list is subject to interpretation as well as further developments in set theory. Note however that there is remarkably little discussion in day-to-day set-theoretic practice about whether a certain axiom counts as a large cardinal axiom or not. For a large cardinal axiom LCA, ZFC + LCA proves the consistency of ZFC, which is famously unprovable from ZFC alone (due to Gödel's incompleteness theorems) - I am assuming here and throughout that ZFC is consistent. The mathematical vernacular is that ZFC + LCA has a higher *consistency strength* than ZFC. Indeed, for any two large cardinal axioms A and B, it holds that the consistency strength of ZFC + A is at least as high as ZFC + B or vice versa. That is, the large cardinal axioms form a linear order along consistency-strength: the so-called *large cardinal hierarchy.* Notice how this formulation suppresses the distinction between a large cardinal, which is a set, and a large cardinal axiom, which states that a certain class of cardinals exists. This is common practice in set theory as it is generally clear from the context what is meant and I follow this convention in this chapter; I here order the large cardinals by consistency-strength. Another common ordering relies on the fact that large cardinal properties carry over. For example, every Woodin cardinal is also a measurable cardinal but not vice versa. In this sense Woodin cardinals are stronger than measurable cardinals. This ordering aligns with the ordering by consistency-strength in all but one case: real-valued measurable cardinals are neither stronger nor weaker than weakly compact cardinals in this ordering. This makes the property-ordering a nonlinear ordering. Real-valued measurable cardinals are stronger in consistency-strength than weakly compact cardinals. Because the large cardinal axioms have been identified in vastly different contexts and their initial formulations share no obvious connections, the fact that they can be linearly ordered along their

consistency strength is a remarkable emergence of regularity and a source of wonder and marvel among many experts; see (Kanamori 2009) for technical results and historical remarks.

The large cardinal axiom stating that a Reinhardt cardinal exists is inconsistent with ZFC. Reinhardt cardinals thus form an upper bound of the large cardinal hierarchy in the context of ZFC, even though Reinhardt cardinals and even stronger large cardinals have been meaningfully studied in the context of ZF, that is, ZFC without the axiom of choice, since no inconsistency of these theories is thus far known; for example, Cutolo (2019). For large cardinals weaker than Reinhardt cardinals, no such inconsistencies with ZFC are (yet?) known. "The investigation of [such] large cardinal hypotheses is indeed a mainstream of modern set theory" (Kanamori 2009, p. XI).

To order undecidability in ZFC, set theorists can calibrate the *degree of undecidability* of (some) set-theoretic propositions using the large cardinal hierarchy. This is sometimes referred to as *degree of unsolvability*, for example, Woodin (2011c). Say P is a set-theoretic proposition which is undecidable in ZFC. Set theorists can then search for the smallest large cardinal axiom LCA such that ZFC + LCA decides (i.e., proves or disproves) P; such searches have sometimes taken years of international research efforts and led to new developments along the way. For example, work on the axiom of projective determinacy took years and contributed to the development of Woodin cardinals. See Martin and Steel (1987) for helpful remarks. A satisfying history of the development of Woodin cardinals is still missing. The degree of undecidability of P is then LCA, which can then be compared to the degree of undecidability of other set-theoretic propositions. This way the large cardinal hierarchy can be used to order undecidability along the large cardinal hierarchy. However, not every set-theoretic proposition is decidable in some large cardinal extension of ZFC. For example, CH is undecidable in all known large cardinal extensions of ZFC.

I now return to Woodin's Ultimate-L argument. As we saw he seeks to tame the specter of undecidability. A prominent way to do so is *Gödel's program*:

> Decide mathematically interesting questions independent of ZFC in well-justified extensions of ZFC. (Steel 2014, p. 153)

Large cardinal extensions of ZFC can answer mathematically interesting questions which are undecidable in ZFC; e.g., they can determine whether projective sets are Lebesgue-measurable. The *inner model program* seeks justifications for such large cardinal extensions. As I will show momentarily, this program has certain limitations (versions of Scott's theorem). Woodin's argument for an Ultimate-L rests on a proposal to overcome these limitations.

The basic idea of the inner model program is to build "mathematically accessible" (i.e., L-like, see below) set-theoretic models in which large cardinal axioms hold.

The completeness theorem implies that ZFC is consistent if and only if there is a model of set theory. In a model of set theory, every set-theoretic proposition is either true or false, that is, undecidability is an issue for theories, not for the models of these

theories. Indeed, a proposition P is undecidable in ZFC if there are two ZFC models, one in which P is true and one in which P is false.

Set-theoretic models tend to be resistant to mathematical exploration. Set theorists have means to construct models in which certain targeted propositions hold; they can construct a model for non-CH, for example. However, it is generally not possible to discern whether some other proposition is true or false in the constructed model. This is different for Gödel's constructible universe L. Assuming the consistency of ZF, Gödel constructed the model L, which, unlike most other models, can be studied with relative ease. Today, the so-called *structure-theory* of L, that is, what is true and false in L, is well understood.

An important result for the inner model program is Scott's (1961) theorem. It states that there are no measurable cardinals in L. Measurable cardinals are in the middle of the large cardinal hierarchy and Scott's result implies that there are also no large cardinals stronger than measurable cardinals in L-here and throughout I employ this visual language to generate "a feel" for the strength of the LCA in question; I rely on Kanamori's (2009, p. 472) chart of large cardinals for my positional judgements. The inner model program seeks to construct models that (a) can accommodate such stronger large cardinals and (b) whose structure theory is similarly accessible to that of L, that is, are L-like.

The inner model program has successfully constructed L-like models for some large cardinal axioms. The general heuristic is to target a large cardinal axiom LCA and assume that ZFC + LCA is consistent. The completeness theorem then ensures that there is a model for ZFC + LCA, but it may not be L-like. The inner model program seeks to construct an L-like model for ZFC + LCA. If this can be done, then this is viewed by the proponents of Gödel's program as justification for the belief that ZFC + LCA is consistent. It may seem as if there is a circularity here: ZFC + LCA is assumed to be consistent in an argument for the consistency of ZFC + LCA. Proponents of the argument hold that the construction of the L-like model adds sufficient evidence to justify the conclusion.

In all these L-like models the inner model program produces, a version of Scott's theorem holds, that is, only the targeted large cardinal axiom LCA but no stronger axiom is part of the structure theory of the constructed model. This has led to a stepwise approach in which the set theorists produce such L-like models for stronger and stronger large cardinal axioms. The currently strongest results are the Mitchel-Steel extender models, which can accommodate Woodin cardinals (Mitchel and Steel 1994); Woodin cardinals are two thirds up in the large cardinal hierarchy.

Woodin shows that "and this is the surprise" (Woodin 2011c, p. 465), if the set theorists would succeed in building an inner model of the desired kind for a supercompact cardinal, then this model could accommodate essentially all large cardinals, even those that are stronger than supercompact cardinals. That is, in such a model, versions of Scott's theorem would not hold. This structure would be L-like, that is, its structure theory would be accessible to mathematical exploration, and it would accommodate essentially all large cardinal axioms. Woodin calls it an *Ultimate-L*.

Supercompact cardinals are about three quarters up in the large cardinal hierarchy. The inner model program has thus far not succeeded in constructing an inner model which could accommodate supercompact cardinals. Thus far no Ultimate-L has been constructed. However, Woodin has produced a family of axioms, each expressing the idea that the universe of sets, called V, is an Ultimate-L, expressed as V=Ult-L.

The V = Ult-L axioms are powerful. ZFC + V=Ult-L + LCAs, whereby V=Ult-L stands for one of the Ultimate-L axioms and LCAs stands for a schema expressing that large cardinals exist, decide all contemporarily known formal questions about sets:

> There is no known candidate for a sentence which is independent from [a version of V = Ult-L] and which is not a consequence of some large cardinal axiom. (Woodin 2011c, p. 470)

To note how strong this is, recall that some propositions, such as the Continuum Hypothesis CH, are undecidable even in large cardinal extensions. In a V=Ult-L model however, CH and all other known propositions like it become decided. In this sense, the V=Ult-L axioms tame the "specter of undecidability."

Two principal issues remain. First, set theorists have not yet succeeded in constructing an L-like model which can accommodate a supercompact cardinal (and it is at least theoretically possible that this cannot be done for reasons as of yet unknown). This point requires further mathematical investigation and I will say no more on it in this chapter. Second, there are multiple versions of the V=Ult-L axioms which contradict each other; for example, CH is true under one version of the axiom but false under another. These multiple versions of the V=Ult-L axioms are the focus of what follows.

Woodin and his close colleague Peter Koellner have proposed the *convergence argument* to choose between the different versions of the V=Ult-L axioms (Koellner 2013). They propose to analyze the structure theories of two models under each of the V=Ult-L axioms. One of these models has already received much attention from set theorists and is well understood; the other model has not yet received much attention. If the structure theories of these two models converge under the given version of V=Ult-L, then this counts as evidence in favor of this version of the axiom; if the structure theories diverge, then this counts as evidence against the axiom.

The first model is L($\mathbb{R}$), a model which is constructed just like Gödel's constructible universe L but instead of starting with the empty set, one starts with the set of all real numbers $\mathbb{R}$. There is some monism lurking in the imperative to start from "all real numbers." For the monist there is only one "true" set of all real numbers. The pluralist Hamkins disagrees. For him each set-theoretic universe comes with its own set of real numbers; see next section. L($\mathbb{R}$) "has figured prominently in the investigation of strong hypotheses" (Kanamori 2009, p. 142) and set theorists have an intimate understanding of it. For example, it has featured in the analysis of Projective Determinacy; see (Rittberg 2015) for details. This exploration has delivered arguments (but no generally accepted proof) that the Axiom of Determinacy, AD, holds in L($\mathbb{R}$). The assumption that AD holds in L($\mathbb{R}$) features below.

The second model is a technical generalization of L(ℝ). Notice that L(ℝ) is set-theoretically the same as $L(V_{\omega+1})$. The second model is $L(V_{\lambda+1})$, whereby λ is a cardinal such that there is an elementary embedding from $L(V_{\lambda+1})$ into itself with critical point below λ. As Koellner remarks, this embedding condition is "the strongest large cardinal axiom that appears in the literature" (Koellner 2013). The model $L(V_{\lambda+1})$ has thus far not received much attention from the general set-theoretic community (but this may be changing with Woodin's work).

Assuming that AD holds in L(ℝ), and thus in $L(V_{\omega+1})$, set theorists can prove that ω^+, that is, the successor cardinal of ω, is a measurable cardinal in $L(V_{\omega+1})$: see (Rittberg 2015) for more technical details. Similarly set theorists can prove that λ^+ is a measurable cardinal in $L(V_{\lambda+1})$ under the embedding condition. Woodin and Koellner now propose to study these two models under the assumption of versions of the V=Ult-L axioms. If further similarities of the structure theories of these models are observed, then this counts in favor of the version of the axiom. Diverging structure theories count as evidence against the version of V=Ult-L under consideration.

A negative convergence result has already been obtained. Woodin could show that under the assumption of one version of V=Ult-L the structure theories diverge (Koellner 2013). According to the convergence argument, this is evidence against this version of V=Ult-L.

The search for similarity leads to, as Koellner terms them, a "list of definite questions" (Koellner 2013). These questions are hard mathematical questions that will require considerable effort to answer, but they are answerable: "independence is not an issue" (ibid.). The "specter of undecidability," which resurfaced in the wake of the choice-problem between the different versions of V = Ult-L, seems finally to be tamed. It is not yet banished, as the analysis of the convergence of structure theories is not yet completed, but the set theorists are now in possession of a method that could potentially banish the specter for good.

This completes my presentation of the Ultimate-L argument. To quickly recap, because Woodin is a monist, the undecidability situation set theorists are currently facing is a "dilemma" for him. Woodin embarks on a research project to overcome this problematic situation. He realizes that an L-like structure that can accommodate a supercompact cardinal can accommodate essentially all large cardinals. Furthermore, the axioms stating that the universe of sets is such an Ultimate-L resolve the undecidability issue. The choice problem between the different versions of V=Ult-L is met with the convergence argument. In short, Woodin's monist understanding of the undecidability situation is a driving force of an elaborate (and by now over 10 years long) research project. Metaphysical belief is a driving force for mathematical activity here.

It is worth remembering at this point that metaphysical thought is not prior to mathematical activity. Woodin has come to his monist point of view in part because of mathematical results; he stresses the astonishing linearity of the large cardinal hierarchy, for example. The claim of this chapter is that metaphysical views and mathematical research develop in tandem and may partially influence each other.

This partial influence of metaphysics and mathematical research takes concrete form in the proposal to employ the V=Ult-L axioms and the convergence argument to decide the set-theoretic pluralism issue. Recall that the convergence argument tests the different versions of the V=Ult-L axioms by picking one such version and, under the assumption that this version of the axiom holds, investigates the structure theories of two set-theoretic structures. This leads to a list of hard-to-answer-but-answerable questions; undecidability is not an issue.

> if the answers converge then one will have strong evidence for new axioms settling the undecided statements (and hence non-pluralism about the universe of sets); while if the answers oscillate, one will have evidence that these statements are "absolutely undecidable" and this will strengthen the case for pluralism. (Koellner 2013)

In (Rittberg 2016a) I called this *mathematical pull*. Mathematics and metaphysics are set up in such a way that by doing more mathematics, namely by answering those hard-to-answer-but-answerable questions, cases for metaphysical positions can be made. The hyperbolic enthusiast might say that we can calculate an answer for the set-theoretic pluralism issue. A more sober formulation would stress that this is Koellner's proposal, that it relies on the V=Ult-L axioms and hence also on the assumptions that are part of the Ultimate-L argument, among which are Woodin's monist inclinations. Relevant for this chapter is that this is an argument by mathematicians, made in the context of a debate had by mathematicians, on how a metaphysical debate might be influenced by mathematical results. My point is that (future) mathematical research is set up to pull metaphysical views here.

In the next section I introduce Hamkins' contribution to the set-theoretic pluralism debate. My aim is again to show the reciprocal relationship between metaphysical beliefs and formal mathematical activity.

3 Hamkins' Multiverse View

Joel David Hamkins is a prolific set theorist and active participant in the debates of the mathematical community; for example, he is the top-rated user by reputation-score on the mathematics community website mathoverflow.net (as of June 2020). He was Woodin's PhD student and has maintained amicable relations with his former supervisor which have resulted in several joint publications, including (Hamkins and Woodin 2005, 2017; Hamkins et al. 2008).

Hamkins' primary published contributions to the pluralism debate are his (2012, 2014, 2015) and he has given numerous presentations and contributed online content on the issue. The 2012 paper begins by acknowledging that set-theoretic monism, or the "universe view" as he calls it, "is the commonly held philosophical position" (Hamkins 2012, p. 416) among set theorists (but he gives no evidence to support this empirical claim). The (2012) paper is an argument against this view. In (2015) and (2014) Hamkins focuses on specific arguments made in (2012) and develops them further.

Hamkins' arguments rely on the experience set theorists have with various models of set theory. Set theorists can build set-theoretic models and study them. The inner model program introduced in the last section does this, for example. The general methodology to build a set-theoretic model is to start from a given model and build a new model from there. Forcing is a well-known method to build such new models, looking at submodels such as L or HOD is another. The structure theory of the new model, that is, those statements that are true in the new model, is usually different from the structure theory of the model one has started with. The prominent example comes from Gödel's proof that CH is true in L and Cohen's proof that CH is false in a forcing extension of L. Indeed, set theorists have become very good at manipulating their constructions of the new models to ensure that certain targeted propositions are true or false in the newly built model. This is how the independence phenomenon is studied: start with a model in which a proposition is true (or false) and build a model in which it is false (or true). As Hamkins (2012, p. 418) puts it, "set theorists build models to order."

The orthodox monist interpretation of these many set-theoretic models is roughly this. The true universe of sets is the model of set theory which contains all sets. Every model of set theory "believes" that it contains all sets, this is part of what it means to be a model of set theory. (It is common practice in set theory to ascribe agency to models; for example, Jech (2006). I follow this convention here, using inverted commas to highlight such use.) However, the monist maintains, there is only one true model, the so-called *universe of sets*. All other models are "mistaken" and "overlook" some sets. For example, Gödel's constructible universe "overlooks" all non-constructible sets. The monist may find relevance and joy in the study of these models, as happens in the inner model program, but they are seen as studies restricted to certain kinds of sets (those that are in the model), rather than all sets (that "live" in the universe of sets). In this sense there are many set-theoretic models that "live" inside the one true universe of sets, according to the monist.

Hamkins argues that because set theorists have such extensive experience with diverse set-theoretic models, the ideological stance that one of these models enjoys the privileged status of being the "true universe of sets," that is, monism, has to be given up. For him all set-theoretic models are universes in their own right. The set theorists' experiences with these models indicate "that there are diverse distinct concepts of set, each instantiated in a corresponding set-theoretic universe, which exhibit diverse set-theoretic truths" (Hamkins 2012, p. 416). The collection of all set-theoretic models/universes constitutes what Hamkins calls the "multiverse". Woodin (2011b), Steel (2014), and Väänänen (2014) have introduced different conceptions of a multiverse, which are more limited than Hamkins'.

To highlight, Hamkins argues that the intimate understanding of certain formal mathematical results, those extensive experiences with diverse set-theoretic models, puts pressure on the monist orthodoxy. The argument is that because of formal mathematical results a certain metaphysical position ought to be given up. Hamkins treats formal mathematical results as evidence and lets them shape his metaphysical views here.

For Hamkins, the set-theoretic knowledge about independence constitutes an answer to questions which are undecidable from the ZFC axioms, such as the CH:

> On the multiverse view the continuum hypothesis is a settled question, for the answer consists of the expansive, detailed knowledge set theorists have gained about the extent to which the CH holds and fails in the multiverse, about how to achieve it or its negation in combination with other diverse set-theoretic properties. Of course, there are and will always remain questions about whether one can achieve CH or its negation with this or that hypothesis, but the point is that the most important and essential facts about CH are deeply understood, and it is these facts that constitute the answer to the CH question. (Hamkins 2012, p. 3)

This reminds of Rouse's point about scientists interpreting the situation they are in differently. Where Woodin sees "specters" and "dilemmas," Hamkins sees answers to set-theoretic questions.

According to Woodin there is a paucity to the current set-theoretic machinery because it cannot fully capture the (one true) universe of sets. For Hamkins there is no paucity in these methods. On the contrary, these methods help to reveal that there are in fact multiple set-theoretic realities. Hamkins needs to argue that all those models the set-theoretic machinery can produce deserve to be considered as such set-theoretic realities- it is worth recalling here from the introduction that both Woodin and Hamkins agree that mathematical objects exist; Hamkins (2012, p. 417) calls it a "Platonic sense," but Hacking (2014) would use lower-case p). Hamkins (2012) starts by engaging with the counterargument that there are no forcing extensions of the true universe of sets V.

For a nontrivial forcing extension of a set-theoretic model M, one needs to find a special set G, called an M-generic *filter*, which is not in M. Therefore, to find a forcing extension of the true universe of sets V, one needs to find a V-generic filter G which is not in V. But V is supposed to be the collection of all sets, thus G is in V by assumption. Hamkins calls it a "catechism" (2012, p. 420) of monism that "there are no V-generic filters." He proposes to view the matter rather like moving from real to complex numbers. There is no solution to the square root of -1 in the real numbers, but once one moves to an extension of the real numbers, the complex numbers, such a square root can be found. There is no (non-trivial) V-generic filter G in the conception of sethood embodied by V, but such a G can be found in a new and broader conception of sethood, that is, a new model. As Koellner (2013) points out, this new model is non-standard because it is either countable, or not two-valued, or neither transitive nor well-founded. Despite this, Hamkins even suggests at times that such bigger models are "better" (2012, p. 438), but his argument for pluralism does not hinge on this value-judgment.

Hamkins also draws an analogy between set theory and geometry. According to his presentation, non-Euclidean geometries were considered to be curiosities and mere reinterpretations of concepts within Euclidean geometry. With time came understanding and experience with non-Euclidean geometries and they became accepted as geometrically meaningful. "The situation with set theory is the same"

Hamkins (2012, p. 426) suggests. For a set theorist disagreeing with this analogy, see (Magidor 2012). For a philosophical assessment of the analogy see Maddy (2011).

We saw in the last section that Woodin's monism is a driving-force for his mathematical research project on Ultimate-L. Similarly, Hamkins' pluralism gives relevance to two mathematical research projects he is engaged in, the *modal logic of forcing* and *set-theoretic geology*. I now turn to a discussion of these projects.

Forcing enjoys a special status among the methods set theorists employ to build set-theoretic models because it has proven to be very fruitful (Džamonja and Kant 2019). Today forcing is widely applied in set-theoretic proofs. The method is even influential enough to warrant the philosophical *Forcing Project*, led by Carolin Antos, which explores the philosophical implications of forcing as a mathematical method. This is much aligned with the tenet of this paper that mathematical activity and philosophical thought intertwine and may inform each other. Via forcing one can build from one model of set theory, the ground model, another model of set theory, the forcing extension. This relates these two models: one is the extension of the other. Relational structures can be studied via modal logics (Blackburn et al. 2001), which gives rise to the investigation of the modal logic of forcing. The set-theoretic geology project inverses the extension-relation of forcing: instead of looking "outwards" to find forcing extensions one looks "inwards" to find ground models.

Modal logics make use of two logical symbols: diamond, often interpreted as "it is possible that. . .," and box, often interpreted as "it is necessary that. . ." Hamkins (2003) proposes to interpret these modal operators as "it is forceable that. . ." and "in every forcing extension. . ." respectively. For example, Cohen showed that there is a forcing extension of Gödel's constructible universe L in which non-CH holds and thus "diamond non-CH" is true in L.

Hamkins (2003) introduced this interpretation of the operators in connection with a certain maximality principle. In 2004 Hamkins and Löwe started their collaboration on the modal logic of forcing (Löwe, personal commutation 2020). They uploaded a preprint to arxive.org in 2005, which was later published as (Hamkins and Löwe 2008), neither of which mentions the term *multiverse*. As Löwe recalls it (personal communication 2020), he started thinking about the topic in multiverse terms only in 2009. A few years earlier Woodin (2004) had introduced what he called the *generic multiverse* (see last section) in a talk at a conference in Copenhagen. Hamkins' first mention of the term *multiverse* in print seems to be (Hamkins 2009), which refers to Woodin's (2004). Hamkins and Löwe (2013) is the first contribution to the modal logic of forcing project that explicitly endorses a multiverse perspective, and they refer to a draft version of (Woodin 2011d) as a source for the terminology.

As this timeline shows, Hamkins and Löwe did not have a fleshed out multiverse view when they began work on the modal logic of forcing project. My claim is that metaphysical positions and mathematical activity develop in tandem and may inform each other. At this point of this chapter, my interest is in making visible how pluralism gives relevance to the set-theoretic geology and the modal logic of forcing projects.

Some set-theoretic propositions can be switched on and off via forcing like a light-switch. The CH is an example. For every model in which CH holds there is a forcing extension in which CH fails, and for every model in which CH fails, there is a forcing extension in which CH holds.

Other set-theoretic propositions can be forced to be false, that is, there is a model in which the proposition holds which has a forcing extension in which the proposition is false, but once they are false they cannot be recovered by forcing. They are like buttons that, once pushed, cannot be unpushed. The axiom stating that every set is constructible in Gödel's sense, $V = L$, is such an axiom. It is true in L but false in, for example, Cohen's forcing extension of L, and there is (provably) no forcing extension of Cohen's forcing extension of L in which $V = L$ is true again.

Hamkins and Löwe (2008) show that every set-theoretic proposition is either such a switch, a button, or the negation of a button. This allows them to prove that forcing forms a modal logic called S4.2, the modal logic describing finite pre-lattices. More precisely, S4.2 is sound and complete with respect to the class of Kripke-frames that form a finite pre-lattice. For a detailed formal presentation of Hamkins' and Löwe's result see (Rittberg 2010). This result, and the methods employed to obtain it, is at the core of the modal logic of forcing project. Further relevant contributions to this project include (Inamdar and Löwe 2016; Block and Löwe 2015; Hamkins et al. 2015; Hamkins and Löwe 2013; Inamdar 2013; Friedman et al. 2018).

This completes my brief presentation of the modal logic of forcing project. I now turn to the set-theoretic geology project.

The set-theoretic geology project inverses the extension-relation of forcing. Where forcing starts with a ground model M and moves to a forcing extension M[G], in the set-theoretic geology project one starts with a model N and asks if there is a ground model M and a forcing such that N is the forcing extension M[G]. Metaphorically speaking, the set-theoretic geology project strips away the forcing-residue from a model (Hamkins 2012, pp. 446–447).

In general, the ground M is not unique, which gives rise to the notion of a *bedrock*, a minimal ground. Reitz (2006, 2007) showed that there are models that do not have a bedrock, which in turn gives rise to what Hamkins (2012, p. 444) calls "the principal new set-theoretic concept in [set-theoretic] geology": the *mantle* is the intersection of all grounds.

A "fundamental question" (Hamkins 2012, p. 445) of the set-theoretic geology project was whether the mantle is a model of ZFC. In what Hamkins calls a "breakthrough" (in numerous talks on set-theoretic geology, abstracts for which can be found on his http://jdh.hamkins.org/), Usuba (2017) proved the strong directed grounds hypothesis, which implies that the mantle is indeed a model of ZFC. Usuba's result answered many, but not all, central questions of the set-theoretic geology project and raised new ones. The mathematical introductory paper to the set-theoretic geology project is Fuchs et al. (2015). Further contributions to the project include Habič et al. (2019), Davis (2016), Reitz (2006, 2007).

Hamkins (2012, p. 440) calls the modal logic of forcing and the set-theoretic geology projects "multiverse-inspired mathematics" and points out that "the multiverse view in set theory leads one to consider how the various models of set theory

interact, or how a particular world sits in the multiverse." Words such as "inspire" and "leading to" suggest that the philosophical view was there before the mathematical projects were started. As pointed out above, this temporal order may not suitably reflect the actual happenings to count as good history-writing. To read Hamkins as a historian doing a questionable job is, I suggest, uncharitable. Instead I propose to read him as pointing out that his pluralist philosophy gives relevance to the two mathematical projects. My argument is that the Hamkins-style pluralist can recognize the value of the two projects in ways monists generally do not.

Johansen and Misfeldt (2016, p. 262) remark on what they call the *value of recognizability*:

> The value of recognizability prescribes that the mathematical community or a relevant subgroup thereof can recognize your work as interesting and important, and it is enacted by choosing problems and areas of work that other mathematicians find attractive and can identify themselves with... The subjection to recognizability contrasts with the idea of mathematics being rational and democratic, and also with the idea that mathematics is a landscape to be investigated. As a mathematician, it is not sufficient to obtain mathematical results; you must obtain results that others in the mathematical community are able to recognize [as interesting and important].

In the context of Hamkins-style pluralism, the two projects are interesting and important. The Hamkins-style pluralist holds that there are many different set-theoretic universes, all of which are members of the set-theoretic multiverse. This raises questions about the structure of this multiverse, that is, about how the set-theoretic universes are related to one another in the set-theoretic multiverse, how they "sit in the multiverse" to cite Hamkins (2012, p. 440) again. One of the many ways these universes relate to each other is via forcing; other examples include inner model constructions. The modal logic of forcing and the forcing geology project study this forcing-relation, which make them interesting and important explorations of the multiverse to the Hamkins-style pluralist.

The monist lacks a clear story why the two projects are interesting or important. Unlike Woodin's hard-to-answer-but-answerable questions, they do not aide monism in finding a resolution to the current undecidability-situation. Arguably, very few set-theoretic questions currently explored have this explicitly monist goal, so this may be too much to ask. But there is also no other obvious connection of the two projects with those set-theoretic problems which are currently regarded as relevant, important, and interesting by the set-theoretic community at large. The point is that the two projects explore new questions which can make it hard to recognize their relevance. The projects do not initially promise new consistency results, insights into structure theories of known set-theoretic models that would generally be considered as deep, or fresh insights into definability for example. Without a Hamkins-style view on mathematics, the two projects may seem neither interesting nor important, that is, they lack in value of recognizability.

As so often with accounts of a mathematical practice, lived realities do not adhere to neat accounts. In an interview with Kant, Džamonja remarked that "one can do the same mathematics independently of one's philosophical view" (Džamonja and Kant 2019, p. 19). She is right about this and some set theorists have contributed to the two projects without drawing on a Hamkins-style outlook on set-theory. For example, Gitman (a former PhD student of Hamkins) pointedly remarks that "It took four decades since the invention of forcing for set theorists and to ask (and answer) what post factum seems as one of the most natural questions regarding forcing. Is the ground model a definable class of its set-forcing extensions?" (Gitman and Johnstone 2014). This question was independently answered by Laver (2007) and Woodin (2004), and Hamkins (2012, p. 443) presents their results as a starting point for the set-theoretic geology project. That is, Laver and Woodin were interested in questions relating to the set-theoretic geology project. Another example is Golshani's and Mitchell's (2016) contribution to the modal logic of forcing project. They explore what they call the "Hamkins-Löwe property," showing that even though the project explores new questions, there are points of connection to issues that are of concern to a general set-theoretic audience.

I believe Džamonja's remark mentioned above to be correct. My point is not that proponents of one philosophical view cannot or do not participate in the mathematical projects driven by differing philosophical views. There is indeed much overlap, such as the already mentioned collaborative works of Woodin and Hamkins. But some projects, such as the search for an Ultimate-L or the modal logic of forcing and set-theoretic geology projects, are driven by philosophical interests in a way that discourages mathematicians of a differing philosophical bent to participate. My point is not that these mathematicians cannot participate in such projects, my point is that they rarely do. (I borrow this formulation from Netz (1999, 2017), who argues against the historiographical strategy of conceptual impossibility: "they did not do X because they did not have the concept." He argues: "it's not that they couldn't: it's just that, more often than not, they didn't" (Netz 2017, p. 38).)

I argued that in the context of Hamkins-style pluralism the two projects are recognizably relevant. And indeed, many of the contributors to the two projects are Hamkins' former students (e.g., Jonas Reitz, Victoria Gitman, George Leibman) or set theorists whose philosophical view aligns with Hamkins' (e.g., Benedikt Löwe)- Löwe's views on set-theoretic pluralism remain unpublished; I know of them mostly through personal communication and presentations he gave, for example, at the *FOMUS* conference in Bielefeld 2018. There are other motivations to pursue these projects as well, but they have led to few contributions to the projects. In this sense then is Hamkins-style pluralism a (but possibly not the only) driving force of the modal logic of forcing and the set-theoretic geology projects.

Hamkins-style pluralism gives relevance to the two projects. That is, a philosophical outlook on set theory shapes what counts as relevant, important, or interesting mathematical research. A philosophical outlook shapes mathematical activity here.

4 Concluding Remarks

The case studies presented in the last two sections have shown some of the reciprocal relationship between the mathematical activity of Woodin and Hamkins and the metaphysical positions the two set theorists endorse. I briefly summarize the relevant points here. I then turn to other mathematical practices, such as geometry, and provide some suggestions why it is reasonable to expect that mathematical activity and metaphysical positions shape each other in these practices as well. If I am right about this, then metaphysical views are an integral part of mathematical practices.

Woodin is a monist. Monism is the orthodox view in set theory, which relieves Woodin from providing detailed argument for the position. He focuses on resolving the problem monists see in the current undecidability situation in set theory. Monism thus motivates Woodin's (current) proposal to engage with this "specter of undecidability." He proposes the Ultimate-L argument. His monist inclinations shape his mathematical activity. The key insight is that if an L-like model which can accommodate a supercompact cardinal could be built, then this model would accommodate essentially all large cardinals. It would be an Ultimate-L. No such model has yet been built, but Woodin has offered many different set-theoretic propositions stating that the universe of sets is such an Ultimate-L: the V=Ult-L axioms. These axioms conflict with one another, and Koellner and Woodin have proposed the convergence argument to decide among them. This convergence argument feeds back into the set-theoretic pluralism debate. If there is convergence (of the relevant structure theories), then this would be evidence for monism. But should there be divergence, then pluralism would score a win. This convergence argument relies on hard-to-answer-but-answerable formal mathematical questions. That is, producing new mathematical results can, via the convergence argument, impact a metaphysical debate the set theorists are currently having. Mathematical activity is set up to pull the metaphysical debate.

Hamkins is a pluralist. He needs to argue against the reigning monist orthodoxy, which he does by pointing to the experience set theorists have with various set-theoretic models. These experiences come from mathematical results and mathematical activity involving these results. Thus, Hamkins' arguments draw on mathematical activity (and its results) to make the case for pluralism; mathematical activity shapes the metaphysical view. Hamkins-style pluralism gives relevance to the *modal logic of forcing* and the *set-theoretic geology* projects because it provides motivations to pursue these projects. Metaphysical thinking gives relevance to mathematical activity here.

As I have stressed repeatedly in this chapter, neither metaphysical belief nor mathematical activity is prior to the other. There is no chicken-or-egg problem here but a reciprocal relationship in which both sides may inform and shape each other. This challenges the distinction between intrinsic and extrinsic justifications for axioms often made in the debate about axiom choice in set theory (Maddy 2011; Parsons 2008; Gödel 1947). Set-theoretic axioms are intrinsically justified, according to the proponents of this distinction, if they "unfold the content of the concept of set" (Gödel 1947). Extrinsic justifications of an axiom are its success, its

fruitful and verifiable consequences. Intrinsic justifications thus play on metaphysical beliefs, extrinsic justifications on the results of mathematical activity. The distinction between intrinsic and extrinsic justification hence relies on a demarcation between metaphysical beliefs and mathematical activity. I have argued in this chapter that rather than a clear demarcation there is a reciprocal relationship between metaphysical beliefs and mathematical activity. They inform and shape each other, and this challenges the intrinsic/extrinsic distinction.

There are many mathematical practices and in this chapter I have thus far only engaged with one, set theory. In fact, engagement with set theory is limited to a detailed study of two set theorists. My hope is that the numerous pointers to other set theorists in this paper provide enough indication on where to look to suggest that my claims extend to the practice as a whole. I now want to briefly suggest that it is reasonable to expect that in mathematical practices other than set theory mathematical activity and metaphysical beliefs inform each other as well.

First off, we may simply take the mathematicians seriously who profess to the relevance of philosophical thought to their mathematical activity. Fields medallist Terrance Tao aptly titled his (2018) "There's more to mathematics than rigour and proof." Tao speaks of a post-rigorous stage in which mathematicians are guided in part by what he calls "solid intuition." Such intuitions, I suggest, are partially shaped by the mathematician's metaphysical views. For example, Hamkins' intuitions about set-hood are different from Woodin's. For Hamkins "every countable transitive model of set theory can be continued to a model of V = L" (2012, p. 439) by extending the class of ordinals. This makes no sense for Woodin-style monism: the class of ordinals (in the true universe of sets) cannot be extended (because there is nothing more than "all of them"). Timothy Gowers, another Fields medallist, speaks of two cultures in mathematics. "I mean the distinction between mathematicians who regard their central aim as being to solve problems, and those who are more concerned with building and understanding theories" (Gowers 2000). There are different philosophies at work here. Even more such philosophies are explored in terms of narrative in (Doxiadis and Mazur 2012); Alexander's contribution to the volume (pp. 20–64) is particularly insightful on the point I am making here.

Understood this broadly it is fairly obvious that philosophical thought is part of mathematical research. In this chapter I have made the narrower claim that metaphysical beliefs about mathematical objects of study, sets in this case, stand in a reciprocal relationship with mathematical activity. A similar claim has been put forward by Chang about scientific activity. He argues that "if we want to engage in a certain epistemic activity, then we must presume the truth of some particular metaphysical principles" (Chang 2008, p. 113). Given that a convincing argument to the effect that there is a relevant difference between science and mathematics is still lacking, I take the points of philosophers of science like Chang to support the expectation that metaphysical beliefs play a role in mathematical practices other than set theory as well.

A promising place to look for a reciprocal relationship between mathematical activity and metaphysical beliefs is geometry. Bos (2001, p. 3) reminds us that in what he calls Classical Greek geometry "geometrical figures were 'known' or

'given' if they could be *constructed* starting from elements that were considered given at the outset." Classically, a geometric object is considered 'constructed' if it can be obtained from some given points by ruler and compass constructions; see Manders' seminal (2008) for detailed discussion of subtleties brushed over here. But other possible ways to construct geometric objects where always known.

> Plato himself reproached Eudoxus and Archytas and Menaechmus for setting out to remove the problem of doubling the cube into the realm of instruments and mechanical devices, as if they were trying to find two mean proportionals not by the use of reason but in whatever way would work. In this way, he thought, the advantage of geometry was dissipated and destroyed, since it slipped back into the realm of sense-perception instead of soaring upward and laying hold of the eternal and immaterial images. (Plutarch 1961, pp. 121–123)

What is at stake here is the metaphysical question when a geometrical object counts as constructed or given. Bos' (2001) is an insightful analysis of some of what Plutarch called "mechanical devices" and their usage. It reveals that some mathematicians have considered geometrical objects as constructed when they can be obtained by means of these "mechanical devices," such as marked ruler or Neusis constructions. Such constructions motivated mathematical activity which was sufficiently different from ordinary ruler-and-compass geometry that the two were kept apart: mainstream geometry constructed via ruler and compass, other geometrical practices allowed other tools of construction. Some mathematicians felt bound to one way of doing geometry, others worked in multiple idioms (Bos 2001).

Geometry has also been done by paper folding (Friedman 2018). This off-loads permissible constructive moves to the physical realities of paper (Friedman and Rittberg 2019). For mathematician Sundara Row, this gives proofs by paper folding an undoubtable immediacy. This reflects in his writing. He relies heavily on pictures of folded paper and uses formal syntactic presentations sparingly. Other mathematicians, such as Piazzolla Beloch, forego actual foldings of paper altogether and rely entirely on formal syntactic representations of folds in their proofs. My suspicion is that Row's and Belloch's metaphysical views shape how they conduct their mathematical activities.

Perhaps geometry is as easy a case as set theory to make visible the reciprocal relationship between metaphysical thought and mathematical activity in a mathematical practice. Much harder, I believe, would it be to make visible how this reciprocal relationship plays out in some other mathematical practices, such as number theory (the work of Ed Nelson may be a starting point). The history of the conception of number as well as historical works on mathematical practices whose motivations are remarkably different to those of contemporary mathematics (Chemla 2012; Kümmerle 2019) may be good starting points. But such studies would still lack some indication of how metaphysical beliefs shape and are shaped by *contemporary* mathematical activity in number theory. I know of no good example. Perhaps there is large-scale agreement among contemporary number theorists on relevant matters. It could be that such a state of affairs simply hides instances where mathematical activity is shaped by metaphysical thought. Or it could be evidence

that research activity in number theory is indeed free of any metaphysical presuppositions. I find the latter highly unlikely.

Metaphysics is not influenced by mathematics, nor is mathematics influenced by metaphysics. There is not one thing influencing the other (as e.g., Hamkins has suggested during his CLMPST presentation in Prague 2019; slides available at http://jdh.hamkins.org/wp-content/uploads/Set-theoretic-mereology-as-a-foundation-of-mathematics-CLMPST-Prague-2019.pdf (accessed June 2020). Rather, metaphysical thinking and mathematical research activity are both ordinary aspects of the participation in a mathematical practice (or at least: set-theoretic practice) and partially shape each other. In this section, I have suggested that we might have expected this, that the message of this chapter is ultimately obvious. The interesting part is, I believe, in the *how* mathematical activity and metaphysical thought interact. In this chapter, I have explored how this happens in set-theoretic practice.

Acknowledgments I would like to thank Carolin Antos, Neil Barton, Deborah Kant, Juliette Kennedy, Daniel Kuby, Benedikt Löwe, and Deniz Sarikaya for helpful remarks on draft versions of this chapter. Research for this chapter has been funded by the Centre for Mathematical Cognition at Loughborough University and the Research Foundation - Flanders (FWO), project G056716N.

References

Antos C, Friedman SD, Honzik R, Ternullo C (2015) Multiverse conceptions in set theory. Synthese 192(8):2463–2488

Baldwin J (2018) Model theory and the philosophy of mathematical practice. Cambridge University Press, Cambridge

Barton N (2019) Forcing and the universe of sets: must we lose insight? J Philos Log. online first

Blackburn P, de Rijke M, Venema Y (2001) Modal logic. Cambridge University Press, Cambridge

Block AC, Löwe B (2015) Modal logic and multiverses. RIMS Kokyuroku 1949:5–23

Bos H (2001) Redefining geometrical exactness. Springer, New York

Centrone S, Kant D, Sarikaya D (2019) Reflections on the foundations of mathematics. Springer, Cham

Chang H (2008) Contingent transcendental arguments for metaphysical principles. In: Massimi M (ed) Kant and the philosophy of science today. Cambridge University Press, Cambridge, pp 113–133

Chemla K (2012) The History of Mathematical Proof in Ancient Traditions, Cambridge University Press, Cambridge

Cutolo R (2019) The cofinality of the least Berkeley cardinal and the extent of dependent choice. Math Log Q 65(1):121–126

Davis J (2016) Universal graphs at $\aleph_ {\omega_1+ 1} $ and set-theoretic geology. arXiv preprint:1605.08811

Doxiadis A, Mazur B (2012) Circles disturbed: the interplay of mathematics and narrative. Princeton University Press, Princeton

Džamonja M, Kant D (2019) Interview with a set theorist. In: Centrone S, Kant D, Sarikaya D (eds) Reflections on the foundations of mathematics. Springer, Cham, pp 3–26

Friedman SD, Fuchino S, Sakai H (2018) On the set-generic multiverse. In: The Hyperuniverse Project and Maximality (pp 109–124). Birkhüuser, Cham

Friedman M (2018) On the history of folding in mathematics: a mathematization of the margins. Birkhäuser, Basel

Friedman M, Rittberg CJ (2019) On the material reasoning of folding paper. Synthese. online first

Friend M (2014) Pluralism in mathematics: a new position in philosophy of mathematics. Springer, Dordrecht
Fuchs G, Hamkins JD, Reitz J (2015) Set-theoretic geology. Annals of Pure and Applied Logic 166 (4):464–501
Gitman V, Johnstone TA (2014) On ground model definability. In: Infinity, computability, and metamathematics: Festschrift in honour of the 60th birthdays of Peter Koepke and Philip Welch. College Publications, Milton Keynes
Gödel K (1947) What is Cantor's continuum problem? Am Math Monthly 54:515–525. reprinted in: Benacerraf P, Putnam H (1983) Philosophy of mathematics, 2nd edn. Press Syndicate of the University of Cambridge, Cambridge
Gödel K (1995) Kurt Gödel, Collected Works Vol III, Feferman, S., Dawson Jr, J. W., Goldfarb, W., Parsons, C., Solovay, R. N. (eds), Oxford University Press
Golshani M, Mitchell W (2016) On a question of Hamkins and Löwe on the modal logic of collapse forcing. Pre-print. Available at: http://math.ipm.ac.ir/~golshani/Papers/On%20a%20question%20of%20Hamkins%20and%20Lowe%20on%20the%20modal%20logic%20of%20collapse%20forcing.pdf
Gowers WT (2000) The two cultures of mathematics. In: Arnold V et al (eds) Mathematics: frontiers and perspectives. American Mathematical Society, Providence, pp 65–78
Habič ME, Hamkins JD, Klausner LD, Verner J, Williams KJ (2019) Set-theoretic blockchains. Arch Math Logic 58(7–8):965–997
Hacking I (2014) Why is there philosophy of mathematics at all? Cambridge University Press, Cambridge
Hamkins JD, (2003) A simple maximality principle. J Symb Logic 68(2):527–550
Hamkins JD (2009) Second order set theory. In: Ramanujam R, Sarukkai S (eds) Logic and its applications, Third Indian Conference, ICLA, Chennai, India, Proceedings. Springer, Berlin
Hamkins JD (2012) The set-theoretic multiverse. Rev Symb Logic 5(3):416–449
Hamkins JD (2014) A multiverse perspective on the axiom of constructiblity. In: Chong C, Feng Q, Slaman TA, Woodin WH (eds) Infinity and truth. Lecture note series, Institute for Mathematical Sciences, National University of Singapore. World Scientific, Singapore
Hamkins JD (2015) Is the dream solution of the continuum hypothesis attainable? Notre Dame J Form Log 56(1):135–145
Hamkins JD, Lowe B (2008) The modal logic of forcing. Transactions of the American Mathematical Society 360:1793–1817
Hamkins JD, Löwe B (2013) Moving up and down in the generic multiverse. In: Lodaya K (ed) Logic and its applications. Springer, Berlin, pp 139–147
Hamkins JD, Woodin WH (2005) The necessary maximality principle for c.c.c.-forcing is equiconsistent with a weakly compact cardinal. Math Log Q 51(5):493–498
Hamkins JD, Woodin WH (2017) The universal finite set. ArXiv e-prints (currently under review)
Hamkins JD, Reitz J, Woodin WH (2008) The ground axiom is consistent with V ≠HOD. Proc Am Math Soc 136(8):2943–2949
Hamkins JD, Leibman G, Löwe B (2015) Structural connections between a forcing class and its modal logic. Israel J Math 207:617–651
Inamdar TC (2013) On the Modal Logics of some Set Theoretic Constructions, MSc Thesis
Inamdar T, Löwe B (2016) The modal logic of inner models. J Symb Log 81(1):225–236
Jech T (2006) Set theory. 4th printing. Springer, New York
Johansen MW, Misfeldt M (2016) An empirical approach to the mathematical values of problem choice and argumentation. In: Larvor B (ed) Mathematical cultures, the London meetings. Birkhäuser, Basel, pp 259–270
Kanamori A (2009) The higher infinite, 2nd edn. Springer, Berlin
Kennedy J (2014) Interpreting Gödel. Cambridge University Press, Cambridge
Kennedy J, Kossak R (2011) Set theory, arithmetic, and the foundations of mathematics. Cambridge University Press, Cambridge
Koellner P (2006) On the question of absolute undecidability. Philos Math 14:153–188

Koellner P (2011) Independence and large cardinals. The Stanford Encyclopaedia of Philosophy. Available at https://plato.stanford.edu/entries/independence-large-cardinals/
Koellner P (2013) Hamkins on the multiverse. Available at: http://logic.harvard.edu/EFI_Hamkins_Comments.pdf
Koellner P, Woodin H (2009) Incompatible Ω-complete theories. J Symb Log 74(4):1155–1170
Koellner P, Woodin H (2010) Large cardinals from determinacy. In: Foreman M, Kanamori A (eds) Handbook of set theory. Springer, New York, pp 1951–2120
Kümmerle H (2019) Die Institutionalisierung der Mathematik als Wissenschaft im Japan der Meiji- und Taishō-Zeit. PhD thesis
Laver R (2007) Certain very large cardinals are not created in small forcing extensions. Ann Pure Appl Log 149(1–3):1–6
Maddy P (1988) Believing the axioms I & II. J Symb Log 53(2):481–511 & 53(3):736–764
Maddy P (1997) Naturalism in mathematics. Claredon Press, Oxford
Maddy P (2011) Defending the axioms: on the philosophical foundations of set theory. Oxford University Press, Oxford
Magidor M (2012) Some set theories are more equal. Pre-print. Available at: http://logic.harvard.edu/EFI_Magidor.pdf
Mander K (2008) The Euclidean diagram. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford
Martin D, Steel J, (1987) A proof of projective determinacy. J Am Math Soc 2(1):71–125
Mitchel W, Steel J (1994) Fine structure and iteration trees. Lecture notes in logic, vol. 3, Springer
Netz R (1999) Shaping of deduction in Greek mathematics. Cambridge University Press, Cambridge
Netz R (2017) Mathematical concepts? A view from ancient history. In: de Freitas E, Sinclair N, Coles A (eds) What is a mathematical concept? Cambridge University Press, Cambridge
Parsons C (2008) Mathematical thought and its objects. Cambridge University Press, Cambridge
Plutarch (1961) Moralia, vol 9. Loeb classical library (trans: Minar Jr EL, Sandbach FH, Helmbold WC). Heinemann, London
Reitz J (2006) The Ground Axiom. PhD Thesis
Reitz J (2007) The Ground Axiom. Journal of Symbolic Logic 72(4):1299–1317
Rittberg CJ (2010) The modal logic of forcing. Diploma thesis. Available at: https://www.uni-muenster.de/imperia/md/content/logik/c.rittberg_the_modal_logic_of_forcing.pdf
Rittberg CJ (2015) How Woodin changed his mind: new thoughts on the continuum hypothesis. Archive for History of Exact Sciences 69(2):125–151
Rittberg CJ (2016a) Mathematical pull. In: Larvor B (ed) Mathematical cultures, the London meetings. Birkhäuser, Basel, pp 287–302
Rittberg CJ (2016b) Methods, goals and metaphysics in contemporary set theory. PhD thesis. Available at: https://uhra.herts.ac.uk/bitstream/handle/2299/17218/13036561%20Rittberg%20Colin%20-%20Final%20submission.pdf?sequence=1
Rittberg CJ, Van Kerkhove B (2019) Studying mathematical practices: the dilemma of case studies. ZDM Math Educ 51(5):857–868
Rouse J (1996) Engaging science: how to understand its practices philosophically. Cornell University Press, Ithaca
Scott D (1961) Measurable cardinals and constructible sets. Reprinted in: Sacks G (ed) Mathematical logic in the 20th century. Singapore University Press. Singapore. pp 411–426
Shelah S (2003) Logical dreams. Bull Am Math Soc 40(2):203–228
Shelah S (2014) Reflecting on logical dreams. In: Kennedy J (ed) Interpreting Gödel. Cambridge University Press, Cambridge, pp 242–255
Steel J (2014) Gödel's program. In: Kennedy J (ed) Interpreting Gödel. Cambridge University Press, Cambridge, pp 153–179
Tanswell FS (2018) Conceptual engineering for mathematical concepts. Inquiry 61(8):881–913
Tao T (2018) There's more to mathematics than rigour and proofs. Available at: https://terrytao.wordpress.com/career-advice/theres-more-to-mathematics-than-rigour-and-proofs/. Accessed May 2020

Ternullo C (2019) Maddy on the multiverse. In: Centrone S, Kant D, Sarikaya D (eds) Reflections on the foundations of mathematics. Springer, Cham, pp 43–78
Usuba T (2017) The downward directed grounds hypothesis and very large cardinals. J Math Log 17(2):1750009
Väänänen J (2014) Multiverse set theory and absolutely undecidable propositions. In: Kennedy J (ed) Interpreting Gödel. Cambridge University Press, Cambridge, pp 180–208
Van Heijenoort J (1967) From Frege to Gödel: a source book in mathematical logic. Harvard University Press, Cambridge, MA
von Neumann J (1925) Eine Axiomatisierung der Mengenlehre. Journal für die reine und angewandte Mathematik 154:219–240
Woodin WH (2004) Recent development's on Cantor's continuum hypothesis. In: Proceedings of the continuum in philosophy and mathematics. Carlsberg Academy, Copenhagen
Woodin WH (2010a) Strong axioms of infinity and the search for V. In: Proceedings of the international congress of mathematicians, Hyderabad
Woodin WH (2010b) Suitable extender models. I. J Math Log 10:101–339
Woodin WH (2010c) Ultimate L, talk given at the University of Pennsylvania. Available at: http://philosophy.sas.upenn.edu/WSTPM/Woodin
Woodin WH (2011a) Suitable extender models II. J Math Log 11(2):115–436
Woodin WH (2011b) The realm of the infinite. In: Heller M, Woodin WH (eds) Infinity. New research frontiers. Cambridge University Press, Cambridge, pp 89–118
Woodin WH (2011c) The transfinite universe. In: Baaz M, Papadimitriou CH, Putnam HW, Scott DS, Harper CL Jr (eds) Kurt Gödel and the foundations of mathematics. Cambridge University Press, Cambridge, pp 449–474
Woodin WH (2011d) The continuum hypothesis, the generic multiverse of sets, and the Ω-conjecture. In: Kennedy J, Kossak R (eds) Set theory, arithmetic, and the foundations of mathematics. Cambridge University Press, Cambridge, pp 13–42

Definitions (and Concepts) in Mathematical Practice

V. J. W. Coumans

Contents

1 Introduction 136
2 The Relation Between Definitions and Concepts 137
3 Theme I: The Nature of Mathematical Definitions 138
4 Theme II: The Relation Between a Term and the Concept It Represents 141
5 Theme III: Concepts and Definitions as Communal Notions 144
5.1 Descriptive Accounts 145
5.2 Normative Accounts 148
5.3 Why Do Mathematicians Use Certain Definitions and Not Others? 149
6 Theme IV: Values Regarding Mathematical Definitions and Concepts 151
6.1 Naturalness 151
6.2 Fruitfulness 152
6.3 Explanatoriness 153
7 Conclusion 154
References 155

Abstract

Definitions are traditionally seen as abbreviations, as tools for notational convenience that do not increase inferential power. From a Philosophy of Mathematical Practice point of view, however, there is much more to definitions. For example, definitions can play a role in problem solving, definitions can contribute to understanding, sometimes equivalent definitions are appreciated differently, and so on. This chapter reviews the literature on definitions and (to a certain extent) concepts in mathematical practice. It is structured according to four themes through which definitions (and concepts) in mathematical practice have been studied. These themes concern (1) the nature of definitions, (2) whether and how concepts evolve, (3) definitions and concepts from a communal perspective, and (4) different values relating to definitions and concepts.

V. J. W. Coumans (✉)
Institute for Science in Society, Radboud University, Nijmegen, The Netherlands

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_94

Keywords

Definition · Concept · Mathematical practice · Philosophy

1 Introduction

Definitions are a central part of mathematics and mathematical practice. Traditionally, they have been studied in logic for their inferential role (see Belnap 1993). However, the role of definitions in mathematical practice goes beyond this. As Tappenden (2008a) remarks, "mathematicians often set finding the 'right'/'proper'/'correct'/'natural' definition as a research objective and success – finding the 'proper' definition – can be counted as a significant advance in knowledge" (Tappenden 2008a, p. 256).

Understanding proofs in practice is a prominent topic in the philosophy of mathematical practice. To give an imcomplete overview, there is literature on the various functions of proofs (De Villiers 1990; Baldwin 2018; Dawson 2006), on the difference between the ideal conceptions of proofs and proofs in practice (Andersen 2018; Larvor 2012; Hamami 2019), on virtues of proofs (Detlefsen 2008; Inglis and Aberdein 2015; Kleiner 1991), and on proofs and the peer review process (Andersen 2018; Frans and Kosolosky 2014; Geist et al. 2010). By contrast, definitions in practice have not been studied systematically as of yet. This does not mean that definitions have not been investigated at all. In fact, there is a reasonable literature on definitions in mathematical practice. Usually, however, the main point of focus in that literature is not to get a better understanding of definitions in practice, but to discuss a different topic in which definitions are a particular case. As a result, definitions in mathematical practice have been discussed, but not systematically, and it is therefore difficult to navigate through the existing literature that relates to definitions in mathematical practice.

This chapter is an effort to make the literature on definitions more accessible by reviewing and organizing the literature on definitions *in mathematical practice* (from now on referred to as simply "definitions"). Such an overview can help structure future research, in the hope of obtaining a better understanding of definitions in practice.

The chapter will also discuss the notion of "concept." This is because definitions and concepts are intricately related. One of the main functions of definitions is to specify or introduce a concept. In that sense, definitions are sometimes evaluated on the basis of the concept they specify. Vice versa, mathematicians try to understand certain concepts, and definitions can serve as means to do that. In that sense, some definitions are considered better than others for understanding a concept. In short, to meaningfully discuss definitions, one also needs to take concepts into account to a certain extent.

This chapter is outlined as follows. It starts with a brief discussion of the relation between definitions and concepts. This discussion serves two purposes. First, it helps to determine to what extent literature on concepts should be included in this review

(as the literature on concepts is vast and goes well beyond the scope of discussing definitions in practice). Second, the term "definition" is used by various authors, possibly meaning different things. Similarly, the usage of the term "concept" varies slightly across authors. By fixing a meaning for these terms, one can compare the usages.

After these initial remarks, the literature on definitions and concepts in mathematical practice will be discussed. The overview is structured according to four themes that can be identified in the literature. These themes concern (1) the nature of definitions, (2) whether and how concepts evolve, (3) definitions and concepts from a communal perspective, and (4) different values relating to definitions and concepts. N.B. Definitions are also studied in the context of mathematics education. However, a discussion of this is beyond the scope of this chapter. The interested reader is referred to (Ouvrier-Buffet 2013), which includes a summary of the mathematics education literature on definitions. This chapter is concluded with some remarks on the relation between the different themes.

2 The Relation Between Definitions and Concepts

Before looking into the literature regarding definitions, it is useful to first get an idea of the distinction between definitions and concepts. Unfortunately, the discussion regarding definitions and concepts *in general* is too diverse to find an uncontroversial description of both (Belnap 1993; Gupta 2019; Robinson 1954; Cappelen and Plunkett 2020; Margolis and Laurence 2019). For the purpose of this chapter, it suffices to take the notion of "definition" to refer to linguistic entities that specify the meaning or usage of a term. When it comes to the ontological status of concepts, there are also various competing theories (Margolis and Laurence 2019). The relation between definitions and concepts as used in this chapter is that *definitions are linguistic entities that give rise to concepts, mental or otherwise, and in doing so they can concur with or differ from pre-existing/intended concepts.*

As an illustration, consider the example of prime numbers, which is also considered in (Tappenden 2008a) for different purposes. For the natural numbers, one can define prime numbers through the following definitions:

- A natural number P is prime if and only if the only divisors of P are 1 and P.
- A natural number P is prime if and only if for every pair of natural numbers a, b we have that if P divides ab, then P divides a or P divides b.

These are different definitions which give rise to the same concept of "prime number."

Another example can be found in Lakatos' *Proofs and Refutations* (2015). In this well-known work, which will be discussed later in more detail, the notion of a polyhedron is investigated: What is a polyhedron? Several definitions are discussed, and they are often evaluated on the criterion of whether they concur with some intended concept of what a polyhedron is.

In the following, the terms "definition" and "concepts" are used in the way described above. However, some of the authors, whose work is discussed in this chapter, also use these terms, possibly meaning different things. What is meant with the term is usually clear from the context. However, when I want to stress the meaning from the above formulation, I will write *definition* and *concept*.

To illustrate the intricacy of the relation between the notions of *definition* and *concept*, and to show the use of distinguishing between *definition* and *concept*, one can look at the history of the topological notion of compactness (Raman-Sundström 2015). Nowadays, compactness usually refers to open-cover compactness (X is compact if each open cover U has a finite subcover). For metric spaces, the notion of open-cover compactness is equivalent to so-called sequential compactness (a space X is sequentially compact if every sequence xn has a subsequence that converges in X). For topological spaces in general, this is not the case. This means that in some contexts, different *definitions* can lead to the same *concept*, whereas their generalizations can give rise to different *concepts*.

Furthermore, Raman-Sundström (2015) describes definitions of compactness that are analogous to the definition of sequential compactness. In other words, these versions are equivalent to open-cover compactness but, in some way, do justice to the formulation of sequential compactness in terms of convergence and sequences. One of these generalizations is in terms of nets. Whereas sequences are usually indexed by the natural numbers (0, 1, 2, 3, etc.), nets allow elements to be indexed by a far broader class of sets. By using a notion of convergence that relies on the notion of net, one can transform the formulation of sequential compactness in such a way that the result is equivalent to open-cover compactness. In conclusion, it can occur that different *concepts* have analogous *definitions* (in this case, the definition of sequential compactness is analogous to the definition of compactness in terms of nets). For more detail on the history of compactness, the reader is referred to (Raman-Sundström 2015).

3 Theme I: The Nature of Mathematical Definitions

As emphasized in the above relation between definitions and concepts, a definition is a linguistic expression, but not every linguistic expression is a definition. Therefore, the question is: What do we expect definitions to do? What differentiates definitions from linguistic expressions in general? In this section, three points of view on this issue are discussed: the traditional, the heuristic, and the work of Ouvrier-Buffet (2013).

The traditional view and the heuristic view are discussed by Cellucci (2018). He describes the traditional view as a combination of five statements (that he traces back to the works of Frege):

> (1) *A definition merely stipulates the meaning of a term* [. . .] (2) *A definition is an abbreviation* [. . .] (3) *A definition is always correct* [. . .] (4) *A definition can always be*

> *eliminated* [. . .] (5) *A definition says nothing about the existence of the thing defined* [. . .] (Cellucci 2018, p. 608).

Claim (3) is intended to mean that definitions are not right or wrong, as they are not assertions to begin with. Claims (2) and (4) are also discussed in the context of definitions in general (as opposed to mathematical definitions). They are referred to as the criterion of eliminability and conservativeness, respectively (Belnap 1993). (Note that, although item (4) uses the word "eliminated," it corresponds to the criterion of conservativeness and not to the criterion of eliminability.) The idea behind eliminability is that a definition can be formulated in terms that are already known. The criterion of conservativeness says that no new conclusions can be derived by introducing a certain definition. More precisely, let L be a language and T a theory. Then a definition consists of expanding the language, resulting in L', and adding sentences to the theory, resulting in T'. The criterion of elimination is then given by: with respect to theory T', every L'-sentence should be equivalent to an L-sentence. Conservativeness can be formulated as: For every L-sentence S, if S follows from T', then it should also follow from T.

Cellucci discusses the traditional view of mathematical definitions in order to refute it and to defend the heuristic conception of definitions. The latter fits within the heuristic conception of mathematics, which, although not a mainstream view of mathematics, has gained some traction (also see Goethe and Friend 2010). In this view, mathematical problem solving is not a matter of deductively proving statements from axioms, but a process in which problems are solved by the analytic method. This method consists of presenting plausible hypotheses that solve the problem. These hypotheses, in turn, are problems that need to be solved as well. This process of providing hypotheses for open problems can continue indefinitely. In that regard, Cellucci argues that definitions are hypotheses too. He formulates this as: "[a] definition is the hypothesis that there exists something which satisfies the condition stated in the definiens" (Cellucci 2018, p. 618).

Cellucci claims that the heuristic view of definitions is preferable to the traditional one because it can describe mathematics better. What he means with mathematics is "the 'real' mathematics of the 'real' mathematicians" (Cellucci 2018, p. 617). He contrasts this with formal conceptions of mathematics in which there is no room for the logic of discovery. According to Cellucci, the heuristic conception of definitions can explain various phenomena that are not explained by the traditional view of definitions, namely that: "(a) Two extensionally equivalent definitions of the same concept may have different heuristic values; (b) Definitions in mathematics are not starting points but arrival points in the solution to problems; (c) There is no circularity of definitions and theorems; (d) Many definitions are proof-generated; (e) Definitions in mathematics can be justified" (Cellucci 2018, p. 620).

Although Cellucci portrays both points of view as competitors, one might argue that they are complementary. The traditional view focuses on the inferential role of definitions and the context of justification, whereas the heuristic view zooms in on the preformal mathematics and the context of discovery. In that sense, both accommodate an investigation of mathematical practice, albeit of different aspects thereof.

Which view is more appropriate differs for each situation and depends on contextual factors.

The third view on the nature of definitions is offered by Ouvrier-Buffet (2013). She aims to understand the act of defining in order to develop didactical situations for mathematics students. To this end, she develops three conceptions of definitions: the Aristotelian, the Popperian, and the Lakatosian. In the following, only sketches of these conceptions are presented. The interested reader is referred to (Ouvrier-Buffet 2013) for a detailed discussion and references to her earlier work.

In the Aristotelian conception, definitions serve as tools for classification. One starts such a classification by looking at the set of objects and for specific differences within this set. Such an analysis should lead to an "aesthetic definition." Some principles for defining are logical, like that there should be no vicious circles, that previous terms should be defined, etc. Other principles in this conception include that redundancies and metaphors should be avoided and that the concepts, which are defined, should actually exist.

Whereas the Aristotelian conception considers definitions as tools for classification, the Popperian conception considers them useful for choosing between competing theories. In that sense, one can try to reduce the number of postulates of a theory and generate counterexamples. The principles that guide the defining process include, for instance, the resistance to refute theories and criteria for when one theory is better than another.

The Lakatosian conception combines aspects from the previous ones. According to this conception, definitions are employed to solve intramathematical problems (i.e., finding the domain of validity for a given theorem) and for classification. Some of the actions that guide the defining process are "generating examples and counterexamples" and consequently dealing with these counterexamples by the monster-barring and exception-barring (see Sect. 4).

These three conceptions are part of a model of the process of defining. This model also consists of different moments in the definition process ("in-action," "zero," "formalized," and "axiomatic") and an account of the type of problems that definitions can help solve. This model is used to create strategies for developing educational contexts for defining.

Finally, Ouvrier-Buffet (2015), on the basis of interviews with mathematicians, differentiates different types of definitions. For instance, she distinguishes between definitions that are temporarily used to shorten a talk from those that "remain and will belong in the public domain" (Ouvrier-Buffet 2015, p. 2216). Another distinction is made between working definitions that enable the mathematician to start investigating the object and the formal definition that one might arrive at later.

In conclusion, we saw various descriptions of what definitions are and what they (ought to) do. What all these views have in common, however, is that they see definitions as a device that links a term to a concept. This connection between a term and its concept is the subject of the following section.

4 Theme II: The Relation Between a Term and the Concept It Represents

Definitions are a way to assign to a term a concept/meaning. (Technically, concepts and meanings are different entities, but for the purpose of this chapter, I group them as one.) In this section, the relation between terms and their associated concepts is discussed. First, two points of view explicated by Schlimm (2012), regarding the status of this relation, are discussed: Is it fixed or is it subject to change? After establishing that, from a particular point of view, this relation can change over time, we turn to the work of Lakatos (2015) and Kitcher (1983), who discuss two ways in which this relation changes. Finally, the work of Tanswell (2018) is discussed. He not only demonstrates that this relation is not as definite as is sometimes assumed, but also suggests a way to steer conceptual change.

Although the title of Schlimm's *Mathematical Concepts and Investigative Practice* (2012) speaks of concepts, the article does not concern *concepts* (cf. Sect. 2), but the relationship between terms and what they refer to. Schlimm describes this in terms of whether "an object *a* falls under a concept *P*" (Schlimm 2012, p. 128). He then discusses two points of view regarding this relation: the Fregean view and the Lakatosian view. (Schlimm stresses that these perspectives are not per se those held by Frege and Lakatos, respectively, but that these are nonetheless inspired by their work.)

The Fregean view sees concepts as *definite* and *fixed*. That a concept is definite means that for every object, it is ontologically determined whether it falls under that concept or not. This does not suggest, however, that mathematicians themselves can immediately determine whether an object falls under that concept or not. That mathematical concepts are fixed means that if a term refers to a specific concept at one time, it does so always.

Mathematical concepts should be fixed, because otherwise logical deduction is not possible (Schlimm 2012). Arguments like "All *A*'s are *B*'s, this *X* is an *A*, therefore *X* is a *B*" might suffer if concept *A* changes between the first and second premise.

According to the Lakatosion view, by contrast, concepts are fluid and inexact: They may change, and sometimes it is not clear whether a certain object falls under the concept or not. One of the main motivations for this view is its ability to capture heuristic aspects of mathematics as well and not just the deductive parts.

Instead of arguing for either the Fregean or the Lakatosian view, Schlimm suggests that both of these descriptions are very useful in making sense of mathematical developments. He argues that "[i]n the history of mathematics many concepts were introduced informally and went through considerable changes, redefinitions, splittings, etc. However, other concepts were introduced by formal, determinate definitions from the start" (Schlimm 2012, p. 134). Accordingly, the Fregean and Lakatosian notion can complement each other.

One example that Schlimm discusses concerns the term "number." With the introduction of new number systems, "number" refers to different concepts over time. From a Lakatosian view, this is unproblematic. However, the Fregean view

would have us write different terms for these different systems. For instance, the natural numbers can refer to the integers greater than or equal to 0, or to the integers greater than 0. These should be indicated by $\text{Numbers}_{\geq 0}$ and $\text{Numbers}_{>0}$, respectively. Schlimm concludes that both views should be used to highlight different parts of mathematical practice, and therefore, he opts for a pluralistic approach.

Although Schlimm does not do this, one might relate these two views to the points of view regarding definitions listed in Sect. 3. Although the Fregean/Lakatosian views concern the relation between a term and a concept, and the traditional/heuristic views concern the nature of definitions, one can argue that the Fregean view corresponds to the traditional view, and the Lakatosian view to the heuristic view. Both the Fregean and the traditional view highlight the inferential role of definitions and concepts, whereas the Lakatosian and the heuristic view explicitly accommodate for the interference of mathematicians in the course of defining/fixing a concept for a term.

Furthermore, analogous to the case of traditional vs. heuristic, one can make similar remarks about the accuracy of the Fregean and the Lakatosian view: It is a matter of what these views try to describe. If we are looking at an idealized version of mathematics, the Fregean version might be more accurate, whereas the Lakatosian view might be more accurate for preformal mathematics or the human side of mathematics. Which view is more appropriate is dependent on the particulars of a situation.

If we do indeed want to approach conceptual change from the preformal perspective, we cannot ignore Lakatos' seminal *Proofs and Refutations* (2015). As is well-known, *Proofs and Refutations* constitutes a rational reconstruction of the research on Euler's polyhedron formula: $V - E + F = 2$ (i.e., for polyhedra, the number of vertices (V) − the number of edges (E) + the number of faces (F) = 2). The format of the book is a hypothetical discussion between a teacher and its pupils and is connected with historical events via its footnotes. The book contains a wealth of insights regarding mathematical practice. Due to space restrictions, this chapter cannot contain a review of the entire book. Nonetheless, what follows is *Proofs and Refutations*' main idea regarding conceptual change. For in-depth discussions of Lakatos' Philosophy of Mathematics, see (Koetsier 1991) and (Larvor 1998).

Proofs and Refutation's discussion starts right after the conjecture $V-E+F=2$ has been formulated, with the teacher presenting a "proof" of the theorem. However, his sense of a proof is not the standard conception of something that establishes the truth of the theorem beyond a doubt, but a "thought experiment – or 'quasi-experiment'- which suggests a decomposition of the original conjecture into sub-conjectures or lemmas" (Lakatos et al. 2015, p. 10). The students express doubt regarding this "proof" and the theorem and suggest several counterexamples, either to the theorem itself (*global* counterexamples) or to one of the lemmas in the proof (*local* counterexamples). It turns out that there is no consensus with respect to what a polyhedron actually is. The students then try to pin down what polyhedra are by presenting various definitions. It turns out that there is a dialectic between the proof and the concept of a polyhedron. What is understood by polyhedron leads to reformulations of the theorem (and its "proof"), and the steps in the proof suggest

properties that polyhedra should have. In short, improving a "proof" is a process that alters the theorem, the "proof," and the concepts involved. A concept that is the result of such a process is referred to as a *proof-generated* concept.

To get an idea of how this change occurs, one can look at the phenomena of "monster-barring" and "exception-barring." During the hypothetical discussion, some students come up with counterexamples to the polyhedra formula, i.e., a polyhedron for which either the proof fails or the theorem does not hold. One possible reaction to this counterexample is to label it a "monster," i.e., not real polyhedron. By doing this, the concept of polyhedron is revised and the validity of the proof and the theorem preserved. This is referred to as "monster-barring." A more nuanced approach is *exception-barring*, where the counterexample is not yet dismissed as a monster, but is investigated to determine the domain of validity for the theorem.

Another account of conceptual change is presented by Kitcher (1983). This description is given in a general discussion of how mathematical practices change (see Sect. 5). Kitcher argues that concepts (in particular, expressions and sentences) get meaning through so-called *initiating events*. In other words, an initiating event is the way in which a *token* of a *type term* gets appointed a referent. Due to this type/token distinction, the same concept (type term) can have multiple initiating events and the set of these events is called the *reference potential*. Ideally, this reference potential is *homogeneous*, i.e., all of its initiating events lead to the same entity.

Sometimes, however, the reference potential is *heterogeneous* and this can lead to refined expressions by adapting the referential potential so to make it homogeneous again. One of Kitcher's examples concerns Cantor's introduction of the transfinite ordinal ω. Cantor's contemporaries protested against Cantor's move from the natural numbers, a sequence that continues indefinitely, to a completion of that sequence in the form of ω. Kitcher suggests that there is a heterogeneous referential potential at work. Cantor describes ω as a number. But what do we consider numbers to be? One might for instance refer to the complex numbers and say that numbers constitute expressions of the form $a + bi$. In that sense, ω has trouble fitting the bill. However, when one considers numbers as entities on which one can perform certain arithmetical operations, then ordinal arithmetic shows that transfinite ordinals are numbers too. Therefore, this episode *stretched* the concept of number.

If one assumes that mathematical concepts are subject to change, then one can investigate how to deal with these changes. This is done by Tanswell (2018). After establishing that mathematical concepts can be inexact, he suggests that *conceptual engineering* can play a valuable role in making concepts more exact. Conceptual engineering is a branch of philosophy that looks at how concepts acquire meaning, how they are used, and how one can steer the usage of certain terms (see Cappelen and Plunkett 2020). A specific mathematical issue for which conceptual engineering can be useful concerns the concept of sets. Some argue that there exists a single set-theoretic universe, whereas others suggest that there exist many. Tanswell argues that "the best hope for a resolution is through the active deployment of the tools of conceptual engineering" (Tanswell 2018, p. 904).

In conclusion, according to some, the relation between terms and their related concepts is fixed and according to others this relation is subject to change. The next section zooms out and looks at concepts and definitions from the perspective of a mathematical community. A single mathematician might investigate a particular concept or use a definition, but what makes that other mathematicians are also interested in that concept or use that definition as well?

5 Theme III: Concepts and Definitions as Communal Notions

Mathematics is often framed as an objective and individual activity. The philosophy of mathematical practice has shown two things: There are many (inter) subjective factors in mathematics, and it is for a large part also a communal activity. In this section, we look at definitions and concepts from a *communal* perspective.

As in many other activities, there is some sense of fashion in mathematics (Wilder 1953; Thurston 1994). Therefore, one question that presents itself is as follows: What makes concepts or definitions fashionable? Corfield (2003) formulates this accurately through a hypothetical case:

> [i]f I define a *snook* to be a set with three binary, one tertiary and a couple of quaternary operations, satisfying this, that and the other equation, I may be able to demonstrate with unobjectionable logic that all finite snooks possess a certain property, and then proceed to develop snook theory right up to noetherian centralizing snook extensions. But, unless I am extraordinarily fortunate and find powerful links to other areas of mathematics, mathematicians will not think my work worth a jot. By contrast, my articles may well be in demand if I contribute to the understanding of Hopf algebras, perhaps via noetherian centralizing Hopf algebra extensions. (Corfield 2003, p. 11)

As the line "mathematicians will not think my work worth a jot" suggests, Corfield talks as if there is some general consensus to what is considered valuable or not. Whether such a consensus should exist is a topic of debate. Nonetheless, the discussion in the literature is framed as though there is such a consensus. Assuming this is indeed the case, the two questions that are discussed in this section are as follows:

1. Why do mathematicians study the concepts that they study and not others?
2. Given a concept with a variety of equivalent definitions, why would/do mathematicians prefer one definition over another?

This section starts by investigating the first question. The literature regarding this question can be split into descriptive and normative accounts. Descriptive accounts describe the processes by which concepts are introduced, by which they are (possibly) accepted, and how they "survive." Normative accounts have the implicit assumption that mathematical concepts are accepted when they are considered to be valuable. These normative accounts therefore consist of general criteria of

valuable mathematics. Finally, this section concludes with a review of the literature on the second question.

5.1 Descriptive Accounts

In discussing how mathematics develops, metaphors are often used. Hacking (2014), for instance, discusses two metaphors that explain how mathematics grows and which concepts are studied: the butterfly model and the Latin model. According to the butterfly model, mathematics develops teleologically, like a caterpillar becoming a cocoon and then turning into a butterfly. In that sense, mathematical developments are predetermined, although failures can hinder its development. By contrast, the cornerstone of the Latin model is that mathematics is not predetermined and that it evolves through contingencies. The comparison is made with how Latin evolved into various other languages.

A more concrete approach can be found in Wilder (1953), which specifically addresses the introduction and lifespan of concepts. Wilder argues that when it comes to the introduction of concepts, the communal or social aspects of mathematics are just as important as the individual aspects, if not more. He makes the case that mathematical concepts are not created out of nothing, but that they are the result of synthesizing established mathematics. "A concept doesn't just pop up full grown "like Venus from the waves," although it may seem to, to the individual mathematician who does the conceiving. Usually its elements are lying in [. . .] the *mathematical culture stream*" (Wilder 1953, p. 426).

Wilder's goal is to "inquire, above the individual level, the manner in which mathematical concepts originate, and to study those factors that encourage their formation and influence their growth" (Wilder 1953, p. 425). To this end, Wilder discusses several case studies of the introduction and evolution of concepts. From these, he identifies factors that can lead to new concepts. First of all, mathematical concepts arise from interaction with the "non-mathematical environment." He points to primitive mathematics as solutions to practical problems, in farming or taxes, but notes that contemporary mathematics is also influenced by practical, e.g., economical, problems.

Second, Wilder argues that while mathematics is partly individual (the final step in the development of a concept is by the individual), mathematics can only flourish by cooperation. In this sentence, cooperation should be read both in the literal and in a broader sense. The literal one referring to when small groups of mathematicians collaborate, the broader one indicating the efforts of previous mathematicians in generating the needed mathematical tools (e.g., "a Newton can carry on only from the level which the mathematics of his time has reached" (Wilder 1953, p. 438)). Therefore, the greatest factor for mathematical growth is what Wilder calls *conceptual contacts*.

Wilder differentiates conceptual contacts on the level of the individual from the level of the group. The first one represents meetings of mathematicians, like colloquia or international collaboration. The second refers to "diffusion of

concepts," like the fusion of two mathematical subfields (he discusses, for instance, the combination of algebra and geometry into algebraic geometry). The crucial difference is that on the group level, these developments depend on a far larger number of mathematicians.

Wilder then suggests two possible ways in which conceptual contacts can lead to new concepts. First, concepts can arise from the observation of similar structures and patterns in different branches of mathematics. He describes this as a mathematical equivalent of the observation of patterns in nature in physics. The second example of concept formation is that tools from one branch of mathematics are imported into another branch.

One can also look at the axiomatic method as a tool for concept formation. Given that a certain structure is characterized by a set of axioms, one can change one of the axioms and get another structure. Wilder mentions, for instance, how non-Euclidian geometry is obtained by substituting a nonparallel axiom for the parallel axiom. The final example of concept formation that Wilder discusses is the "well known and much abused generalization" (Wilder 1953, p. 442). He remarks that this does not only work at the individual level, but also at the group level, namely via *concept splitting*. He argues that mathematicians sometimes gradually investigate certain objects from particular points of view, each leading to different generalized versions of that concept.

In addition to these various ways in which concepts are introduced, a problem-solving oriented model for concept introduction is discussed in (Carter 2013) and (Kjeldsen and Carter 2012). Using the example of Riemann surfaces, Carter (2013) sets out to study how mathematicians handle mathematical objects, and how they prove, but ends up suggesting a model for how mathematical objects are introduced. Note, however, that where Carter writes 'objects', we can replace this with *concepts* (cf. Sect. 2). This model is crystallized and tested on the introduction of convex bodies in (Kjeldsen and Carter 2012). Despite several differences in comparison to Riemann surfaces, the introduction of convex bodies is seen as evidence for the validity of the model. The model consists of three stages.

Stage 1: The new object is defined "*with reference to already accepted objects*" (Kjeldsen and Carter 2012, p. 360). In Carter's case study on Riemann surfaces, these surfaces are used to study Abelian functions.

Stage 2: Correspondence theorems between the already accepted objects and the new object are obtained, and the already accepted objects are studied with the use of the new object and these correspondence theorems.

Stage 3: The new object is defined without reference to the already accepted objects and studied for its own sake.

The essence of this model is a motif that can be found in several other sources as well. The idea is that in order to solve a particular problem, mathematicians might go in a certain direction and that that direction turns into an established part of mathematics itself. For instance, Corfield (2003) differentiates between concepts that are *fundamental*, *convenient,* or *pointless*. The first type corresponds to concepts

that are "intrinsically worth studying" (Corfield 2003, p. 228) and the second to concepts that are "valuable as a tool" (Corfield 2003, p. 228). According to Corfield, the status of mathematical entities can change from useful tools to intrinsically worth studying. In effect, this is the same message as the model by Kjeldsen and Carter.

A general account of mathematical practices, which can also accommodate the aforementioned model, is presented in Kitcher's *The Nature of Mathematical Knowledge* (1983). Even though the main goal of this book is to show that mathematical knowledge is not a priori, this book is also well known for the introduction of the notions of *mathematical practice* and *interpractice transitions*. A mathematical practice is a five-tuple (L, M, Q, R, S), "where L is the language of the practice, M the set of metamathematical views, Q the set of accepted questions, R the set of accepted reasonings, and S the set of accepted statements" (Kitcher 1983, pp. 163–164). An interpractice transition is a transition from a practice (L, M, Q, R, S) to a practice (L′, M′, Q′, R′, S′).

For our purposes, it is interesting that, with this theory, Kitcher describes three ways in which expressions get introduced. In the first, definitions are abbreviations of other expressions and therefore serve a pragmatic purpose. This concurs with the traditional account of mathematical definitions as described in (Cellucci 2018). The second type is where a particular notion does not have a clearly defined referent and the community of mathematicians looks for a suitable interpretation of that term. This concurs with the account described by Lakatos (2015). As an example, Kitcher discusses how the term "function" evolved. The third type is Kitcher's description of conceptual change in which a heterogeneous reference potential is transformed into a homogeneous one (as described in Sect. 4).

In general, Kitcher presents five patterns of how practices change, namely question-answering, question-generating, generalization, rigorization, and systematization. The motif that some mathematical concepts that were introduced as a means to study another concept become the subject of investigation themselves arises via the patterns of question-answering and question-generating. Concretely, mathematical concepts are sometimes introduced by question-answering. However, these mathematical concepts can lead to questions about those concepts, i.e., a case of question-generating.

These patterns of change are also reviewed by Schlimm (2012). He uses them for the construction of five patterns of *conceptual* change: "Clarifications of informal concepts, systematizations of concepts and results, investigations of sharp concepts (defined by axiom systems), and generalizations and abstractions" (Schlimm 2012, p. 138). He suggests that these can lead to new or modified mathematical concepts. However, Schlimm only mentions these patterns and does not investigate them any further.

Next to looking at how certain concepts can become topics of study, one can also inquire into what Wilder (1953) calls the "life span of a concept." He argues that usually, concepts become fashionable and are studied intensely until they are no longer fashionable. When that happens, the concept's life span has ended. However, concepts can become very old or even immortal when they are connected to other mathematical concepts. Unfortunately, Wilder cannot give examples of concepts that

have died out completely. He remarks that mathematics has a tendency to reorganize itself, making various distinct parts of mathematics coherent. In that way, concepts never really die out but survive by being absorbed by a general theory.

5.2 Normative Accounts

Several authors tacitly assume that valuable mathematics becomes part of the shared mathematical theories and mathematics with little to no value does not. Muntersbjorn (2003) discusses this relation explicitly. She argues that the history of mathematics is sometimes wrongfully portrayed as teleological: Earlier mathematics is developed to facilitate later mathematics. Nowadays, mathematicians might value historical mathematical theories because of the possibilities they provide for our contemporary mathematics. However, that need not be why the theory was appreciated originally. She writes:

> Mathematics is a goal-oriented activity. In general, mathematicians strive to solve problems as quickly, accurately, and generally as possible. But the immediate goals of early individuals cannot be identified solely with the particular accomplishments of later individuals. [. . .] An important measure of mathematical success is the extent to which a particular result enables subsequent results. Good mathematics is fruitful. Yet, mathematical results must have immediate value for their community of origin before they can become candidates for further elaboration, as results without immediate value do not get passed down to the next generation. (Muntersbjorn 2003, p. 163)

Therefore, Muntersbjorn claims that there is a connection between mathematical concepts being accepted and cultivated, on the one hand, and being deemed valuable, on the other. Regarding these values, Muntersbjorn suggests "that the value of a mathematical innovation depends on its provenance and potential. Good mathematics originates from the identification of tacit insights responsible for prior success and generates further success" (Muntersbjorn 2003, p. 175). In addition to provenance and potential, she also discusses the importance of psychological factors such as brevity and clarity.

The quality or value of concepts is also discussed by Carballo (2020). His main question is as follows: In virtue of, what is a concept, better or worse, epistemically? Carballo argues that several claims have been made that some concepts are better because they are more "natural," "sparse," "joint-carving," etc. He refers to these terms as *elite* properties. Carballo first puts forward the hypothesis that concepts are epistemically better if and because they are more elite than others. Then Carballo tries to argue against that hypothesis. He presents several arguments. For instance, good concepts figure in good explanations. However, good explanations do not necessarily need elite concepts. Therefore, eliteness is not necessary for good concepts. Whether Carballo makes a convincing case is beyond the scope of this chapter, but the fact is that when it comes to evaluating concepts (or definitions), mathematicians often refer to values like naturalness or fruitfulness. We will come back to this in Sect. 6.

One can also look at the success of mathematical developments in general. Corfield (2001, 2003), for instance, argues that mathematicians often have to present the value of their work. He presents five criteria for valuable mathematical developments:

1. when a development allows new calculations to be performed in an existing problem domain, possibly leading to the solution of old conjectures;
2. when a development forges a connection between already existing domains, allowing the transfer of results and techniques between them;
3. when a development provides a new way of organising results within existing domains, leading perhaps to a clarification or even a redrafting of domain boundaries;
4. when a development opens up the prospect of new, conceptually motivated domains;
5. when a development reasonably directly leads to successful applications outside of mathematics. (Corfield 2001, pp. 508–509)

Developments that score high on all criteria are considered to be uncontroversially valuable. When developments do not score high on all criteria, the value depends on a tacit weighing of the criteria.

As these criteria are formulated quite generally, they do not only apply to concepts, but also to definitions. Some definitions can be valuable because they provide "a new way of organising results within existing domains" (see point 3) or because they forge "a connection between already existing domains" (see point 2). However, there is also literature that addresses the value of definitions directly.

5.3 Why Do Mathematicians Use Certain Definitions and Not Others?

An approach that addresses the value of definitions can be found in (Werndl 2009). Her main question is: "in what ways are definitions in mathematics justified, and are these kinds of justification reasonable?" (Werndl 2009, p. 313). Werndl argues that it is important to justify definitions, since unjustified ones are less meaningful to "us." Note that Werndl's use of the word "us" suggests some form of consensus on what are meaningful definitions. Therefore, she concludes, justified definitions are the ones we should focus on.

The first type of justification she identifies is the Lakatosian *proof-generated definition*. Note, however, that in *Proofs and Refutations*, proof-generated *definitions* are seldom discussed; Lakatos mainly discusses proof-generated *concepts*. Nonetheless, Werndl suggests that "[Lakatos'] main idea is the notion of a *proof-generated definition*. Here his main example is definitions of polyhedron which are justified because they are needed to make the proof of Eulerian conjecture work" (Werndl 2009, p. 315). Extrapolating the meaning of proof-generated concepts to proof-generated definitions, one might argue that the latter are definitions that are the result of improving a proof. Werndl phrases this as "a definition which is needed in order to prove a specific conjecture regarded as valuable" (Werndl 2009, p. 315). Note that in suggesting that definitions can be justified by being proof-generated, the

term "definition" should be interpreted as *concept*. A part of proof-generated definitions is that they capture the right domain of validity for the theorem. This then involves not the *definitions* themselves, but the *concepts* they correspond to.

Next to the justification of definitions being proof-generated, Werndl identifies three other types of justifications of definitions:

- Natural-world justified definitions: definitions that "*capture a preformal idea regarded as valuable for describing or understanding the natural world*" (Werndl 2009, p. 320)
- Condition-justified definitions: "*a definition … justified by the fact that it is equivalent in an allegedly natural way to a previously specified condition which is regarded as mathematically valuable*" (Werndl 2009, p. 326)
- Redundancy-justified definitions: a definition that "*eliminates at least one redundant condition in an already accepted definition*" (Werndl 2009, p. 330)

In these cases, the term "definition" does refer to *definitions*. Natural-world justifications are about what *captures* a preformal idea, condition-justified definitions show that a certain expression captures a relevant condition as well, and redundancy-justified definitions are more economical than others.

Furthermore, one can examine how certain definitions developed over time. As mentioned, Raman-Sundström (2015) does this by looking at "the origin and development of open-cover and sequential compactness" and "how and why open-cover compactness came to be favoured" (Raman-Sundström 2015, p. 619). She shows how in general topology, open-cover and sequential compactness are not equivalent and that the open-cover version was preferred because it is "more general and applicable" (Raman-Sundström 2015, p. 629).

In the final comments, she remarks that "the story of how open-cover compactness came to be seen as the right one is a story of developing mathematics without always knowing where it was going, how important terms should be defined and how widely they might be applied" (Raman-Sundström 2015, p. 633). Even though the mathematicians involved might not always have known where the path would lead them, Raman-Sundström also shows that many choices were based on values. This becomes clear from the following excerpts:

- "[Frechet] preferred definitions that had an intuitive feel rather than analytic power" (Raman-Sundström 2015, p. 625)
- "Alexandroff claimed the accumulation point characterization was most important initially, due to the dominance of the Bolzano–Weierstrass property, but after some years it became clear that the open-cover property was more fruitful" (Raman-Sundström 2015, p. 625)
- "some mathematicians find nets more intuitively appealing and useful, while others prefer filters" (Raman-Sundström 2015, p. 631).

The next section discusses several values regarding definitions and concepts.

6 Theme IV: Values Regarding Mathematical Definitions and Concepts

As the above section shows, values play an important role when it comes to concepts and definitions. Not only is mathematics often evaluated on the basis of values (see Sect. 5.2), mathematical developments are also often steered by these values (see Sect. 5.3). Usually, these values are referred to in an imprecise and informal way. As a response, there are several efforts to make these notions more precise. In this section, philosophical accounts of three values are discussed: naturalness, fruitfulness, and explanatory value.

6.1 Naturalness

Mathematicians often describe constructions and concepts as "natural." Although there are methodological difficulties involved in studying an informal concept like "natural" (Mauro and Venturi 2015), several authors have reflected on this term. For instance, Corfield (2003) briefly differentiates between three types of naturalness. The first interpretation sees natural as the opposite of artificial. Natural concepts are those that we can encounter "in nature." This, obviously, does not refer to the same nature in which one encounters plants and birds, but to "reasonably well frequented regions of mathematics" (Corfield 2003, p. 224). In the second interpretation, naturalness is linked to inevitability. Natural concepts are those concepts that mathematicians will discover at some point. Formulated more strongly, natural concepts "will necessarily appear and that, by virtue of their nature, they rather than the user will determine their use" (Corfield 2003, p. 226). If mathematicians from various backgrounds discover the same concept, then this indicates that the concept might be a natural one. Lastly, Corfield discusses a narrow interpretation of naturalness, based on origins of category theory. In this interpretation, naturalness constitutes the absence of arbitrary choices. For instance, the proof that every vector space is isomorphic with its double dual is considered natural as it requires no choice of basis. Where Corfield discusses naturalness of concepts, constructions, proofs, and a variety of mathematical phenomena, Tappenden (2008a) zooms in on concepts and definitions. He describes naturalness in terms of "carving at the conceptual joints." Then he looks at two examples, the Legendre symbol and prime numbers. Regarding the first, he asks (1) whether this concept is natural and (2) how we can establish its naturalness. At first sight, the Legendre symbol seems to be an artificial concept due to the case distinction in the definition (The Legendre symbol is defined by: for an odd prime p, an integer a, $\left(\frac{a}{p}\right) = 1$, if a is a quadratic residue modulo p; $\left(\frac{a}{p}\right) = -1$, if a is a nonquadratic residue modulo p; and $\left(\frac{a}{p}\right) = 0$ if $a \equiv 0 \pmod{p}$.). However, it also appears to unify the theory of quadratic reciprocity. Tappenden remarks this and concludes that there are two possible perspectives on the symbol: Either it is a useful notion which helps to streamline the theory, in which case it is mathematically natural,

or it is only useful because it is a gerrymander in a particular context and not generally useful, and therefore, not natural. Tappenden argues for the first approach by showing that the Legendre symbol is useful in various different mathematical contexts and that it proves to be a key notion in various generalizations. He therefore concludes, regarding questions (1) and (2) that "[i]t is a mathematical question whether the Legendre symbol carves mathematical reality at the joints and the verdict is unequivocally yes" (Tappenden 2008a, p. 264).

Tappenden also considers the carving at the conceptual joints in the case of prime numbers. He provides two definitions:

1. A natural number P is prime if and only if the only divisors of P are *1* and *P*.
2. A natural number P is prime if and only if for every pair of natural numbers a, b we have that if P divides ab, then P divides a or P divides b.

Although these definitions are equivalent in the context of natural numbers, the definition that is taught to high school students is the first. In contexts like algebraic number theory, however, these definitions are no longer equivalent. In order to generalize the notion of primeness, mathematicians have to choose between these definitions. To make this decision, Tappenden argues, mathematicians look at the original theorems involving prime numbers to see which property was actually doing the heavy lifting so that these theorems can be generalized. It turns out that in the general context, the high school definition is referred to as *irreducibility* whereas the other is referred to as *primeness*. The second definition therefore carves at the conceptual joints of primeness.

6.2 Fruitfulness

Fruitfulness is seen by Tappenden (2012) as a bridge between the foundational works of Frege and the study of mathematical practice. The main connection is Frege's view that "logical reasoning needs 'fruitful concepts' to extend knowledge, and that identifying and studying the fine structure of such concepts can produce further discoveries" (2012, p. 205). Several authors have tried to make sense of Frege's usage of the term fruitful for definitions (Shieh 2008; Tappenden 1995, 2012; Horty 2007). The consensus seems to be that some definitions are merely collections of unordered characteristics, whereas others display a logical structure That, when analyzed, leads to new knowledge. In that sense, fruitful definitions are those that enable the extension of knowledge. Horty (2007), on the other hand, takes this notion and contrasts it with other parts of Frege's theory of concepts. He concludes that there is a friction between the conception of definitions as abbreviations (cf. Sect. 3) and definitions as being fruitful. If definitions should not increase the inferential power of a logical system, how can they lead to new knowledge? Horty then assesses Frege's work to explain this seeming contradiction, the details of which are beyond the scope of this text.

After discussing Dedekind's remarks on fruitfulness, which are similar to Frege's, Tappenden (2012) asks: "how can Frege/Dedekind's recognition of fruitfulness as a practical basis for choice of concepts and definitions in mathematics helpfully guide us?" (2012, p. 211). He hopes to show that by looking at fruitfulness, we can better understand the reliance of mathematicians on "beauty." More generally, Tappenden argues that many judgments like "elegance" or "simplicity" have a guiding role in science, even though it is not clear why these notions would lead to truth. Therefore, he concludes, we should look at those notions in detail. In Tappenden (2012), he makes a start by exploring a connection between beauty and fruitfulness. He argues that seeing the way one result might lead to other, can lead to an aesthetic reaction.

Fruitfulness is also discussed by Yap (2011). She first tries to grasp the notion by looking at remarks by Gauss and provides a mathematical interpretation of those remarks. In this interpretation, mathematical calculi are fruitful if they can be integrated in existing methods, if they help to split mathematical problems into subproblems, and if they help to capture the core of the issues we frequently encounter. Note, however, that a "calculus" is not the same as a "concept," in general. However, one of the cases that Yap discusses is the Legendre symbol which is also discussed in Tappenden (2008a, b). Therefore, Yap's article on the fruitfulness of calculi is also discussed here.

She then compares the Legendre symbol with Gauss' congruence notation, i.e., $a \equiv b \pmod{n}$. She argues that Gauss' congruence notation and not the Legendre symbol was useful for *proving* the quadratic reciprocity theorem. Yap presents three arguments for the usefulness of Gauss' congruence notation: It rather straightforwardly leads to partitioning the integers in equivalence classes, equivalence behaves much like equality, and this calculus enabled the induction strategy that is essential in Gauss' first proof of the theorem.

Yap argues that the Legendre symbol is not as fruitful as the congruence notation, because it is too specific. She concludes that the congruence notation has the right balance between generality and specificity. Nonetheless, the Legendre symbol has proved to be useful, but that usefulness is "tied to the truth of the reciprocity theorem itself, whereas the usefulness of congruences, being more general, is wider" (Yap 2011, p. 414).

6.3 Explanatoriness

It has been established that many mathematicians refer to proofs as being explanatory. Consequently, there have been many efforts to describe what mathematical explanation actually is (see D'Alessandro 2019). Another question regarding mathematical explanation is about which mathematical techniques are (or can be) explanatory or not. Lange (2009), for instance, argues that mathematical induction, in general, is not explanatory. Lehet (2019) disagrees and looks at induction in mathematical definitions to argue that it can be explanatory in some cases.

To do this, Lehet first expands the theory of mathematical explanation to cover explanatory definitions as well. "A mathematical definition is explanatory if it makes

the mathematical concept being defined more accessible—i.e., it explains some feature or property of the relevant concept" (Lehet 2019, p. 6). She gives two examples of explanatory definitions. The first concerns the derivative. She focuses on which definition might be best suitable for an introductory calculus course. She argues that a definition of the derivative, as a limit of secants and tangents, is more explanatory than one in which the algebraic manipulations are highlighted (e.g., $\frac{d}{dx}x^n = nx^{n-1}$). For those only interested in calculating the derivatives, the algebraic example is preferable. However, for a more profound understanding, the limit example is preferable, and as the limit definition makes the "deeper meaning accessible" (Lehet 2019, p. 8), it is explanatory.

The second example concerns CW-complexes. The details are beyond the scope of this text. The crux is that CW-complexes are larger spaces, constructed out of smaller spaces. By doing this, mathematicians can reduce complex properties of these larger spaces into a number of simpler properties of smaller spaces. In effect, these larger spaces become accessible by reducing them to a number of accessible smaller spaces.

Although Lehet (2019) was the first to explicitly discuss explanatory definitions, the aforementioned article by Raman-Sundström (2015) in a sense already shows how explanatoriness is an important factor for definitions. Given the vast literature on mathematical explanation in general, there are many opportunities for further investigating explanatory definitions.

7 Conclusion

Definitions (and concepts) in mathematical practice is not a standard topic of investigation in the philosophy of mathematical practice. Usually, they are discussed in light of others matters; for example, Kitcher (1983) discusses conceptual change in his description of mathematical practices. In the above, the literature on definitions (and concepts) is structured using five themes. Whereas definitions and concepts are intricately related, the first theme focuses on definitions: What is the nature/role of definitions? As one important function of definitions is specifying the meaning of a term, the relationship between terms and concepts constitutes theme two. From a broader perspective, one can look at how definitions and concepts are accepted and studied by mathematical communities, which is theme three. Within this theme, one can further differentiate between descriptive and normative accounts of how concepts enter the shared mathematical scope and literature on why certain definitions and not others are used. Theme four is about various values that are attributed to concepts and definitions. This is summarized in Fig. 1.

These themes are not disjoint. For instance, the work by Tappenden (2008a) shows why the mathematical community might choose to work with a particular definition (cf. theme III) as well as the role of the naturalness considerations in this decision (cf. theme IV). Moreover, some types of conceptual change (cf. theme II)

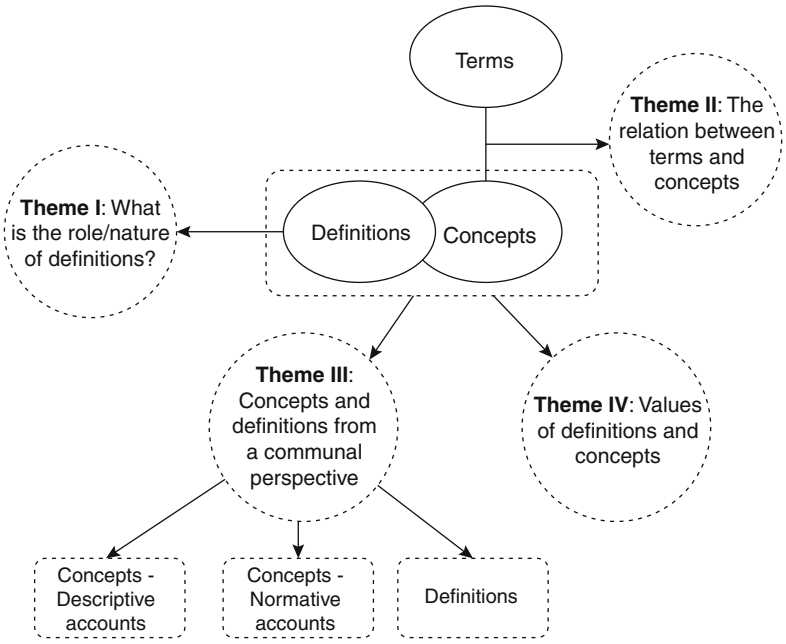


Fig. 1 The four themes and their relations

can also be considered cases of certain concepts becoming fashionable and others falling out of fashion (cf. theme III).

The themes represent different paths that lead to the discussion of definitions and concepts in mathematical practice. However, these paths are connected. For instance, beneficial properties of definitions (cf. theme IV) are dependent on the expectations and aims of definitions (cf. theme I). Different values (cf. theme IV) can guide the process of choosing which concept to relate to a term (cf. theme II). Definitions can become fashionable because they capture a relevant concept in a natural way (cf. themes II, III, and IV). Studying the relations between the different themes is therefore a promising avenue for future research.

References

Andersen LE (2018) Acceptable gaps in mathematical proofs. Synthese:1–15. https://doi.org/10.1007/s11229-018-1778-8

Baldwin J (2018) The explanatory power of a new proof: Henkin's completeness proof. In: Piazza M, Pulcini G (eds) Truth, existence and explanation: FilMat 2016 studies in the philosophy of mathematics. Springer International Publishing, Cham, pp 147–162. https://doi.org/10.1007/978-3-319-93342-9_9

Belnap N (1993) On rigorous definitions. Philos Stud 72(2):115–146. https://doi.org/10.1007/BF00989671

Cappelen H, Plunkett D (2020) Introduction: a guided tour of conceptual engineering and conceptual ethics. In: Conceptual engineering and conceptual ethics. Oxford University Press, Oxford, pp 1–26. https://doi.org/10.1093/oso/9780198801856.003.0001
Carballo AP (2020) Conceptual evaluation: epistemic. In: Conceptual engineering and conceptual ethics. Oxford University Press, Oxford, pp 304–332. https://doi.org/10.1093/oso/9780198801856.003.0015
Carter J (2013) Handling mathematical objects: representations and context. Synthese 190(17): 3983–3999
Cellucci C (2018) Definition in mathematics. Eur J Philos Sci 8(3):605–629
Corfield D (2001) The importance of mathematical conceptualisation. Stud Hist Philos Sci Part A 32(3):507–533. https://doi.org/10.1016/S0039-3681(01)00007-3
Corfield D (2003) Towards a philosophy of real mathematics. Cambridge University Press, Cambridge
D'Alessandro W (2019) Explanation in mathematics: proofs and practice. Philos Compass 14(11): e12629. https://doi.org/10.1111/phc3.12629
Dawson JW Jr (2006) Why do mathematicians re-prove theorems?†. Philos Math 14(3):269–286. https://doi.org/10.1093/philmat/nkl009
De Villiers M (1990) The role and function of proof in mathematics. Pythagoras 24:17–24
Detlefsen M (2008) Purity as an ideal of proof. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 179–197
Frans J, Kosolosky L (2014) Mathematical proofs in practice: revisiting the reliability of published mathematical proofs. Theoria Int J Theory Hist Found Sci 29(3(81)):345–360
Geist C, Löwe B, Van Kerkhove B (2010) Peer review and knowledge by testimony in mathematics. In: PhiMSAMP: philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London, pp 155–178
Goethe NB, Friend M (2010) Confronting ideals of proof with the ways of proving of the research mathematician. Stud Logica 96(2):273–288. https://doi.org/10.1007/s11225-010-9284-0
Gupta A (2019) Definitions. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. Winter 2019 edn. Metaphysics Research Lab, Stanford University
Hacking I (2014) Why is there philosophy of mathematics at all? Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781107279346
Hamami Y (2019) Mathematical rigor and proof. Rev Symb Log:1–41. https://doi.org/10.1017/S1755020319000443
Horty JF (2007) Frege on definitions: a case study of semantic content. Oxford University Press, Oxford
Inglis M, Aberdein A (2015) Beauty is not simplicity: an analysis of mathematicians' proof appraisals. Philos Math 23(1):87–109. https://doi.org/10.1093/philmat/nku014
Kitcher P (1983) The nature of mathematical knowledge. Oxford University Press, New York
Kjeldsen TH, Carter J (2012) The growth of mathematical knowledge—introduction of convex bodies. Stud Hist Philos Sci Part A 43(2):359–365. https://doi.org/10.1016/j.shpsa.2011.12.031
Kleiner I (1991) Rigor and proof in mathematics: a historical perspective. Math Mag 64(5): 291–314. https://doi.org/10.2307/2690647
Koetsier T (1991) Lakatos' philosophy of mathematics: a historical approach. Studies in the history and philosophy of mathematics, 0928–2017, vol 3. North-Holland, Amsterdam
Lakatos I, Worrall J, Zahar E (2015) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge
Lange M (2009) Why proofs by mathematical induction are generally not explanatory. Analysis 69(2):203–211
Larvor B (1998) Lakatos: an introduction. Routledge, London & New York
Larvor B (2012) How to think about informal proofs. Synthese 187(2):715–730. https://doi.org/10.1007/s11229-011-0007-5
Lehet E (2019) Induction and explanatory definitions in mathematics. Synthese. https://doi.org/10.1007/s11229-019-02095-y

Margolis E, Laurence S (2019) Concepts. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. Summer 2019 edn. Metaphysics Research Lab, Stanford University

Mauro LS, Venturi G (2015) Naturalness in mathematics. In: Lolli G, Panza M, Venturi G (eds) From logic to practice: Italian studies in the philosophy of mathematics. Springer International Publishing, Cham, pp 277–313. https://doi.org/10.1007/978-3-319-10434-8_14

Muntersbjorn MM (2003) Representational innovation and mathematical ontology. Synthese 134(1/2):159–180

Ouvrier-Buffet C (2013) Modeling of the defining activity in mathematics and of its dialectic with the proving process epistemological study and didactical challenges. Université Paris-Diderot – Paris VII, Paris

Ouvrier-Buffet C (2015) A model of mathematicians' approach to the defining processes. In: CERME 9 – ninth congress of the European Society for Research in Mathematics Education, Prague, Czech Republic, 2015-02-04. pp 2214–2220

Raman-Sundström M (2015) A pedagogical history of compactness. Am Math Mon 122(7):619–635. https://doi.org/10.4169/amer.math.monthly.122.7.619

Robinson R (1954) Definition. Clarendon Press, Oxford

Schlimm D (2012) Mathematical concepts and investigative practice. In: Scientific concepts and investigative practice. De Gruyter, Berlin/Boston, pp 127–148. https://doi.org/10.1515/9783110253610

Shieh S (2008) Frege on definitions. Philos Compass 3(5):992–1012. https://doi.org/10.1111/j.1747-9991.2008.00167.x

Tanswell FS (2018) Conceptual engineering for mathematical concepts. Inquiry 61(8):881–913. https://doi.org/10.1080/0020174X.2017.1385526

Tappenden J (1995) Extending knowledge and 'fruitful concepts': Fregean themes in the foundations of mathematics. Noûs 29(4):427–467. https://doi.org/10.2307/2216281

Tappenden J (2008a) Mathematical concepts and definitions. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 256–275. https://doi.org/10.1093/acprof:oso/9780199296453.003.0010

Tappenden J (2008b) Mathematical concepts: fruitfulness and naturalness. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 276–301

Tappenden J (2012) Fruitfulness as a theme in the philosophy of mathematics. J Philos 109(1/2):204–219

Thurston WP (1994) On proof and progress in mathematics. Bull Am Math Soc 30(2):161–177. https://doi.org/10.1090/S0273-0979-1994-00502-6

Werndl C (2009) Justifying definitions in mathematics – going beyond Lakatos. Philos Math 17(3):313–340

Wilder RL (1953) The origin and growth of mathematical concepts. Bull Am Math Soc 59(5):423–448

Yap A (2011) Gauss' quadratic reciprocity theorem and mathematical fruitfulness. Stud Hist Philos Sci Part A 42(3):410–415. https://doi.org/10.1016/j.shpsa.2010.09.002

Mathematics and the Method of Abstraction

Kit Fine

Abstract

I provide a brief history and some philosophical reflections on the development of the method of abstraction in mathematics during the late nineteenth and early twentieth century.

Keywords

Abstraction · Frege · von Neumann · Cantor · Locke · Berkeley · Russell · Number

It is a familiar thought that mathematics derives from abstraction. We observe three rocks, three stars, three French hens and from these respective observations we are able to derive the concept of the number 3; or we observe a line in the sand and from this observed line, thick and wiggly as it may be, we are able to derive the concept of a line, perfectly indivisible in extent, absolutely straight, and extending infinitely in either direction; or we observe a car in motion, the rising temperature of the air, or the increasing volume of the water in the tub and from these sundry observations we are able to derive the concept of the rate of change of one quantity with respect to another.

Abstraction may be a familiar idea but it is also vague. What exactly is abstraction? *From* what do we abstract? *To* what do we abstract? And *how* do we get from the one to the other?

This chapter is based upon an address I gave to a general audience in 2013, under the auspices of the Humboldt Foundation.

K. Fine (✉)
New York University, New York, NY, USA
e-mail: kf14@nyu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_91

These are questions which have been discussed since the beginning of mathematics and philosophy. But what I want to consider is the discussion of the questions in a particular period and by a particular group of people. The period was the end of the nineteenth century and the beginning of the twentieth century. And the people in question were mathematicians who had a strong interest in foundational issues. They were not only concerned to *do* mathematics but also to put the subject on a firm foundation.

The period was, I believe, a time when discussion of abstraction came to fruition. And this, certainly, was no accident. For this was a time when many of the most significant advances in mathematics depended upon extending its ontology. Mathematicians could no longer be content with the familiar numbers of arithmetic or the familiar points of Euclidean geometry; they also had to traffic in complex numbers, points at infinity, transfinite numbers, and the like. The question therefore arose as to what might justify these extensions of the familiar ontology; and to many mathematicians it seemed that the answer lay in the method of abstraction. These new objects were to be introduced and legitimated, if at all, on the basis of abstraction.

When one examines the mathematical writings of this period, then there are, I think, three main conceptions of abstraction that emerge. I will call them abstraction via *extension*, via *subtraction*, and via *representation*. Each of these three conceptions is associated with particular authors – extensional abstraction with Frege and Russell, subtractive abstraction with Dedekind and Cantor, and representational abstraction with Zermelo and von Neumann. And so what I would like to do is to discuss each conception in turn, more or less as it was developed by the respective authors.

In each case, I shall focus on the case of natural numbers – the numbers 0, 1, 2, . . . and the question of how they are to be obtained by abstraction. But it is important to bear in mind that the case of numbers is merely illustrative. The method of abstraction is meant to have more general application – to the rational numbers, for example, or to the reals of analysis, or to directions in geometry, or to free groups in algebra, and so on and so forth.

In each case, I shall emphasize the historical and philosophical milieu in which the investigation into abstraction took place. For it is my view that this investigation did not take place within some isolated mathematical bubble but was heavily informed by previous philosophical discussion of the issues, dating all the way back to Euclid. Sometimes this influence was conscious – as in the case of Frege and Russell and Cantor – but, even when it was not conscious, there is no doubt of its presence.

I begin with Frege and his account of abstraction as a form of extension. Gottlob Frege was a German mathematician, who was born in 1848 and died in 1925 and who spent most of his working life in Jena. He is best known for his work in logic, the philosophy of mathematics, and the philosophy of language and he is commonly regarded as the founder of analytic philosophy.

Although Frege seems to have received very little formal training in philosophy, there is no doubt that he was very well read in the subject. His work is replete with references to philosophers, both past and present; and he himself is credited with having said “Every good mathematician is at least half a philosopher, and every good

philosopher is at least half a mathematician." (If only more philosophers and more mathematicians would heed his words!)

His principal contribution to the philosophy of mathematics may well be regarded as a response to Kant. Kant had thought that arithmetical truths were synthetic apriori. They were a priori in that they could be known independently of experience – no experiments are required, for example, in order to know that $2 + 2 = 4$; and they were synthetic in that their truth could not, in general, be revealed through analysis alone. Whereas the truth of "every bachelor is unmarried" could be made evident through the analysis of the subject term "bachelor" as "unmarried man," no such analysis would reveal the truth of $2 + 2 = 4$. Frege, by contrast, held that all that arithmetical truths were analytic, that, just as in the bachelor case, a proper analysis of "2" and "+" and "4" would make evident the truth of $2 + 2 = 4$; and a large part of his detailed work in the philosophy of mathematics was devoted to showing how this was so.

To do this, Frege had to give an analysis of "2" and of other number concepts; and it is here that he appeals to the idea of extensional abstraction. Consider all pairs of objects: the pair of cities, Hamburg and Berlin; the pair of celestial bodies, the sun and the moon; or the pair of politicians, Trump and Merkel. Despite their vast differences, they all have something in common – duplicity or being double. A natural thought is that one might then take this common element of duplicity to be the number 2.

But what does one mean by saying that all these pairs have duplicity in common? What is this duplicity and what is it for a pair to have it? Frege's answer is beautifully simple. Just as we can collect objects into pairs, so we can collect all pairs. Thus we may take duplicity to be the collection of all pairs – or perhaps more simply still, it is the set of all sets with two members. But how can I say that a set has two members without presupposing the number 2? Simply by saying that there is a member x and a distinct member y which are such that very member of the set is either identical to x or identical to y. In this way, we can provide a purely logical definition of the number 2 and any more substantive conception of number simply disappears in a puff of logical smoke.

Moreover, the account generalizes. For given any two collections we can say that they are "equinumerous" if they can be paired one to one. The collection of boys and the collection of girls at a party, for example, will be equinumerous (in this purely logical sense) if they can be paired off, with each boy going with one girl and each girl with one boy. Given a collection of objects, we can then say that its number is the collection of all equinumerous collections; and a number, in general, will be the number of some collection.

Frege's account of number did not appear out of the blue. It owed a great deal to the discussion of the role of ideal elements in mathematics earlier in the century. Frege himself credits one critical aspect of his account to Hume. In his Treatise (Part 3 of Book 1), Hume had written:

> When two numbers are so combined, as that the one has always a unit answering to every unit of the other, we pronounce them equal.

And it was this "standard of equality," as Hume called it, that provided Frege with his general account of number. But Frege went one step further. For he did not simply adopt one-one pairing as a "standard of equality" but took numbers themselves to be the embodiment of that standard, with each number being a totality of collections that conformed to the standard.

I turn to Cantor and Dedekind and what I have called their subtractive account of abstraction. I shall focus on Cantor, although Dedekind held very similar views.

Cantor was a German mathematician, who lived from 1845 to 1918 and who spent most of his working career in Halle. He is best known for founding set theory and developing the theory of transfinite number. David Hilbert, in a rare accolade, called his work in this area "the finest product of mathematical genius and one of the supreme achievements of purely intellectual human activity."

I do not know what formal training Cantor received in philosophy but there is no doubt that he took his work on the transfinite to require philosophical defense. He was especially concerned with investigating previous philosophical work on the infinite to see if it provided any good arguments for or against his conception of actual infinity. Indeed, at one point in his career, when he was especially depressed about the reception of his mathematical work, he asked if he could lecture on philosophy instead! He was also an accomplished violinist and spent a large amount of time studying Elizabethan literature in the attempt to show that Francis Bacon wrote the plays attributed to Shakespeare. Perhaps we should say, not that he was half mathematician and half philosopher, but quarter mathematician, quarter philosopher, quarter musician, and quarter literary critic.

I cannot deal with his theory of transfinite number here. But I do want to look at one aspect of that theory which is invariably overlooked in contemporary expositions of his work but which bears directly on the question of abstraction. Cantor wanted to extend numbers from the finite to the infinite. He did not simply want to use the natural numbers to count finite collections; he also wanted to count the natural numbers themselves. And not just the natural numbers but all sets of natural numbers. And all sets of such sets. And so on.

To this end, he required a uniform account of number, one which would apply equally well to the finite natural numbers and to the transfinite cardinals, by which infinite collections of objects might also be counted. Take the number 2, as an example. What is it? Like Frege, Cantor thought that the number 2 could be obtained by abstraction from all pairs. But his conception of abstraction was quite different. Instead of collecting all pairs of objects together to form a single object, the number 2, Cantor focused on a particular pair – it does not matter which, say the pair consisting of Anthony and Cleopatra – and then stripped the two objects of all their defining features – the color of their hair, the softness of their skin, even their humanity – until we were left with two objects which differed merely in being distinct. The resulting pair of objects was then taken to be the number 2. Thus the number 2 was itself a pair of objects, but no ordinary pair, but a pair of objects that had been stripped of all individuality beyond their mere numerical difference. A similar account would then apply to all finite collections of objects or all collections whatever, whether finite or infinite.

Cantor likened his account to Euclid's well-known account of number as a multiplicity of units, where he took the units to be what is left of the concrete individuals once they have been stripped of their individuality. But one might also compare Cantor's account of number to Locke's account of abstract ideas. One obtains an abstract idea of a triangle by stripping away the particular idea of a triangle of all those features – such as being right angled or obtuse or acute – that might distinguish it from one kind of triangle as opposed to another. Clearly, the process by which Cantor's units are obtained is similar, though in this case the stripping away is much more drastic, since only the mere identity of the particular is left. A further difference is that a unit is not left, in a solitary denuded state, out in the cold; it has similar denuded companions and these may combine with one another to form sets or other more complex structures. One might say that Cantor's conception of abstraction is the same as Locke's but with a mathematical twist, making it more rigorous and more suited to the needs of mathematics.

Bishop Berkeley subjected to Locke's account of abstract ideas to scathing criticism; and Frege, in his turn, subjected Cantor's account of numbers as sets of units to what was possibly even more scathing criticism. It seems that, in either case, we can form no intelligible conception of what an abstract idea of a triangle or of what a unit might be; and even Locke, in what appears to be an unguarded moment, admits that the abstract idea of a triangle "must be neither equilateral, nor equicrural, nor scalenon; but none and all of these at once." A strange thing, indeed!

My view is that the criticisms of Berkeley and Frege can be met, that the Cantorian view can be put on a firm mathematical foundation, and that it thereby offers a much more satisfactory and fruitful account of number than the Fregean account. I cannot argue this here but let me note that the Cantorian view is much more *natural* than the Fregean view. If we want the number two to be what is *common* to all pairs, then it cannot sensibly be regarded as the set of all pairs but is much more naturally regarded as the template, so to speak, from which all pairs might then be derived. But the set of two units provides such a template. Substitute Anthony and Cleopatra, or any other two objects, for the units and one thereby obtains the pair of those objects; the particular pairs are, in this way, derived from the generic set that is common to all pairs.

The Frege-Russell account of number, by contrast, is highly artificial – instead of telling us what different pairs have in *common* it simply piles them one on top of the other within an over-arching set. Imagine, upon being asked what all cocktails have in common, one was simply given more cocktails! Moreover, in standard axiomatizations of set theory, the account is inconsistent. To put it crudely, the totality of *all* pairs, including pairs of objects that include this totality itself, is too large to constitute a well-defined set and, in conjunction with other plausible assumptions, the assumption of its existence will lead to contradiction.

I turn finally to Zermelo and von Neumann and their development of the representational account of abstraction. I shall focus on von Neumann rather than Zermelo since, in his hands, the representational account of number received its most developed form.

John von Neumann lived from 1903 to 1957. He was an infant prodigy able to perform complex mathematical calculations at lightning speed. He made significant

contributions to a wide number of different fields – including game theory, quantum mechanics, the theory of computing, and the foundations of mathematics – and was perhaps the most versatile mathematician who ever lived.

My concern, as before, is with his account of number and with its significance for the concept of abstraction. We have already encountered two accounts of number. According to the first, Frege's, the number 2 is the set of all pairs. According to the second, Cantor's, the number 2 is a generic pair, the pair of two units. Von Neumann's account is different still. Like Cantor, he takes the number 2 to be itself a pair. But he does not take it to be a pair of units but a pair of ordinary objects. But which objects might these be? Surely, we cannot take the number 2 to be the pair Anthony and Cleopatra. For Anthony and Cleopatra, intense as their love may have been, are only one pair among many; and why choose Anthony and Cleopatra, rather than Trump and Merkel or the sun and the moon – two other remarkable pairs?

Von Neumann realized that there was a very natural choice for the pair. For we might take the number 2 to be the set of its predecessors. 2 has two predecessors, 1 and 0; and so the number 2 will be the set whose members are 1 and 0; a two-membered set. But what of 1 and 0? A similar account may be given. 1 will be the set of its predecessors. But it has only one predecessor, viz. 0, and so 1 will be the set whose sole member is 0. 0 itself has no predecessors and so it will be the null set, the set with no members. Thus 2 will turn out to be the set whose members are 1, that is, the set whose sole member is the null set, and 0, that is, the null set. And similarly for all other natural numbers.

One nice feature of von Neumann's account is that it readily extends to the transfinite. According to Cantor, there should be a transfinite ordinal ω that occurs after all the natural numbers. But this can now be taken to be the set of natural numbers. Likewise ω + 1, the next ordinal after ω, can be taken to be the set whose members consist of the natural numbers and the set of all natural numbers. And so on, through ever higher levels of infinity.

Another nice feature is that it does not lead to paradox and conforms to a well-founded conception of sets. If I take a number, it will be the set of preceding numbers; each of those, in its turn, will be the set of preceding numbers; and so on. Eventually, the whole construction will bottom out in the empty set, that is, the number 0.

There is little doubt that von Neumann saw his account of number as an alternative to the Frege-Russell view and perhaps to the Cantorian view as well. He says of his account "What we really wish to do is to take as the basis of our considerations the proposition: 'Every ordinal is the type of the set of all ordinals that precede it.' But in order to avoid the vague notion 'type,' we express it in the form: 'Every ordinal is the set of the ordinals that precede it'" (von Neumann 1923, p. 347 of the English translation). But what is this "type" that is said to be vague? Von Neumann could only have had two models in mind: the Frege-Russell type which is obtained by extensional abstraction; or the Cantorian type which is obtained by subtractive abstraction. But he would have regarded both as unsatisfactory – the first because it leads to inconsistency and the second because it seems hopeless obscure.

In the face of these difficulties, he made the bold move of replacing the type with a particular instance of that type, the set of preceding ordinals.

Those familiar with the history of philosophy will recognize that he makes the same move as Berkeley made in response to Locke. As I mentioned, Berkeley criticized Locke's theory of abstract ideas as incoherent, but he had a positive view of his own, which Hume described as "one of the . . . most valuable discoveries that has been made . . . in the Republic of Letters." His view was that the role played by Locke's abstract idea should instead be played by a particular idea that served to represent all similar ideas or, to state the view in less psychological terms, the type was to be replaced with a particular instance of the type but used in a generic way to represent all other instances of the type. Thus just as Cantor is Locke with a mathematical twist, so von Neumann is Berkeley with a mathematical twist. For von Neumann did not simply choose any particular instance to represent the type but a very particular instance, one that was especially well suited to meet the mathematical demands placed upon an adequate theory of number within the context of set theory.

I do not want to say that Cantor or Zermelo or von Neumann explicitly had Locke or Berkeley in mind. Indeed, I doubt that they did. But I find it hard to believe that, as well educated individuals, they were not to some extent familiar, either directly or indirectly, with the empiricist debate over abstract ideas and were not to some extent influenced by it.

Let me conclude with two general points. The first is that the work of these mathematicians was not only relevant to the foundations of mathematics but has become assimilated into general mathematical practice. Any mathematics undergraduate is likely to encounter some form of these different methods of abstraction in his or her study. An example is arithmetic modulo 2 (or, more generally, modulo n). Consider again the natural numbers and suppose we ignore the difference between all of the even numbers, on the one hand, and all of the odd numbers on the other hand. Then abstraction gives us two numbers, so to speak, the Odd and the Even. The numbers can then be added and multiplied. Odd plus Even will be Odd, for example, Odd plus Odd will be Even, while Odd times Even will be Even. But mathematically speaking, what are these numbers, the Odd and the Even? Here the method of extensional abstraction may be used. For we can identify the Even with the set of all even numbers and the Odd with the set of all odd numbers. Moreover, this procedure can be generalized to cover many other cases in which we are given some elements, such as the numbers, some operations on these elements, such as addition and multiplication, and wish to identify certain of these elements, just as we have identified all odd numbers and all even numbers (another nice example is Conway's theory of surreal numbers, in which each surreal number is identified with a set of surreal "forms").

The methods of extensional and representational abstraction have been thoroughly absorbed into the mathematical literature. But my own view is that the potential of subtractive abstraction has not been fully realized and that there is a great deal more that could be done, both in making the idea mathematically precise and in applying it across a wide range of different fields.

The second point is that this work on abstraction would not have been possible – or so readily possible – in the absence of an intellectual milieu in which there was a free interplay between mathematical and philosophical ideas and that such a milieu would not itself have been possible if the mathematicians had not themselves had some kind of philosophical background or had been receptive to philosophical ideas. As we look back, we can see how beneficial it was for them to be working within such a milieu; and, as we look forward, we should be aware, even in this age of increased specialization, of how important it is that the interplay between mathematical and philosophical ideas be sustained.

1 Cross-References

▶ Platonism, *De Re*, and (Philosophy of) Mathematical Practice
▶ The Nature of Mathematical Objects

References

Berkeley G (1710) A Treatise concerning the Principles of Human Knowledge. Dublin: Aaron Rhames

Cantor G (1955) Contributions to the Founding of the Theory of Transfinite Numbers, translation of 'Beiträge zur Begrűndung der Transfiniten Mengenlehre', Mathematische Annalen, volume 46, pp. 481–512 [1895] and with an introduction and notes, by P. E. B. Jourdain, New York: Dover

Conway J (2001) On Numbers and Games (2nd edition), New York: CRC Press

Dedekind R (1888) Was sind und was sollen die Zahlen?, Braunschweig: Vieweg; English translation 'What Are Numbers and What Should They Be?' (H.Pogorzelski, W. Ryan & W. Snyder, eds. and trans.), Orono, ME: Research Institute for Mathematics

Frege G (1960) The Foundations of Arithmetic: A Logico-Mathematical Enquiry Into the Concept of Number, translated by Austin, J. L. (2nd ed.). Evanston, Illinois: Northwestern University Press

Locke J (1690) An Essay Concerning Human Understanding, Book II, London: Thomas Basset

Russell B (1919) Introduction to Mathematical Philosophy, London: George Allen & Unwin

Volume I: Set Theory, Miscellanea' [2010] H.-D. Ebbinghaus and A. Kanamori (eds.), with a facing-page English translation, and an Introduction by Ulrich Felgner 189–228. Berlin: Springer

von Neumann J (1923) "On the introduction of transfinite numbers", in Jean van Heijenoort (ed.), 'From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931' (3rd ed.), Harvard University Press, pp. 346–354, English translation of von Neumann (1923), "Zur Einführung der Transfiniten Zahlen", Acta Litterarum ac Scientiarum Ragiae Universitatis Hungaricae Francisco-Josephinae, Sectio scientiarum mathematicarum 1:199–208

Zermelo E (1908) Untersuchungen über die Grundlagen der Mengenlehre, I, Mathematische Annalen 65: 261–281; reprinted in 'Zermelo: Collected Works

Leibniz on Number Systems

Lloyd Strickland

Contents

1 Introduction 168
2 Who Influenced Leibniz? 170
3 Early Writings on Binary 172
4 The Development of Binary 177
5 Binary Fractions and Expansions 181
6 Binary and Philosophy 184
7 Binary and The Yijing 187
8 Sedecimal 188
9 Other Nondecimal Number Systems 191
10 Conclusion 193
Bibliography 195

Abstract

This chapter examines the pioneering work of Gottfried Wilhelm Leibniz (1646–1716) on various number systems, in particular binary, which he independently invented in the mid-to-late 1670s, and hexadecimal, which he invented in 1679. The chapter begins with the oft-debated question of who may have influenced Leibniz's invention of binary, though as none of the proposed candidates is plausible, I suggest a different hypothesis that Leibniz initially developed binary notation as a tool to assist his investigations in mathematical problems that were exercising him at the time, namely those concerning the divisibility of composite numbers, primality, and perfect numbers. The chapter then explores Leibniz's development of binary, his little-known work on binary fractions and expansions, his use of binary as a symbol for creation in his philosophical-theology, and his response to the suggestion that there was a correlation between binary numeration and the hexagrams of the ancient Chinese divinatory text, the Yijing. The chapter

L. Strickland (✉)
History, Politics, and Philosophy, Manchester Metropolitan University, Manchester, UK
e-mail: L.Strickland@mmu.ac.uk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_90

then focuses on Leibniz's work on other number systems, in particular his invention and exploration of hexadecimal as well as his work on duodecimal. The chapter concludes by revealing a hitherto unknown practical application of binary that Leibniz devised in the last year of his life.

Keywords

Leibniz · Binary · Non-decimal · Number systems · Hexadecimal

1 Introduction

Although the German polymath Gottfried Wilhelm Leibniz (1646–1716) made original contributions to a wide range of disciplines, such as law, philosophy, politics, languages, and many areas of science, his fame and renown have always rested principally on his pioneering mathematical work, in particular his independent invention of the calculus in 1675. Another of his enduring mathematical contributions was his invention of the binary number system, the basis for today's world of digital computing and communications. In this system, which uses only two digits – 0 and 1 – every position starting from the right represents a successive power of 2, such that the rightmost place is the 2^0 (or 1's) column, the next-rightmost place the 2^1 (or 2's column), and so on. While Leibniz had a particular fascination with binary, his work on nondecimal number systems covered all the bases, or at least bases 2–16. Particularly notable is his invention of base 16, or "hexadecimal" in modern parlance (Leibniz called it "sedecimal" and "sedenary"), another number system used in modern computing.

128's	64's	32's	16's	8's	4's	2's	1's
1	0	0	1	0	0	1	0

Despite the proliferation of nondecimal weights, measures, and currencies in the seventeenth century, the idea of exploring nondecimal number bases, as opposed to tallying systems, was not commonplace. While counting in fives, twelves, or twenties was well understood and widely practiced, the idea of counting in bases 5, 12, or 20 was not. The modern idea of a base as a positional numbering system was still coalescing. But it was glimpsed by a few. Indeed, the binary number system was independently invented at least three times in the seventeenth century: by Thomas Hariot (1560–1621) near the turn of the century (see Strickland 2023c), by Juan Caramuel y Lobkowitz (1606–1682) in 1670, and again by Leibniz c.1677–1678. Moreover, Erhard Weigel (1625–1699) proposed the quaternary or base-4 number system in 1673, and Joshua Jordaine vigorously defended the duodecimal or base 12 system in 1687 (see Weigel 1673 and Jordaine 1687). But if Leibniz's willingness to explore nondecimal number systems was not entirely unprecedented in the seventeenth century, the depth of his explorations was. He wrote in excess of a

hundred manuscripts on binary, and at least as many more in which binary is mentioned alongside other matters. There are manuscripts dealing with other nondecimal bases too. Of these riches, Leibniz chose to publish just a single paper in his lifetime: "Explanation of Binary Arithmetic," which appeared in 1705. This was enough to bind his name to the binary system in the centuries that followed, such that it was often incorrectly stated that he was its sole inventor. But of the three mathematicians who had devised binary in the seventeenth century, only Leibniz would influence the future development of base 2 and its eventual adoption by early computer engineers (see Strickland and Lewis 2022, 16–21): Hariot's manuscripts on binary arithmetic remained unknown until the 1920s (see Morley 1922) and have been published only in the last decade, while Caramuel y Lobkowitz's book was widely overlooked, being cited more in catalogues of book collections than in the work of other mathematicians.

Leibniz's decision not to publish more of his work on nondecimal number systems limited his impact; he is not, for example, generally given credit for the invention of base 16 (an honor usually given to the nineteenth-century Swedish-American engineer John Nystrom; see Knuth 1998, 202). Even today, more than 300 years after Leibniz's death, the vast majority of his writings on binary and other nondecimal number systems remains unpublished. As a consequence, many of the texts cited in this chapter are unpublished manuscripts. Since these are not generally accessible, I have quoted from them liberally. Given that the full depth of Leibniz's writings on binary and other nondecimal number systems remains to be explored, the account of them in this chapter will necessarily be incomplete but nevertheless should go some way toward showing the depth of Leibniz's engagement with nondecimal number systems.

The plan of this chapter is as follows. Section 1 takes up the oft-debated question of who may have influenced Leibniz's invention of binary. As none of the proposed candidates is plausible, I turn in Sect. 2 to the question of why Leibniz invented binary, considering both his own answer and a rather different one suggested by his earliest writings on the subject. Section 3 concerns Leibniz's subsequent development of binary, in particular his methods for converting decimal to binary and for performing the four basic arithmetic operations in base 2, as well as his ideas for binary-calculating machines and his fascination with the periodic recurrence of digits of the numbers of different arithmetic and geometric sequences expressed in binary. Section 4 focuses on Leibniz's little-known work on binary fractions and expansions, and his hope that these would throw light on irrationals and transcendental quantities. In Sect. 5, I outline the symbolic use of binary in Leibniz's philosophy, where 1 and 0 are envisaged as analogues of God and nothingness, such that binary numeration is seen as a representation of the theologically orthodox doctrine of creation *ex nihilo*. Having learned of Leibniz's binary system, a Jesuit missionary in China, Joachim Bouvet (1656–1730), suggested a correlation between binary numeration and the hexagrams of the ancient Chinese divinatory text, the Yijing. This correlation – and Leibniz's enthusiastic reaction thereto – is the subject of Sect. 6. In Sect. 7, I turn to Leibniz's invention and exploration of base 16, or sedecimal, while Sect. 8 concerns Leibniz's work on other nondecimal number

systems, in particular bases 4 and 12. The chapter concludes with an account of a hitherto unknown practical application of binary that Leibniz devised in the last year of his life. Let us begin, then, with Leibniz's potential influences.

2 Who Influenced Leibniz?

The question of who influenced Leibniz's invention of binary has been often asked and just as often has been answered inadequately. Four candidates have been proposed: John Napier (1550–1617), Francis Bacon (1561–1626), Erhard Weigel, and Juan Caramuel y Lobkowitz.[1] While none of these suggestions is tenable, examining them is worthwhile for the light it sheds on seventeenth-century exploration of different number bases and of ways to simplify calculation.

In 1617, Napier published his "location arithmetic" which employed a form of binary *reckoning* based on the powers of 2 geometric sequence – 1, 2, 4, 8, 16, 32, 64, etc. – carried out on a chessboard. The successive powers of 2, identified by letters from the classic Latin alphabet (so no j, u, or w) and then the Greek letters, are given in the margins. When multiplying two numbers, each had to be decomposed into its respective powers of 2, e.g., 37 into 32, 4, and 1, with these then converted into letters which determined where counters would be placed on the board. By moving the counters according to fixed rules, one could carry out the multiplication sum without effort, aside from having to reconstruct the product by adding up its various powers of 2, the form in which the answer was given (see Napier 1617, 115–154).

Napier's binary reckoning was shortly followed by a form of binary *information code*. In 1623, Bacon published his *De dignitate et augmentis scientiarum* [On the Dignity and Advancement of the Sciences] which contained a description of a cipher employing a biliteral alphabet, where aaaaa = A, aaaab = B, aaaba = C, aaabb = D, etc. Bacon (1623, 279) believed these cipher-characters opened the way for people to express their thoughts "by objects which may be presented to sight or hearing, provided those objects are capable of a twofold difference only, as by bells, trumpets, torches, the report of muskets, and any other things whatsoever." The suggestion was not pursued, but the principle survives in the binary code used in the Braille reading system and in the ASCII (American Standard Code for Information Interchange) character table.

The first publication of binary *notation* for numbers occurred almost fifty years after Bacon. In his *Mathesis biceps* of 1670, Caramuel y Lobkowitz included brief chapters on bases 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, and 60, in each case providing a table of numbers, details of who (if anyone) had used that system, and examples of how each number is embedded or reflected in the natural or supernatural world (Caramuel y Lobkowitz 1670, XLV–LXI). In the chapter entitled "On Binary Arithmetic," he presents a table using 0 and the letter *a* to form the sequence of numbers up to 32 expressed in binary (for example, in his notation, 17 is a00a). Despite the chapter

[1] See Couturat (1901, 473); Zacher (1973, 9–33); Tropfke (1980, 12); and Ingaliso (2017, 111–112).

title, no attempt is made to show how arithmetic would work in base 2; as such, Caramuel y Lobkowitz did not venture beyond binary notation.

However, a fully developed nondecimal *number system*, complete with arithmetic algorithms, was not long in coming. In 1673, Weigel published *Tetractys*, a small book in which he advocated for the widespread adoption of the quaternary, or base-4 system, loosely inspired by the Pythagorean tetrad, a triangular figure with four rows of 1, 2, 3, and 4 points, respectively. In a relatively detailed treatment, Weigel (1673, 16–23) outlines methods of performing all four arithmetic operations in base 4, as well as methods for the extraction of square and cube roots.

It is easy to see why the ideas of many of these thinkers have been considered as a possible influence on Leibniz's invention of binary, since all share structural features with the binary number system. But as superficially tempting as it might be to consider one or more of these thinkers as Leibniz's influence, there is no evidence that any of them were, and good reason to think that two of them were not. Manuscript evidence indicates that Leibniz became aware of Weigel's quaternary system only in 1683 (see A VI 4, 1162–1163), too late for it to have influenced his development of binary, while there is evidence that Leibniz never saw Caramuel y Lobkowitz's *Mathesis biceps* at all.[2] As for Napier and Bacon, Leibniz knew of their work but did not mention it in connection with binary. It was unlikely he was influenced by either. The similarities between Bacon's biliteral alphabet and the binary number system might be obvious to us, as we are in a position of being able to compare both, but to a seventeenth-century mind knowing only the former there would have been little reason to think of the latter, as the biliteral alphabet hardly suggests or implies a positional number system employing two digits. Indeed, Bacon's biliteral alphabet is sufficiently different from a binary number system in both form and function that the passage from one to the other would have required an enormous leap of imagination, almost on a par with that required to arrive at binary from any of the routine examples of duality or twoness with which Leibniz would have been familiar, such as matter and spirit, good and evil, and light and dark. As for the specific powers of 2 geometric sequence used by Napier, this was a common feature of mathematical works: Adam Ries (1529, n.p., chapter "Progressio") referred to it as the progression "in which one number exceeds another twofold"; Peter Ramus (1599, 177) called it the "subduple ratio," while Marin Mersenne (1644, n.p., "preface") dubbed it the "double progression" (Leibniz's own term was "double geometric progression"). In many cases, the sequence was used in case studies of how to sum portions of progressions (see Schott 1658, 636), a staple part of the medieval and early modern mathematical curriculum. But some mathematicians went further. Michael Stifel (1544, 237, cf. 31), for example, treated the sequence as an object of study in its own right, observing that addition and

[2] In 1711, decades after devising binary and only 5 years before his death, Leibniz (1768, V: 418) wrote: "Regarding Caramuel's *Mathesis biceps* [*Old*] *and New*, for which he asks ten thalers, I am unable to judge well because I have not yet seen it, and I fear it may contain vain subtleties, which is not unusual for Caramuel."

subtraction in the arithmetic sequence of natural numbers correspond to multiplication and division in the powers of 2 geometric sequence:

> Whatever the arithmetic progression does by addition and subtraction, such are the things the geometric progression does by multiplication and division...

0.	1.	2.	3.	4.	5.	6.	7.	8.
1.	2.	4.	8.	16.	32.	64.	128.	256.

Moreover, as we shall see in the next section, the powers of 2 geometric sequence was also the cornerstone of seventeenth-century investigations into both prime and perfect numbers, and some of Leibniz's earliest writings involving binary are on those subjects. The sequence was also well known to assayers, who routinely used weights in the proportion 1:2:4:8..., in order to weigh money and metals with as few weights as possible, as Leibniz often noted (Strickland and Lewis 2022, 130, 139, 195, 202). Accordingly, as the powers of 2 geometric sequence was in sufficiently wide use by mathematicians and assayers, Leibniz is no more likely to have drawn it from Napier than anyone else.

Since no plausible case can be made for thinking that any of the aforementioned thinkers influenced Leibniz's invention of binary, and there are no other likely candidates, we ought to be open to the idea that "who influenced Leibniz?" might be the wrong question to ask. A better question would be: Why did he invent it? There are two ways to answer this question: either by looking at Leibniz's own answer to it, or by looking at his earliest manuscripts on binary and seeing how it was used. In the next section, we shall do both.

3 Early Writings on Binary

In his own *ex post facto* narrative, written decades after his invention of binary, Leibniz claimed that it was conscious reflection on other nondecimal number systems that led him to think of binary, as "the simplest and most natural" base (Strickland and Lewis 2022, 138). He sometimes identified the duodecimal and the quaternary as these other nondecimal bases. While there is no evidence that he paid any attention to the quaternary system prior to encountering Weigel's defense of it in 1683 (see A VI 4, 1162–1163), there is one early writing in which Leibniz treats binary as an alternative to duodecimal: a manuscript entitled "Thesaurus mathematicus" [Mathematical Thesaurus], probably written in either 1678 or 1679. In this text, Leibniz works through various topics in arithmetic, geometry, and mechanics; near the end, he outlines how positional notation works in the decimal and duodecimal number systems before identifying binary as an alternative:

> From this outline it is clear that only these ten digits are needed: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Those in the first position signify the equivalent number of 1s, namely no 1s, one 1, two

> 1s, three, four, five, six, seven, eight, nine 1s. Those in the second position signify the equivalent number of *10s*, that is, 1s taken ten times; in the third position, the equivalent number of *100*s, that is, 10s taken ten times, or the squares of 10; in the fourth position, the equivalent number of *1000s*, that is, 100s taken ten times, or the cubes of 10, and so on. And in place of 10 one would be able to put any other number, for example, 12. For just as when the base a is 10, the square a^2 signifies 100 and the cube a^3 signifies 1000, so when a is 12, a^2 will be 12 times 12, that is, 144, and a^3 will be 12 times 144. But on this method, instead of the digits mentioned above—0, 1 etc. 9—two new digits would be needed in addition, one which would represent ten, the other which would represent eleven; but [the digits] 10 would signify twelve, and 100 would signify one hundred and forty four. And there are some who prefer to use this method of calculating over the common method, because 12 can be divided by 2, 3, 4, and 6; in addition, a calculation is completed with fewer digits. But the difference is not so great as to be worth abandoning the decimal progression. If anyone should want to use the binary progression, he would not need any digits except 0 and 1... But the calculation would be longer, albeit easier. (LH 35, 1, 25 Bl. 3v)

This is the only writing from the 1670s in which Leibniz mentions both binary and duodecimal. And his remarks in it do not support his later narrative that he devised binary after reflecting upon duodecimal; rather, binary is simply presented as an alternative to bases 10 and 12. Moreover, at the time of writing this manuscript, Leibniz must have developed a good understanding of binary already, given his remarks about the length and ease of calculations in base 2. Quite how good his understanding was at the time is unclear: in a marginal comment, Leibniz added, "We shall say more things about the binary progression below," though this promise is not fulfilled and the manuscript contains nothing further on the subject.

The lack of supporting evidence is one good reason to be cautious about Leibniz's *ex post facto* narrative. A second reason is that the narrative presents his invention of binary as detached from and independent of the other mathematical problems with which he was dealing at the time, as though binary emerged *sine contextu*, in what can only be described as a eureka moment. As little more can be said about Leibniz's narrative, let us consider an alternative hypothesis, that Leibniz's invention of binary – or at least binary *notation* – occurred in response to his work on three problems that exercised him in the mid-to-late 1670s, namely, devising methods and formulae to determine the divisibility of composite numbers, primality, and perfect numbers.[3] This hypothesis not only has the virtue of showing that binary developed in response to and as an aid for mathematical work with which he was already engaged, but also is supported by manuscript evidence. To illustrate, we shall consider three manuscripts, all written c.1677–1678, featuring some of Leibniz's earliest uses of binary notation. At the heart of all three manuscripts is the powers of 2 geometric sequence, which was at the forefront of Leibniz's mind at the time because of his investigations into the problems of number composition and primality. In the late 1670s, these would lead him to formulate and prove the primality test known as Fermat's little

[3] This work yielded his first publication in mathematics in February 1678, namely, the short journal article "A new observation about the way of testing whether a number is prime"; see Leibniz 1678/ http://www.leibniz-translations.com/prime.htm. For details of Leibniz's work on primes, see Mahnke (1912–1913).

theorem: $2^{n-1} - 1 \equiv 0 \pmod{n}$.[4] On his way to that, Leibniz investigated numbers of the form $2^n - 1$ (these would later be dubbed "Mersenne numbers"), as in the first of our three early manuscripts, a series of remarks about the divisibility of numbers written on the back of an envelope probably in 1678. Leibniz begins by stating the formulae $2^n - 1$, $3^n - 2$, $4^n - 3$, etc.:

> Every number exactly divides any number of the double geometric progression diminished by one.
> Every number exactly divides any number of the triple geometric progression diminished by one or two.
> Every number exactly divides any number of the quadruple geometric progression diminished by one or two or three.
> And so on ad infinitum.

(To the first statement, Leibniz later added: "false, except about a prime [number].") The manuscript continues:

> The numbers of any geometric progression can be expressed generally thus: 1, 10, 100, 1000, 10000 etc. Putting 10 means either 2 for the double [geometric progression] or 3 for the triple [geometric progression] etc. Therefore the numbers of the double geometric progression *diminished by one* can be expressed thus: 1 or 11 or 111 or 1111 etc.
> And of the triple [geometric progression] thus: 2 or 22 or 222 etc.
> And of the quadruple [geometric progression] thus: 3 or 33 or 333 etc.
> And of the decuple [geometric progression] thus: 9 or 99 or 999 etc. (LH 35, 12, 2 Bl. 4r)

Here, Leibniz uses binary numeration as an alternative way of expressing the numbers of a base (or "geometric progression"), noting that the numbers 1, 10, 100, 1000, 10000, etc. have different values depending on which base one is using.[5]

[4] In the manuscript entitled "Formarum reductio ad simplices," written 12 September 1680: "Therefore, $2^{z-1} - 1$ will be divisible by z, if z is prime" (LH 35, 3 A 4 Bl. 14r). During these investigations, Leibniz also glimpsed Wilson's theorem: $(n-2)! \equiv 1 \pmod{n}$, or in Leibniz's words: "The product of continuous [integers] up to the number which anteprecedes the given integer, when divided by the given integer, leaves 1, if the given integer is prime. If the given integer is derivative, it will leave a number which, since it has a common measure with the given integer, is greater than one" (LH 35 3 B 11 Bl. 21r). However, when testing his articulation of the theorem, Leibniz made a miscalculation that led him to add the false statement "(or the complement of 1)" after "leaves 1."

[5] In a slightly later manuscript, from 1679 (LH 35, 8, 30 Bl. 148), Leibniz again uses binary notation to illustrate his (still immature) formulation of a prime number theorem, namely, $2^n - 1$: "Let ⊙ = 111111, ☽ = 1111, ♀ = 111, ♂ = 11. ⊙ = ♀A, ⊙ = ♂B. Now ♀ and ♂ are prime among themselves, because their exponents are such, that is, their indices, or the numbers 2 and 3. Therefore A and B are not prime among themselves and necessarily will become A = ♂C and B = ♀C, and will become ⊙ = ♀♂C. $Z^z - 1$ is divisible by Z if Z is prime, and I have demonstrated this as follows: $2^2 - 1$ by 3 and $2^4 - 1$ by 5, therefore 1111 is divisible by 5 and 11 by 3 and 111111 by 7. But 11111 cannot be divided by 6, for since 11 can be divided by 3, 11111 and 11 have a common divisor, yet they are prime among themselves. And hence we have the sought-for demonstration of a reciprocal property of a prime number."

His use of binary here is purely illustrative, and binary plays no further role in his analysis (in the rest of the manuscript, Leibniz concerns himself with devising a method to determine which decimal number ending . . .991 to . . .999 is divisible by 17).

The second manuscript deals with a related matter: perfect numbers, that is, positive integers that are equal to the sum of their divisors (including 1 but excluding the number itself), such as 6 (= 3 + 2 + 1) and 28 (= 14 + 7 + 4 + 2 + 1). From the time of Euclid, investigations in this area of number theory have concentrated upon the powers of 2 geometric sequence, since Euclid's theorem of generating perfect numbers (or rather, even perfect numbers) is this: "if any number of numbers are set out successively in a double proportion [starting] from one, until the whole sum becomes prime, and this sum multiplied into the last [number] makes some number, then the [number] made will be perfect."[6] That is, if $2^n - 1$ is prime, then $2^{n-1}(2^n - 1)$ is perfect. Leibniz was reminded of this theorem when reading a 1678 dissertation on perfect numbers by Johann Wilhelm Pauli (1658–1723), who had used Roman letters to stand for the numbers of the powers of 2 geometric sequence in order to investigate the parts of perfect numbers via a sort of rudimentary algebraic calculus. Upon reading Pauli's book, Leibniz copied out Euclid's theorem then illustrated it in binary notation: "If 111 is prime, 11100 (that is, 111 by 100) is perfect (assuming 10, 100, 1000 are 2, 4, 8), production of the same thing by a different route" (LH 35 3 B 17, Bl. 2r). There is no indication here of Leibniz using binary notation to compute unknown perfect numbers, nor would that have been a realistic prospect in any case: the first five perfect numbers are 6, 28, 496, 8128, and 33550336, so using binary notation to represent any except the first two would obviously require unmanageably long strings of digits.[7]

Given the central role of powers of 2 in Leibniz's work on number composition, primes, and perfect numbers in the mid-to-late 1670s, it is easy to see why binary notation would have appealed to him. Take, for example, the number 7 expressed in binary: 111. A striking feature of this notation is that it displays the component powers of 2, as there is a 1 in the 4's position, a 1 in the 2's position, and a 1 in the 1's position. Thus, the binary value 111 is not just an alternative way of writing the decimal number 7, it is one that shows the three calculation steps $-(1 \times 2^2) + (1 \times 2^1) + (1 \times 2^0)$, that is, 4 + 2 + 1 – required to produce the number through powers of 2. In this vein, in 1686 Leibniz claimed that binary numeration is more perfect than decimal because in decimal there is no way to demonstrate from the digits 3 and

[6] Euclid, *Elements*, IX.36. That is, if p is a positive integer and $2^p - 1$ is prime, then $2^{p-1}(2^p - 1)$ is perfect.

[7] Perhaps because of this, in another manuscript on the subject, probably written in 1678, Leibniz sought to demonstrate perfect numbers using a mixture of binary numeration and algebra based thereon (so unrelated to Pauli's algebra), eventually reaching the conclusion "$2^{2z+1} - 2^z$ will be a perfect number if $2^{z+1} - 1$ is prime. Likewise, $2^{z-1} - 1$ will be divisible by z, if z is prime" (LH 35, 3 B 17 Bl. 1).

9 that $3 \times 3 = 9$, whereas in binary it is apparent (with a little working out) that $11 \times 11 = 1001$ (A VI 4, 800). And in the 1690s, when Leibniz started to provide detailed accounts of binary to some of his mathematician correspondents, he often made the point that a unique feature of binary notation is that it shows how any number can be represented as the sum of distinct powers of 2 (see A II 3, 452; Strickland and Lewis 2022, 111; 195).

Moreover, binary notation is convenient for expressing the numbers of the powers of 2 geometric sequence because these always consist in a 1 followed by 0s, e.g., $2 = 10$, $4 = 100$, $8 = 1000$, etc. And as Leibniz acknowledged in the first of our three early manuscripts above, binary notation is just as convenient for expressing the values of the Mersenne numbers, $2^n - 1$, because these always consist in 1s with no 0s, e.g., $2^1 - 1 = 1$; $2^2 - 1 = 11$; $2^3 - 1 = 111$; etc.[8] (In contemporary terminology, every Mersenne number is a binary repunit.) Therefore, binary numeration serves as a very convenient and very informative shorthand if one is working with various powers of 2, or the values of $2^n - 1$, especially for the lower values thereof, the very ones Leibniz tended to favor when illustrating his formulae.

And indeed, in our third manuscript, entitled "Calculus" [Calculation], written sometime after 1676 (but probably before 1680), Leibniz uses binary notation as one of the various ways of expressing the powers of 2. In this manuscript, Leibniz outlines elementary arithmetic terms and operations, and when he comes to "powers" he offers this:

Powers. bb is b^2, bbb is b^3, etc. $b^r b^z$ is b^{r+z}; thus, $b^2 b^3$ is b^{2+3}, that is, b^5.

1	2	4	8	16	32	64	128	256
b^0	b^1	b^2	b^3	b^4	b^5	b^6	b^7	b^8
1	b	bb	bbb	b^3b		b^{3-2}	b^{4+3}	
			b^2b	b^2b^2	b^3b^2	b^3b^3	b^4b^2	
1	10	100	1000	10000	100000	1000000	10000000	100000000

(LH 35, 4 11 Bl. 8r)[9]

The three manuscripts surveyed share a number of notable features. First, while binary *notation* may be present in all three, binary *arithmetic* is not; instead, Leibniz performs all his arithmetic in decimal or algebra. Second, binary notation is used for illustration rather than for calculation. These common features suggest that at the time of writing these three manuscripts, Leibniz was not thinking beyond the notation, and this in turn would support the hypothesis that he developed this

[8] Leibniz explicitly acknowledged these features, at least in the 1680s onward. For example: "The very last digits of a number of the double geometric progression can be easily obtained like this: if 1 is subtracted from it, then it is written in binary: etc.1111111, that is, 1 + 2 + 4 + 8 + 16 + 32 etc." (LH 35, 3 B 11 Bl. 10r; cf. LH 35, 8, 30 Bl. 75; LH 35, 13, 3 Bl. 33; LH 35, 15, 5 Bl. 10r; and LH 35, 3 B 5 Bl. 51r).

[9] A similar table is found in other manuscripts, such as LH 35, 4, 11 Bl. 10r, though this was likely written in 1681.

notation as a tool to assist his investigations in mathematical problems that were exercising him at the time.[10] Of course, given the difficulty in dating Leibniz's earliest writings on binary with great accuracy,[11] the hypothesis just sketched about binary's origins must remain speculative. But the three manuscripts are probably among his earliest writings on binary, predating the developed treatment in his earliest dated writing on the subject, "On the Binary Progression," written 15 March 1679 (on the Julian calendar, 25 March 1679 on the Gregorian).

As for when Leibniz invented binary, a precise date is difficult to fix, but the three early manuscripts outlined above were probably written in 1677 or 1678. He was certainly in possession of the idea of a binary progression, i.e., base 2, no later than February 1678, the date he wrote on a manuscript in which he floated the idea of working with various nondecimal number systems, including binary. The manuscript begins with Leibniz giving himself instructions: "Attend to other progressions, such as nines, eights, etc., starting from the lowest—twos, threes, and so on going upwards—for thus only then will a progression be apparent and even the method of confirmation, which otherwise will be difficult to procure from decimal alone" (LH 35, 3 A 24 Bl. 6r).[12] (He subsequently struck this passage out, and the rest of the manuscript concerns the reduction of fractions to decimal numbers.) An earlier date for the invention of binary is possible, since another manuscript containing a table of the numbers 0–8 in binary along with a few simple binary sums has been dated to October 1674 (see Strickland and Lewis 2022, 30). However, the material in binary – written in a different ink from the rest of the contents – was probably added later, though how much later is not known. Hence on the question of dating the invention of binary, two hypotheses may be formed: first, that Leibniz invented binary in 1674 or thereabouts, but did no further work on it or with it for another 3 or 4 years, or second, that he invented it in 1677 or early 1678.

4 The Development of Binary

Although Leibniz probably invented binary notation in response to, and as an illustrative aid for his work on divisibility, primes, and perfect numbers, it did not take him long to develop a full-blown base-2 number system. The evidence suggests this occurred in stages rather than all at once. There are many manuscripts containing rudimentary explorations of binary, such as tables of numbers and simple sums, sometimes scrawled in the margins of manuscripts on nonmathematical topics (for an example of which, see Strickland and Lewis 2022, 4). In his earliest *dated* writing

[10] Another writing in this vein is printed in Strickland and Lewis (2022, 41–43).

[11] Leibniz did not affix a date to them, and the paper contains no watermarks that could be used to determine a dating, so they have to be dated using internal evidence.

[12] Leibniz also refers to "the binary progression, where only ones and 0 express a number," in a manuscript on the construction of a universal language, tentatively dated to February 1678 (A VI 4, 68).

on binary, "On the Binary Progression," written 15/25 March 1679, Leibniz outlined algorithms for conversion between decimal and binary notation and for addition, subtraction, multiplication, and division in binary numerals. He also devised a binary algebra in order to work out the square of a 6-digit number and to extract square roots (Strickland and Lewis 2022, 48–59). Similar ground is covered in later systematic treatments, such as "Essay on a New Science of Numbers" from 1701, and "On Binary" from 1705 (see Strickland and Lewis 2022, 138–144, 206–212). Rather than attempt to plot the steps of Leibniz's increasing facility with binary, I shall focus on his methods for converting decimal to binary along with the algorithms for the four arithmetic operations, his attempts to mechanize binary calculation, and his investigations into the periodicity of arithmetic progressions expressed in binary.

Leibniz's original method for converting decimal numbers to binary was to start with a given number, for example, 47, then successively halve it, at each step noting whether 0 or 1 remains (see Strickland and Lewis 2022, 50). When the process is finished, the digits can be written in a row, where the bottommost digit is the leftmost, and the topmost digit the rightmost, i.e., in the 1s position: 101111.

47	
23	1
11	1
5	1
2	1
1	0
	1

Leibniz later devised a variation of this method, whereby one arranges the successive halves of the given decimal number in columns and then simply writes a 1 underneath an odd number and a 0 underneath an even number to get the binary equivalent:

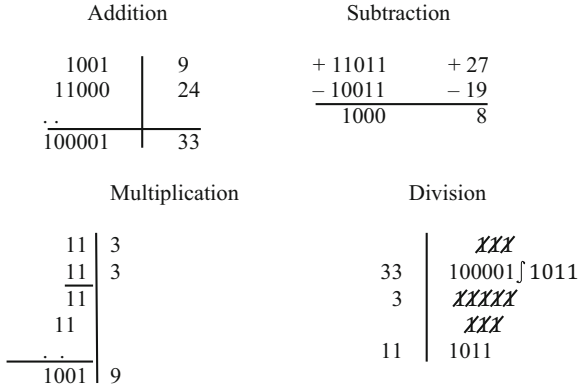
1	2	5	10	21	42	85	171	343	
1	0	1	0	1	0	1	1	1	343

(LBr 705, 69r)

The algorithms Leibniz developed for the four basic arithmetic operations are straightforward. He claimed that to perform these operations, one need know only the following:

> For *addition*, 0 plus 0 is 0. And 0 plus 1 is 1, and 1 plus 1 is 10.
> For *subtraction*, 0 minus 0 is 0, 1 minus 0 is 1, and 1 minus 1 is 0.
> For *multiplication*, only this: 0 times 0 is 0, 0 times 1 is 0, and 1 times 0 is 0, and 1 times 1 is 1.
> For *division*, I have 1 in 1 once, and I have 0 in 0 once. (Strickland and Lewis 2022, 92)

Typical examples given were:



(Strickland and Lewis 2022, 96–97)

Note here that the dots in the examples of multiplication and addition indicate a carry of 1. Note also that the example of division follows the so-called "galley" method, popular in Europe in the seventeenth and eighteenth centuries.

Leibniz later devised a shortcut when adding multiple numbers, such as:

```
    1
   11
  100
 1001
 1011
11100
```

(example simplified from Strickland and Lewis 2022, 140)

The first column on the right has four 1s, but instead of making two carries to the second column as per the standard algorithm, Leibniz realized that one could instead make a single carry to the third column (see Strickland and Lewis 2022, 140, 146, 208–209). When dealing with subtraction, Leibniz also anticipated the two's complement operation, where x – y is computed by determining the two's complement of y, i.e., the number that when added to y gives 0:

> It should also be noted that for the subtrahend number, its complement [up] to 100000 etc. can be placed, and then subtraction is not needed but only addition... As, for example:

The subtrahend is	11 0101
Supposing it should be subtracted from	100 0000
It is evident that this gives	00 1011
For if you add this:	11 0101
It gives us	100 0000

(Strickland and Lewis 2022, 51)

During his early explorations of binary, Leibniz envisaged the mechanization of the four basic arithmetic operations, jotting down ideas for binary-calculating machines of

two different designs. The first was based on his decimal-calculating machine that he had had built in 1673 (though it never worked properly). Instead of being driven by a series of cylinders with nine teeth, as per his decimal machine, the binary version was envisaged as having cylinders with four teeth, each alternately designated as 0 or 1, so that "one revolution will cause two carries" (Strickland and Lewis 2022, 45). Leibniz acknowledged that, with this machine, "the greatest difficulty is to convert a binary number into an ordinary one or vice versa" (Strickland and Lewis 2022, 45), but from his description it is clear that he envisaged conversion going only one way, for while the machine's user would enter the sum in decimal, the result would be given in binary (for more details on this machine's design, see von Mackensen 1974). The second proposed binary machine, conceived as carrying out only the operation of multiplication, is described in "On the Binary Progression." This design is of a sliding box or carriage with shutters on the underside that can be set open or closed, the whole contraption sitting atop various fixed channels representing the columns of a binary number. If a shutter is set open, then a ball is released into the channel below it, and the box is then moved by an unspecified mechanism from one channel to the next, dropping a ball into the channel below if the shutter is open and nothing if not. The process is repeated for each bit of the multiplier, and if a channel accumulates two balls, another unspecified mechanism passes one of these to the next channel on the left while the other is removed (Strickland and Lewis 2022, 52). Although Leibniz's description of this machine lacks detail, and it is doubtful he would have been able to have had it constructed, from his sketchy outline a functional replica was built in 2003/2004 (see Stein et al. 2006).

While Leibniz certainly enjoyed a flirtation with mechanization during his exploration of binary, his true infatuation was with arithmetic and geometric progressions, on account of the recurring periods of the individual digits being significantly shorter in binary than in decimal. Take, for example, the sequence of square numbers: 1, 4, 9, 16, 25, 36, etc. In decimal numeration, the last digit of the successive square numbers recurs periodically: 1, 4, 9, 6, 5, 6, 9, 4, 1, 0; switch to binary numeration and the period of the last digit of successive square numbers is drastically reduced to just two terms: 0, 1 (see below). The other columns are periodic too: Leaving aside the second column, which consists only of 0s, the third column has the period 0010, the fourth 0001 0100, etc. Leibniz's hope was that working with the shorter periods in binary notation would make it easier to determine formation laws for the periods, making it possible to calculate any term of a periodic sequence.

0	0000000
1	0000001
4	0000100
9	0001001
16	0010000
25	0011001
36	0100100
49	0110001

From the late 1670s onward, Leibniz intermittently investigated column periodicity in binary, filling many dozens of manuscript pages, often using algebraic terms to stand for period-sequences or even entire columns. To assist his efforts further, he occasionally devised new notation, though it was often abandoned immediately afterward after it failed to produce the desired breakthrough. In 1699, he introduced the parenthetical term (b) to signify logical complement, calling it "anti-b," "so that if b is 1 then (b) is 0; and if b is 0 then (b) is 1" (LH 35, 3 B 6 Bl. 1r). Two years later, he devised the notation a $\dot{+}$ b, which

> signifies the digit which is produced from the digits a + b added into one in the same column. For example, if a is 0 and b is 0, a $\dot{+}$ b will be 0. If a is 0 and b is 1, or the other way around, a $\dot{+}$ b will be 1; if a is 1 and b is 1, a $\dot{+}$ b will be 0, for a + b is 10, where the 1 is carried into the next column (LH 35, 3 B 5 Bl. 87).

So stated, Leibniz's a $\dot{+}$ b clearly anticipates the boolean exclusive-or operator $\oplus$.

Despite his efforts, the anticipated number-theoretic rewards of investigating column periodicity eluded Leibniz. He did, however, have some successes. In 1698, he observed that in the natural and triangular numbers, the first half of any period is always the complement of the second half, such that if the first half is 0000, then the second half will be 1111 (LH 35, 3 B 5, Bl. 10). And in 1701, he established that if two or more periodic sequences are added together, the resulting sequence will itself be periodic, its period determined by the periods of the two or more generating sequences (Strickland and Lewis 2022, 158–160). When Leibniz began to provide details of the binary system to mathematicians from the late 1690s onward, much of the information he provided was about column periodicity (see A II 3, 450–452; Strickland and Lewis 2002, 111, 122–124, 128–130). His hope was that others would assist in the investigations for which he had little time himself, but there were few takers. The exceptions were Theobald Overbeck (1672–1719), who corresponded with Leibniz about column periodicity from 1709 onward (see LBr 705 Bl. 23–26, 71–73; LH 35, 15, 5 Bl. 10–11), and Pierre Dangicourt (1664–1727), who at Leibniz's behest published an essay entitled "On the Periods of Columns in a Binary-Expressed Sequence of Numbers of an Arithmetic Progression" in 1710 (see Dangicourt 1710).

5 Binary Fractions and Expansions

While the majority of Leibniz's writings on binary focused on integers, he also explored binary fractions and their expansions, especially but not exclusively in the late 1670s and early 1680s. Leibniz quickly discovered that just like decimal fractions, binary fractions are either terminating, i.e., expanding to a finite number of digits, or repeating or recurrent, i.e., expanding to infinity, with a finite set of digits (the repetend, or period) repeating forever. For example:

Decimal fraction	Binary equivalent	Binary expansion
$\frac{1}{2}$	$\frac{1}{10}$	0.1
$\frac{1}{3}$	$\frac{1}{111}$	$0.\overline{01}$
$\frac{1}{4}$	$\frac{1}{100}$	0.01
$\frac{1}{5}$	$\frac{1}{101}$	$0.\overline{0011}$

Leibniz's preferred method for working out binary expansions was the galley method (mentioned above in relation to division in integers), but this often proved challenging. In one manuscript (LH 35, 13 2 B Bl. 155), Leibniz focuses on the binary expansions of the decimal fractions $\frac{1}{13}$ and $\frac{1}{17}$. While he gets the correct result – $0.\overline{000100111011}$ – for the former, for the latter he first gets the incorrect result 0.0000111 (which is in fact the binary expansion for $\frac{7}{128}$) before arriving at the correct $0.\overline{00001111}$ on the second attempt. In another manuscript, two out of seven attempted binary expansions arrive at the wrong result: for the expansion of $\frac{1}{6}$ Leibniz gets 0.001 (which is the expansion for $\frac{1}{8}$), while for $\frac{1}{9}$ he gets 0.00101, which is the expansion for $\frac{5}{32}$ (see Strickland and Lewis 2022, 62). Such mistakes are common in Leibniz's use of the galley method to determine binary expansions, and perhaps for this reason he developed various other methods, of which I shall mention two.[13]

The first is found in a manuscript from c.1679, where Leibniz describes an algorithm for finding the repetend in the binary expansion of those reciprocals whose denominator is odd, and this with "no need for actual division":

Take, for example, the fraction $\frac{1}{7}$, expressed in binary as $\frac{1}{111}$.
Add 1 to the numerator = 1001.
Take half of this = 100; this number serves as the "constant multiplier."
The last digit of the period is always 1, so write 1.
Multiply this by the constant multiplier: $100 \times 1 = 100$. Write 0 as the penultimate digit of the period, and carry the remaining 10.
Multiply the last digit obtained – 0 – by the constant multiplier: $100 \times 0 = 0$. To this, add the carry of 10: $0 + 10 = 10$. Write 0 as the antepenultimate digit of the period, and carry the remaining 1.
Multiply the last digit obtained – 0 – by the constant multiplier: $100 \times 0 = 0$. To this, add the carry of 1: $0 + 1 = 1$. So write 1. Since the remainder is now 0, the digit just obtained is the last one of the next iteration of the repetend (see LH 35, 3 B 3 Bl. 12r).

As the algorithm reveals that the period is 001, the binary expansion of $\frac{1}{111}$ is $0.\overline{001}$. After obtaining this result, Leibniz seeks to confirm it using the galley method of division, but as if trying to demonstrate the need for the alternative method just sketched, his calculation goes awry and he comes up with $0.\overline{100}$ instead.

[13] A third is outlined in Strickland and Lewis (2022, 40n2).

A second method for working out binary expansions comes from a manuscript written in 1680 (LH 35, 3 B, 5 Bl. 89). Here, Leibniz seeks to multiply known binary expansions to determine a new one. Beginning by multiplying the binary expansions of $\frac{1}{3}$ and $\frac{1}{5}$ to determine the binary expansion of $\frac{1}{15}$, Leibniz multiplies $0.\overline{0011}$ by $0.\overline{01}$ and ends up with $0.\overline{00001111}$, triumphantly proclaiming "Therefore, two periodic sequences have been added into one," before realizing the calculation has gone astray (as $\frac{1}{15}$ in fact expands to $0.\overline{0001}$). In a second attempt, Leibniz first computes the binary expansion of $\frac{1}{9}$, namely, $0.\overline{000111}$, then tries to reach the same result by multiplying the binary expansion of $\frac{1}{3}$ by itself. This time, he gets the correct result, stopping at 0.000111000111.

Some of Leibniz's earliest explorations with binary fractions and expansions concern irrationals. In "On the Binary Progression," he computes the first digits of $\sqrt{2_{10}}$, getting the result $1.011010100..._2$. In order to confirm his result, Leibniz then multiplies that number by itself, obtaining the value 01.111111111110100, noting that as the infinite numeral [0].1111111... represents 1, so 1.1111111 represents 2 (Strickland and Lewis 2022, 58). In another manuscript, probably from 1680, he repeats the process, this time obtaining several more digits of $\sqrt{2_{10}}$: 1.0110111101 (LH 35, 8, 30 Bl. 125v). The aim of such exercises was not to find a periodic sequence of digits – Leibniz was well aware that irrationals did not have such[14] – but to show that irrationals could be "determined as far as possible" through binary (Strickland and Lewis 2022, 54). His hope was to find "laws of sequences which not only seem not to have a period but also really do not have one. Such are those which correspond to irrationals. For not all laws of sequences are periods" (LH 35, 13 3 Bl. 33r).

In a similar vein, Leibniz hoped that binary could offer insight into transcendental quantities, most notably the alternating converging series that now bears his name: $\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \frac{1}{13} - \frac{1}{15} + \frac{1}{17} - \frac{1}{19}$, etc. The series, discovered in 1674, was the key result of his investigation into the classic problem of squaring the circle, that is, determining the area of a circle enclosed by a 1×1 square. After discovering the series in 1674, Leibniz for a time entertained hopes that it might prove to have a finite sum (as some infinite series of reciprocals do, for example, that of triangular numbers $\frac{1}{1} + \frac{1}{3} + \frac{1}{6} + \frac{1}{10} + \frac{1}{15}$, etc. sums to 2), but these had dissipated by the time he had invented binary. Nevertheless, he believed that investigating the periodicity of these fractions in binary could yield a rule that would enable one to continue the expansion of the digits of π: "Having obtained the period from a given number, one will also obtain the method of expressing by a certain rule the sequence of digits expressing the size of a circle, that is, $\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9}$ etc." (Strickland and Lewis 2002, 59). As such a rule eluded him, in August 1680 Leibniz tried out a different approach in a manuscript entitled "Attempted Expression of the Circle in Binary Progression" (Strickland and Lewis 2022, 61–62). Here, Leibniz considers the powers of 2^{-1} geometric sequence, namely, 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{64}$, etc. (unhelpfully

[14] Although in one early writing he exclaims that binary will be "remarkable for periodic progressions in expressible quantities which are not whole or rational" (Strickland and Lewis 2022, 32).

also called "binary progression"), comparing various terms of this sequence with sequences of terms from his circle formula. However, a calculation error early in the process led him to incorrectly conclude that $\frac{1}{2}+\frac{1}{4}$ is greater than the area of the circle (which, if true, would make π less than 3). Failing to spot the mistake, which subsequently gets compounded, Leibniz eventually gave up.[15]

In the last twenty years of his life, Leibniz outlined binary notation and arithmetic to many of his mathematician correspondents, but he was not similarly forthcoming about his work on binary fractions and expansions, electing to keep this to himself. However, he did suggest to a few correspondents the idea of exploring π using binary. In 1704, Leibniz informed Jakob Hermann (1678–1733) of his hope that a rule would be found to enable the digits of π to be computed, describing this as something "we will obtain more easily through binary" (GM IV 266), though he did not elaborate on this remark. The following year, he made the same suggestion to Jacob Bernoulli (1655–1705), who took the bait and presented Leibniz with a 118-digit binary representation of the first 36 decimal digits of π:

$$11110010000001001111110110011000110110110110011000010011110010010110010101101101101010001101110010101000000011110100000_0^1$$

There are a number of problems with this result. First, Bernoulli, apparently not realizing the significance of the radix point, simply converted the integer 31415926535897932384626433832795028 into binary, rather than the fraction $\frac{31415926535897932384626433832795028}{10000000000000000000000000000000000}$, so he did not express π in binary at all. Second, his 118-digit binary number is actually equal to the integer 31415926535897932384626433832792980 8, which is short by 20480, indicating that he dropped some bits in his calculation (for more details, see Knuth 1983, 238). Bernoulli did at least get one thing right, noting of his binary number that "no period of digits is apparent here, nor any law of the progression" (GM III/1, 97). In his response, which made no mention of Bernoulli's errors, Leibniz clung to the view that the law of progression of π was simply hidden and probably more easy to detect in binary than in decimal (GM III/1, 101).

6 Binary and Philosophy

Leibniz's faith in the illuminatory power of binary is nowhere more evident than in his use of it in philosophical theology. In 1694/1695, when sketching notes on one of Weigel's books, Leibniz devised an analogy between the representation of all numbers by 1 and 0 and the theologically orthodox idea of creation *ex nihilo*, that is, the creation of all things out of nothing by God. Treating God as the analogue of 1 and nothingness as the analogue of 0, Leibniz took the origination of all numbers

[15] For further details of Leibniz's work on the binary expression of his circle formula, see Strickland (2023a).

from 1 and 0 in binary as a reflection of the doctrine that all created things originate from God and nothingness. This analogy would play a pivotal role in Leibniz's willingness to inform correspondents about binary from 1696 onward (and as we shall see in Sect. 6, this in turn played a pivotal role in his decision to publish an essay on binary in 1705).

To understand Leibniz's binary-creation analogy, note that in his metaphysics, all created beings contain the same perfections as God (namely, power, wisdom, and goodness), albeit not to the same degree, because, as created beings, they have limitations or privations. In this sense, all created beings originate both from God, the source or principle of their perfections, and from nothingness, the source or principle of their privations or imperfections. Hence:

> According to the binary [progression], all numbers are expressed by the digits 0 and 1 alone, by unity and nothing, a remarkable analogy of the origin of created things from God and nothing, with creatures having their own perfections from pure, positive actuality, or God, and their imperfections or limits from the negative, or nothing. (Strickland and Lewis 2022, 85)

Such remarks can easily be read as saying that created things are literally composed of two distinct but seemingly mutually exclusive elements, such that it is difficult to see how they could coexist in one and the same being. Leibniz clarifies:

> there are in him [God] three primacies: power, knowledge and will; the result of these is the operation or creature, which is varied according to the different combinations of unity and zero; or rather of the positive with the privative, for the privative is nothing other than limits, and there are limits everywhere in a creature, just as there are points everywhere in the line. However, a creature is something more than limits, because it has received some perfection or power from God, just as the line is more than points. For ultimately the point (the end of the line) is nothing more than the negation of the progress beyond which it ends. (Leibniz 2006, 39)[16]

Leibniz thus envisaged the creation of something from nothing as involving God divesting a finite portion of his essence to the creature and then *stopping*, the stopping point marking the limit of the creature. As such, Leibniz should not be understood as claiming that nothingness is an element of the creation process; rather, the reference to nothingness is a poetic way of referring to a limit, the idea that a creature contains some of God's essence but not all. A created thing should not be construed as literally a mixture of God's essence and nothingness, but rather as a finite instantiation (or dilution) of the former.[17]

[16] Also: "Without doubt boundaries or limits are of the essence of creatures, but limits are something privative and consist in the denial of further progress." Leibniz (2006, 38), cf. A I 15, 369; Leibniz (2006, 102–103); and Strickland (2014, 22).

[17] Or as he puts it in a manuscript written perhaps in 1695 or possibly in 1702, "things are educed from God's active power—as Julius Scaliger said—rather than from passive nothing. My invention of the composition of numbers from 0 and 1 would wonderfully support this doctrine" (LH 4, 3, 5e Bl. 5). See Scaliger (1557, fols. 16–17).

Clarifying the analogy in this way does not prevent it from being inapt, however. In binary numeration, all numbers are composed from two elements (0 and 1) by compounding them in ever longer strings, whereas Leibniz envisaged all created things arising from God through an entirely different process, in which God divests some of his own essence to a creature, so that in terms of essence, or perfection, each creature is greater than nothingness but less than God, i.e., it has the same qualities as God but not to the same degree as God. Because of this, it would make more sense to say that all created things are analogues of binary fractions, each having a unique expression of their perfection and limitations in the same way that all fractional numbers between 0 and 1 are uniquely expressed in binary. Leibniz did not frame his analogy in this way, however.

Once the binary-creation analogy had been devised, Leibniz quickly recognized its philosophical value. It next appeared in December 1695, in the middle of a brief sketch of a planned book entitled *Theodicy*, envisaged as a defense of God's justice in response to the world's evil. The sketch reveals Leibniz's belief that the binary-creation analogy could help explain why created beings are liable to sin, i.e., because, by virtue of being limited, they fall short in the perfections of power, wisdom, and goodness:

> New studies: concerning the original imperfection of creatures before sin, from which the cause of sin is to be sought. All things are good but nothing is absolutely perfect. Concerning the composition of things from perfection and limitation, i.e. from action and privation, and concerning the origin of things from God and nothing, in which concerns the origin of numbers from 1 and 0, i.e. from unity and nothing. (LH 1, 6, 2 Bl. 20r/http://www.leibniz-translations.com/theodicaea.htm)[18]

With the binary-creation analogy now envisaged as a key part of his philosophical theology, Leibniz sought to promulgate both it and the number system on which it was based. In early May 1696, he discussed it with Rudolph August, Duke of Brunswick and Lüneburg (1627–1704). The duke's warm reception of the analogy prompted Leibniz to pitch the idea of commemorating it on a medal, to be struck at the duke's command. Various designs were sketched (for which see Strickland and Lewis 2022, 103), but the duke decided to commemorate the binary-creation analogy in a different way: by commissioning three wax seals (for which see Strickland and Lewis 2022, 106).

Believing that the binary-creation analogy had the potential to make the Christian doctrine of creation *ex nihilo* intelligible and even plausible to non-Christians, Leibniz communicated it – with attendant details of the mathematics of binary – to two Jesuit missionaries in China, Claudio Filippo Grimaldi (1638–1712) in 1697 and Joachim Bouvet in 1701 (see Strickland and Lewis 2022, 108–115, 126–134). While the former did not respond, the latter famously suggested a correlation between Leibniz's binary numeration and the sixty-four hexagrams of the ancient Chinese

[18] When Leibniz did come to write the *Theodicy* more than a decade later, the binary-creation analogy was not mentioned. The *Theodicy* was published in 1710. See Leibniz (1985).

divinatory text, the Yijing (or Book of Changes), a suggestion that would prompt Leibniz to go public with his invention.

7 Binary and The Yijing

The Yijing consists of sixty-four individually named hexagrams, each accompanied by a short, enigmatic essay. The hexagrams themselves consist of six stacked horizontal lines, each line being either unbroken (representing yang) or broken (representing yin). From these two elements, arranged in all possible groupings of six, emerge the sixty-four hexagrams. Bouvet had been a student of the Yijing for many years before he learned of Leibniz's binary arithmetic, and had already suspected that it somehow encapsulated "a numerary metaphysics" that, when properly deciphered, would reveal the perfected science of the ancient Chinese (A I 20, 573). Upon learning of Leibniz's binary arithmetic, Bouvet hypothesized that the hexagrams had been intended as a binary number system.

To see what Bouvet saw, and thus to understand his hypothesis, we need to consider the hexagram-sequence he sent to Leibniz. This is the so-called "Fuxi sequence," named after the mythical Chinese philosopher-emperor traditionally believed to be the author of the Yijing. The first sixteen hexagrams of the Fuxi sequence are:

䷀ ䷁ ䷂ ䷃ ䷄ ䷅ ䷆ ䷇ ䷈ ䷉ ䷊ ䷋ ䷌ ䷍ ䷎ ䷏

Let us start with the component lines of each hexagram. Bouvet equated – albeit without justification – the broken line with 0 and the unbroken line with 1. If we replace the lines with binary digits in the way Bouvet suggested, the hexagrams are converted thus:

000000 000001 000010 000011 000100 000101 000110 000111 001000 001001 001010 001011 001100 001101 001110 001111

Rotating the digits 90° clockwise gives us the decimal numbers 0–15 in binary notation. From this, it is clear that Bouvet read each hexagram from the bottom line upward, such that he took the bottom line to represent the leftmost digit (32's place) of a six-digit number in binary, the next line up to represent the leftmost-but-one digit (16's place) and so on, such that the hexagram ䷆ would be equivalent to 000110 in binary (6 in decimal).

Various problems have been noted both with Bouvet's approach to reading the hexagrams and with his hypothesis more broadly (see Zacher 1973, 112; Zhonglian 2000; Swetz 2003, 288–289; Sypniewski 2005). Here it will be sufficient to mention just one, namely, that the sequence of hexagrams upon which Bouvet based his hypothesis – the so-called "Fuxi sequence" – is a relatively modern rearrangement (by the eleventh century philosopher Shao Yong) of a much older and more

traditional sequence, namely the King Wen sequence. Bouvet neglected to mention this to Leibniz, though as a long-time student of the Yijing, he would at least have been aware of the King Wen sequence. Bouvet's omission is problematic because while the arrangement of hexagrams in the Fuxi sequence may support the idea that the hexagrams were intended as a binary number system, that of the King Wen sequence does not. The first sixteen hexagrams of the King Wen sequence are:

䷀ ䷁ ䷂ ䷃ ䷄ ䷅ ䷆ ䷇ ䷈ ䷉ ䷊ ䷋ ䷌ ䷍ ䷎ ䷏

Here, there is no prospect of seeing an orderly sequence of numbers expressed in binary. If we again take an unbroken line to be 1 and an unbroken line to be 0, and thus treat each hexagram as representing a number expressed in binary, the order of numbers in the King Wen sequence would be 63, 0, 34, 17, 58, 23, 16, 2, 59, 55, 56, 7, 47, 61, 8, and 4. Little wonder that Leibniz, seduced by Bouvet's hypothesis of a correlation with binary arithmetic, concluded that the King Wen sequence – which he had seen in a Latin book about Confucius (namely, Intorcetta et al 1689, xliv) – was a scrambling of the true order (Strickland and Lewis 2022, 187).[19]

Leibniz was so intoxicated by the idea of his binary system being the key to deciphering an ancient Chinese enigma that he related Bouvet's hypothesis to many of his correspondents, even informing Pope Clement XI (see A I 23, 353). Bouvet's hypothesis also prompted Leibniz to go public with the binary system at long last. Days after receiving Bouvet's letter, Leibniz wrote a short paper, "Explanation of Binary Arithmetic," in which he outlined binary numeration and arithmetic as well as Bouvet's hypothesis (Strickland and Lewis 2022, 195–197).[20] The paper was read at the Académie Royale des Sciences in Paris, and published in its annual periodical for 1703 (though the volume did not appear until 1705). This paper was the only one of Leibniz's hundreds of writings on binary to be published in his lifetime.

8 Sedecimal

Leibniz's "Explanation of binary arithmetic" is also notable for containing the first published reference to base 16, though this is often overlooked. Leibniz's experiments with base 16 in fact began in 1679, giving him clear priority over the English schoolmaster Thomas Wright Hill (1763–1851) and the Swedish-American engineer John Nystrom (1825–1885), both of whom independently reinvented it in the nineteenth century (see Hill 1850, Nystrom 1862, 1863).[21]

[19] There are other sequences of hexagrams apart from the two mentioned already, such as the Eight Palaces sequence, which predates the Fuxi sequence. These sequences have elicited many studies of their structure, interrelation, and history, though such matters are outside of the scope of this chapter.

[20] For more information on Bouvet's hypothesis and Leibniz's reaction thereto, see Kempe (2022, 141–174).

[21] For more information about the history of base 16, and Leibniz's place therein, see Strickland and Jones (2023).

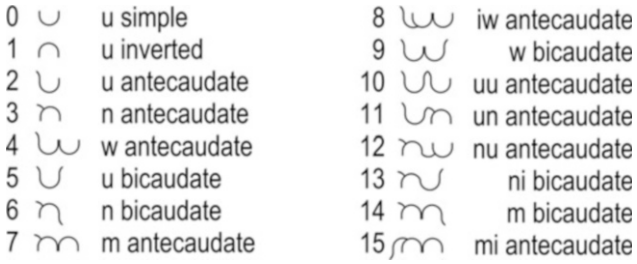
Initially, Leibniz's term for base 16 was "sedecimal," though on occasion he called it "sedenary" (the former means "sixteen," the latter "sixteen-fold," or "in sixteens"). He devised no fewer than six different forms of notation for sedecimal. In his first writing on the subject (Strickland and Lewis 2022, 64–68), he used the Roman letters m, n, p, q, r, and s for the six extra digits, before abandoning them in favor of the six Aretinian syllables ut, re, mi, fa, sol, and la, abbreviated to u, r, m, f, s, and ℓ. By combining these syllables with the German words for numbers, he created an entirely new set of terms for values expressed in sedecimal. For example, "utzwanzig," a combination of the syllable "ut," standing for 10, and the German word "zwanzig," which traditionally meant 20 but was repurposed by Leibniz to stand for 32, was Leibniz's term for the decimal number 42 expressed sedecimally.

Leibniz's later experiments with sedecimal notation were more elaborate. In one manuscript, he stacked dots and dashes, using a dot for each 0 bit and a dash for each 1 bit, with the most significant bit at the top and the least significant bit at the bottom:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
•	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	– •
		•	–	•	•	–	–	•	•	•	•	–	–	–	–	
				•	–	•	–	•	•	–	–	•	•	–	–	
								•	–	•	–	•	–	•	–	

(Strickland and Lewis 2022, 70)

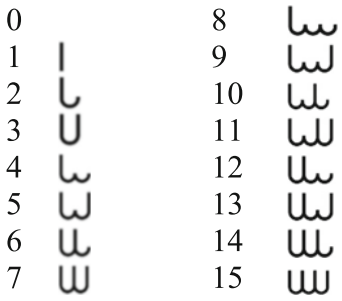
A further set of novel sedecimal digits were scrawled on the back of an envelope c.1682. This time, he starts with a concave-up semicircle to denote 0, a concave-down semicircle to denote 1, and then modifies these characters to create the remaining digits:



(Strickland and Lewis 2022, 78)

The accompanying descriptions (or instructions) indicate that Leibniz based each digit on Roman letters, either single or joined-up, with some of them being "bicaudate," that is, extended by a tail on either side, and others "antecaudate," that is, extended by a tail on the left-hand side alone.

Almost twenty years later, Leibniz devised yet another form of sedecimal notation, with the characters loosely based on the Roman letters i, u, and w, some extended with one or two right-hand tails:



(Strickland and Lewis 2022, 130)

Leibniz's sixth and final form of notation for sedecimal appears in a manuscript featuring a table of decimal numbers 0–40 represented in every base from 2 to 16 (LH 35, 3 B 11 Bl. 11v). Here, Leibniz uses the Arabic numerals 0–9 and lower-case Roman letters a, b, c, d, e, and f for the six extra digits, anticipating the modern convention, which employs the same six letters, usually in upper-case.

Shortly after having invented base 16, Leibniz arrived at the view that while binary is for theory, sedecimal is for practice (Strickland and Lewis 2022, 70). Thereafter, he occasionally suggested that sedecimal would be the ideal choice if practice were to be changed. To justify this, he used what may be termed an argument from economy, namely, that as values expressed in sedecimal are shorter than those expressed in decimal, sedecimal is more convenient. The argument is illustrated only in his very first writing on sedecimal, where he converts the decimal number 1679 (representing the date of writing) into the sedecimal number 68ℓ (equivalent to 68F in modern notation), leading him to conclude: "From this it is apparent that sedecimal expressions are much shorter" (Strickland and Lewis 2022, 64). Subsequent conversions, such as 10000 to 2710_{16} and 100000 to $186u0_{16}$, secure the point that sedecimal expressions tend to be shorter than decimal, though Leibniz's conclusion that they are "much shorter" is difficult to draw from the examples he gave, each of which compact the expression by a single digit.[22] Nevertheless, he continued to use the argument from economy to support the practical advantage of sedecimal, most notably in his published essay "Explanation of Binary Arithmetic," where he noted in passing that decimal has an advantage over binary because its numbers are not as long, but "if we were accustomed to proceed by twelves or sixteens, there would be even more benefit" (Strickland and Lewis 2022, 196). Hill and Nystrom, later advocates of base 16, also used the argument from economy, but both believed the chief advantage of base 16 lay in the fact that 16 allowed for more convenient bisections of the sort that were routinely required in shops and trade, an argument Leibniz himself did not articulate (though he could not have been unaware of it, since in the passage from "Thesaurus mathematicus" quoted in Sect. 2 above, he noted that an advantage of duodecimal is that it allows for divisions by 2, 3, 4, and 6). Moreover, while these later advocates of base

[22] The same conversion of 100000 to $186u0_{16}$ is also found in a contemporaneous text: LH 35, 8, 30 Bl. 148r.

16 envisaged and argued for a wholesale switch away from decimal, with weights, measures, coinage, and (in the case of Nystrom) even the day and the year being divided into 16 units, Leibniz's own preference for sedecimal did not extend beyond wistful thinking about reforming "practice" without indicating how far this should go. Consequently, his advocacy of base 16 was neither forthright nor vociferous, as Hill and Nystrom's would be, and Leibniz conceded that since decimal was so widely accepted, we should be satisfied with it (Strickland and Lewis 2022, 85).

9 Other Nondecimal Number Systems

Leibniz's flirtation with sedecimal shows that his engagement with nondecimal number systems was not exhausted by binary, even if that was his predominant focus. But in contrast to his writings on binary, Leibniz's writings on other nondecimal systems are occasional, scattered, and unsystematic. I have already mentioned Leibniz's 1703 sketch of table of decimal numbers 0–40 represented in every base from 2 to 16 (LH 35, 3 B 11 Bl. 11v). The table may have been intended as a ready reckoner of sorts, and no doubt came in useful as such when in the same year he sought to apply his formulation of Fermat's little theorem, $2^{n-1} - 1$, in various other number bases, such as the septenary, octonary, and nonary (LH 35, 3 B 11 Bl. 9–10). However, such ventures outside of decimal and binary were rare, and aside from binary (and to a lesser extent, sedecimal), most of Leibniz's other references to nondecimal number systems were to bases 4 and 12.

References to base 4 start to appear in Leibniz's writings from 1683 when he read Weigel's *Tetractyn*. Leibniz made brief reading notes on Weigel's book; these consist mostly of verbatim reproductions of Weigel's reasons for thinking four the most natural base but conclude with a cryptic assessment that refers to bases 2, 3, and 8: "I think the binary is best absolutely, the ternary in planes, and the octonary in solids" (A VI 4, 1163). In critical reading notes on another of Weigel's books in which the quaternary system is outlined, Leibniz offers this assessment:

> As far as quaternary arithmetic is concerned, I think that if anything were to be changed in practice, it would be to use the duodecimal or sedecimal instead of the decimal, for the larger the numbers used by a progression, the more convenient the calculation... As for theory and the discovery of extraordinary truths in arithmetic, which may likewise be very useful for practice, I think that not only should the quaternary be preferred to the decimal, but also that preferable to the quaternary in turn is the binary, which is the most perfect of all, nor does it suppose anything but instead completely resolves numbers. (Strickland and Lewis 2022, 85)

This is as close as Leibniz came to a comparison of the relative merits of the binary and quaternary systems. Thereafter, he often mentioned Weigel's advocacy of base 4, but invariably in passing when writing about binary (see Zacher 1973, 353; Strickland and Lewis 2022, 110). It would be fair to say that he had little interest in the quaternary system itself, opting to mention it only as a pretext for discussing and extolling the merits of binary.

To some extent, the same is true of the duodecimal system which, like Weigel's quaternary, is often mentioned in passing in the preamble when Leibniz outlines binary

to a correspondent for the first time (see Zacher 1973, 353; Strickland and Lewis 2022, 110). However, Leibniz's engagement with duodecimal goes back further and is occasionally deeper. The earliest mention of duodecimal in Leibniz's writings is in a preface he wrote in 1670 to an edition of the writings of the Italian humanist Marius Nizolius (1498–1576). There he considers the view of those who would have it that truth depends upon the definitions of terms, and definitions of terms in turn upon the human mind (in other words, that truth is arbitrary). To this, Leibniz (1969, 128) responds: "In arithmetic, and in other disciplines as well, truths remain the same even if notations are changed, and it does not matter whether a decimal or a duodecimal number system is used." Unfortunately, Leibniz does not reveal where he had learned about duodecimal, though his use of it suggests he thought it was sufficiently well known that educated readers would be able to understand his argument.

Other early writings hint at explorations of duodecimal. In a manuscript written around 1678, Leibniz noted that if the duodecimal system were used, the arithmetic checking method known as casting out nines "could become the proof by casting out elevens" (LH 35, 4, 13 Bl. 21). In the spring of 1680, Leibniz met the Dutch mathematician Johann Jakob Ferguson (1630–1706) and must have mentioned both binary and duodecimal to him, as in the scratch paper upon which both recorded their ideas, Leibniz wrote out a table showing the values of the decimal numbers 0–8 in binary, and a set of duodecimal digits, with the two extra digits given as ~ and $ (A III 3, 137). (In passing, it should be noted that on the same scratch paper Leibniz also worked out $\frac{1}{2}$ in base 11: "$\frac{1}{2} = \frac{555\ 555\ \text{etc.}}{1\ 000\ 000}$"; A III 3, 138). And in August of 1680, Leibniz noted: "It is well known that all fractions can be expressed by an infinite sequence of integers of a certain progression, for example, the decimal, or even the duodecimal, or the one I prefer overall, the binary" (LH 35, 13, 3 Bl. 33). The remainder of this manuscript is concerned with binary fractions, but given the confidence of his remark here, it is likely that he had already undertaken some investigation of duodecimal fractions. If he had, his work on that remains to be discovered.

Leibniz often claimed that some people had preferred duodecimal to decimal, but it is not certain who he had in mind. In a manuscript from 1712, he claimed that a proponent of duodecimal had been identified by the German mathematician Daniel Schwenter (1585–1636): "In the German *Deliciae mathematicae*, someone is reported to have given preference to the duodecimal progression, in which eleven digits will be needed, namely 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, δ, ε, where δ is 10 and ε is 11" (LBr 705 Bl. 93r).[23] Having identified an unnamed proponent of duodecimal, Leibniz proceeds to describe algorithms for converting decimal numbers into duodecimal:

> Lay out the twelvefolds: 1, 12, 144, 1728, 20736 etc. Now, the aforementioned number, 1712,[24] is to be expressed duodecimally. The largest of the twelvefolds lower than 1712 itself, namely 144, is subtracted as many times as possible, that is, it [i.e. the number 1712] is

[23] Leibniz is probably thinking of Schwenter's *Deliciae physico-mathematicae* [The Charms of Physico-Mathematics] of 1636, which was posthumously revised and expanded by Georg Philipp Harsdörffer (1607–1658) in 1651 and again in 1653. However, as far as I have been able to tell, in none of those works is there any mention of duodecimal, let alone any report of anyone endorsing it.

[24] Leibniz actually wrote "1728," but this is clearly a mistake as his example uses 1712.

> divided by 144; the quotient will be 11, or ε. The remainder is 128, and that is divided by the next lowest twelvefold, 12, and the quotient will be 11 with a remainder of 6. And 1712 will be εε6 = 6 + ε ×12 + ε. However, if anyone wants to accomplish this without the list of twelvefolds, divide the number 1712 by 12, then divide the quotient, 142, by 12; the quotient, 11, cannot likewise be divided by 12, therefore the largest twelvefold is 12 × 12 or 144, by which 1712 is divided to produce 11 with a remainder of 128. Divide this [remainder] in turn by 12 to produce 11 with a remainder of 6. This [remainder] cannot be divided by 12, therefore 1712 = 6 + 11 × 12 + 11 × 144 = εε6 (LBr 705 Bl. 93r).[25]

This passage represents Leibniz's deepest engagement with duodecimal aside from a single manuscript exclusively devoted to it, probably written in 1695 (LH 35, 12 1 Bl. 40r/Strickland 2022). That manuscript begins with the observation that, in the decimal system, the digital root (that is, the digit sum) of multiples of nine is always nine, e.g., 9 × 3 = 27, and 2 + 7 = 9. Leibniz then generalizes this to any number base, or "progression," supposing that for any base n, the digital root of multiples of $n - 1$ is always $n - 1$. He illustrates this in duodecimal, showing that the digital root of any multiple of 11 is always 11, or rather, since he here uses the Greek letters χ and φ for 10 and 11 respectively, any multiple of φ is always φ. To secure the point, Leibniz draws a large table of duodecimal numbers in which his "new notation" for duodecimal is shown alongside the "old meaning" (i.e., the decimal equivalents).

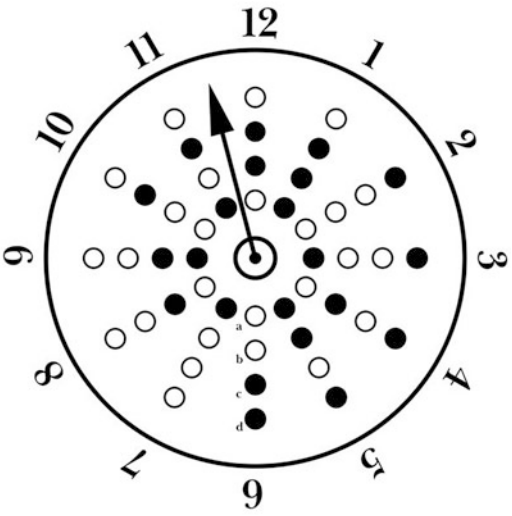
10 Conclusion

The depth as well as the breadth of Leibniz's explorations of nondecimal number systems is quite remarkable in itself, and even more so against the backdrop of a widespread lack of interest in – and awareness of – such systems among the mathematicians of his time. This backdrop no doubt contributed to Leibniz's unwillingness to go public with many of the ideas discussed in this chapter, most of which were groundbreaking at the time. However, it is worth noting that he saw fit to publish only a tiny fraction of what he wrote in any field; his writings on nondecimal number systems were no exception in that regard.

Leibniz never gave up his hope that binary would help unlock the innermost secrets of numbers. Success eluded his efforts. Yet while he cleaved fast to the belief that binary would eventually prove to be an invaluable tool for number-theoretic investigations, he did not dismiss the possibility of practical applications altogether. Indeed, in 1716, the last year of his life, he hit upon an innovative application of binary that has remained unknown till now. In a single-page manuscript (LBr 916 Bl. 44r), Leibniz sketched out a plan for a tactile clock based on binary reckoning that would enable one to tell the time at night. His illustration shows a single hour hand that moves around the clock face. Between each numeral and the center is a row of four points, each of which is either hollow (represented by an empty circle) or bumpy

[25] After this example, Leibniz turned to the quaternary system, efficiently sketching out the steps required to convert the decimal value 1712 to 122300_4 (LBr 705 Bl. 93r). He then did the same with binary, before turning to the periodicity of columns in binary.

(represented by a filled circle). Hollow points denote 0, and bumpy points denote 1 if they occur in the first position, 2 in the second, 4 in the third, and 8 in the fourth, with positions being counted from the center outward to the edge. The idea is that you feel for where the hand points, then trace your finger from the tip of the hour hand down toward the center and then further down along the same line until you encounter a row of four points, which are then read by touch. Hence in the diagram, a is 0, b is 0, c is 1, and d is 1, and since c and d are in the third and fourth positions, they denote 4 and 8, respectively. Add them together, and one knows that the time is 12 o' clock.



A casual glance at the clock face will reveal a mismatch between numerals and binary representations, as the row of holes and bumps leading from each numeral to the center contains the binary representation not of that numeral, but of the one on the opposite side of the clock (e.g., the row between the numeral 7 and the center is ooo•, which is the binary representation for 1, and vice versa). The reason for this, of course, is that the clock face is not meant to be *seen*, only touched – the row of dots one reads is the one directly opposite where the hour hand happens to be pointing, not the one that would be underneath the hand (which would not be accessible to touch anyway, because the hour hand would be over it).

Leibniz's tactile clock is as simple as it is ingenious, and the idea is all the more remarkable because it precedes the cognate inventions of night writing and Braille by more than a century.[26] In the manuscript, Leibniz says nothing about the clock's mechanism, but this would have been easy enough for a horologer of the day to construct.[27] Yet Leibniz appears not to have commissioned the clock, despite its

[26] For further details of Leibniz's tactile binary clock, and an English translation of the manuscript, see Strickland (2023b).

[27] Leibniz himself had devised a spring-driven pocket watch in the mid-1670s and continued to improve it even four decades later; see Leibniz 1675, and LH 38 Bl. 274–275.

potential to benefit not only those in ill-lit rooms at night but also the blind and visually impaired. Accordingly, his tactile binary clock is much like the majority of his other ideas about binary and other nondecimal number systems, in that it was destined only for the admiration of future generations rather than the benefit of his own.

Acknowledgments The author would like to express his gratitude to the Gerda Henkel Stiftung, Düsseldorf, for their award of a research scholarship (AZ 46/V/21), which made this chapter possible. The author would also like to thank Daniel J. Cook, Donald E. Knuth, and Harry R. Lewis for their helpful comments on an earlier version of this chapter; Siegmund Probst and Julia Weckend for their assistance in transcribing manuscripts and their illuminating discussion thereof; and Lisa Silva – @lisa_cs8 – for her illustration of Leibniz's tactile clock.

Bibliography

Bacon F (1623) De dignitate et augmentis scientiarum, libri IX. Haviland, London

Caramuel y Lobkowitz, Juan (1670) Mathesis biceps, vetus et nova. Annison, Campagna

Couturat L (1901) La logique de Leibniz. Félix Alcan, Paris

Dangicourt P (1710) De periodis columnarum in serie numerorum progressionis arithmeticae dyadice expressorum. Miscellanea Berolinensia 1:336–376

Glaser A (1981) History of binary and other nondecimal numeration, 2nd edn. Tomash, Los Angeles

Harsdörffer GP (1651) Delitiae mathematicae et physicae der Mathematischen und Philosophischen Erquickstunden. Zweiter Theil. Dümler, Nuremberg

Harsdörffer GP (1653) Delitiae mathematicae et physicae der Mathematischen und Philosophischen Erquickstunden. Dritter Theil. Endters, Nuremberg

Hill TW (1860) Selections from the papers of the Late Thomas Wright Hill, Esq. F.R.A.S. John W. Parker and Son, London

Hochstetter E (1966) Herrn von Leibniz' Rechnung mit Null und Eins. Siemens, Berlin

Ineichen R (2008) Leibniz, Caramuel, Harriot und das Dualsystem. Mitteilungen der deutschen Mathematiker-Vereinigung 16(1):12–15

Ingaliso L (2017) Leibniz e la lettura binaria dell'I Ching. In: Lio GMS, Ingaliso L (eds) Alterità e Cosmopolitismo nel Pensiero Moderno e Contemporaneo. Rubbettino, Soveria Mannelli, pp 109–132

Intorcetta, Prospero, ChristianWolfgang Herdtrick, François de Rougemont, Philippe Couplet, eds. and trans. (1687) Confucius Sinarum philosophus. Daniel Horthemels, Paris.

Jordaine J (1687) Duodecimal arithmetick. John Richardson, London

Kempe M (2022) Die beste aller möglichen Welten: Gottfried Wilhelm Leibniz in seiner Zeit. S. Fischer, Frankfurt

Knobloch E (2018) Determinant theory, symmetric functions, and dyadic. In: Antognazza MR (ed) The Oxford handbook of Leibniz. Oxford University Press, Oxford, pp 225–246

Knuth DE (1983) History of binary and other nondecimal numeration. By Anton Glaser. Los Angeles, CA (Tomash Publishers). 1981. 218 + xiii pp." Historia Mathematica 10: 236–243.

Knuth DE (1998) The art of computer programming. Volume 2: Seminumerical algorithms, 3rd edn. Addison–Wesley, Upper Saddle River

Kolpas S (2019) Napier's Binary chessboard calculator. Math Horiz 27:25–27

LBr = Manuscript held by Gottfried Wilhelm Leibniz Bibliothek – Niedersächsische Landesbibliothek, Hanover, Germany. Cited by shelfmark and Blatt [sheet]

Leibniz GW (1675) Extrait d'une lettre de Mr. Leibniz à l'auteur du journal, touchant le principe de justesse des horloges portatives de son invention. J des sçavans:93–96

Leibniz GW (1678) Extrait d'une lettre ecrite d'Hanovre par M. de Leibniz à l'Auteur du Journal, contenant une observation nouvelle de la maniere d'essayer si un nombre est primitif. J des sçavans:75–76
Leibniz GW (1768) Opera Omnia. Edited by Louis Dutens. 6 vols. Fratres de Tournes, Geneva
Leibniz GW. (1849–1863) Leibniziens Mathematische Schriften. Edited by C. I. Gerhardt. 7 vols. H. W. Schmidt, Halle
Leibniz GW (1923–.) Sämtliche Schriften und Briefe. Edited by Berlin-Brandenburgische Akademie der Wissenschaften. 8 series, each divided into multiple volumes. Akademie Verlag, Berlin
Leibniz GW (1969) Philosophical papers and letters, ed. and trans. Leroy E. Loemker. D. Reidel, Dordrecht
Leibniz GW (1971) Calcolo con zero e uno. Translated by Lia Ruffino. Etas Kompass, Milan
Leibniz GW (1985) Theodicy. Translated by E. M. Huggard. Open Court, Chicago
Leibniz GW (2006) Shorter Leibniz texts. Edited and translated by Lloyd Strickland. Continuum, London
LH = Manuscript held by Gottfried Wilhelm Leibniz Bibliothek – Niedersächsische Landesbibliothek, Hanover, Germany. Cited by shelfmark and Blatt [sheet]
Mahnke, Dietrich. 1912–13. Leibniz auf der Suche nach einer allgemeinen Primzahlgleichung," Bibliotheca Math. 13:29–61.
Mersenne M (1644) Cogitata physico mathematica. In quibus naturae quam artis effectus admirandi certissimis demonstrationibus explicantur. Bertier, Paris
Morley FV (1922) Thomas Hariot—1560–1621. Sci Mon 14(1):60–66
Napier J (1617) Rabdologiæ. Andreas Hart, Edinburgh
Nystrom JW (1862) Project of a new system of arithmetic, weight, measure and coins, proposed to be called the tonal system, with sixteen to the base. J. B. Lippincott & Co, Philadelphia
Nystrom JW (1863) On a new system of arithmetic and metrology, called the tonal system. J Franklin Institute 76:263–275, 337–348, 402–407
Pascal B (1665) Traité du triangle arithmetique, avec quelques autres petits traitez sur la mesme materiere. Guillaume Desprez, Paris
Pauli JW (1678) Disputatione mathematica numerum perfectum perfectissimi. Georg, Leipzig
Ramus P (1599) Arithmetica et geometria. Claude Marnium and Johannes Aubrium, Frankfurt
Riese A (1529) Rechnung auff der Linihen und Federn, n.p., Erfurt
Scaliger, JC (1557) Exotericarum exercitationum liber quintus decimus, de subtilitate, ad Hieronymum Cardanum. Michael Vasocsani, Paris
Schott G (1658) Magia universalis naturae & artis. Operis quadripartiti. Tomus tertius & quartus. Johann Gottfried Schönwetter, Frankfurt
Schwenter D (1636) Deliciae physico-mathematicae oder Mathematische und philosophische Erquickstunden. Dümmler, Nuremberg
Serra Y (2010a) Le manuscrit «De Progressione Dyadica» de Leibniz. http://www.bibnum.education.fr/sites/default/files/leibniz-analyse.pdf
Serra Y (2010b) Traduction du fac-similé du manuscrit de Leibniz du 15 mars 1679, «De Progressione Dyadica». http://www.bibnum.education.fr/sites/default/files/leibniz-numeration-texte.pdf
Stein E, Kopp FO, Wiechmann K, Weber G (2006) Neue Forschungsergebnisse und Nachbauten zur Vier-Spezies- Rechenmaschine und zur Dyadischen Rechenmaschine nach Leibniz. In: *VIII*. Internat. Leibniz-Kongress, Vorträge 2. Teil, 1018–1025. Hannover
Stifel M (1544) Arithmetica integra. Johan. Petreium, Nuremberg
Strickland L (2014) Leibniz's monadology: a new translation and guide. Edinburgh University Press, Edinburgh
Strickland, L (2022) An unpublished manuscript of Leibniz's on duodecimal. The Duodecimal Bulletin 54: 20_z–26_z
Strickland L (2023a) How Leibniz tried to tell the world he had squared the circle. Hist Math 62:19–39. https://doi.org/10.1016/j.hm.2022.08.004

Strickland L (2023b) Leibniz's tactile binary clock. L.I.S.A. Wissenschaftsportal Gerda Henkel Stiftung. https://lisa.gerda-henkel-stiftung.de/leibnitzstactileclock
Strickland L (2023c) Why did Thomas Harriot invent binary? Math Intell. https://link.springer.com/article/10.1007/s00283-023-10271-9
Strickland L, Jones OD (2023) F things you (probably) didn't know about hexadecimal. Math Intell. https://doi.org/10.1007/s00283-022-10206-w
Strickland L, Lewis H (2022) Leibniz on binary: the invention of computer arithmetic. MIT Press, Cambridge, MA
Swetz, Frank J. 2003. Leibniz, the Yijing, and the Religious conversion of the Chinese. Math Mag 76 (4): 276–291.
Sypniewski, Bernard Paul. 2005. China and universals: Leibniz, binary mathematics, and the Yijing Hexagrams Monumenta Serica 53:287–314.
Tropfke J (1980) Geschichte der Elementarmathematik. Band 1: Arithmetik und Algebra, 4th edn. De Gruyter, Berlin
von Mackensen L (1974) Leibniz als Ahnherr der Kybernetik—ein bisher unbekannter Leibnizscher Vorschlag einer 'Machina arithmeticae dyadicae'. In: Akten des II Internationalen Leibniz-Kongresses, II:255–268. Steiner Verlag, Wiesbaden
Weigel E (1673) Tetractys, Summum tum Arithmeticae tum Philosophiae discursivae compendium, artis magnae sciendi. Meyer, Jena
Zacher HJ (ed) (1973) Die Hauptschriften zur Dyadik von G. W. Leibniz. Vittorio Klostermann, Frankfurt
Zhonglian S (2000) Leibniz's binary system and Shao Yong's Xiantiantu. In: Li W, Poser H (eds) Das Neueste über China: G.W. Leibnizens Novissima Sinica von 1697. Franz Steiner Verlag, Stuttgart, pp 165–169

Fermat and Descartes in Light of Premodern Algebra and Viète

Jeffrey A. Oaks

Contents

1 Introduction 200
2 Premodern Monomials, Polynomials, and Equations 202
2.1 Premodern Numbers, Premodern Polynomials 203
3 Modifications to Premodern Algebra 208
3.1 Divisions and Roots in Equations 208
3.2 Irrational Coefficients 209
3.3 Independent Unknowns 210
4 François Viète's Geometrical Algebra 211
4.1 Preliminary Remarks on Viète 211
4.2 Overview of Viète's Program 212
4.3 Viète's Magnitudes and Numbers 213
4.4 Viète's Two Algebras 217
4.5 Summary of Viète's *logistice speciosa* 221
5 Fermat's Algebra 222
5.1 Fermat, Student of Viète 222
5.2 Fermat's Interpretation of the Species 223
6 Descartes's Algebra 226
7 Synthesis 230
References 232

Abstract

There were two radically different algebras practiced in Europe in the time of Fermat and Descartes. One was the traditional cossic algebra grounded in arithmetic whose roots stretch back to al-Khwārizmī and Diophantus, and the other was the new geometrical algebra of François Viète. Both Fermat and Descartes chose to work in the latter. To understand their choice, and to determine how they understood their algebraic terms, we first outline the conceptual foundations of the two algebras, one that regards a monomial as a premodern number in which

J. A. Oaks (✉)
Department of Mathematical Sciences, University of Indianapolis, Indianapolis, IN, USA
e-mail: oaks@uindy.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_8

the coefficient is a multitude and the power is a kind, and the other that regards a monomial as the product of two magnitudes, one known and the other unknown. We find that cossic algebra could not have been modified to include undetermined coefficients, and thus would have been inadequate for the purposes of Fermat and Descartes. Further, both authors, when solving problems in geometry, worked with precisely the same non-arithmetized algebra as Viète, without a unit measure, respecting dimension, and using magnitudes of arbitrarily high dimension for which no justification is provided.

Keywords

Algebra, premodern algebra · Modern algebra · Michael Stifel · François Viète · Pierre de Fermat · René descartes

1 Introduction

Algebra played a limited role in mathematics prior to the seventeenth century. Along the path from Greek to Arabic to European algebra, it had remained a numerical technique for finding unknown numbers in problem-solving. Algebraic language and notation were not used to state arithmetical identities like $(a+b)^2 = a^2 + b^2 + 2ab$, formulas like $A = \pi r^2$, curves like $y = x^2 + 1$, or physical laws like $d = \frac{1}{2}gt^2$. For those purposes, arithmetical or geometrical language was used. Apart from rare applications in proofs in some Arabic authors (Oaks 2019), algebra concerned itself solely with the setting up and solving of equations to find the values of unknowns. While several authors such as al-Karajī and Bombelli had studied different aspects of algebra from a more theoretical viewpoint, these investigations remained internal to algebra as a problem-solving technique. This started to change in the first half of the seventeenth century, when Pierre de Fermat and René Descartes independently began to study curves through equations. Each of them discovered independently how curves can be represented by equations in two unknowns, and they both found ways to apply algebra to problems of determining normal and tangent lines to curves. These were among the first steps in what would eventually lead to the wholesale algebraization of mathematics, and by association of science as well.

In this chapter, I examine the ontology of the knowns and unknowns in the algebra practiced by Fermat and Descartes. The algebraic landscape of their time offered two basic alternatives: the traditional numerical algebra familiar to readers of such figures as Rafael Bombelli, Simon Stevin, and Christoph Clavius, and the new, geometrical algebra of François Viète and his followers. As the familiar story goes, Fermat and Descartes both built on the latter. To understand the nature of their algebraic terms and the rationale behind their choice, we should summarize the conceptual foundations of both kinds of algebra, especially because for each of them recent research has modified old views.

We begin with an overview of the numerical algebra that Europeans had acquired from the Arabic world. This premodern algebra was not simply an early version of

the numerical algebra that we currently teach. It was grounded in a corresponding premodern notion of number, so that the monomials, polynomials, and equations we find in Diophantus, Abū Kāmil, Cardano, and Clavius were conceived in a radically different way from their modern counterparts. Very briefly, the core of the difference lies with the fact that the powers of the unknown were kinds of number, and not values as they are today. Our x is the value of an unknown, while Clavius's cossic 2℘ is a kind of number distinct from other kinds (powers). Only with a coefficient, like 12℘, does it become a value. Consequently, coefficients held a different relationship to the powers, and polynomials were aggregations, free of arithmetical operations. The different notations devised for working out problems in premodern algebra across all languages also reflect this conception, so that reading it as if it were some version of modern notation distorts its meaning.

Next we review Viète's *algebra nova*, where knowns and unknowns, instead, are values deriving from geometry. More precisely, they are the non-arithmetized measures of magnitudes, possessing dimension but unfettered by the limitation to line, surface, and solid, and equations between them arise out of ratio and proportion. They are not, as many historians since Jacob Klein have maintained, uninterpreted symbols that transcend number and magnitude. Viète's program was to develop propositions in geometry modeled on those in Euclid and other classical Greek authors, that is, in geometry without numerical measure, to serve as foundations for calculating the entries in astronomical tables. Ptolemy, after all, had based the calculations for his tables in Euclidean geometry, and Viète wished to respect this tradition. But Viète diverged from astronomers in the Ptolemaic tradition by reworking the "computational geometry" of the Greeks into a new algebra that would greatly facilitate calculation. Because his end goal was to use the geometric models for numerical calculation, he had no reason to restrict himself to three geometric dimensions – not just in his algebra, but even in his geometry. Instead of justifying these higher dimensions, he explained how they are useful. Not surprisingly, the notation Viète devised for his *logistice speciosa*, as he called his new algebra, is also a radical departure from that of contemporary algebraists.

When we finally turn to Fermat and Descartes, we find that in practice they both worked with the same non-arithmetized geometrical algebra as Viète, again with no justification for the higher-dimensional magnitudes that come with it. Fermat adopted Viète's system without modification, and even kept the same notation. And although Descartes began his 1637 *La Geometrie* by describing an algebra of line segments that bypasses the problem of dimensionality, he did not apply it in his algebraic solutions to problems. Instead, he worked in the same non-arithmetized algebra as Viète and Fermat, in which the dimensions of the terms must be taken into account.

Last, we inquire what led Fermat and Descartes to choose Vietan algebra over premodern, cossic algebra. Two obstacles in particular prevented the latter option from being chosen. The most obvious is that it is a numerical art, while Fermat and Descartes worked in geometry. But even if the terms could have been reinterpreted geometrically, there was no natural way that premodern algebra could have been modified to accommodate undetermined knowns, which are necessary for the

general solutions that Fermat and Descartes sought. The cause of this limitation lies with the conceptual foundation of premodern algebra in kinds, not values, and this foundation also made premodern algebra ill-equipped for other applications outside problem-solving. It was the shift from types to values, not the shift to abstract magnitudes, that allowed algebra to begin its path of overrunning the rest of mathematics.

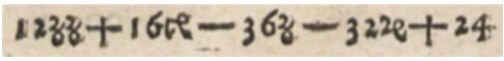
2 Premodern Monomials, Polynomials, and Equations

European algebra in the time of Viète can be traced back via Italian and Latin texts to medieval Arabic algebra, and although we lack a textual connection, there is clearly a historical link between Arabic algebra and the algebra practiced in the *Arithmetica* of Diophantus of Alexandria (ca. third century CE) (Christianidis and Oaks 2013). It is the seemingly minor yet consistent differences in practice, wording, and notation between pre-Vietan algebra and our own elementary algebra that testify to a different way of conceiving of monomials, polynomials, and equations. Because my aim in this section is to provide a backdrop for comparison with the algebra of Viète, Fermat, and Descartes, I will describe the main features of premodern algebra using examples leaning toward their time. Many will be taken from Michael Stifel's 1544 *Arithmetica Integra*, though just about any other sixteenth-century text could have been chosen. Stifel's cossic notation was about as standard as one could find at the time, and it was the notation later adopted by Christoph Clavius, whose 1608 *Algebra* was studied by Descartes.

Stifel's notation uses the standard cossic signs that had been common among German algebraists since the previous century. The names given to the powers of the unknown in this system were adapted from Latin terms which were themselves ultimately translations of the corresponding terms in Arabic. In the cossic system, the first degree unknown is called a *radix*, the Latin word for "root," which was translated from the Arabic first-degree term *jidhr* ("root"). In notation its sign is ʓ. The square of a *radix* is called a *zensus*, deriving from the Latin *census*, translated from *māl*, the name of the second degree unknown in Arabic. Like the Latin *census*, the common meaning of *māl* is "sum of money" or "wealth." The cossic sign for a *zensus* is ʒ. The cube of a *radix* is called a *cubus*, the Latin word translated from the Arabic *kaʿb* ("cube"), and its sign is ƈ. And just as in pre-Renaissance algebra, most higher powers are named with combinations of the terms for the second and third powers. The fourth power was called a *zensizensus*, which ultimately comes from the Arabic *māl māl*, and its sign is ʒʒ. The European powers diverge from the Arabic after this point. Stifel's fifth power is called a *surdesolidus* (ß), his sixth is a *zensicubus* (ʒƈ), and higher powers follow without limit. The constant term in a polynomial, or what we call the zero-degree term, is called a *numerus*, and in Arabic, these numbers are usually counted in dirhams, a denomination of silver coin. Because algebra before Viète was founded in arithmetic, Stifel's unknown *radix* is always found to be a number, and numbers in arithmetic from the earliest Arabic texts can take on any positive value, including fractions and irrational roots.

In arithmetic, the cossic sign for square root is √ȝ, for cube root √cȝ, etc., so that for instance $\sqrt{72}$ was shown as √ȝ72.

In notation, this cossic algebra can look deceivingly modern. Take for example this polynomial from Stifel's book (1544, f. 238b):



It is tempting to read this as being identical to $12x^4 + 16x^3 - 36x^2 - 32x + 24$ and indeed if one follows Stifel's calculations the modern reading makes sense. But there are many indications in the book that he is not thinking like a modern mathematician and these differences are shared with algebraists all the way back to al-Khwārazmī and still farther back to Diophantus

Before turning to the premodern arithmetical foundation of algebra, I will mention two differences between Stifel's notation and ours. One is that when there is one of a term, he always writes the "1." So where we write $x^2 + 8x$, Stifel shows "1ȝ + 8ȝȝ" (1544, f. 277b.2), or where we write $x + \sqrt{205 - x^2}$ Stifel has "1ȝȝ + √ȝ.205 − 1ȝ." (the periods serve as parentheses) (1544, f. 281a.5). This inclusion of the "1" is present in every premodern notation from Diophantus to Arabic algebra to Latin, Italian, and the various notations in Renaissance Europe. Where we would write simply x, Scheubel (1550) wrote "1 ra.", Bombelli (1572) "1 $\overset{1}{\smile}$," Stevin (1585) "1 ①," and de Billy (1643) "1R" (Oaks 2017).

Another difference is that with a few exceptions beginning in Stifel's century, premodern algebraists insisted on keeping their coefficients rational. For example, Stifel writes in one place, "I multiply √ȝ $501\frac{9}{11}$ by √1ȝȝ, or √ȝ1ȝ, which makes √ȝ $501\frac{9}{11}$ ȝ." ("√ȝ $501\frac{9}{11}$ multiplico per 1ȝȝ, seu per √1ȝ, facit √ȝ $501\frac{9}{11}$ ȝ") (Stifel 1544, f. 292b.12). In modern notation, he multiplies $\sqrt{501\frac{9}{11}}$ by x to get $\sqrt{501\frac{9}{11}x^2}$, where we would write instead $\sqrt{501\frac{9}{11}}x$. In another instance he writes "I then multiply 1ȝȝ by √ȝȝ864 − 1ȝȝ, which makes √ȝȝ864ȝȝ − ȝ.") ("Multiplico igitur 1ȝȝ in √ȝȝ 864 – 1ȝȝ, facit √ȝȝ864ȝȝ − 1ȝ.") (Stifel 1544, f. 288a.2). In modern notation, he multiplies x by $\sqrt[4]{864} - x$ to get $\sqrt[4]{864x^4} - x^2$, again keeping the power under the root. Algebraists starting with the earliest extant Arabic texts in the ninth century similarly ensured that their coefficients remained rational by taking the root of the coefficient *and* the power (Diophantus did not work with irrational numbers). We discuss the sixteenth-century exceptions in Sect. 3.2 below.

2.1 Premodern Numbers, Premodern Polynomials

These curiosities, along with other consistent differences in the wording of certain steps and in the overall approach to problem-solving, are reflections of a different way of understanding the relationship between the coefficient and the power, and this relationship derives from the way that numbers themselves were conceived. Today mathematicians typically think of numbers as being elements of some abstract set equipped with the binary operations of addition and multiplication that satisfy the

ordered field axioms. Others might view numbers as being constructed from scratch, where zero is identified with the empty set. Many people with no university training in mathematics retain the idea that numbers are points on some imagined number line, or at least that the units in which one counts or measures (e.g., apples, pesos, meters, pounds) are not part of what constitutes a number.

Such was not the case with premodern numbers. I will first describe the setting in Arabic, from which European premodern algebra originated. (What I describe here applies also to numbers in ancient Greece (Klein 1968, 46ff; Mueller 1981, 59).) Numbers in Arabic, whether they arose in practical calculation or in the tradition of Aristotle and Euclid, were conceived as being numbers *of something*. A six was not simply a six, but must be six of some kind, like six geese, six men, six dirhams, six mithqals (a unit of weight) of gold, six sevenths ($\frac{6}{7}$, which in turn might be a number of some other unit, like mithqals), six degrees (6°), or six (abstract) units. In other words, numbers were not objects in themselves, but they necessarily modified – i.e., they counted or measured – a multitude of some material or intelligible unit. Some kinds of unit were divisible, such as units of weight and length, while others were indivisible, such as geese and the Euclidean unit. In an arithmetic book composed in 1301, the Moroccan polymath Ibn al-Bannāʾ wrote about the two aspects of a number, its "meaning" (*maʿnan*) and its "term" (*lafẓ*) (1994, 255.3). For "three men," the meaning is "three" and the term is "men." Of course, the term could be ignored when working out operations. And because numbers are amounts of some unit, they were necessarily positive, though sometimes "nothing" or "zero" entered into calculations.

Niccolò Tartaglia (1556, f. 2a-b), to pick a European example, frames these two aspects of number in Aristotelian terms while appealing to Euclid's definitions of "unit" and "number." Units, he tells us, come in two kinds, "Natural" ("*Naturale*") and "Mathematical" ("*Mathematico*"). "The Natural [unit] considers the things according to their being, as according to reason conjoined with some sensible material." ("Il Naturale considera le cose si secondo lesser, come secondo la ragione congionte con qualche materia sensibile" (Tartaglia 1556, f. 2a.29).) He gives many examples of Natural units, such as one gold ducat, one scudo, one fiorino (three denominations of currency), one yard of cloth, one pound of silk, one ounce of saffron, one pace, one yard, one degree, one minute (sexagesimal units), etc. Because they are rooted in the material world, such units are "infinitely divisible" ("divisibile in infinito"). "The Mathematical [unit], then, also considers the things conjoined, according to their being, with such sensible material (as with the Natural), but one takes them, or considers them as being abstracted from such sensible material according to reason." ("Il Mathematico poi considera le cose pur congiunte secondo lessere, con tal materia sensibile, (si come fa anchora il Naturale) Ma le piglia, over considera poi si come astratte, da tal materia sensibile secondo la ragione" (Tartaglia 1556, f. 2a.45).) Unlike Natural units, Mathematical units are indivisible ("indivisibile"). After covering these two kinds of unit, he turns to the definition of "number": "The number (as Euclid defined it in the second definition of the seventh [book]) is none other than a multitude composed of units," ("Il Numero (come diffinisse Euclide nella seconda diffinitione dil settimo) non è altro, che una

Moltitudine composta dalle unitade." (Tartaglia 1556, f. 2b.14).) where these units can be either Natural or Mathematical. As in Aristotle and in medieval Arabic, numbers in Tartaglia are always multitudes of some unit, whether that unit is sensible or noetic. Even if some sixteenth-century mathematicians may have advocated for a different meaning of "number," it is the common understanding exhibited in Ibn al-Bannāʾ and Tartaglia that informed the way algebraic monomials were conceived.

Premodern algebra was strictly an arithmetical art. As al-Khwārazmī wrote in the beginning of his seminal *Book of Algebra* (early ninth century), "When I considered what people need in calculation I found that it is always a number" ((al-Khwārazmī 2009, 97.1), my translation). Algebra could be used to solve problems in arithmetic, geometry, metrology, and even horology, as long as the parameters are given numerical measure. Algebra was not used, for instance, to find a geometric magnitude in itself, but only its numerical size.

The names of the powers in algebra are kinds of number in the same sense that geese, dirhams, and degrees are kinds of number. Again quoting al-Khwārazmī, "I found that the numbers which are necessary for calculation in algebra are of three kinds (*ḍurūb*), which are roots, *māl*s, and a simple number unrelated to a root or a *māl*." ((al-Khwārazmī 2009, 97.9), my translation. The plural of *māl* is *amwāl*, but we write it with the English suffix "s.") So in addition to men, degrees, and mithqals, we can also count or measure roots, *māl*s, and simple numbers. Al-Khwārazmī's algebraic terms "eight roots" (2009, 147.1) and "four ninths of a *māl*" (2009, 181.11) are multitudes of different kinds just as "eight men" and "four ninths of a mithqal" are multitudes of different kinds. (The modern versions of these monomials are $8x$ and $\frac{4}{9}x^2$.) Sixteenth-century algebraists likewise often called these algebraic kinds "numbers" or "quantities," like Tartaglia, who writes just before listing the names of the powers of the "different kinds of quantities considered in Algebra" ("diverse spetie de quantità, considerati in Algebra") (Tartaglia 1560, f. 1a). Just five examples of the kinds as "numbers" are Stifel (1544, f. 228a), Scheubel (1550, 1ff), Peletier (1554, 5), Gosselin (1577, f. 54a), and Clavius (1608, 7). The algebraic kinds, or powers, were also called by words meaning "name" or "denomination" in every language from Greek to European vernaculars (Oaks 2018, 269).

The premodern concept of number just described was formed without consideration of irrational roots, most likely before irrational numbers entered into calculations. Although irrational numbers are abundantly common in Arabic arithmetic and algebra, they did not participate in this idea of a number as a multitude. One can have three loaves of bread or two and a half loaves of bread, but it makes no sense to have $\sqrt{10}$ loaves of bread. For this reason, the "number," i.e., the coefficient, of a term in algebra was required to be rational. Stifel's $\sqrt{\text{ }501\frac{9}{11}\text{ }}\left(\sqrt{501\frac{9}{11}x^2}\right)$ is acceptable because $501\frac{9}{11}$, the number of *zensi*, is rational, while the whole term is acceptable as a potentially irrational number (Oaks 2017). Words meaning "number" from al-Khwārazmī to Clavius were used in more than one sense. The whole term "five cubes" ($5x^3$ in modern notation) could be called a "number," and at the same time the "number" of the term (5 in this example) meant its multitude.

Let us compare Stifel's term "16cc" from the polynomial shown above with its modern equivalent, $16x^3$. (The analysis would be no different if we were to choose in

place of Stifel's notation Cardano's "16 cu," Bombelli's "16$\overset{3}{\smile}$," or any other pre-Vietan version of this term.) In our $16x^3$, the 16 and the x^3 are both numbers, one known and the other unknown, that are understood to be multiplied together by the concatenation of their signs. But in Stifel's 16ƈɛ, as in the rhetorical "sixteen cubes," the "16" and the "ƈɛ" are two aspects of a single number, one telling the amount and the other the kind. It is like saying "16 bottles" or "16 euros." With regard to this last example, "euros" is a kind of currency, while "one euro" is a value or an amount. If in place of "1ȝ + 8²ɛ" Stifel had written "ȝ + 8²ɛ," it would have been read as "*zensi* and eight roots," which leaves unanswered how many *zensi* there are, much like "euros and 8 dollars" leaves unanswered the number of euros. This is why, whenever there is one of a term in premodern algebra, the coefficient of "1" is always written in notation. In modern algebra, we can write simply x^3 because the notation for the power already stands for a value. Here is another way of putting it: Imagine "sixteen euros" as a pile of sixteen one-euro coins. The "sixteen" and the "euros" are not two separate entities brought together to form the term, but are different aspects of a single amount. The same applies to Stifel's "16ƈɛ": the ƈɛ is the kind, and the 16 is how many there are. While we can today also interpret our monomials in this premodern manner, it only works for simple expressions. Reading modern notation as a number-kind pair breaks down when we stray just a little from simple monomials to terms like πr^2 or $3xy$.

The addition of algebraic terms works like the addition of other kinds of number, which helps shed light on how polynomials were conceived. Adding, say, two cubes to three cubes gives five cubes, but when the kinds (powers) are different, the conjunction "and" or some equivalent is used. Sibṭ al-Māridīnī (fifteenth c.) gives the example "if you add two dirhams to three things, the answer is two dirhams and three things" ("thing" (*shay'*) was another name for the first power in Arabic algebra). Written in our notation, this does not seem to do anything: $2 + 3x = 2 + 3x$. Unlike our equation, the Arabic is stated as an operation and its outcome: one *adds* the two terms *to get* the algebraic binomial "two dirhams and three things." And where our $2 + 3x$ is written with the sign for "plus," a word that functions as a quasi-preposition to indicate the arithmetical operation of addition, the Arabic "and" (*wa*) is simply the conjunction. There is no more addition in "two dirhams and three things" than there is in "two geese and three chickens."

Some algebraists compared the names of the powers to denominations of coins to help explain the addition of the different kinds. In Europe, these include Luca Pacioli (1494, f. 112a), Michael Stifel (1544, f. 231a), and Marco Aurel (1552, f. 71a). To add three ducats to two crowns, Aurel explains, "I cannot said that they are five ducats, nor five crowns: in reality they are three ducats and two crowns more, or two crowns and three ducats more. In this same way, I want to sum 3²ɛ with 2ȝ: I cannot say that they are 5²ɛ, nor 5ȝ: we are compelled to say that they are 3²ɛ and 2ȝ more, or 2ȝ and 3²ɛ more." (". . . no pueden dezir que son cinco ducados, ni menos cinco coronas: mas realmente pueden dezir que son tres ducados, y mas dos coronas, o dos coronas, y mas tres ducados. Assi mesmo quiero summar 3²ɛ, con 2ȝ: no podemos

dezir que son 5ce, ni 5g: mas forçadamente, diremos que son 3ce, y mas 2g: o 2g, y mas 3ce" (Aurel 1552, f. 71a).)

Unlike English, Arabic has a word that serves as the negative counterpart to "and." This word is *illā*, which I translate as "less," though outside mathematics "except" is usually better. For example, one can say "all the students except (*illā*) Joan." In arithmetic, the apotome that we write as $5 - \sqrt{3}$ is stated by Ibn al-Bannāʾ as "five less a root of three." (Ibn al-Bannāʾ 1994, 288.2), while he writes the corresponding binomial as "five and (*wa*) a root of three" (p. 287.21). As with the example sentence, the word *illā* indicates that what follows it is lacking from what precedes it. In algebra, the same "less" is used to state an amount in which one term has been removed from another term. Al-Karajī writes in one problem "If you cast away the thing from fourteen, it leaves fourteen less a thing" (Saidan 1986, 193.15). Like with the addition of different kinds, here the "fourteen less a thing" is the *result* of the operation. It is a deficient 14, specifically a 14 that is lacking a thing. The word *illā* does not mean the same as the English "minus," which, like "plus," is a quasi-preposition indicating the arithmetical operation. Incidentally, the word *al-jabr* ("restoration"), from which our word "algebra" originates, is used in the simplification of equations to "restore" such diminished amounts (Oaks and Alkhateeb 2007).

In Italian abacus texts of the fourteenth and fifteenth centuries, "less" was rendered as "*meno*" ("less"), while "and" became either "*et*/*e*" ("and") or "*più*" ("more"). The words *più* and *meno* correspond to the Latin *plus* and *minus*, where they also mean "more" and "less." (In my translation of the passage of Marco Aurel above, the "and two coronas more" is from "y mas dos coronas," where "y mas" is the Spanish version of the Latin "et plus," or the Italian "e più.") These two words appear in arithmetical and algebraic expressions in several sixteenth-century books where they also take the same meaning as the Arabic "and" and "less." In notation *plus* and *minus* are either abbreviated as "p̃" and "m̃," or they are shown with the signs "+" and "−". It is clear by the way authors before ca. 1600 described the terms and how they were applied in context that neither the Latin words nor the signs "+" and "−" were used in the modern sense. The words and the signs began to assume their modern meanings indicating the operations of addition and subtraction only with Viète (Oaks 2018, 253–4).

So Stifel's polynomial 12ggg + 16cc − 36g − 32ce + 24 is not a linear combination of the powers constructed from operations. It is a 12, a 16, and a 24 of three different kinds that have been diminished by a 36 and a 32 of two other kinds. The signs ce, g, cc, and ggg represent kinds, not values like our powers of x, and the terms are joined together like we join any collection of multitudes of different kinds. (Stifel does not have a sign for units, though other cossic algebraists did.) There is no scalar multiplication, no addition, and no subtraction intended in the notation. Said another way, the polynomial is an aggregation, or a kind of inventory, listing how many of each kind are present, and how many are lacking. It is thus fundamentally different from our $12x^4 + 16x^3 - 36x^2 - 32x + 24$.

Premodern polynomials were not thought of as functions or as elements of some abstract ring. They were ways of expressing unknown numbers using the same kind of language for saying known numbers. In Arabic one said "a hundred and seven and

forty" (147), the same way one said "a *māl* and forty-nine dirhams and fourteen things" (in modern notation, $x^2 + 49 + 14x$). (The number and the polynomial are from (Abū Kāmil 2012, 449.9, 451.14).) Several Arabic algebraists called polynomials "composite numbers" to distinguish them from monomials, or "simple numbers." In the sixteenth century polynomials were often referred to by terms like "quantities" or "cossic numbers." Caspar Peucer writes about three kinds of cossic number (*numerorum cossicorum*): simple (like 20ϩ, 24ϧ, and 30ⲥϩ), composite (like 1ϩ +12, 1ϧ + 1ϩ, and 2ϧ + 8), and diminished (like 2ϩ − 8 and 4ϧ − 6ϩ) (Peucer 1556, f. Mviija). Then, on the next page, he gives the example of the "number" (*numerum*), 12ϧϧ + 16ⲥϩ − 36ϧ − 32ϩ + 24, which he borrowed from Stifel.

Algebraic equations ideally expressed the equality of two cossic numbers, without any unresolved operations hanging over them. Jacques Peletier compared a two-term equation with an "equation" in currency (he writes "ç" for "ϧ," and "ℝ" for "ϩ"):

> Thus, an equation is an equality of value between numbers denominated differently, such as when we say, 1 écu is worth 46 sous: it is an equation between 1, with its denomination of écu, and 46, with its denomination of sous. So when we say, 1ç, equal to 4ℝ, it is an equation between 1 with its denomination of ç, and 4 with its denomination of ℝ.
>
> ("Equacion donq, ét une equalite de valeur, antre nombres diversemant denommez. Comme quand nous disons, 1 Ecu valoèr 46 Souz: il y à Equacion antre 1 avec sa denominacion d'Ecu: e 46 avec sa denominacion de Souz. Einsi, quand nous disons, 1ç, egal a 4ℝ: il y à Equacion antre 1, avec sa denominacion de ç: e 4 avec sa denominacion de ℝ" (Peletier 1554, 22).)

Because equations state the equality of numbers, however, represented, premodern algebraists routinely worked out all operations before stating equations. Thus, all equations should be polynomial equations, free of any unresolved operations. Four articles that deal with this topic for Arabic, Italian, Greek, and Renaissance European algebra, respectively are Oaks (2009), Oaks (2010), Christianidis and Oaks (2013), and Oaks (2017).

3 Modifications to Premodern Algebra

3.1 Divisions and Roots in Equations

Adding, subtracting, and multiplying polynomials results in another polynomial, but this does not hold for division and taking roots. When confronted with division by a polynomial, Arabic algebraists usually found some way of reasoning through the operations to construct expressions without unresolved divisions. The simplest way was to preemptively multiply all terms that would eventually form the equation by the divisor. Abū Kāmil was particularly adept at finding other ways around divisions, sometimes applying arithmetical identities, and sometimes giving names to the results of divisions and forming equations using them along with the names of the

powers. Two names he and other algebraists gave are *dīnār* and a *fals*, which are denominations of gold and copper coin. In one problem, for example, he writes: "we divide a thing by ten less a thing, so it results in a *dīnār*, and we divide ten less a thing by a thing, so it results in a *fals*" (Abū Kāmil 2012, 397.2). He then works with these names along with dirhams, things, and *māl*s, until multiplications can be performed to eliminate them. But some authors, notably al-Karajī, occasionally worked with terms of the form "A divided by B" in equations until the appropriate multiplications could be performed to cancel the division. Treating a phrase like "a thing divided by ten less a thing" (like our $\frac{x}{10-x}$) as if it were a term in a polynomial violated the idea that expressions should be free of operations, but it also made the solutions to many problems much simpler (Oaks 2009).

Resorting to terms of the form "A divided by B" never became common in Arabic algebra. But in medieval Italian abacus books, divisions occur more frequently because they were written in rhetorical solutions with the division bar, along with Arabic numerals, as in this example from Maestro Biagio as related by Benedetto da Firenze (1463): "$\frac{180 \text{ chose e } 120}{1 \text{ censo } 3 \text{ chose}}$, e questo è iguali a 10." (In modern notation, $\frac{180x+120}{x^2+3x} = 10$ (Biagio 1983, 117.7).) Divisions continued to be common in sixteenth-century Europe, still with the division bar. A simple example in Michael Stifel is "Aequantur igitur $\frac{10}{1+1\text{2}\text{ℓ}}$ & 7, facit $1\text{2}\text{ℓ}.\ \frac{3}{7}$" (f. 266b) ("Equating, then, $\frac{10}{1+1\text{2}\text{ℓ}}$ and 7, makes $1\text{2}\text{ℓ}\ \frac{3}{7}$").

Roots were treated similarly. As mentioned above, irrational roots did not fit the premodern way of conceiving of numbers as multitudes, yet they became an integral part of arithmetic anyway. Similarly, roots of monomials and polynomials were accepted in algebra.

The introduction of divisions and roots to equations did not open the door for other operations. This is true for addition and subtraction, but it is most evident for multiplication. No premodern text will admit an unresolved multiplication, such as "a root multiplied by ten less a root," to an equation, since multiplications can always be performed beforehand. The two sides of an equation remained ideally polynomials, even if they may have become compromised in some solutions with divisions and roots.

3.2 Irrational Coefficients

Before ca. 1500, notation was primarily a tool for working out problems. To communicate this work in a book or to explain the general rules, a rhetorical version, sometimes with a little notation included, was composed. As long as the notation was somehow tied to speech, it had to make sense, so irrational coefficients were forbidden. After 1500, we find notation in books becoming increasingly independent of the spoken word. An early example is Christoff Rudolff's 1525 *Coss*, where algebraic calculations are presented without the intervention of any verbalization of the equations. Rudolff was the first of a small group of algebraists of his century who saw an advantage in allowing irrational coefficients in notation. Rudolff used them hesitatingly, then they become more common in Girolamo Cardano (1539), Johann

Scheubel (1550), Rafael Bombelli (1572), and Simon Stevin (1585), with some others following suit through copying these authors. But many more algebraists continued to respect the semantics of the expressions by not allowing for irrational coefficients, including Michael Stifel (1544) (who criticized Rudolff's coefficients), Jacques Peletier (1554), and Niccolò Tartaglia (1560).

In the following century, irrational coefficients in premodern algebra became rarer. Benedetto Maghetti (1639) set off a controversy when he used them in his solutions, and authors like Christoph Clavius (1608), Jacques de Billy (1643), and Jacob R. Brasser (1663) followed tradition by ensuring that their coefficients remain rational. (See Oaks (2017, 161ff) for Maghetti's story. De Billy, for example, multiplies "1R − ℞5 − 3" by "℞5 + 3" to get "℞5Q + 3R − 14 − ℞180" (He writes R and Q in place of ²ℓ and ȝ, and ℞ for √). In modern notation, he multiplies $x - \sqrt{5} - 3$ by $\sqrt{5} + 3$ to get $\sqrt{5x^2} + 3x - 14 - \sqrt{180}$ (De Billy 1643, 269). Brasser, for example, squares "√1920–1²ℓ" (our $\sqrt{1920} - x$) to get "1920 − √7680ȝ + 1ȝ" ($1920 - \sqrt{7680x^2} + x^2$) (Brasser 1663, 301.11).) Like the introduction of divisions to equations, the admission of irrational coefficients by a handful of authors was not indicative of any conceptual shift. All other characteristics that identify a piece of text as being premodern are still present in these authors. The irrational "numbers" of terms stems from a convenient preference for notational syntax over rhetorical semantics. (See Oaks (2017) for a full discussion. There is one significant typographical error in the published version of the article. On page 154, line 19, the polynomial is shown incorrectly as "1ȝ + 2²ℓ + √2²ℓ." It should be "1ȝ + 2²ℓ + √2ȝ.")

3.3 Independent Unknowns

Different names for independent unknowns are attested already in the ninth century in Abū Kāmil's *Book of Birds*, and by the sixteenth century, several European algebraists were working with independent unknowns in conjunction with different notations to express them. Christoff Rudolff (1525), Girolamo Cardano (1539), Jacques Peletier (1554), and Simon Stevin (1585) are just some of the algebraists whose independent unknowns are discussed in Heeffer (2010). We review here the notation of Stifel's *Arithmetica Integra*, which was later adopted in the *Algebra* of Christoph Clavius.

Stifel introduces his unknowns on folio 251b. Recall that the signs for the powers of the (primary) unknown are ²ℓ, ȝ, cℓ, ȝȝ, etc. The powers of the second unknown are designated A²ℓ, Aȝ, Acℓ, Aȝȝ, etc., for the third unknown B²ℓ, Bȝ, Bcℓ, Bȝȝ, etc. He writes that one can simply write "A" for "A²ℓ," "B" for "B²ℓ," etc. To multiply powers of different unknowns, one concatenates the signs, so, for example, the product of 2cℓ by 4Aȝ is 8cℓAȝ. By this rule, 3Aȝ by 2²ℓ by 4Bȝȝ would be 24²ℓAȝBȝȝ: the signs for the unknowns are arranged in order from ²ℓ to A, B, etc. This is a notation for multiplying kinds, not values. The traditional cossic names already employed this idea, where for example ȝcℓ is the kind that results from multiplying *zensi* (ȝ) by cubes (cℓ). If Stifel meant the notation to be the

multiplication of values, then the result of 1q by 1A would be 1q1A instead of 1qA. Here "qA" is the kind that results from multiplying the kinds "q" and "A." Stifel, like other premodern algebraists, has no notation for the multiplication of values.

Simon Stevin also devised a notation for multiplying kinds. Where Stifel multiplied 3A by 2B to get 6AB, Stevin multiplies 3 *sec.*① by 2 *ter.*② to get 6 *sec.*① M *ter.*②. Here the kind "*sec.*①" is multiplied by the kind "*ter.*②" to get the new kind "*sec.*① M *ter.*②" (Stevin 1585, 259). Again, this is not the multiplication of values. The earliest use of concatenation for the multiplication of values I have found is in Vaulezard's 1630 translations of Viète's *Isagoge* and *Zeteticorum*. Concatenation appeared soon after in other algebraists, all of them indebted to Viète's new algebra.

4 François Viète's Geometrical Algebra

4.1 Preliminary Remarks on Viète

This section is condensed from my article (Oaks 2018), where I investigate the nature of Viète's species. In some places below I give the corresponding section of that article with a reference like "(§2.3)."

Below is a list of Viète's works cited in the present chapter. All were reprinted in his collected works, which were published in 1646 by Frans van Schooten, titled *Opera Mathematica*. After each title, I give the date of the original printing and the page number in the 1646 *Opera Mathematica*.

In Artem Analyticem Isagoge (1591, p. 1) (henceforth *Isagoge*)

Variorum de Rebus Mathematicis Responsorum, Liber VIII (1593, p. 347) (henceforth *Variorum*)

Effectionum Geometricarum Canonica Recensio (1593, p. 229) (henceforth *Effectionum Geometricarum*)

Zeteticorum (1593/1600, p. 42) (The first 16 folios were printed in Tours in 1593, and the remainder in Paris in 1600. Folio 16 ends in the middle of Zetetic IV.6 (Van Egmond 1985, 362))

Supplementum Geometriæ (1593, p. 240)

Ad Problema Quod omnibus Mathematicis totius orbis construendum proposuit Adrianus Romanus (1595, p. 305) (henceforth *Ad Problema*)

Pseudo-Mesolabum & Alia Quædam Adiuncta Capitula (consisting of the two works *Pseudo-Mesolabum* and *Adiuncta Quædam Capitula*. 1595, pp. 258, 275)

Ad Angularium Sectionum Analyticen (1615, p. 287)

De Æquationum Recognitione et Emendatione Tractatus Duo (consisting of the two works *De Recognitione Aequationum* and *De Æquationum Emendatione*. 1615, pp. 82, 127)

Ad Logisticem Speciosam, Notæ Priores (1631, p. 13) ((Freguglia 2008, p. 53) dates the composition to before 1593. I have not located any copy of the 1631 printing, so I refer only to the 1646 *Opera Mathematica*.)

J. Winfree Smith's English translation of Viète's *Isagoge* is printed in (Klein 1968), and T. Richard Witmer's English translations of several of Viète's works are published in (Viète 1983). I have taken or adapted translations of many passages

from these sources. Translations without a reference are mine. Sometimes I give Witmer's modern rendering of Viète's notation to clarify for the reader how that notation functions.

4.2 Overview of Viète's Program

There were two main currents in sixteenth-century mathematics, which were at once conflicting and entwined. One stemmed from the classical Greek texts of Euclid, Apollonius, Nicomachus, and others, in which the arithmetical unit is indivisible and geometric magnitudes are free of numerical measure. The other was the practical arithmetic, algebra, and geometry that was known to have come from medieval Arabic, and which in fact traces back to earlier Old World cultures. Numbers in this current can be partitioned indefinitely and were routinely applied to lines, surfaces, and solids.

The main interest of François Viète lay with Ptolemaic astronomy, in which these two currents met but did not mingle. Hipparchus of Bithynia (active 147–127 BC) is credited with being the first astronomer to apply Greek geometrical models of planetary motion to observational data. The models are cast in the non-arithmetized geometry of Euclidean magnitudes, while the observations consist of numerical measurements often recorded using physical instruments and arranged in tables. Hipparchus's approach was adopted by Ptolemy (second century CE), whose authority guaranteed its continuation down to the early seventeenth century, when the models, based as they were in uniform circular motion, were made obsolete with the work of Johannes Kepler.

Viète was one of the last astronomers to work in this Greek tradition. From his first published book of 1579, covering geometric models and the trigonometric tables derived from them, to the draft on planetary astronomy that he was working on at his death in 1603, his mathematical interests lay squarely with Euclidean geometry in the service of astronomical calculation. His publications in the intervening years deal directly or indirectly with the geometry and arithmetical calculations that underlie Ptolemaic astronomy, including his theorems on triangles, continued proportions, angular sections, and the numerical solutions to equations. Throughout these works, Viète remained faithful to the style and precepts of Greek authors.

One Greek work particularly influential for Viète's approach was Pappus of Alexandria's *Collection*, which was finally published in Commandino's Latin translation in 1588. Book VII covers the "domain of analysis," a topic that is barely touched on in other extant Greek works. In it, Pappus lists several books on, or important for, analysis by different Greek authors, and he describes their contents, including some of the propositions they contained (Pappus 1986, 66ff, 84). Most of these books are lost, and Viète was perhaps the first to compose a restoration of one of them. Based on the description given by Pappus (1986, 90–94), he wrote his *Apollonius Gallus* as a reconstruction of the *Tangencies* (*Ἐπαφῶν*, sometimes translated as *On Contacts*) of Apollonius. But more importantly, Pappus's Book VII likely played a role in inspiring Viète to transform the geometrical equalities prevalent in Greek works into a geometrical algebra that parallels the numerical

algebra practiced in his time. He introduced his *logistice speciosa* in his 1591 *Isagoge*, and that book opens by situating the new algebra in the context of ancient analysis.

4.3 Viète's Magnitudes and Numbers

The letters that constitute the objects of Viète's algebra will not make sense until we first establish how he understood geometric magnitudes and numbers, and for that we need to begin with the views of his contemporaries. The term *magnitudo* in sixteenth-century mathematics meant any geometric magnitude of dimension 1, 2, or 3, to the exclusion of numbers, time, angles, and other quantities (§2.3). Greek authors stated and proved propositions about these magnitudes without recourse to numerical measure, and this is naturally respected in medieval and early modern translations. But commentators often assigned numbers to the magnitudes in diagrams to illustrate the relationships between them, and many of these numbers are fractions or irrational. For example, in his commentary to Proposition II.12 of Euclid's *Elements*, Federico Commandino makes line AC 16 and line AB 20, and then approximates $\sqrt{231}$, the length of line AD, as $15\frac{1}{5}$, and with that he finds the area of triangle ABD to be approximately $98\frac{4}{5}$. In his commentary to Proposition XIII.1 he keeps the irrational roots, measuring lines with numbers like $5-\sqrt{5}$ and $\sqrt{20}+2$ (Euclid 1572, ff. 34b, 230a). The numbers that these mathematicians assigned to magnitudes were not intended as absolute measures. Rather, they illustrate the relative sizes of the magnitudes in a diagram, much like the numbers 1, 2, and $\sqrt{3}$ that we assign to the sides of the 30°–60°–90° triangle (§3.7).

Viète respected the ancient tradition of stating and proving propositions in non-arithmetized geometry, and like his contemporaries, he often assigned numbers to the lines after finishing his proof to illustrate their relative sizes. His numbers, too, are sometimes shown as approximations and sometimes as irrational roots. Figure 1 shows two examples, the first from *Variorum* in which he characteristically approximates the lengths of sides not with fractions, but with large integers. He gives three different arithmetizations, where each side in turn is set to be 100,000 and the other sides are calculated from it (Viète 1593a, f. 42a; 1646, 420). In the second, from *Zeteticorum*, he shows the sides of a triangle as $\frac{153}{61}$, $\frac{158}{61}$, and $\sqrt{13}$ (the latter as "*l* 13," where "*l*" stands for *latus* (side)) (1593c, f. 16a; 1646, 67) (§3.8).

In some settings in sixteenth-century mathematics, the word *numerus* took its discrete meaning after Aristotle and Euclid, so that numbers consist only of positive integers. In other settings, the word took its continuous meaning that was common in practical arithmetic, algebra, and astronomy, where numbers include fractions and irrational roots. In Viète's works, all numbers are regarded as being the measures of magnitudes, so they are of the latter, continuous type. The underlying context of all of Viète's mathematical works is geometry, and because there is no canonical geometric unit, the numbers with which he measures his lines express the sizes of the magnitudes relative to each other, like those in Commandino and other geometers (§3.8).

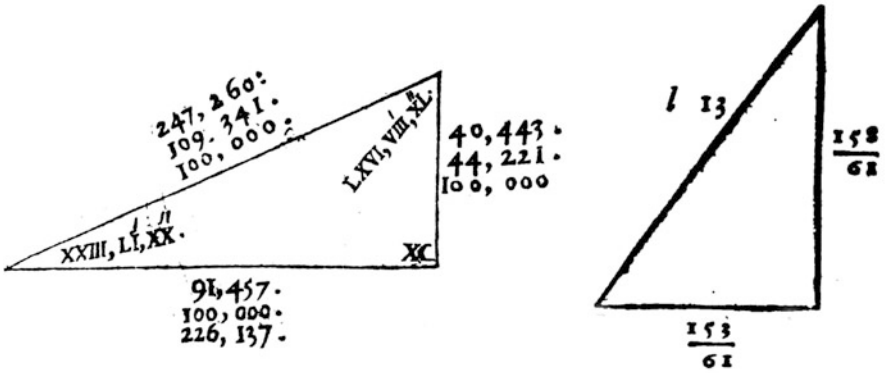


Fig. 1 Two of Viète's triangles, one showing approximations, the other fractions and an irrational square root

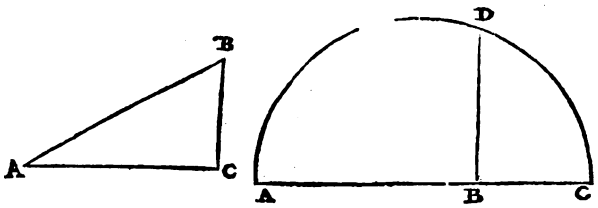


Fig. 2 The diagram for Viète's Proposition XIIII in *Effectionum Geometricarum*

One major way that Viète's geometry diverges from sixteenth-century practice is that he stepped beyond the third dimension. This occurs most clearly in Propositions XIIII and XV of *Effectionum Geometricarum* (1593), where he works with four-dimensional magnitudes based in the diagram shown in Fig. 2. He begins with right triangle ABC with right angle at C, and next to it he shows the same lines AB and BC laid out on a line that becomes the diameter if a semicircle. From the proportion $AB : BD :: BD : BC$, he gets $AB^2 : BD^2 :: BD^2 : BC^2$, and substituting $AC^2 + BC^2$ for AB^2, he arrives at $(AC^2 + BC^2) : BD^2 :: BD^2 : BC^2$. He then converts the proportion into an equality, first in *logistice speciosa*, and then again in terms of the magnitudes in the diagram, which gives $BC^4 + BC^2 \cdot AC^2 = BD^4$. This is expressed in purely geometric language and notation unrelated to any algebra:

> Quadrato-quadratum ex BC plus planoplano sub Quadrato ex AC & Quadrato ex BC aequetur Quadrato-quadrato ex BD. (Viète 1593b, f. 4b; 1646, 235)

He does not address the issue of the existence of these magnitudes, nor of the four-dimensional magnitudes in Proposition XV (§3.3).

In another direction, Viète found himself deriving relationships between magnitudes of arbitrarily high dimension in the setting of the trigonometry of multiple angles. He published his theorems on angular sections in two works: Chapter 10 of *Variorum* and Chapter 9 of *Ad Problema*. After Viète's death, Alexander Anderson added proofs and other theorems and published them in *Ad Angularium Sectionum*

Analyticen (1615). The main theorem on angular sections is in some sense a corollary of the corollary stated by Ptolemy to what is called Ptolemy's Theorem (Ptolemy 1998, 51ff). Converting Ptolemy's corollary from chords of arcs to sines and cosines, we can write it as:

$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$$
$$\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b).$$

Viète expressed these rules in terms of the ratios of the sides of right triangles, and by applying them repeatedly for angles $\theta + \theta$, $2\theta + \theta$, $3\theta + \theta$, etc., he derived rules for what we would write as $\cos(n\theta)$ and $\sin(n\theta)$ in terms of $\cos(\theta)$ and $\sin(\theta)$. These rules constitute Viète's main theorem, which is Theorem III in *Ad Problema* and *Ad Angularium Sectionum Analyticen*, and Theorem I in *Variorum*. In all three books, he first states the rules verbally, then in terms of *logistice speciosa*, and this is followed with several numerical examples. In *logistice speciosa*, he calls the hypotenuse of the single-angle triangle Z, the base D, and the height B. If we call the base of the triangle with angle $n\theta$ D_n, its altitude B_n, and its hypotenuse Z_n, then his proportions can be written as

$$D_2 : Z_2 :: D^2 - B^2 : Z^2$$
$$B_2 : Z_2 :: 2DB : Z^2$$
$$D_3 : Z_3 :: D^3 - 3DB^2 : Z^3$$
$$B_3 : Z_3 :: 3D^2B - B^3 : Z^3$$
$$D_4 : Z_4 :: D^4 - 6D^2B^2 + B^4 : Z^4$$
$$B_4 : Z_4 :: 4D^3B - 4DB^3 : Z^4$$
$$D_5 : Z_5 :: D^5 - 10D^3B^2 + 5DB^4 : Z^5$$
$$B_5 : Z_5 :: 5D^4B - 10D^2B^3 + B^5 : Z^5.$$

The rules for $D_2 : Z_2$ and $B_2 : Z_2$ are his versions of the rules for $\cos(a + b)$ and $\sin(a + b)$ shown above. Because the succeeding calculations are an easy corollary to those known rules, Viète does not give a proof or show a diagram. I drew the diagram in Fig. 3 to conform to the diagrams he shows for his numerical calculations. In the diagram, $Z = Z_n$, but this is not required in the theorems.

Here is how Viète states the theorem for the base of the quadruple angle triangle in all three books:

> In ratione quadruplâ, Hypotenusa secundi [Z_4] fit similis Quadrato-quadrato-Hypotenusæ primi [Z^4]. Basis [D_4], Quadrato-quadrato basis primi, minùs plano-plano sexies sub quadrato perpendiculi primi & quadrato basis eiusdem, plus quadrato-quadrato perpendiculi [$D^4 - 6B^2D^2 + B^4$]. (Viète 1593a, f. 15b; 1595a, f. 15b; 1615a, 12; 1646, 289, 316, 372)

The language here belongs to geometry, not to arithmetic or algebra. The meanings of the terms "*hypotenusa*" and "*basis*" are "hypotenuse" and "base," respectively, and "*plano-plano*" indicates a four-dimensional geometric magnitude just as it does in the example from *Effectionum Geometricarum* shown above.

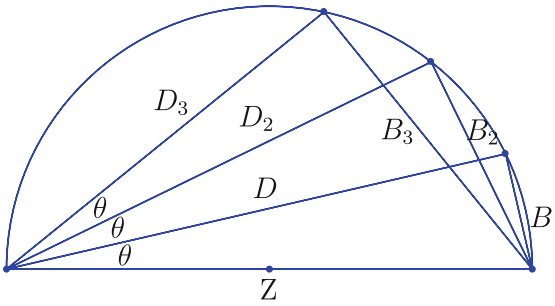


Fig. 3 Reconstructed diagram for the theorem on angular sections

B E	184	160	
B F	164	953	6
B G	139	159	056
B H	107	778	549 76
B I	72	096	901 529 6
B K	33	531	337 238 016

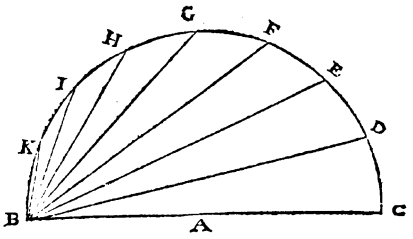


Fig. 4 Some numerical calculations with angular sections (Viète 1595, f. 11b; 1646, 319) Ad Problema

Figure 4 shows one of the sets of calculations that follow. Viète began by assigning the radius AC of a semicircle to be "100 000" and BD to be "196 000," from which he calculated BE through BK. Here, too, it is only the relative sizes of the numbers that matter. Viète's higher-dimensional magnitudes may have no perceptible existence, but the rules in which they appear give the correct proportions for the sides of triangles when converted to numbers (§3.5)

Just as Viète gave no explanation for his magnitudes in *Effectionum Geometricarum*, he gives none for these higher-dimensional magnitudes, either. But he does address magnitudes of dimension greater than three in points 26 and 27 of Chapter VIII in *Isagoge*:

> 26 Since all magnitudes are either lines or surfaces or solids, of what earthly use are proportions above triplicate or, at most, quadruplicate ratio except, perhaps, in sectioning angles, so that we may derive the angles of figures from their sides or the sides from their angles?
>
> 27 Hence the mystery of angular sections, perceived by no one up to the present either arithmetically or geometrically, is now clear, and shows
> How to find the ratio of the sides, given the ratio of the angles;
> How to construct one angle [in the same ratio] to another as one number is to another. ((Viète 1591, f. 9a; 1646, 12). Translation adapted from Viète (1983, 32).)

Instead of justifying the existence of higher dimensional magnitudes, Viète explains that they are *useful*, and the use to which he puts them is in numerical calculation.

Just as Ptolemy applied his version of the corollary to construct his table of chords, Alexander Anderson later remarked that Viète's theorems on angular sections can be applied to construct trigonometric tables in increments of $1'$ (Viète 1615, 47; 1646, 303–304) (§3.6).

To summarize, Viète's geometric magnitudes, like those of other mathematicians of his era, conform to classical Greek works in that they do not possess any inherent numerical size. But they are metrizable, and the numbers frequently assigned to them illustrate only their relative sizes, and not any absolute size. Viète parts from his contemporaries by freely admitting magnitudes of dimension greater than three. He gives no explanation of what kind of existence these magnitudes might have, but that would not have mattered to him because his interest lay with numerical computation. Viète's numbers are the continuous, practical numbers of the astronomers (and the algebraists) that include fractions and irrational roots, and their only purpose in his works lies with the measurements of geometric magnitudes. They have no independent existence.

In the sixteenth century, geometric magnitudes could be measured in two different ways. Classical Greek magnitudes are measured through ratio and proportion, and Viète and other Europeans of his time set up and manipulated ratios and proportions in their geometric propositions. But these same early modern European mathematicians also routinely assigned numerical measures to magnitudes, so one could also perform arithmetical calculations that parallel those non-arithmetized proportions. This gives magnitudes a dual nature when it comes to measurement, and these two ways correspond to the two algebras that Viète presents in his books.

4.4 Viète's Two Algebras

Viète called his new algebra *logistice speciosa* (specious logistic) to distinguish it from traditional numerical algebra, which he called *logistice numerosa* (numerical logistic). He borrowed the notation for the latter from Xylander's 1575 translation of Diophantus's *Arithmetica*. Translating the terms from Diophantus, Xylander called the first degree unknown a "Numerus," which he abbreviated as "N." The second degree is a "Quadratum," abbreviated as "Q," and the third degree is a "Cubus," abbreviated as "C." Higher powers, still translated from Diophantus, are "Quadrato-quadratum" (QQ), "Quadratocubus" (QC), and "Cubocubus" (CC). Here is an example equation in this notation from Viète's *De Emendatione Aequationum*: "30Q. − 156 N. − 1C., *æquatur* 440." ((1615b, 123.7; 1646, 155.30). The 1646 printing is identical except that it omits the periods.) Distorted into modern notation, it would be $30x^2 - 156x - x^3 = 440$. The recurring coefficient of 1, the absence of operations, and other features of *logistice numerosa* throughout Viète's works confirm that it belongs to the premodern algebra of the sixteenth century. And because all numbers in Viète are the measures of lines, planes, solids, etc., his *logistice numerosa* functions as an algebra for arithmetized geometry (§§4.1, 5.6).

Logistice speciosa, by contrast, is the algebra that Viète developed for geometry in which the knowns and unknowns are the relative sizes of the magnitudes themselves, and not numbers that measure them. This truly geometrical algebra was inspired in part by the computational geometry he read in Greek authors. Some propositions in Greek geometry concern spatial relationships of the magnitudes, while others specify only metrical relationships. (See, for example, Marie (1884, 4) and Panza (2006, 279).) In the latter case, it often happens that some of these magnitudes cannot exist in position. For example, Apollonius could write in proposition I.37 of his *Conics* that the rectangle contained by ΔZ, ZE is equal to the sum of the rectangle contained by ΔE, EZ and the square on ZE, even if no such rectangles can be drawn because the points Δ, Z, and E are colinear (Apollonius 1891, 112.5). Impossible magnitudes like these enter into the arguments only as (abstracted) sizes relative to other magnitudes deriving from the diagram. (Others, too, have made this observation, such as Mueller (1981, 42).) Moreover, the spatial relationships of some magnitudes are not relevant at all to the proposition, and in those cases they are drawn apart from the other magnitudes and are labeled with single letters. Euclid's Proposition VI.17, for example, deals with proportionality and shows four individual lines labeled A, B, Γ, and Δ. In other propositions, it is an area that is designated by a single letter, like regions Γ and Δ in *Elements* Proposition VI.29.

Viète reworked these geometrical equalities into an algebra, partly modeled on *logistice numerosa*. The knowns and unknowns cannot be based in any particular diagram showing spatial relationships, so he chose single capital letters to represent them, and he followed the order of the transliterated Greek alphabet, calling his "species" A, B, G, D, Z, etc. This notation would have made sense to mathematicians of the time, given the presence and purpose of single-letter magnitudes in Greek geometry (§§5.4–5).

Viète presented the outlines of his *logistice speciosa* in the short yet enigmatic 1591 book *In Artem Analyticem Isagoge* (*Introduction to the Analytic Art*). After three chapters covering preliminary stipulations and laws, he introduces the species:

On the rules of calculating with species. Chapter IIII.

Numerical logistic employs numbers, specious logistic employs species or signs for things, as, say, the letters of the alphabet. There are four basic rules for calculating with species just as there are for calculating with numbers.

RULE I.
To add a magnitude to a magnitude.

Let there be two magnitudes A and B. One is to be added to the other. ((Viète 1591, f. 5a; 1646, 4). Translation adapted from (Viète 1983, 17–18).)

Viète calls the species A, B, etc. of his algebra "magnitudes" (*magnitudines*) here and throughout his books. Unknown magnitudes are designated with vowels, and known magnitudes with consonants. These species, or magnitudes, can be raised to different powers, and these powers possess dimension. A simple "A," "B," "G," etc. is of the first dimension and is called a "side" (*latus*). (In rare cases when it does not

designate the measure of a line, it is called a "root" (*radix*), such as we see in Proposition LII in *Notæ Priores* (Viète 1646, 37).) The two-dimensional squares of these sides are called "A quadratm," "G quadratum," etc. Their cubes are called "A cubus," "G cubus," etc., their fourth powers, which are naturally four-dimensional, are "A quadrato-quadratum," "G quadrato-quadratum," etc., and higher powers continue to follow the pattern for the names in Xylander, translated from Diophantus: "quadrato-cubus," "cubo-cubus," "quadrato-quadrato-cubus," quadrato-cubo-cubus," etc. There is another series of terms used to label magnitudes whose lower-dimensional versions do not enter into the problem. A two-dimensional magnitude whose side does not appear is called "A planum," "B planum," etc. A three-dimensional magnitude is called "A solidum," "B solidum," etc., continuing with the four-dimensional "plano-planum," the five-dimensional "plano-solidum," etc., following the same pattern as the powers (§4.1).

The sample equation in *logistice numerosa* given above is a numerical version of this equation in *logistice speciosa*, here reproduced as shown in the 1615 printing (Viète 1615b, 122):

$$\left.\begin{array}{l} \text{B in A quad 3} \\ -\text{D plano in A} \\ -\text{A cubo.} \end{array}\right\} \text{æquetur} \left\{\begin{array}{l} \text{B cubo bis.} \\ -\text{D plano in B.} \end{array}\right.$$

In the 1646 *Opera Mathematica*, the same equation is written on one line: "B in A quadr. 3 − D plano in A − A cubo, æquetur B cubo 2 − D plano in B" (Viète 1646, 155.23). Witmer transforms it into more modern notation as $3BA^2 - D^PA - A^3 = 2B^3 - D^PB$ (Viète 1983, 303).

Certain features suggest that this notation should be read differently than the premodern notations we described earlier. First, a "1" is not written when there is only one of a term. In this equation, the "A cubo" has no coefficient that could count or measure the term, and this applies to every example in Viète's works. This means that the letters are not kinds, but values. Second, these values are sometimes multiplied together. In the first term, "B" is expressly multiplied by "A quad" by the presence of the preposition *in* ("by"). Similarly, "D plano in A" expresses the multiplication of the two-dimensional "D plano" by the one-dimensional "A." Perhaps the most noticeable difference is that known values are designated with letters and not numbers, which makes them provisionally undetermined. This feature has been cited as a great advance by historians dating back to Montucla's 1758 *Histoire des Mathematiques* (1758, 488). The "3" and the "bis" ("2"), on the other hand, function more like premodern "numbers" of terms to indicate how many of that term there are.

Proposition XVI of *Supplementum Geometriae* nicely illustrates the similarities and differences between three ways of comparing magnitudes. There a particular relationship is first written as a geometrical equality, then as an equation in species, and finally as a numerical equation. For the equality, A, B, C, and E are points in the diagram (Viète 1593d, f. 17b; 1646, 249):

> cubus ex AC minus triplo solido sub AC & quadrato ex AB, æqualis est solido sub CE & quadrato ex AB

For the equation in species, the unknown line AC is called A, and lines CE and AB are taken to be equal in length and are called Z:

> A cubus minus Z quadrato ter in A, æquetur Z cubo

For the numerical equation, Viète sets Z to be 1 and A to be 1N:

> 1C − 3N, æquatur 1

The equality is naturally stated using geometrical language, including the phrase we can translate as "triple the solid under CE & the square on AB." In the equation in species, the language of arithmetical operations is substituted, where this becomes "thrice Z squared [multiplied] by A."

We saw above how Viète produced proportions involving magnitudes of arbitrarily high dimension. It is by virtue of Viète's "supreme stipulation," in Chapter II of *Isagoge*, that these magnitudes can be freed from their ratios and enter into equations as independent objects:

> But the supreme stipulation of equations and proportions and which is all-important in analysis is:
>
> 15 If there be three or four magnitudes and the result of the multiplication of the extreme terms is equal to the result of the multiplication of the mean by itself or to the product of the means, then those magnitudes are proportional. And conversely,
>
> 16 If there be three or four magnitudes, and the first is to the second, so that second, or else some third, is to another, the product of the extremes will be equal to the product of the means.
>
> And so, a proportion can be called the composition (*constitutio*) of an equation, an equation the resolution (*resolutio*) of a proportion. (Translation adapted from Klein (1968, 324))

By this stipulation, the proportion $a : b :: c : d$ is equivalent to the equality $ad = bc$. It generalizes Euclid's *Elements* Propositions VI.16 and XI.34 by removing any restriction on the dimension of the terms in the equality. It is what allowed Viète to resolve the proportion in Proposition XIIII from *Effectionum Geometricarum* into a geometric equality with four-dimensional magnitudes.

In numerical logistic, one can add 1Q to 3N to get 1Q + 3N because numbers are homogeneous. But in geometry, magnitudes of different dimension are heterogeneous, so one cannot add, say, A quadratum to A. Viète solved this problem by multiplying the lower degree terms by known magnitudes of such dimension that all terms become homogeneous. For some given B, one can add "A quadratum" to "B in A" because they are both now two-dimensional. Thus his "Law of Homogeneity,"

laid out in Chapter III of *Isagoge*, states that only terms of the same dimension (power) can be added or subtracted.

Specious logistic was designed for solving problems in geometry. Given magnitudes are named with consonants, an unknown is named with a vowel, and then an equation or proportion is set up in terms of these species and simplified. The resulting equation or proportion can then be read as a formula from which one can either perform a geometric construction to show the unknown magnitude, or numbers can be assigned to the givens and the unknown can be calculated numerically (Viète 1591, f. 8a; 1646, 10). These two ways of exhibiting the solution are a reflection of the dual nature of geometric magnitudes. In his *Zeteticorum*, Viète solves 82 problems with his new algebra, giving only numerical solutions. Two of his successors, Jean-Louis Vaulezard and Marino Ghetaldi, later took up the task of producing answers by means of geometrical construction (Ghetaldi 1630; Viète 1630).

4.5 Summary of Viète's *logistice speciosa*

In creating an algebra for the non-arithmetized sizes of geometric magnitudes, Viète caused a radical reconception of polynomials, and of algebraic expressions in general. Now coefficients (knowns) and unknowns are both distinct objects of the same kind – measures of geometric magnitudes – that are necessarily multiplied together. And because monomials in *logistice speciosa* are now often stated with multiplications, and the final equation or proportion in a solution serves as instructions for either performing a geometrical construction or an arithmetical calculation, the signs "+" and "−" should be reinterpreted as operations, too.

Given quantities can only be undetermined in such a system. While one can name specific numbers, like 6 or $4\frac{4}{5}$, one cannot name the measure of a specific magnitude that is distinct from other magnitudes. Thus, given magnitudes are necessarily undetermined and are shown as consonants. The simplified equation or proportion obtained in solving a problem now conveniently becomes a formula stated in terms of the givens, so one can see precisely how the solution is obtained from the original parameters.

Two other features of this geometrical algebra differentiate it from numerical algebra. Because there is no canonical unit, all terms possess dimension, and this dimension must be respected when forming expressions. One can add and subtract only homogeneous terms, and multiplication results in a term of higher dimension than the multipliers. Because of this, the minimum dimension of a term is 1, so there can be no zero degree terms or reciprocals of terms as there are in numerical algebra. Division is thus possible only of a higher degree term by a lower degree term, as Viète tells us in Chapter IIII, Rule IIII in *Isagoge* (Viète 1591, f. 6a; 1646, 6). (Thus the expression shown in (Klein 1936, 180; 1968, 173) is Klein's fabrication, since it is a third degree term divided by a third degree term). Also, roots must have positive

integer degree. While Abū Kāmil could write "a root of three cubes" (in modern notation, $\sqrt{3x^3}$) (2012, 449.8), Viète could not take the square root of "A cubus 3" because the dimension would be fractional.

The letters of Viète's algebra are not symbolic entities that transcend arithmetic and geometry, as Jacob Klein argued (1936, 1968). Klein read Viète's rules for operating on the species in Chapter IIII of *Isagoge* as being axioms governing the relationships between the species. These rules, according to Klein, are general enough to cover both arithmetical and geometrical calculation, so that the species A, B quad., etc. in an algebraic expression can represent either numbers or magnitudes. The species themselves "are, therefore, comprehensible only within the language of *symbolic formalism*" (Klein (1968, 175), his emphasis). By this view, one can solve either arithmetic or geometry problems by translating the knowns and unknowns into *logistice speciosa*, and once the equation or proportion has been simplified, the solution can be found by reinterpreting the letters in terms of the original numbers or magnitudes.

Klein was apparently unaware that Viète worked with higher dimensional geometric magnitudes, and he passed over the several passages in Viète's works that unequivocally associate the species with magnitudes and not with numbers. His error was taken up by many historians, including even Michael Mahoney, who writes that Viète's *logistice speciosa* "is, to use modern terms, a language of uninterpreted symbols" (1994, 39), and Henk Bos, who saw it as "a method of symbolic calculation concerning abstract magnitudes" (Bos 2001, 147). I address Klein's account more thoroughly in 2018, §§6.2–3.

While readers of Viète in the past century have had a difficult time understanding the nature of the species, his immediate successors seem to have had little problem. Viète is quite explicit that his letters are magnitudes, and contemporary geometers all understood that magnitudes can enter into calculations via ratio and proportion, and they can also be assigned numerical values. It is not the species, but magnitudes themselves that can be measured in two ways.

5 Fermat's Algebra

5.1 Fermat, Student of Viète

When Pierre de Fermat (ca. 1607–1665) was in his early 20s, he traveled to Bordeaux, where he met students of Viète and studied Viète's works. As Mahoney observes, Viète "shaped Fermat's mathematical career as decisively as if the two men had stood in a living master-student relationship," and that "everything that Fermat did in mathematics, however original, bears Viète's imprint" (Mahoney 1994, 26). In Bordeaux, Fermat quickly and wholeheartedly embraced Viète's new algebra, which, like Viète, he saw as a tool with which to restore the analysis practiced by ancient geometers. Fermat appropriated precisely Viète's notation, too, which he

continued to use to the end of his career even after the notation of Descartes had eclipsed it. This notation is already the main tool of analysis in Fermat's *Method for Determining Maxima and Minima and Tangents to Curved Lines*, completed by 1629. (*Methodus ad Disquirendam Maximam et Minimam*, published in Fermat (1891, 133–179). The title and its translation are taken from Mahoney (1994, 55).) For example, to solve the problem of finding the cone of maximum surface that can be inscribed in a given sphere, he draws the semicircle shown in Fig. 5, calling line AB *B* and line BC *A*. (Following the edition of his works (Fermat 1891), we write Fermat's species in italic font, and points in diagrams in regular font.) From there, he determines that he needs to maximize the fourth-degree amount "*B* cub. in *A* + *B* quad. in *A* quad. − *B* in *A* cub. − *A* quad.quad." (Fermat 1891, 157).

It is also with this algebra that Fermat developed his analytic geometry . Like Viète, he undertook the restoration of a lost work of Apollonius described by Pappus in Book VII of the *Collection*. Fermat chose Apollonius's *Plane Loci*, and it was during his struggle to prove Proposition II.5 of his restoration, and through his subsequent investigations of loci generally, that he hit upon the idea of transforming the *symptômata* of conic sections, that is, geometrical equalities that characterize the curves, into Vietan equations in two unknowns (Mahoney 1994, 112ff). Fermat completed his restoration of *Plane Loci* around 1635, and the following year he outlined a classification of equations and their corresponding curves in *Introduction to Plane and Solid Loci*. (*Ad Locos Planos et Solidos Isagoge*, published in Fermat (1891, 91–110).)

Where Viète had been concerned with developing geometrical models for astronomical calculation, Fermat's focus was on the geometry itself. So unlike Viète, Fermat never assigns numerical measures to the species, and in work after work he relates the algebraic expressions and equations to the magnitudes in geometric diagrams.

5.2 Fermat's Interpretation of the Species

With about four decades separating them, and given their different interests, we might wonder if Fermat interpreted the species differently than their creator did. Reading over Fermat's works, one finds that he rigorously observed the law of homogeneity and that he followed Viète in working with magnitudes of arbitrarily high dimension. At least for the mechanics of the algebra, there is no indication that Fermat's understanding of the species diverged from that of Viète. Mahoney, however, argued that Fermat viewed his algebraic terms as all being homogeneous:

> Although Fermat paid lip service to Viète's Law of Homogeneity, his use of algebra in geometric situations clearly shows that he no longer attached dimension to the degree of an expression. Fermat operated as if all combinatory operations on geometrical magnitudes yielded homogeneous products. At no time, however, does he justify his behavior. (Mahoney 1994, 42)

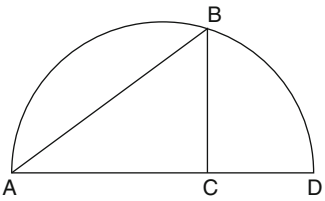


Fig. 5 From Fermat (1891, 156)

Mahoney came to this conclusion based in part on this claim:

> . . .every equation of Viète's analytic art had dimension, and the dimension was directly linked to the degree of the equation. For example, the solution of a cubic equation corresponded to a solid construction in three-dimensional space . . . (Mahoney 1994, 42)

Viète never solves a problem in his *Zeteticorum* with a geometrical construction, so such solutions to equations will not be found among his worked-out problems. But he does give the constructions necessary for solving particular problems in two other works: *Effectionum Geometricarum*, which covers basic ruler-and-compass constructions, and *Supplementum Geometriae*, which covers propositions on the trisection of an angle. In 7 of the 45 propositions presented in these 2 books, Viète states algebraic equations that can be solved by the given construction, and of those 7, the degree of the equations in 4 of them do not match the dimension of the construction. (Propositions IX, X, XI, XIIII, and XV in *Effectionum Geometricarum*, and Propositions XVI and XXIIII in *Supplementum Geometriae*. Only the first three in the first work show a match between degree and dimension.)

In the introduction to *Supplementum Geometriae*, Viète postulates a neusis construction: "To draw a straight line from any point to any two given lines, the intercept between these being any possible predefined distance" (Viète (1593, f. 13a; 1646, 240), translated in Bos (2001, 168)). In Proposition IX, he applies this construction to the problem of trisecting a given angle in the context of a two-dimensional diagram. Then, in Proposition XVI, with essentially the same diagram, he gives a third-degree algebraic equation that corresponds to that solution: "A cubus minus Z quadrato ter in A, æquetur Z cubo." Viète also arrives at cubic equations in Proposition XXIV, for the problem of inscribing a regular heptagon in a circle. In both propositions, the degree is 3 and the dimension is 2.

A similar situation arises in Propositions XIIII and XV in *Effectionum Geometricarum*. These are the propositions in which Viète works with four-dimensional magnitudes, discussed above in Sect. 4.3. The diagrams are again two-dimensional, and this time the equations are of the fourth degree. After showing the relationship between the lines as a proportion involving ratios of two-dimensional magnitudes, Viète gives a corollary showing the fourth-degree equation that this construction will solve. In Proposition XIIII, this is "A quadrato-quadratum, plus B quadrato in A quadratum, æquetur D quadrato-quadrato" (Viète 1593b, f. 5a). Given the equation, one can appeal to Theorem I in Chap. XII of *De*

Recognitione Æquationum to make a substitution that transforms it into a quadratic equation, whose solution is constructible (Viète 1615b, 27–28; 1646, 99; 1983, 191). The equation in Proposition XV can be similarly reduced to a quadratic equation by Theorem III in the same chapter. There is in fact no rule in Viète linking the degree of an equation with the dimension of the solution. We can only infer from the few examples in these two works that it is the dimension of the original problem that corresponds to the dimension of the solution.

Mahoney also misunderstood Viète's higher-dimensional magnitudes. To account for their geometrical meaning, Mahoney wrote: "their interpretation was possible, though it required the introduction of 'mechanical' procedures like the *Pseudo-Mesolabe* for generating fourth, fifth sixth, etc., proportionals to two given line segments" (Mahoney 1994, 42). This *Pseudo-Mesolabe*, described in Proposition II in Viète's *Pseudo-Mesolabum* (1595b, f. 2b; 1646, 259), serves specifically to produce four lines in continued proportion. The propositions that follow likewise all deal with exactly four lines. There is no mention of producing five or more lines in continued proportion, so the mechanical procedure only gives lines whose relationship can be expressed with a cubic equation. Nor is there mention in any other work of Viète's of a "mechanical" procedure to produce a series of five or more magnitudes in continued proportion (though Descartes later would do it with his Mesolabe (Descartes 1637, 317ff, 370)). Instead, Viète generates proportional series in terms of the higher powers of his species. Proportional series are introduced in Proposition II of *Notæ Priores* (1646, 13). Two magnitudes A and B are followed by "$\frac{B\ \ quadr.}{A}$," "$\frac{B\ \ cubus}{A\ \ quadr.}$," "$\frac{B\ \ quad.quadratum}{A\ \ cubo}$," etc. He finds the proportionals between two given magnitudes A and B in Proposition V (1646, 15), with the example of quintuplicate ratio. He first makes the series of five-dimensional species in which the extreme terms are "A quadrato-cubus" and "B quadrato-cubus." Witmer translates this series as A^5, A^4B, A^3B^2, A^2B^3, AB^4, B^5 (1983, 37). Viète then takes the fifth roots of each term so that the extreme magnitudes are A and B. In Viète, higher dimensional magnitudes are used to produce the continued proportions, not the other way around. As I explained in the previous section, Viète did not give any geometric meaning to his magnitudes of dimension greater than three. They arise naturally in the theorem on sectioning angles, and he worked with them because they were useful for numerical calculation.

With the misunderstandings that the degree of the equation corresponds to the dimension of the problem and that Viète's higher dimensions are artifices that are justified through proportional lines, Mahoney concluded that

> both Fermat and Descartes saw clearly the difficulties that Viète's translation of geometry into algebra caused if one tried to apply it to a realm of problems which themselves involved dimensionality, e.g. locus problems. If one wishes to determine the nature of a locus, which itself may be dimensionless or have a dimension depending on the conditions, one cannot use a vehicle of analysis that is already burdened by dimensionality. (Mahoney 1994, 42)

But when viewed against Viète's actual practice, there is nothing about his species that would prevent them from being employed to solve locus problems, and

consequently there is nothing in Fermat's writings to suggest that he held a different view of the species than Viète. But one should wonder, then, how Fermat understood higher-dimensional terms like the four-dimensional expression quoted above, or the terms in an equation like "*Aqq*. æquale *Zpl*. in *Aq*. − *Zs*. in *D*" from *Introduction to Plane and Solid Loci*. ((Fermat 1891, 109). Fermat abbreviates *quadratum*, *planum*, and *solidum* as *q*, *pl*, and *s*.) Characteristically, he does not offer an answer. All I can suggest is that like Viète, he saw higher-dimensional magnitudes in his intermediate calculations as useful, since they give the correct results when the simplified equations are interpreted in two- or three-dimensional space.

6 Descartes's Algebra

Descartes published his only mathematical treatise, *La Geometrie*, as the third appendix to his *Discourse on Method*. The book appeared in 1637, the same year that Fermat shared his *Introduction to Plane and Solid Loci* with the mathematicians in Paris (Mahoney 1994, 72). Descartes, like Fermat, was interested specifically in geometric problem-solving (Bos 1990; 2001, 229), and like Fermat, his analytic geometry was inspired at least in part by his solution to a locus problem originating in Book VII of Pappus's *Collection*. This problem had been proposed to him in 1631 by Jacobus Golius, a professor at University of Leiden, and Descartes had solved it by January of the following year (Bos 2001, 271). It asks for a point C such that lines extending from C to given lines and forming given angles with them make certain products equal (Sasaki 2003, 206; Bos 2001, 313ff; Pappus 1588, f. 165a; 1986 I, 120). The solutions to this indeterminate problem form a curve, and Descartes's solution to the general case of four lines, given in Books I and II of *La Geometrie*, illustrates his method of representing loci by equations in two unknowns. The method is equivalent to that of Fermat: one unknown is measured along a given line from a given point, and the other is measured at some angle from it. He sets up equations for other locus problems, too. One of these is the equation that describes the curve traced by what Bos calls the "turning ruler and moving curve" procedure (Bos 2001, 278), which Descartes employs later to trace the solutions to equations. He also forms an equation for the solution to a special case of the five-line Pappus problem, and algebraic expressions and equations are also instrumental in his solutions to the other problems in Book II, both for finding normal lines to curves and for constructing ovals satisfying particular optical properties. (The equations for the three loci are given at Descartes (1637, 322.5, 325.-1, 337.10).)

For his version of specious algebra, Descartes uses lowercase letters, usually with x and y as unknowns and other letters as knowns, and he shows multiplication by concatenation. He also introduces superscripts for exponents. What was "E cubo-cubus" in Viète becomes "y^6" in Descartes, though for the square he preferred yy to y^2. Here are two sample equations, the first one describing a curve traced by the turning ruler and moving curve procedure from Book II, and the second an example provided for a rule for manipulating equations from Book III (Descartes 1637, 322, 379):

$$yy = cy - \frac{cx}{b}y + ay - ac$$

$$x^3 - \sqrt{3}xx + \frac{26}{27}x - \frac{8}{27\sqrt{3}} = 0.$$

There was precedence for indicating multiplication by juxtaposing species. Jean-Louis Vaulezard did so in his 1630 French translations of Viète's works, and we find it also in the arithmetical versions of *logistice speciosa* in Thomas Harriot (1631), William Oughtred (1631), and James Hume (1636). These last three authors had no need for the law of homogeneity, since their species are now numbers. Harriot, for example, could write an equation like "$aaa - 127296. a = 85700000$" (Harriot 1631, 145), and Oughtred could divide "BA + A" by "A" to get "B + 1" (Oughtred 1631, 11).

In the beginning of his book, Descartes tells us how to interpret his letters. He famously outlines an algebra of line segments by "taking one line which I shall call unity in order to relate it as closely as possible to numbers, and which can in general be chosen arbitrarily . . .," and then redefining multiplication, division, and root extraction so that the outcome is always another line segment. For multiplication, "having given two other lines, to find a fourth line which shall be to one of the given lines as the other is to unity" (Descartes (1637, 297), translated in Descartes (1954, 2)). He defines division and square root similarly, and he gives ruler-and-compass constructions for these operations on the next page. As Henk Bos has stressed, "in Descartes' interpretation the algebraic operations as applied in geometry did not concern numbers but geometric magnitudes, namely, the segments. Although he did introduce a unit line segment, he did not identify line segments with their numerically expressed lengths." (Bos 2001, 296). By redefining multiplication, division, and square root, Descartes created an algebra of lines whose operations agree with the corresponding operations of numbers in practical algebra. With it he could make sense of powers greater than three, and he could add and subtract what would otherwise have been nonhomogeneous terms. (See Bos (2001, 293ff) for a discussion on Descartes's redefinitions of the operations.) He was thus free to work with equations like the second example shown above.

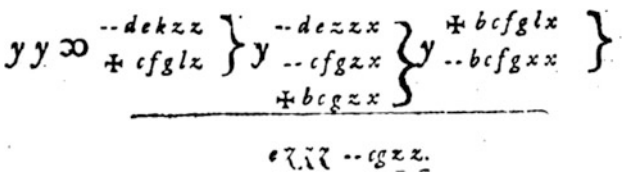
But with this apparent abandonment of homogeneity, it is important to consider what he writes next:

> It should be noted that all parts of a single line should always be expressed by the same number of dimensions, provided unity is not determined by the conditions of the problem. Thus, a^3 contains as many dimensions as abb or b^3, these being component parts of the line which I have called $\sqrt[3]{3a^3 - b^3 + abb}$. (Descartes (1637, 299), translated in Descartes (1954, 6).)

So if no unit is declared, then the operations of multiplication, division, and root extraction necessarily revert to their traditional geometrical meanings, and the law of homogeneity must be obeyed. Indeed, apart from one particular instance of the Pappus problem, no unit is posited in any problem posed in Book II. All expressions

and equations in those solutions are homogeneous, like the first equation shown above. This does not stop him from working with higher-degree equations, however. For example, one homogeneous equation in Book II is of the sixth degree, beginning "$y^6 - 2by^5 \ldots$" (Descartes 1637, 344).

The Pappus problem deserves special attention. Descartes also works with higher powers there, too, though his general equation is only of the second degree. He sets up this equation in unknowns x and y, and in known b, c, d, e, f, g, k, l, and a parameter z used in defining ratios:



In more modern format, it is this:

$$y^2 = \frac{(cfglz - dekz^2)y - (dez^2 + cfgz - bcgz)xy + bcfglx - bcfgx^2}{ez^3 - cgz^2}.$$

After making some simplifying assignments, he solves the equation for y in terms of x. From there, he gives detailed instructions on how to construct solutions, and he explains based on whether different givens are added, subtracted, or "nothing" (*nulle*) whether the locus of solutions is a line, circle, ellipse, parabola, or hyperbola. All this is done with homogeneous expressions without any declared unit. Then, after completing his description of the constructions of solutions, Descartes writes:

> And if one wants to show all the given quantities by numbers, by making, for example, EA = 3, AG = 5, AB = BR, BS = $\frac{1}{2}$BE, GB = BT, CD = $\frac{1}{2}$CR, CF = 2CS, CH = $\frac{2}{3}$CT, and that angle ABR be of 60 degrees, and finally that the rectangle of the two, CB, & CF, be equal to the rectangle of the two others CD & CH ... (Descartes 1637, 333; 1902, 405)

Making EA = 3 establishes the unit. Making AG = 5 is equivalent to saying that AG = $\frac{5}{3}$EA, and making angle ABR 60° serves to make EB = 2BS. It is from the last condition that the equation "$yy = 2y - xy + 5x - xx$" is established, which describes a circle (the axes are at a 60° angle). It was not necessary for the construction of the solution for Descartes to have assigned a unit. Had he written "AG = $\frac{5}{3}$EA" in place of "EA = 3, AG = 5," then since EA = k, the unitless, homogeneous equation would have been $yy = \frac{2}{3}ky - xy + \frac{5}{3}kx - xx$, and the construction of solutions would have been exactly the same. Descartes may have posited EA = 3 to illustrate his algebra of line segments, or perhaps to give meaning to the higher degree terms in the derivation of the general equation that preceded these assignments.

In Book III, where Descartes investigates the roots and transformations of equations and how to construct solutions, he usually works with examples of

heterogeneous equations for which some unit has been chosen. These examples are not given in the context of any particular problem, however. Bos has noted that Descartes works with these two interpretations of the operations, one from his algebra of line segments and the other from traditional geometry that takes into account dimension. Descartes himself does not address the issue of the compatibility of these interpretations, which suggests to Bos "a certain carelessness about the matter" (Bos 2001, 301). (Bos notes that they are in fact compatible, as one early commentator showed.)

Perhaps Descartes's statement that the unit "can in general be chosen arbitrarily" should be understood to mean that it is *potentially* present in problems where it is not declared. No unit is declared in the general solution to the Pappus problem, but afterward Descartes's choice of a unit for one particular arrangement of the lines can be interpreted as retroactively giving meaning to all the terms of degree greater than three in the general solution leading to that specific setup. But that is all it would do, because the selection of a unit is unnecessary for constructing such particular solutions. Further, Descartes does it for only one case of one problem. No other solution in Book II is followed by a solution "by numbers." We can infer, then, that Descartes preferred to find general solutions, after which a solution "by numbers" was merely an option. Because of this, he effectively worked in the same non-arithmetized geometry as Viète and Fermat. Every term in Descartes's expressions arising from general solutions have positive integer degree, and the law of homogeneity is observed. As in Viète, magnitudes in the problems solved by Descartes are without any intrinsic numerical measure, but one can later assign numbers to them for whatever purpose one has in mind.

There is one minor formal development in Descartes's notation for homogeneous expressions that may have been inspired by his geometry of line segments. Where Viète would have written $\frac{\text{B in A}}{\text{Z}}$, a dimension 2 term divided by a dimension 1 term, Descartes writes many homogeneous terms like $\frac{a}{z}x$ (1637, 328), where the letters are arranged so that it appears that the $\frac{a}{z}$ has some geometrical meaning by itself. But this way of writing, the term only functions as a notational device to separate the unknown x from the rest of the expression. The whole term is one-dimensional, like the corresponding term in Viète.

In solving problems in geometry, then, Descartes works with essentially the same algebra as Viète and Fermat. This is an algebra rooted in the non-arithmetized sizes of geometric magnitudes, which can, if one chooses, be given numerical measure (Fermat never takes this option, but it is prevalent in Viète). All terms have positive integer degree, only homogeneous terms can be added and subtracted, and powers greater than three are given no geometric meaning. Descartes's algebra of line segments, which appears to be a creative solution to the problem of magnitudes of higher dimension, comes into play only when numbers are assigned to lines. Such arithmetizations can help illustrate the general solutions or they may be applied to real-life problems, but they are not a necessary part of geometrical problem-solving.

7 Synthesis

One important, but ultimately inconsequential, difference between premodern algebra and Viète's *logistice speciosa* is that the unknowns of the former are numbers, while Viète's species belong to geometry. We have seen that Harriot, Oughtred, and Hume all reinterpreted the species as numbers, and in the other direction, numerical algebra could be applied to problems in classical geometry by positing a unit line, as al-Khayyām had already done in the eleventh century. What definitively differentiates premodern algebra from Viète's specious logistic is how expressions were conceived. Premodern algebra is based in multitudes of kinds of numbers and their aggregations, while Viète's algebra is based in values assembled via operations. We can ask, then, what advantage Viète's algebra had over premodern algebra for the work of Fermat and Descartes. How is it that they chose Viète's species, and could some version of premodern algebra have sufficed?

Descartes learned cossic algebra from Clavius's 1608 textbook, and it is that algebra that he employed from the late 1610s down to at least 1628 (Rabouin 2010). There was at least one time in this period that he could have benefitted from undetermined coefficients. In notes he took in 1619–1620 on cubic equations in relation to his mesolabe compass, he sometimes expressed a particular type of equation by putting a circle, or an "O," in place of the coefficients, such as an "equation between Oꝛꝛ & O²ꝛ + ON" ("aequationem inter Oꝛꝛ & O²ꝛ + ON" (Descartes 1908, 244)). We would write this in modern notation as an equation between ax^3 and $bx + c$. The fact that three copies of the same sign "O" stand in the places of three potentially different amounts indicates that they are more like circles in which one could fill in numbers than letters standing for undetermined values. Clavius expressed an equation type like this by simply not writing anything for the coefficients, like his generic equation "between ȝ + ꝛꝛ, & N" ("inter ȝ + ²ꝛ, & N" (Clavius 1608, 52.6)) for what we would write as an equation between $ax^2 + bx$ and c. This is a notational example of the rhetorical equation types that go back to al-Khwārazmī, who wrote this particular type as "[some] *māl*s and [some] roots equal a number" before giving the specific example "a *māl* and ten roots equal thirty-nine dirhams" ((al-Khwārazmī 2009, 101.2), my translation).

No premodern algebraist is known to have devised either words or notation for undetermined coefficients. One could suggest that Stifel or Stevin were on the verge of such a notation, because they could have regarded their independent unknowns as being given, so that our $ax^2 + bx = c$, for example, could have been written by Stifel as "1ȝA + 1²ꝛB equals 1C." But this construction would have been decidedly inconvenient, because the coefficients become embedded in the kinds, which obliterates the fundamental premodern distinction between the coefficient (i.e., the multitude) and the power (i.e., the kind). Multitudes and kinds cannot just be merged together into new, composite kinds that themselves are assigned a multitude. For one thing, one should be able to replace the letters with determinate numbers, but in this scheme that cannot happen without dismantling the types. I had suggested in Oaks (2018, 287) that independent unknowns could have served as undetermined givens, but now I see that it could not have worked.

Or one might wonder why they could not have just written something like "A𝒿 + B²𝓆 equals C," where A, B, and C now stand for the values of numbers such as we find in Book II of Pappus's *Collection* or in *De Numeris Datis* of Jordanus de Nemore (early thirteenth c.). Such letters as values were known in Stifel's time, though they served as labels in premodern arithmetic and are not found together with algebraic terms. If such a contrivance was ever considered, there would have been a problem as soon as there was a need to operate on the terms, since the letters, as labels, were not operated on to produce more complex expressions (§5.4). Xylander (1575) and Gosselin (1577) took small steps in introducing operations to letters that designate values, though they did not mix this notation with algebraic notation. Those steps might have eventually led to another modification of premodern algebra, but they came too late (Oaks 2018, §5.4).

With nine different known magnitudes in Descartes's general solution to the Pappus problem, it would have been impossible for him to have written the equation in cossic algebra even if he had reinterpreted the kinds as being magnitudes. Descartes is silent about when he first saw examples of *logistice speciosa*, but it was likely to have occurred between 1628, when Beeckman recorded samples of cossic algebra from Descartes's hand, and January 1632, when Descartes solved the Pappus problem, presumably with essentially the same algebra that he shows in *La Geometrie*. (Sasaki (2003, 235ff) reviews the evidence regarding when Descartes may have first become aware of Viète's algebra.)

It is no surprise that Fermat chose to adopt Viète's algebra and its notation, given that he studied both Viète and Pappus, and that he, too, sought to restore Greek analytic geometry. Numerical algebra would have held no appeal for him when he had at hand an algebra designed for the geometer. His work, too, necessarily required undetermined givens, so cossic algebra would have also been inadequate for him.

Premodern algebra had always been an excellent arithmetical technique for solving problems, but it was limited by the identification of expressions as numbers, which were by their nature amounts of some unit or units. Geometric magnitudes are not multitudes of units, so when Viète set about making an algebra for geometry, it was naturally based in values, and thus expressions were necessarily formed through operations, and givens could only be undetermined. These last two features of *logistice speciosa* were somehow accidental, but they offered critical advantages. We have seen that in Viète, solutions to equations in problem-solving become for the first time formulas expressed in terms of given parameters, and in Fermat and Descartes general locus problems and problems of finding tangent and normal lines, with unspecified givens, can be solved algebraically. These accidental features were also necessary for later applications of algebra. Try to write an identity like $(a+b)^3 = a^3 + b^3 + 3a^2b + 3ab^2$ in any premodern notation and you will understand why Cardano states the rule rhetorically (Cardano 1545, f. 16a). Likewise, how could $A = \pi r^2$, $f(x) = 3x^3\sqrt{x^2+6}$, or $d = \frac{1}{2}gt^2$ be expressed if not with letters that stand for values that are joined through operations?

We can credit Book VII of Pappus's *Collection* with playing a significant role in inspiring both Viète's renovation of algebra and Fermat's and Descartes's application

of that algebra to the study of curves. But even if these three authors worked in the sphere, or at least in the shadow, of classical Greek mathematics, it was they who paved the way for the eventual abandonment of geometry in favor of the new algebra for conducting original work in mathematics, and also for the slow demise of premodern algebra in favor of a numerical version of *logistice speciosa* for practitioners. Viète's algebra was supposed to be a restoration of ancient analysis. Instead, ultimately it made obsolete both the geometrical analysis of Pappus and the arithmetical analysis (premodern algebra) of Diophantus.

References

Abū Kāmil (2012) Algèbre et analyse Diophantienne (ed and trans: Rashed R). De Gruyter, Berlin

Apollonius (1891) Apollonii Pergaei quae Graece exstant cum commentariis antiquis, edidit et Latine interpretatus est I.L. Heiberg. in aedibus B.G. Teubneri, Stutgardiae

Aurel M (1552) Libro primero de arithmetica algebratica. Joan de Mey Flandro, Valencia

Biagio M^{o} (1983) Chasi exenplari alla regola dell'algibra nella trascelta a cura di M ^{o} Benedetto: dal Codice L.VI. [i.e. IV] 21 della Biblioteca Comunale di Siena, ed Pieraccini L. Servizio Editoriale dell'Università di Siena, Siena

Bombelli R (1572) L'algebra parte maggiore dell'arimetica divisa in tre libri. Giovanni Rossi, Bologna

Bos H (1990) The structure of Descartes' *Géométrie*. In: Belgioioso G, Cimino G, Costabel P, Papuli G (eds) Descartes: il metodo e i saggi: Atti del convegno per il 350° anniversario della pubblicazione del *Discours de la Méthode* e degli *Essais*. Istituto della Enciclopedia Italiana, Roma, pp 349–369

Bos H (2001) Redefining geometrical exactness: Descartes' transformation of the early modern concept of construction. Springer, New York

Brasser JR (1663) Regula cos, of algebra. Gerrit van Goedesbergh, Amsterdam

Cardano G (1539) Practica arithmetice, & mensurandi singularis in qua que preter alias cotinentur, versa pagina demonstrabit. Bernardini Calusci, Mediolani

Cardano G (1545) Artis magnæ, sive de regulis algebraicis lib. unus. per Ioh. Petreium excusum, Norimbergæ

Christianidis J, Oaks JA (2013) Practicing algebra in late antiquity: the problem-solving of Diophantus of Alexandria. Hist Math 40:127–163. https://doi.org/10.1016/j.hm.2012.09.001

Clavius C (1608) Algebra. Apud Bartholomaeum Zannettum, Romae

De Billy J (1643) Nova geometriæ clavis algebra. Michaelem Soly, Paris

Descartes R (1637) La geometrie. Ian Maire, Leiden

Descartes R (1902) Oeuvres de Descartes, vol VI (eds: Adam C, Tannery P). Léopold Cerf, Paris

Descartes R (1908) Oeuvres de Descartes, vol X (eds: Adam C, Tannery P). Léopold Cerf, Paris

Descartes R (1954) The geometry of René Descartes. Translated from the French and Latin by David Eugene Smith and Marcia L. Latham. Dover, New York

Euclid (1572) Euclidis elementorum libri XV (trans: Commandino F). Iacobus Chriegher German, Pisauri

Fermat P (1891) Œuvres de Fermat, tome premier (eds: Tannery P, Henry C). Gauthier-Villars et Fils, Paris

Freguglia P (2008) Viète reader of Diophantus. An analysis of *Zeteticorum Libri Quinque*. Bollettino di Storia delle Scienze Matematiche 28:51–95. https://doi.org/10.1400/94823

Ghetaldi M (1630) De resolutione & compositione mathematica libri quinque. Opus posthumum. Typographia Reuerendae Camerae Apostolicae, Romæ

Gosselin G (1577) De arte magna. Aegidium Beys, Paris

Harriot T (1631) Artis analyticae praxis. Apud Robertum Barker, Londini

Heeffer A (2010) From the second unknown to the symbolic equation. In: Heeffer A, van Dyck M (eds) Philosophical aspects of symbolic reasoning in early modern mathematics. College Publications, London, pp 57–101
Hume J (1636) Algebre de Viete d'une methode nouvelle, claire, et facile. Chez Louis Boulenger, Paris
Ibn al-Bannāʾ (1994) Rafʿ al-ḥijāb ʿan wujūh aʿmāl al-ḥisāb (ed: Aballagh M). Jāmiʿat Sīdī Muḥammad ibn ʿAbd Allāh, Fās
al-Khwārazmī (2009) Al-Khwārizmī: the beginnings of algebra. Edited, with translation and commentary by R. Rashed. Saqi, London
Klein J (1936) Die griechische Logistik und die entstehung der Algebra. II. Teil. Quellen un Studien zur Geschichte der Mathematik Astronomie und Physic 3(2):122–235
Klein J (1968) Greek mathematical thought and the origin of algebra (trans: Brann E). MIT, Cambridge
Maghetti B (1639) Analisi o risolutione de quesiti altre volte stampati. Marco Saluioni, Ancona
Mahoney MS (1994) The mathematical career of Pierre de Fermat 1601–1665, 2nd edn. Princeton University Press, Princeton
Marie M (1884) Histoire des Sciences Mathématiques et physiques, Tome III. De Viète a Descartes. Gauthier-Villars, Paris
Montucla J-É (1758) Histoire des Mathematiques, vol I. Ch. Ant. Jombert, Paris
Mueller I (1981) Philosophy of mathematics and deductive structure in Euclid's *elements*. MIT, Cambridge
Oaks JA (2009) Polynomials and equations in Arabic algebra. Arch Hist Exact Sci 63:169–203. https://doi.org/10.1007/s00407-008-0037-7
Oaks JA (2010) Polynomials and equations in medieval Italian algebra. Bollettino di Storia delle Scienze Matematiche 30:23–60
Oaks JA (2017) Irrational 'coefficients' in renaissance algebra. Sci Context 30:141–172. https://doi.org/10.1017/S0269889717000096
Oaks JA (2018) François Viète's revolution in algebra. Arch Hist Exact Sci 72:245–302. https://doi.org/10.1007/s00407-018-0208-0
Oaks JA (2019) Proofs and algebra in al-Fārisī's commentary. Hist Math 47:106–121. https://doi.org/10.1016/j.hm.2018.10.008
Oaks JA, Alkhateeb HM (2007) Simplifying equations in Arabic algebra. Hist Math 34:45–61. https://doi.org/10.1016/j.hm.2006.02.006
Oughtred W (1631) Arithmeticæ in numeris et speciebus institutio. Apud Thomam Harperum, Londini
Pacioli L (1494) Summa de arithmetica geometria proportioni & proportionalita, continentia de tutta lopera. Paganino de Paganini, Venetijs
Panza M (2006) François Viète: between analysis and cryptanalysis. Stud Hist Phil Sci 37:269–289
Pappus (1588) Pappi Alexandrini mathematicae collectiones (trans: Commandino F). Apud Franciscum de Franciscis Sensem, Venetiis
Pappus (1986) Book 7 of the *Collection*, 2 vols, ed. with trans. and comm. by Alexander Jones. Springer, New York
Peletier J (1554) L'algebre. Ian de Tournes, Lyon
Peucer C (1556) Logistice astronomica hexacontedon et scrupulorum sexagesimotum, quam algorythmum minutiarum physicalium vocant, regulis explicata & demonstrationibus. Item: logistice regulae arithmeticæ, quam cossam & algebram quadratam vocant, compendio tractata & illustratata exemplis, ut scholarum usui sit accommodata. Excvdebant haeredes Georgii Rhavv, Vitebergae
Ptolemy (1998) Ptolemy's almagest. Trans. and annotated by G. J. Toomer. Princeton University, Princeton
Rabouin D (2010) What Descartes knew of mathematics in 1628. Hist Math 17:428–459. https://doi.org/10.1016/j.hm.2010.04.002

Rudolff C (1525) Behend und hübsch Rechnung durch die künstreichen regeln Algebre, so gemeincklich die Coss genent werden. Jung, Argentorati
Saidan AS (1986) Tārīkh ʿilm al-jabr fīʿālam al-ʿArabī. (History of algebra in medieval Islam). al-Majlis al-Waṭanī lil-Thaqāfah wa'l-Funūn wa'l-Ādāb, Qism al-Turāth al-ʿArabī, Kuwait
Sasaki C (2003) Descartes's mathematical thought. Kluwer Academic, Dordrecht
Scheubel J (1550) Evclidis Megarensis, philosophi & mathematici excellentissimi, sex libri priores, de geometricis principijs . . . Algebrae Porro Regvlae . . . Hervagius, Basileae
Stevin S (1585) L'arithmetique. De l'imprimerie de Christophe Plantin, Leyde
Stifel M (1544) Arithmetica integra. Iohan. Petreium, Norimbergæ
Tartaglia N (1556) La prima parte del general trattato di numeri, et misure. Curtio Troiano, Venice
Tartaglia N (1560) La sesta parte del general trattato di numeri, et misure. Curtio Troiano, Venice
Van Egmond W (1985) A catalog of François Viète's printed and manuscript works. In: Folkerts M, Lindgren U (eds) Mathemata: Festschrift für Helmuth Gericke. F Steiner Verlag Wiesbaden, Stuttgart, pp 359–396
Viète F (1591) In artem analyticem isagoge. Apud Iametium Mettayer, Turonis
Viète F (1593a) Variorum de rebus mathematicis responsorum, Liber VIII. Apud Iamettium Mettayer, Turonis
Viète F (1593b) Effectionum geometricarum canonica recensio. (Published as the second part of (Viète 1593a))
Viète F (1593c) Zeteticorum. (Published as the third part of (Viète 1593a))
Viète F (1593d) Supplementum geometriæ. Excudebat Iametius Mettayer, Turonis
Viète F (1595a) Ad problema quod omnibus mathematicis totius orbis construendum proposuit Adrianus Romanus. Apud Iametium Mettayer, Parisiis
Viète F (1595b) Pseudo-mesolabum & alia quædam adiuncta capitula. Apud Iamettium Mettayer, Parisiis
Viète F (1615a) Ad angularium sectionum analyticen. Apud Oliverivm de Varennes, Parisiis
Viète F (1615b) De æquationum recognitione et emendatione tractatus duo. Ex Typographia Ioannes Laquehay, Parisiis
Viète F (1630) Les cinq livres des zetetiques de Francois Viette, trans. and comm. by Jean-Louis Vaulezard. Iulian Iacquin, Paris
Viète F (1646) Opera mathematica (ed: van Schooten F). Ex Officinâ Bonaventuræ & Abrahami Elzeviriorum, Lugduni Batavorum
Viète F (1983) The analytic art: nine studies in algebra, geometry, and trigonometry from the opus restitutae mathematicae analyseos, seu, algebrâ novâ by François Viète (trans: Richard Witmer T). Kent State University, Kent

Descartes' Transformation of Greek Notion of Proportionality

Piotr Błaszczyk

Contents

1 Introduction 236
2 Euclidean Proportions 236
2.1 Foundations 237
2.2 The Use of Proportions 238
3 Descartes' Algebra of Line Segments 239
3.1 From Greek Proportion to Descartes' Algebra of Line Segments 242
4 Arithmetic of Line Segments 4.1 Defining Descartes' Operations 242
5 The Field of Line Segments 244
5.1 True and False Roots 244
5.2 Zero and Sign Rules 245
5.3 Constructing False Roots 245
5.4 Ordered Field of Line Segments 246
6 Applications 247
6.1 Pythagorean Theorem 247
6.2 Formula for a Curve 249
7 Pappus Problem 251
7.1 Formula for *Locus* 251
7.2 Constructing *Locus* 253
7.3 Numerical Example 255
7.4 Magnitudes, Numbers, and Quantities Versus Elements of an Ordered Field 256
8 Ancient Greek Numerical Practice Based on Geometry 258
8.1 Euclidean Arithmetic and Beyond 258
8.2 Heron's Formula for the Area of a Triangle 260
9 Euclidean Versus Descartean Plane 263
10 Summary 265
References 266

The author is supported by the National Science Centre, Poland grant 2018/31/B/HS1/03896.

P. Błaszczyk (✉)
Institute of Mathematics, Pedagogical University of Cracow, Cracow, Poland
e-mail: piotr.blaszczyk@up.krakow.pl

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_16

Abstract

In ancient Greek mathematics, the general term μέγεθος stands for geometric magnitudes of all kinds. Euclid's book V develops a theory of magnitudes, their ratios, and proportions. Viewed from the modern perspective, the theory of proportions is a technique of processing ratios. In early modern mathematics, it was replaced by implicit rules of an ordered field, then, in the nineteenth century, by the arithmetic of real numbers. We formalize Euclid's theory from book V and show how it was reshaped in Descartes' *La Géométrie.*

Keywords

Magnitude · Ratio · Proportion · Algebra of line segments · Arithmetic of line segments · Ordered field · Real closed field · Descartes' plane

1 Introduction

Descartes' 1637 *Discours de la Méthode pour bien conduire sa raison, et chercher la vérié dans les sciences, plus la Dioptrique, les Météores et la Géométrie qui sont des essais de cette Méthode* (Descartes 1637b) consists of the philosophical essay *Discourse on Method* and three scientific treaties, as we would could call them today. Each of them contains spectacular discovery. In *Meteores,* Descartes provides widely accepted explanation of the rainbow phenomenon. In *Dioptrics,* he derives the law of refraction, currently known as Snell's law, as well as fundamental laws concerning elliptic and hyperbolic lenses. In *Geometry,* he presents a solution to the Pappus problem for any number of lines. The problem itself and its general solution suggested by Descartes are already forgotten. Yet, it was a novel method of solving problems in geometry that has had a far-reaching impact on mathematics and made *La Géométrie* a masterpiece of mathematical writing. Descartes' method is commonly referred to as an analytic geometry. In fact it took centuries to turn his original procedures into the modern analytic form. One aspect of that method includes operations on line segments – operations which replaced the ancient technique of proportions.

2 Euclidean Proportions

In ancient Greek mathematics, the general term μέγεθος (magnitude) covers line segments, triangles, convex polygons, circles, angles, arcs of circles, and solids. Euclid's book V develops rigorous treatment of magnitudes, their ratios, and proportions. It is one of the most thorough theories making up the *Elements.* At the same time, it calls for a formal examination. Technical terms like *greater, lesser, multitude,* and *proportion* are consistently applied throughout the course.

2.1 Foundations

We formalize magnitudes of the same kind (line segments being of one kind, triangles being of another, etc.) as an additive semigroup with a total order, $(M, +, <)$, characterized by the following four axioms:

$$\begin{aligned}
&\text{E1} \quad (\forall x, y)(\exists n \in \mathbb{N})(nx > y),\\
&\text{E2} \quad (\forall x, y)(\exists z)(x < y \Rightarrow x + z = y),\\
&\text{E3} \quad (\forall x, y, z)(x < y \Rightarrow x + z < y + y),\\
&\text{E4} \quad (\forall x)(\forall n \in \mathbb{N})(\exists y)(x = ny).
\end{aligned}$$

The term nx as applied in these formulas is defined by $nx = \underbrace{x + \ldots + x}_{n \text{ times}}$

The relation *greater than* is characterized by transitivity and the law of trichotomy; both Euclid's and Aristotle's texts provide evidence for it. Equality of line segments, angles, and arcs means congruence. Equality of figures should be interpreted in Euclid's theory of area, as explained in (Błaszczyk 2018); equality of solids involve an analogous theory developed in book XII of the *Elements.*

The symbol + interprets arguments like the following: "And since AG [a] is equal to E [a], and CH [c] to F [c], thus AG, F [$a + c$] is equal to CH, E [$c + a$]" (V.25), that is[1]

$$a + c = c + a.$$

Thus, in the text, instead of *AG* + *F*, or *AG and F*, we find the term *AG, F.* Moreover, we interpret the equality *AG, F* = *CH, E* as the commutative property. In a similar way, we could show that the addition of magnitudes is an associative operation. By alike textual analysis we could certify axioms E2 and E3.

Note that in book V, magnitudes are represented by line segments, hence the term *AG, E* finds an obvious interpretation as a concatenation of two lines. Yet the results of book V are applied to magnitudes of any kind. In general, viewed from the methodological perspective, the addition of magnitudes should be treated as a primitive concept. Nonetheless, within Euclid's theory of equal figures, the sum of squares could be defined as a square determined on the ground of proposition I.47, and the sum of similar figures as a figure similar to the ones being added, determined on the ground of proposition VI.31.[2]

E1 interprets definition V.4, the so-called Archimedean axiom: "Magnitudes [x, y] are said to have a ratio with respect to one another which, being multiplied *[nx]* (πολλαπλασιαζόμενα), are capable of exceeding one another [$nx > y$]."

The term *nx* stands here for the phrase *multiple of the magnitude.*

Axiom E1 is applied in book V only once: in the proof of proposition V.8. Then, in a modified form, it is applied in proposition X.1 on incommensurable lines.

[1] English translations of the *Elements* after (Euclid 2007). The accuracy of this edition is easy to verify since it also includes the Greek text of the classic Heiberg's edition (Heiberg 1883).

[2] See (Błaszczyk 2018; Błaszczyk and Petiurenko 2020) respectively.

Geometry developed in books I-IV of the *Elements* does not rely on E1 and can be modeled on a non-Archimedean plane. It can be shown that without E1 proposition V.8, crucial for the theory, does not hold. Yet, the theory of similar figures as developed in book VI builds on results of book V, hence Euclidean similar figures should be modeled on an Archimedean plane.

E4 is implicitly applied in proposition V.5. This axiom is not essential, as it can be derived from the remaining four, i.e., E1-E3, and E5 as presented below.

Next to sums and relation *greater than,* book V introduces yet another relation on magnitudes, namely proportion. In Greek manuscripts, it is rendered by the schematic phrase τὸ A πρὸς τὸ B, οὕτος τὸ Γ πρὸς τὸ Δ ("as A is to B, so is C to D"). For a symbolic representation of proportion, we adopt the term x: y:: z: υ. In fact, this symbol was introduced already in the seventeenth century.[3] Descartes did not apply any symbolic shortcuts for proportions and mimics Euclid's phrase.

Definition V.5 clearly brings in the proportion in a verbal form. We interpret it by the following formula:

$$x : y :: z : \upsilon \Leftrightarrow_{df} (\forall m, n \in \mathbb{N})[(nx > 1my \Rightarrow nz >_2 m\upsilon) \wedge \\ \wedge (nx = my \Rightarrow nz = m\upsilon) \wedge (nx <_1 my \Rightarrow nz <_2 m\upsilon)].$$

The assumption regarding magnitudes *x, y,* on the one hand, and *z, υ,* on the other, being of the same kind is formalized by $x, y \in \mathcal{M}_1 = (M_1, +<_1)$ and $z, \upsilon \in \mathcal{M}_2 = (M_2 + <_2)$.

Definition V.7 introduces a relation *greater than* between pairs of magnitudes. We interpret it as follows:

$$x : y \succ z : \upsilon \Leftrightarrow_{df} (\exists m, n \in \mathbb{N})[(nx >_1 my) \wedge (nz \leq_2 m\upsilon)].$$

Thus, the structure of magnitudes is an ordered semigroup equipped with the four-ary relation of proportion, *(M, +, ::, <)*. Due to the proportion, we need to amplify its characteristic by another axiom, namely

$$\text{E5}\ (\forall x, y, z)(\exists \upsilon)(x : y : z : \upsilon).$$

E5 represents the so-called fourth proportional axiom. In book V, it is applied in proposition V.18. Next, it is explored in the method of exhaustion (see e.g. XII.1).

2.2 The Use of Proportions

Propositions of book V provide the bases for the theory of similar figures as developed in book VI. Descartes applies it all throughout *Geometry, Dioptrics,* and in his essay on the rainbow included in *Meteores.*

In book VI, there is only one explicit reference to definition V.5, namely, in the proof of proposition VI.1. Other propositions apply VI.1 and results of book V. To

[3] See (Cajori 2007), p. 197.

explain how they enable to process ratios, we present a sample stylized on algebra. This modern attire is to reveal similarities between proportions and the arithmetic of fractions.

V.9 $a: c :: b : c \Rightarrow a = b.$
V.12 a: b :: $c : d$, $a : b :: e : f \Rightarrow a : b :: (a + c + e) : (\mathrm{b} + d + f)$.
V.16 a: b :: $c : d \Rightarrow a : c :: b : d$.
V.22 $a: b :: d : e$, $b : c :: e : f = a : c :: d : f$.
V.23 $a: b :: e : f$, $b : c :: d : e \Rightarrow a : c :: d : f$.
V.24 $a: c :: d : f$, $b : c :: e : f \Rightarrow (a + b) : c :: (d + e) : f$.

When proportion $a : b :: c : d$ is replaced with the equality of fractions $\frac{a}{b} = \frac{c}{d}$, the above propositions will turn into simple rules of the arithmetic of fractions. However, not the other way around, e.g., the seemingly obvious rule: if $\frac{a}{b} = \frac{c}{d}$, then $ad = bc$, finds no counterpart in book V. The term $ad = bc$ could be interpreted in terms of figures rather than proportions (see. VI.16). Note also that ratio $a : b$ is not an individual object as it has to be a part of the 4-ary relation like $a : b :: c : d$ (see Błaszczyk and Petiurenko 2019).

In the context of transformation of Euclidean proportionality, one crucial difference is this: while Euclid's V.9 relies on the Archimedean axiom, its ordered field counterpart holds in any field, be it Archimedean or non-Archimedean.

Summarizing this section, the structure $(M, +, ::, <)$ as characterized by the above five axioms E1-E5 represents the ancient Greek magnitudes of the same kind, especially line segments.[4]

3 Descartes' Algebra of Line Segments

On the very first pages of (Descartes 1637a), Descartes introduces brand new operations on line segments: multiplication, division, and the square root.[5] They are defined in such a a way that $a \cdot b$, $\frac{a}{b}$,, and $\sqrt{a}$ are line segment, given a, b are line segments. Hence, for example, a^2, a^3, a^4 etc., or $\sqrt{a^3 - b^2}$, are lines segments. This bold innovation allows Descartes to generalize the Pappus problem to any number of lines, while Greeks could consider only four, five, or six lines. He can also consider "curve[s] of degree next higher than that of the conic sections" (Descartes 1637a, 307), in fact, of any degree, i.e., curves characterized by sums of terms such as $ax^n y^m$, for any $n,m \geq 0$.

[4] For detailed analysis leading to axioms for book V, see (Beckmann 1967), or (Błaszczyk and Mrówka 2013); the other study provides also independence results. Commonly referred (Mueller 2006) basically follows Beckmann's reading. Beckmann (1967) and Błaszczyk and Mrówka (2013) differ on some issues, yet these disagreements are irrelevant in the context of this chapter.

[5] Descartes scholars use to call it algebra and in this section we adopt this label.

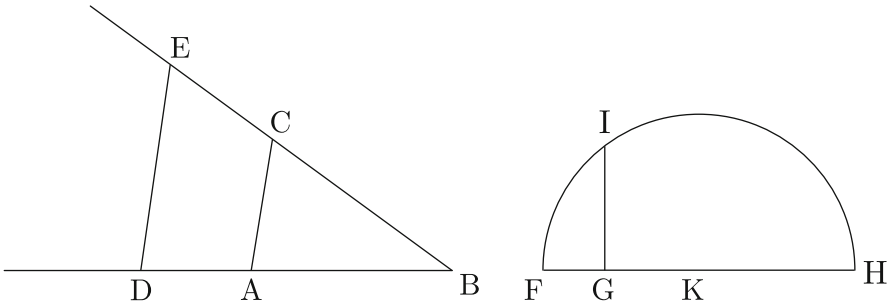


Fig. 1 *La Géométrie,* p. 297

Descartes' definitions (of a kind) are diagram based (see Fig. 1). They read[6]: "let AB be taken as unity, and let it be required to multiply BD by BC. I have only to join the points A and C, and draw DE parallel to CA; then BE is the product of BD and BC. If it be required to divide BE by BD, I join E and D, and draw AC parallel to DE; then BC is the result of the division. If the square root of GH is desired, I add, along the same straight line, FG equal to unity; then, bisecting FH at K, I describe the circle FIH about K as a center, and draw from G a perpendicular and extend it to I, and GI is the required root" (Descartes 1637a, 298).[7]

In Fig. 1, the diagram on the left refers to Euclid's proposition VI.12 on finding the fourth proportional to three given lines. Next to lines BD, BC which are to be multiplied, Descartes assumes the unit line $AB = 1$, "line which I shall call unity in order to relate it as closely as possible to numbers, and which can in general be chosen arbitrarily" (Descartes 1637a, 297). The diagram on the right refers to proposition VI.13 on finding the line in mean proportion to two given straight lines. In this case, Descartes applies Euclidean construction to the line GH and the unit line FG.

Since Euclid's geometry provides the context for these definitions, the line AB has to be arbitrary. Contrary to hyperbolic geometry, there is no basis for a unique line segment in Euclidean geometry. The angle $\angle EBD$ is also an arbitrary choice.

For comparison sake, Hilbert (Hilbert 2015) develops an algebra of line segments in a separate way.[8] It relies on the parallel axiom but not on the Archimedean axiom.

[6] English translations of (Descartes 1637a) after (Smith and Lathan 2007). Occasionally we modify Smith's version, still, these amendments are easy to verify, since (Smith and Lathan 2007) also includes facsimile of the first edition of *La Géométrie*.

[7] The passage continues as follows: "I do not speak here of cube root, or other roots, since I shall speak more conveniently of them later." Descartes' mesolabium (Descartes 1637a, 317–319) provides means to construct cube and other roots of a line segment; it also enables one an alternative construction of the square root.

[8] In this section, we adopt a methodological perspective. Crippa (Crippa 2017) takes Hilbert's arithmetic as an opportunity for far-reaching conclusions. He points out that while in Hilbert's system the unit line is fixed once for all, in *La Géométrie*, it varies from a problem to a problem. As a result, he argues, the generality in Descartes' system "is achieved by means of analogy, in a similar manner to Euclid's theory" (Crippa 2017, 1255). The Greek term.ἀναγογία literally means analogy, technically – proportion. We argue that Descartes transforms the ancient technique of proportion into operations in an ordered field. In this context, *analogy* as used by Crippa is a meaningless term.

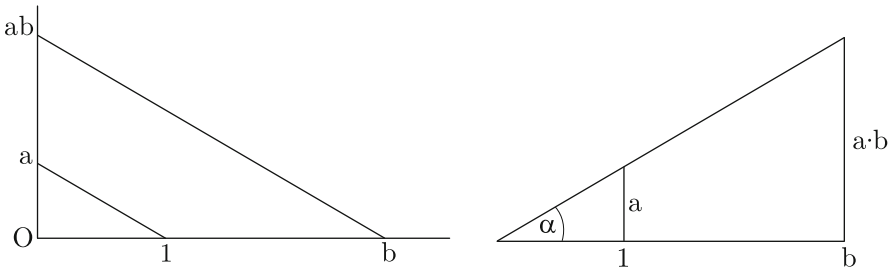


Fig. 2 Hilbert's *Grundlagen,* p. 33 (left). An alternative view on Hilbert's product (right)

It also applies an arbitrary unit segment, yet the choice of the angle depends on the line a in such a way that tg $\alpha = a$ (see Fig. 2, the diagram on the right). Definition of the product enables to determine the segment $\frac{1}{a}$, hence the quotient $\frac{b}{a}$. Hilbert's algebra, though, includes line segments of the form $\sqrt{1+a^2}$, rather than square roots $\sqrt{a}$.

As a result, there are models of Hilbert's plane which are not models of Euclidean plane.[9]

Adopting standards of the modern methodology, Hilbert's definition of the line segment *ab* is correct. As for Descartes, the status of lines introduced through his diagrams is not that clear. His statements to the effect "BE is the product of BD and BC," "BC is the result of the division," and "GI is the required root" are ambiguous. The term "is the product/quotient/root" can mean "is the product etc., based on a definition" or "is the product etc., based a theorem"? Moreover, all throughout *La Géométrie,* Descartes implicitly applies rules such as.[10]

$$ab = ba,\ \ a(b+c) = ab + ac,\ \ \sqrt{a^2} = a,\ \ \left(\sqrt{a}\right)^2 = a,\ \ a \cdot 1 = a,\ \ a < b \Rightarrow ac < bc. \tag{1}$$

They seem obvious, though, from the perspective of the modern theory of ordered fields, or just the arithmetic of real numbers. However, in view how Descartes introduces new operations, these rules are by no means *clear and obvious,* and they can hardly be taken for granted.[11]

[9] See (Hartshorne 2000), p. 187.

[10] Bos (Descartes 1637a), in a footnote on page 300, includes such a strong comment: "The interpretation of multiplication and division by means of a unit was not used in *Geometry* at all." We argue that via the rules of an ordered field, like these listed in (1), multiplication, division, and square root are applied in processing of formulas all throughout *Geometry.*

[11] Mancosu (Mancosu 1996), in regard to Descartes' *definition* of the product, states: "Then construction of *ab* is as follows [...] This is done by joining A and C and drawing the line DE parallel to AC. Then $BE = BD \cdot BC = ab$. The claim is easily verified by exploiting the proportionality between the triangles ABC and DBE" (Mancosu 1996, 67). The similarity (or less formally, proportionality) of triangles *ABC* and *DBE* results in the following proportion *BE*: *BC*:: *BD*: 1. In the next section, we show how to derive from it the equality $BE = BD \cdot BC$. Yet, to this end we need a definition of the product of line segments.

3.1 From Greek Proportion to Descartes' Algebra of Line Segments

Next to the implicit use of the rules of an ordered field, the most intriguing part of Descartes algebra is the implicit way how he turns Euclid's proportions into his new operations. In symbols, it could be rendered as follows:

$$a : b :: c : d \Rightarrow a \cdot d = c \cdot b, \quad a : b :: c : d \Rightarrow a = \frac{c}{d} \cdot b. \tag{ED}$$

Below we provide textual evidence. The following passage refers to Fig. 3.

"For, putting a for BD or CD, and c for EF, and x for DF, we have CF $= a - x$, and since CF or $a - x$ is to FE or c, as FD or x is to BF, which is consequently $\frac{cx}{a-x}$" (Descartes 1637a, 389).

In this passage, Descartes explains how to find roots of a fourth-degree polynomial. Details of his solution are irrelevant to the discussed issue, so we skip them. To simplify the view, we mark in blue parts referring to Euclid's theory of similar triangles, and in red those referring to Descartes' algebra. Thus, since $\Delta CFE \sim \Delta DFB$, the following proportion holds CF : FE :: FD : BF, or, adopting Descartes' alternative notation, $(a - x) : c :: x : BF$. Then, all of a sudden, Descartes declares that $BF = \frac{cx}{a-x}$. The rule (ED) enables one to justify that step, however, in *La Géométrie* there is no hint concerning that transformation.

The discussed passage is by no means accidental. The revealed argument, as encoded by the rule (ED), is applied whenever Descartes invokes Euclid theory of similar figures, that is, all throughout *La Géométrie.*

In sum, given that M stands for the set of line segments, Descartes' algebra is represented by the structure $\left(M, +, \cdot, \div, \sqrt{}, 1, <\right)$. In Sect. 5 below, we will show how it expands by including *negative* line segments.

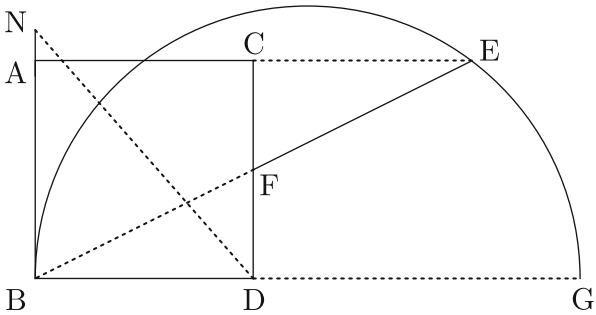
4 Arithmetic of Line Segments 4.1 Defining Descartes' Operations

Going back to Descartes' definitions and rules exemplified in (1), they could be justified in a purely geometrical manner (via proposition VI.2) or on the grounds of book V. Below, we sketch the second option starting with definitions, see formulas (2) and Figs. 4, 5, and 6 respectively.

$$\begin{aligned} &x = a \cdot b \text{iff} 1 : b :: a : x, x = \frac{a}{b} \text{iff} a : b :: x : 1, \\ &x = \sqrt{a} \text{iff} 1 : x :: x : a. \end{aligned} \tag{2}$$

Following these definitions, we can put lines a and 1 on any arm of the angle. Then, instead of $1 : b :: a$: x, by definition V.5, we can also take $b : 1 :: x$: a. Now, the product ab is given by the proportion b : 1 :: ab : a, while ba is given by the

Fig. 3 *La Géométrie,* p. 389



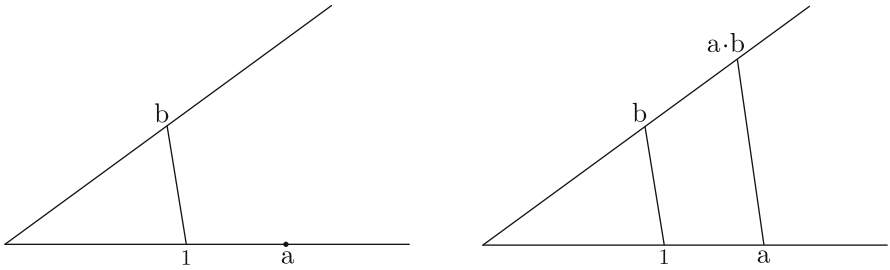


Fig. 4 Definition of multiplication

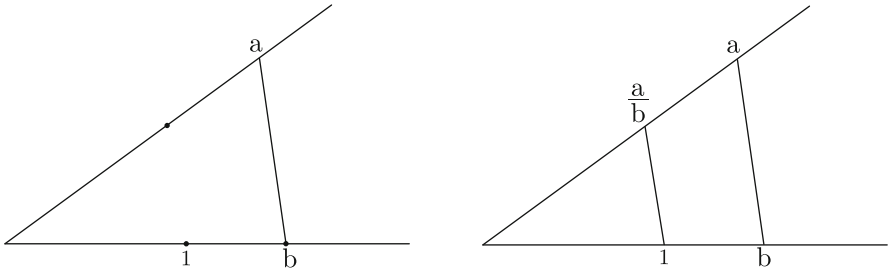


Fig. 5 Definition of division

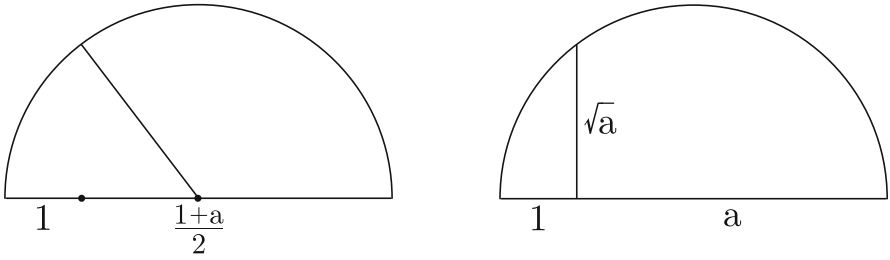


Fig. 6 Definition of square root

proportion $a : 1 :: ba : b$. From these proportions, by V.23, it follows that $ba : 1 :: ab : 1$. Hence, by V.9, $ba = ab$.[12]

For the rule (ED), suppose the proportion $a : b :: c : d$ obtains. According to the definition of division, we have

$$a : b :: \left(\frac{a}{b}\right) : 1, \quad c : d :: \frac{c}{d} : 1.$$

These three proportions give $\left(\frac{a}{b}\right) : 1 :: \left(\frac{c}{d}\right) : 1$. Hence, again by V.9, we get $\frac{a}{b} = \frac{c}{d}$. Finally, since $a = \frac{a}{b}b$, the equality $a = \frac{c}{d}b$ obtains.

5 The Field of Line Segments

The idea of a *negative* line segment in *La Géométrie* is related to roots of polynomials under the cover of *false* roots. Yet, Descartes' concept of a root evolves throughout the essay from an equation between lines to a polynomial equation $\varphi(x) = 0$.

5.1 True and False Roots

In book I, Descartes shows how to construct roots of the equation $z^2 = az + bb$. Since the solution segment line z can be found by straightedge and compass, he writes: "this root (*racine*) or unknown line can easily be found."[13] The discussed equation is presented as the equality between lines, rather than in the polynomial form $z^2 - az - b^2 = 0$. Then, Descartes considers two other versions of this equation, namely $z^2 = az - bb$ and $z^2 = -az + bb$. Minus sign as it appears in these equations should be interpreted from the perspective of the axiom E2, i.e., equations make sense whenever $az \geq b^2$ and $az \leq \mathrm{b}^2$, respectively. Descartes does not consider the case $z^2 = -az - bb$ at all, for the line segment z^2 would be equal to a *negative* line segment $-(az + bb)$.

In book III, he adopts a more liberal concept of a root: it is a line segment x satisfying a polynomial equation $\varphi(x) = 0$. At the same time, his concept of a root of a polynomial appears to be close to the modern one as he writes:

"the sum of an equation $[\varphi(x)]$ having several roots is always divisible by a binomial consisting of the unknown quantity diminished by the value of one of the true *(vrayes)* roots $[(x - a)]$, or plus the value of one of the false (*fausse*) roots $[(x + a)]$. [. . .] On the other hand, if the sum of the terms of an equation is not divisible by a binomial consisting of the unknown quantity plus or minus some other quantity, then this latter quantity is not a root of the equation" (Descartes 1637a, 372). That is,

$$\varphi(a) = 0 \Leftrightarrow (x - a) \mid \varphi(x),$$

[12] See (Hilbert 2015), p. 34 to check how Hilbert proves the law $ab = ba$.

[13] For details see Sect. 6.1 below.

assuming a could be positive or negative. Thus, false roots are line segments taken along with a special sign "–."

False roots are more widely discussed in the following passage:

"It often happens, however, that some of the roots are false (*faussent*) or less than nothing (*moindres que rein*). Thus, if we suppose x to represent the defect (*defaunt*) of a quantity 5, we have $x + 5 = 0$ which, multiplied by $x^3 - 9xx + 26x - 24 = 0$, yields $x^4 - 4x^3 - 19xx + 106x - 120$, an equation having four roots, namely three true roots, 2, 3, and 4, and one false root" (Descartes 1637a, 372).

Despite linguistic maneuvers, the mathematical meaning of the false root -5 comes straight from the following factorization of the polynomial

$$(x + 5)(x - 2)(x - 3)(x - 4) = x^4 - 4x^3 - 19xx - 120.$$

Thus, algebra of polynomials ensures meaning to *negative* line segments.

5.2 Zero and Sign Rules

These specific objects, i.e., *negative* line segments, are included into the algebra of line segments due to implicit sign rules as explained below.

"Every equation can have as many distinct roots *(racines),* that is values *(valures)* of the unknown quantity as the number of dimensions of the unknown quantity in the equation. Suppose, for example, $x = 2$ or $x - 2 = 0$, and again, $x = 3$, or $x - 3 = 0$. Multiplying together the two equations $x - 2 = 0$ and $x - 3 = 0$, we have $xx - 5x + 6 = 0$, or $xx = 5x - 6$. This is an equation in which x has the value 2 and at the same time x has the value 3. If we next make $x - 4 = 0$ and multiply this by $xx - 5x + 6 = 0$, we have $x^3 - 9xx + 26x - 24 = 0$ another equation, in which x, having three dimensions, has also three values, namely, 2, 3, and 4" (Descartes 1637a, 397). Descartes does not justify relations like $2 - 2 = 0$, $3 - 3 = 0$, or just the model one.

$$1 + (-1) = 0 \tag{3}$$

They are part of his practice of processing polynomials. In the like manner, while multiplying binomials, he implicitly adopts the following sign rules:

$$(-1) \cdot (-1) = 1, \quad (+1) \cdot (-1) = -1. \tag{4}$$

Indeed, rules (3) and (4) are enough to process *true* and *false* line segments.

5.3 Constructing False Roots

In Descartes' theory of polynomials, the final step in finding roots is their construction. Like common roots, false roots also have to be constructed. A solution to angle trisection illustrates how it is possible. Descartes managed to reduce the problem to

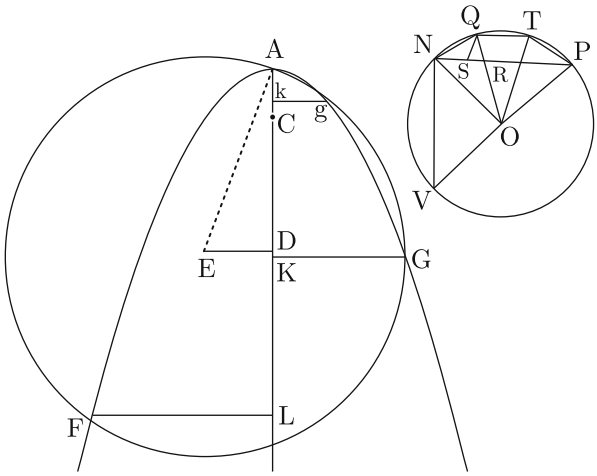


Fig. 7 *La Géométrie,* p. 399. Construction of roots (left). Forming an equation (right)

finding roots of the polynomial $z^3 - 3z + q = 0$ (see Descartes 1637a, 396–397). These roots are determined by intersections of a circle and parabola (see Fig. 7). Lines *kg* and *KG* are true roots, while *FL* is the false root, i.e., $(-FL)^3 - 3(-FL) + q = 0$. For a verification, one has to specify the characteristics of the circle and parabola, and apply rules (3) and (4).

5.4 Ordered Field of Line Segments

Let us assume M is the set of line segments. Let $-M = \{-x : x\ G\ M\}$ be the set of negative line segments, and $\mathbb{F} = -M \cup \{0\} \cup M$ be the set of common and *negative* line segments. Hence, adopting the above rules (3) and (4), the algebra $(\mathbb{F}, +, \cdot, 0, 1)$ is the field of line segments. Moreover, following the ideas of Artin and Schreier (Artin and Schreier 1926), since the set M is closed under sums and products, and $0 \notin M$, the order defined, by

$$x < y \Leftrightarrow y - x \in M,$$

makes the field $(\mathbb{F}, +, \cdot, 0, 1)$ an ordered field $(\mathbb{F}, +, \cdot, 0, 1, <)$. This simple observation justifies the actual practice of processing line segments throughout book III of *La Géométrie.*[14]

[14] Wallis (Wallis 1685), chapter LXVI, introduces negative numbers via the so-called number line. From the mathematical perspective, it builds on an ordered group $(G, +, 0, <)$. Descartes' way of introducing negative numbers builds on algebraic group $(G, +, 0)$, while the total order is generated by the above definition.

6 Applications

6.1 Pythagorean Theorem

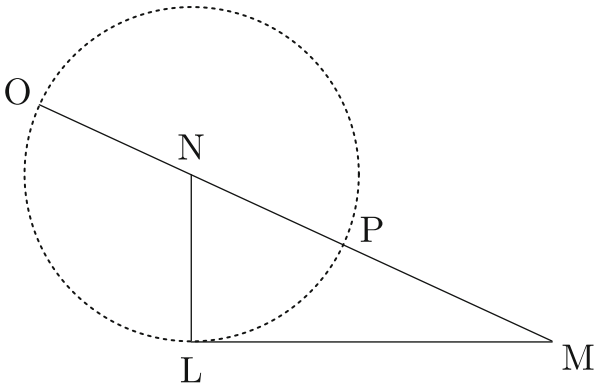
Euclid's proposition I.47 provides ancient Greek version of the Pythagorean theorem. It reads: "The square on (ἀπό) the side subtending the right-angle is equal to the [sum of] squares on (ἀπό) the sides surrounding the right-angle." The proposition ἀπό refers to the construction included in I.46, which reads: "[D]escribe a square on (ἀπό) a given straight line."

While the interpolation *sum of* characterizes modern translations, in the Greek mathematics, instead of *sum of* squares, just squares were considered, as exemplified by the *diorismos*' part of I.47: "the square on BC is equal to the squares on BA, AC."

Clavius (Clavius 1589), one of the most popular seventeenth-century editions of the *Elements,* includes essentially the same version of I.47.[15] Therefore we adopt the perspective that the following scheme □ , □ = □ accurately represents pre-Cartesian form of the Pythagorean theorem. In *La Géométrie,* however, it gets a new, modern-like form, namely $a^2 + b^2 = c^2$, where *a,b* are legs, and *c* is the hypotenuse of a right angle triangle. Symbols a^2, b^2, c^2 stand for line segments, and the Pythagorean formula makes sense within the framework of Descartes' arithmetic.[16] Here is textual evidence (the following passage refers to Fig. 8).

"this root *(racine*) or unknown line can easily be found. For example, if I have $z^2 = az + bb$, I construct a right triangle NLM with one side LM, equal to b, the square root of the known quantity bb, and the other side LN, equal to $\frac{1}{2}a$, that is, the half of other known quantity, which was multiplied by *z,* which I supposed to be the unknown line. Then prolonging MN, the base of this triangle, to O, so that

Fig. 8 *La Géométrie,* p. 302



[15] See (Clavius 1589), 229.

[16] In the modern mathematics, letters *a, b, c* as applied in the Pythagorean formula represent lengths of sides of a triangle, i.e., real numbers, rather than sides of the triangle.

NO is equal to NL, the whole line OM is the required line z. It is expressed in the following way (Descartes 1637a, 302)":

$$z = \frac{1}{2}a + \sqrt{\frac{1}{4}aa + bb}.$$

In this passage, Descartes shows how to construct roots of the equation $z^2 = az + bb$. Line segment $OM = z$ is to be the required root. It is composed of $ON = \frac{1}{2}a$ and the line segment NM identified by the formula $\sqrt{\frac{1}{4}aa + bb}$. Descartes' argument breaks at this stage. We could easily support it by the equation $\left(\frac{1}{2}a\right)^2 + b^2$.[17]

We can verify Descartes' solution algebraically, in a contemporary way, since by substitution the equation

$$\left(\frac{1}{2}a + \sqrt{\frac{1}{4}aa + bb}\right)^2 = a\left(\frac{1}{2}a + \sqrt{\frac{1}{4}aa + bb}\right) + bb,$$

is reduced to the identity $0 = 0$, given that

$$\left(\sqrt{\frac{1}{4}aa + bb}\right)^2 = \frac{1}{4}aa + bb.$$

Indeed, that is how, i.e., *par le calcul,* Descartes verifies that specific lines are roots of the equation $z^3 = 3z - q$ (see Descartes 1637a, 397).[18]

For another example, let us return back to Fig. 3. The following passage resumes the argument and presents the follow up:

"For, putting a for BD or CD, and c for EF, and x for DF, we have CF $= a - x$, and since CF or $a - x$ is to FE or c, as FD or x is to BF, which is consequently $\frac{cx}{a-x}$. Now, since in the right triangle BDF one side is x, and other is a, their squares, which are *(leurs carrés, qui sont)* $xx + aa$, are equal to the square of the base, which is $\frac{ccxx}{xx-2ax+aa}$" (Descartes 1637a, 389).

[17] Bos (Bos 2001), p. 300, interprets Descartes' solution as follows: "[the solution of the equation $z^2 = az + bb$] involved the root $\sqrt{\frac{1}{4}aa + bb}$. Descartes did not interpret this root by means of the unit, but classically (and indeed much more naturally) as the hypotenuse of a right-angled triangle with sides 1 $2a$ and b." Whatever means *classically*, or *naturally* in this context, the formula $\frac{1}{4}aa + bb$ refers to the Pythagorean theorem, while the formula $\sqrt{\frac{1}{4}aa + bb}$ identifies the hypotenuse of the triangle ΔNLM.

[18] Bos (Bos 2001), p. 305 suggests geometric verification of Descartes' solution. Yet, Descartes refers back to that solution in (Descartes 1637a), p. 326. Then, a formula under the square root has no other geometrical meaning but a line segment. See Sect. 7.2 below.

Thus, in the right angle triangle *BFD,* with $FD = x$, and $BD = a$, the square of the hypotenuse BF (Descartes names it *the base*) is clearly identified with the line $x^2 + a^2$, that is

$$x^2 + a^2 = BF^2.$$

Moreover, in this passage, we can also spot a word pattern *leurs carrés, qui sont* consistently used in *La Géométrie.*[19] It is also used in Latin translation of Pappus, where the phrase "if there be given a ratio (λόγος;) of the rectangle contained by the two lines so drawn to the square on the third" is interpreted as follows: "let there be a proportion (*proportio*) of the rectangle contained by the two lines so draw to the square of the third (ad *quadratum reliquæ*)."[20]

Thus, the phrase *square of* which makes sense within the framework of *La Géométrie* replaced *square on* which refers back to Euclidean paradigm. On the level of symbols, the ancient schematic phrase "the square on BC is equal to the squares on BA, AC" was replaced by the formula $a^2 + b^2 = c^2$.

6.2 Formula for a Curve

Arithmetic of line segments plays the crucial role in how Descartes derives the formula for a curve consisting of intersections of moving lines. In the Fig. 9, the straight line GCL turns around the pivot point G, while the line CNK moves along some other straight line (the vertical one). The line CNK is identified by its slope *KL/NL;* since $KL = b$, $NL = c$, in the formula of the curve, it is represented by the term $\frac{b}{c}$. The intersection point $C = C(x,y)$ is identified by two line segments $BA = x$, $CB = y$. Moreover, $GA = a$. A formula for the resulting curve is a polynomial relation between x and y. The relevant passage of *Geometry* reads as follows:

"Then I say that as NL is to LK, or as c is to b, so CB, or y, is to BK, which is therefore equal to $\frac{b}{c}y$. Then BL is equal to $\frac{b}{c}y - b$, and AL is equal to $x + \frac{b}{c}y - b$. Moreover, as CB to LB, that is, as y is to $\frac{b}{c}y - b$, so AG or a is to LA or $x + \frac{b}{c}y - b$. Multiplying the second by the third, we get $\frac{ab}{c}y - ab$ equal to $xy + \frac{b}{c}yy - by$, which is obtained by multiplying the first by the last. Therefore, the required equation is.

$$yy = cy - \frac{cx}{b}y + ay - ac" \quad \text{(Descartes 1637a, 321–322).}$$

[19] Here is other occurrence of this standard: "*ss, qui est le carreé de la baze*, *qui sont les carées des deux* ([Descartes 1637a], 342).

[20] See (Descartes 1637a), 305. Descartes cites, probably, (Pappus 1588), p. 165. It is puzzling what Commandino meant by saying "square of the third." Oaks (Oaks 2018), p. 251–252, referring to another passage of (Pappus 1588), observes: "Squares and cubes were worded with the proposition *ex* ("of")," yet this phrase stood for squares on. For the Greek text of Pappus see (Pappus 1877), vol. II, 676.

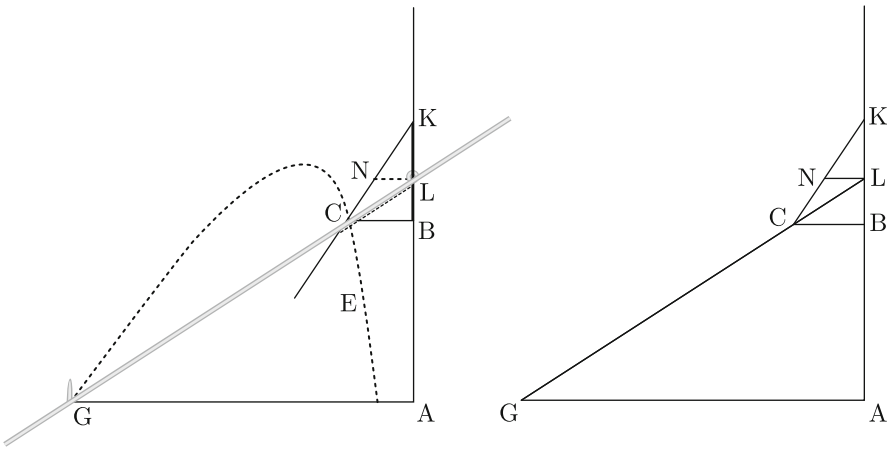


Fig. 9 Descartes's hyperbola. On the left, diagram from *La Géométrie,* p. 320

Derivation of the formula for the line EC (actually it is a hyperbola) initially employs Euclid's theory of similar figures. Since $\Delta NKL \sim \Delta CKB$, the proportion $NL : KL :: CB : BK$ obtains. It is simplified to the following form $c : b :: y : BK$. Yet, Descartes' conclusion is far more reaching, namely

$$c : b :: y : BK \Rightarrow BK = \frac{b}{c}y. \tag{5}$$

As a result, $LB = \frac{b}{c}y - b$, and $LA = x + \frac{b}{c}y - b$.

In the next step, since $\Delta CLB \sim \Delta GLA$, the proportion $CB : LB :: AG : LA$ obtains, or

$$y : \left(\frac{b}{c}y - b\right) :: a : \left(x + \frac{b}{c}y - b\right).$$

Again, by the implicit use of the rule (ED), Descartes could turn that proportion into the following equation

$$a\left(\frac{b}{c}y - b\right) = y\left(x + \frac{b}{c}y - b\right). \tag{6}$$

However, he actually writes

$$\frac{ab}{c}y - ab = xy + \frac{b}{c}yy - by. \tag{7}$$

In order to simplify (6) to (7), one needs to apply rules of an ordered field. And finally, to get from (7) the equation of the hyperbola, namely

$$yy = cy - \frac{cx}{b}y + ay - ac,$$

one should apply, again, rules of on ordered field.

7 Pappus Problem

7.1 Formula for *Locus*

The Pappus problem is the main theme of *La Géométrie.* It is a question of *locus.* For four lines, it requires to find the *locus* of such points C that lines drawn from C meet the given lines at given angles and satisfy (in a simplified version) the following condition: the rectangle contained by two of them is equal to the rectangle contained by the other two. Accordingly, a required condition for six lines takes this form: the (right rectangular) prism contained by three lines is equal to the prism contained by the other three. However, Pappus believed that the problem for more than six lines was meaningless, "since there is no figure of more than three dimensions."[21]

Descartes interprets the Pappus problem in his arithmetic of line segments, while the *locus* of the point C – in terms of a formula for a curve. He discusses the four lines problem in great detail. At first, he derives a formula for the *locus*, then draws (constructs) that *locus*, and finally, explains "all the given quantities by numbers."

Thus, Descartes assumes that straight lines EF, AD, AG, and GF are given "in position," which actually means that angles $\angle DAB$, $\angle FEA$, and $\angle AGH$ are given, moreover k,l are also given, where $EA = k$, $AG = l$ (see Fig. 10). He seeks a curve consisting of points $C = C(x, y)$, where $AB = x$, $BC = y$, such that segment lines CB, CD, CF, and CH, drawn from C to the given lines, make given angles α, β, γ, and δ respectively, and moreover they satisfy the following condition

$$CB \cdot CF = CD \cdot CH \tag{8}$$

The assumption on angles, specifically on $\angle DAB$ and α, amounts to knowing all angles in the triangle ΔARB (see Fig. 12). Then, Descartes considers one by one triangles

$$\Delta DRC, \Delta ESB, \Delta FSC, \Delta BTG, \Delta CTH$$

and, similarly, determines their angles. In the next stage, he solves these triangles in order to determine lines involved in Eq. (8) in terms of x, y and the given quantities. As for the triangle ΔARB, he writes:

"Now, since all the angles of the triangle ARB are known, the ratio between the sides AB and BR is also known. I assume it is z to b. Since AB is x, then RB

[21] That problem was originally stated in book VII of Pappus' *Collection.* In (Descartes 1637a), pages 304–306, Descartes cites, probably, (Pappus 1588). The English translation of Pappus' phrase concluding this paragraph, after (Smith and Lathan 2007), p.

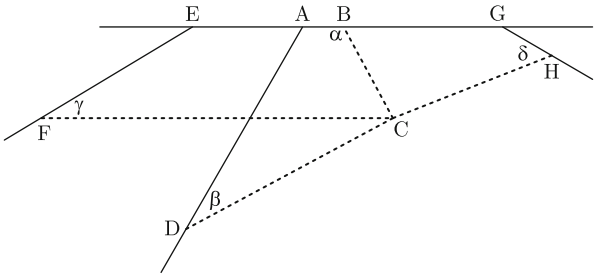


Fig. 10 The four lines Pappus problem

will be $\frac{bx}{z}$; and the whole CR will be $y+\frac{bx}{z}$, since B lies between C and R" (Descartes 1637a, 310).[22]

In short,

$$AB = x, AB : BR :: z : b \Rightarrow RB = \frac{bx}{z} \& CR = y + RB.$$

Indeed, the rule (ED) is implicitly used all through derivation of the formula for the sought *locus* (Fig. 11).[23]

Since all the angles of considered triangles are known, the following proportions are also known[24]

$$AB : BR :: z : b, CR : CD :: z : c, BE : BS :: z : d,$$
$$CS : CF :: z : e, ; BG : BT :: z : f, TC : CH :: z : g.$$

In result, the condition (8) is turned into the following formula

$$yy = \frac{(cfglz - dekzz)y + (-dezzx - cfgzx + bcgzx)y + bcfglx - bcfgxx}{ezzz - cgzz} \quad (9)$$

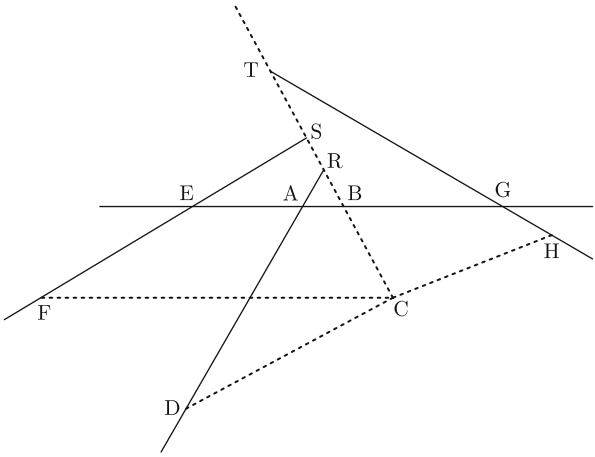
It is "the locus generated by infinity of points C, such that, if four lines CB, CD, CF, and CH be drawn through it making given angles with the given lines, the product of CB and CF is equal to the product of CD and CH" (Descartes 1637a, 325).

[22] (Smith and Lathan 2007) translates this passage as follows: "Now, since all the angles of the triangle ARB are known, the ratio between the sides AB and BR is known. If we let AB: $BR = z$: b, since $AB = x$, we have $RB = \frac{bx}{x}$; and since B lies between C and R, we have $CR = y + \frac{bx}{x}$" (Smith and Lathan 2007, 29). Symbols applied here, and in translation of pages 310–311 alike, do not distinguish between proportion and equality.

[23] Bos (Bos 2001), p. 315, footnote 6, also discusses this passage, yet does not identify the rule (ED). Indeed, he does not identify any implicit rule in (Descartes 1637a) at all. He also does not discuss the role of Euclid's proportion in theory of similar triangles applied by Descartes.

[24] In a numerical explanation provided in (Descartes 1637a), p. 333, Descartes assumes $z = 1$. Then parameters b to g have to be determined based on the axiom E5. Generally, if z is given, b to g exist, however, they are not given in the same sense as, say, $EA = k$. In Euclid's *Data*, when two lines a, b are given, their ratio is also given, which means a: b:: a: b (see Taisbak 2003, 37).

Fig. 11 La Géométrie, p. 311



We can turn Eq. (9) into the form $\psi(x,y) = 0$, where $\psi(x,y)$ is a polynomial of two variables. Otherwise, the *locus* consists of points $C = C(x,y)$, where coordinates x, y are in the relation $\psi(x,y) = 0$.

When the variable y is set, say $y = y_0$, the second coordinate, say x_0, is a root of the polynomial equation $\varphi(x) = \psi(x,y_0) = 0$. In this way, the second coordinate of point C is determined, i.e., $C = C(x_0,y_0)$.

"Since this condition can be expressed by a single equation in two unknown quantities [i.e., $\psi(x, y)$], we may give any value we please to either x or y and find the value of the other from this equation" (Descartes 1637a, 313).

This procedure enables one to determine as many points $C = C(x_0,y_0)$ as one likes and draw the curve $\psi(x, y)$ point wise. Thus, in general, a solution to Pappus problem for any number of lines is approximated rather than drawn by an instrument[25]:

"If then we should take successively an infinite number of different values for the line y, we should obtain an infinite number of values for the line x, and therefore an infinity of different points, such as C, by means of which the required curve could be drawn" (Descartes 1637a, 313).

Nevertheless, the locus for the four lines problem can be drawn, and Descartes studies it carefully.

7.2 Constructing *Locus*

For the economy of calculations, Descartes introduces new parameters,

[25] Note that according to modern standards, a solution is reached when a formula for a curve is given, since we have techniques of studying curves based on their formulas alone. However, as Descartes discovered a formula for a curve, it could hardly be an accepted solution at that time.

$$2m = \frac{cfglz - dekzz}{ez^3 - cgzz}$$
$$\frac{2n}{z} = \frac{dezz + cfgz - bcgz}{ez^3 - cgzz}$$
$$o = -\frac{2mn}{z} + \frac{bcfgl}{ez^3 - cgzz},$$
$$-\frac{p}{m} = \frac{nn}{zz} - \frac{bcfg}{ez^3 - cgzz}$$

As a result, the Eq. (9) takes a simplified form

$$yy = 2my - \frac{2n}{z}xy + \frac{bcfglx - bcfgxx}{ez^3 cgzz} \tag{10}$$

He treats (10) as a quadratic equation in variable y – that is $y^2 = ay + b^2$ – and determines its root by the method exposed in Sect. 6.1 above, that is $y = \frac{1}{2}a + \sqrt{\frac{1}{4}a^2 + b^2}$. In this way[26]

$$y = m - \frac{n}{z}x + \sqrt{mm + ox - \frac{p}{m}xx}. \tag{11}$$

Then Descartes constructs such points *K, I, L* that *K* lies on the line *BC*, and $KB = m$, $KI = x$, $KI \,||\, AB$, and *I* lies on the line *BK* and (see Fig. 12)

$$IK : IL :: n : a.$$

and $KL = \frac{n}{z}x$. When the triangle ΔIKL is given, the following proportion is also given

$$KL : IL :: n : a.$$

Hence (by the same magic trick applied all throughout *La Géométrie*),[27]

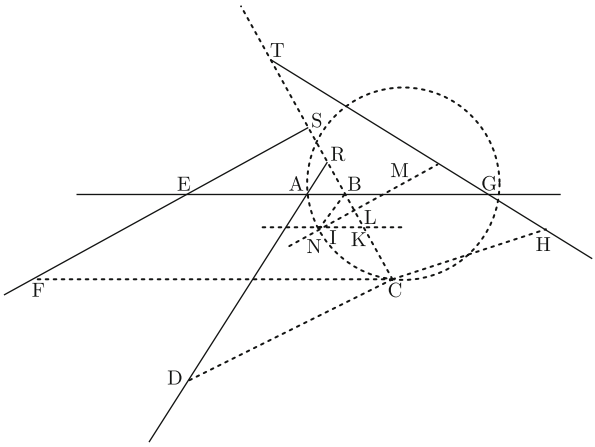
$$IL = \frac{a}{z}x.$$

Given *LC* is not equal zero "the point C always lies on one of the three conic sections, or on a circle" (Descartes 1637a, 328), namely when

[26] Bos (Bos 2001) advocates the thesis that Descartes' square root should be interpreted geometrically rather than via some definition, e.g., our definition (2). While it is easy to read the formula $\sqrt{\frac{1}{4}a^2 + b^2}$ in this way (see Bos 2001, 300, 305), the formula $mm + ox - \frac{p}{m}xx$, as it occurs in (11), has no geometrical interpretation – we can only state that it represents a line segment.

[27] Bos (Bos 2001), p. 322, processes these proportions as if they were just fractions.

Fig. 12 *La Géométrie,* p. 334



$$y = m - \frac{n}{z}x + \sqrt{mm + ox},$$

then the *locus* is a parabola; when

$$y = m - \frac{n}{z}x + \sqrt{mm + ox + \frac{p}{m}xx},$$

it is a hyperbola; and when

$$y = m - \frac{n}{z}x + \sqrt{mm + ox - \frac{p}{m}xx},$$

it is an ellipse; when moreover, $\angle \mathrm{I}LC = \frac{\pi}{2}$ and $a^2m = pz^2$, then the *locus* is a circle.

Descartes ends this lengthy argument with an example that refers to the circle and "explains given quantities by numbers" (see Descartes 1637a, 333–334).

7.3 Numerical Example

Descartes' numerical explanation refers to the solution represented in Fig. 12. At first, he assumes

$$EA = 3, AG = 5, AB = BR, BS = \frac{1}{2}BE, GB = BT,$$
$$CD = \frac{3}{2}CR, CF = 2CS, CH = \frac{2}{3}CT, \angle ABR = \frac{\pi}{3}$$

Under these assumptions, the equation of the *locus* is as follows:

$$yy = 2y - xy + 5x - xx,$$

and

$$y = 1 - \frac{1}{2}x + \sqrt{1 + 4x - \frac{3}{4}xx}.$$

Descartes easily finds that $n : z = \frac{1}{2}$, and "KL must be equal to one-half KI." As a result, the triangle ΔIKL is a right angle triangle. Since $KL = \frac{x}{2}$, then $IL = \sqrt{x^2 - \frac{1}{4}x^2} = x\sqrt{\frac{3}{4}}$. By putting these results in the proportion $IK : IL :: z : a$, he obtains $x : x\sqrt{\frac{3}{4}} :: z : a$.

Assuming "the quantity represented by z above is 1," he gets

$$a = \sqrt{\frac{3}{4}}, m = 1, o = 4, p = \frac{3}{4}.$$

Hence $IM = \sqrt{\frac{16}{3}}$, and $NM = \sqrt{\frac{19}{3}}$. "And since *aam* (which is $\frac{3}{4}$) is equal to *pzz*, and the angle ILC is the right angle, it follows that the curve *NC* is the circle" (Descartes 1637a, 333–334).

7.4 Magnitudes, Numbers, and Quantities Versus Elements of an Ordered Field

In the sixteenth century, the Greek μέγεθος was interpreted by the Latin *magnitudo* or *quantity*. Yet the other term was ambiguous as it could mean a line segment or its measure.[28] In mathematical practice, numbers were assigned to line segments, and then they were terms in expressions derived on a geometric basis.[29] Thus, for example, the term $\sqrt{101}$ stood for the hypotenuse of a triangle with legs 10 and 1.[30]

[28] The ambiguity of *quantity* could be traced back to Aristotle's definition (*Categories* 6, 4b20–25), where quantities (ποσόν) are exemplified by numbers and line segments.

[29] See (Oaks 2018), § 3.7.

[30] Malet (Malet 2006) studies the conceptual shift between numbers and magnitudes based on the sixteenth- and seventeenth-century editions of the *Elements*. He specifically observes: "not only was the neat and consistent separation between the Euclidean notions of numbers and magnitudes preserved in Latin medieval translations [. . .], but these notions were still regularly taught in the major schools of Western Europe in the second half of the fifteenth century. By the second half of the seventeenth century, however, the distinction between the classical notions of (natural) numbers and continuous geometrical magnitudes was largely gone, as were the notions themselves" (Malet 2006, 64–65). While Malet explores bridging the conceptual gulf between numbers and magnitudes, we explore basically the same issue, however, from the perspective of mathematical techniques.

It was also a common sixteenth-century practice to conclude geometric propositions with numerical examples. Viete's practice was even more radical: "All numbers with which Viete operates in his works are the measures of geometric magnitudes. They can be any positive number that arises in geometric calculation, including fractions and irrational roots, and they are never zero, negative, or complex" (Oaks 2018, 268).

It was also a common seventeenth-century practice to turn proportion $a : b :: c : d$ into the equation $ad = bc$. Then it was interpreted via proposition VI.16 of the *Elements* as an equality of rectangles with sides *a, d,* and *b, c.* Viète, again, applied it "regardless of the dimensions of the related magnitudes" (Oaks 2018, 247), and it is considered to be a bold achievement. Bos identifies a pre-Cartesian author who applied a unit line in multiplication of line segments.[31] All these findings may raise doubts whether is there a germ of novelty in *La Géométrie*?

In Descartes' numerical example, the really striking idea is the assumption $z = 1$. If z can be 1, what is the reason for Euclidean proportion at all. And in fact, that is how Newton, a dedicated student of Descartes' geometry, defines number:

"By Number we understand not so much a Multitude of Unites, as the abstracted Ratio of any Quantity to another Quantity of the same kind which we take for Unity. [Number] is threefold: integers, fracted, and surd, to which last Unity is incommensurable" (Newton 1720, 2).

Let us note also that when Descartes applies surds like $\sqrt{3}$ – in the above numerical example but also as coefficients of polynomials – they do not need to be interpreted as a hypotenuse of a triangle. His algebra of line segments provides a coherent foundation for multiplication, division, and square roots for numbers and line segments. In general, it provides a framework for processing both numerical and geometrical terms. Another Newton's definition embodies that idea:

"For by a Square, or Rectangle, or a Quantity of two Dimensions, we do not always understand a surface, but most commonly a Quantity of some other Kind, and very often a Line which is produced by the Multiplication of two other lines" (Newton 1720, 5).

And indeed, a multiplication, division, and square root of line segments is explained further based on a diagram alike Descartes' Fig. 1 (see Newton 1720, 5–7).

There is also an ongoing debate, whether Descartes applied arithmetic operations to line segments, or geometric operations to numbers. That controversy rests on a contrast between numbers and magnitudes. From the perspective of an ordered field, Descartes developed a system of processing formulas which combines numbers and line segments, both positive and negative.[32]

[31] See (Bos 2001), 154–156.

[32] Schubring (Schubring 2005) reaches another conclusion, namely: "A separation of the concept of number from the concept of quantity was not realized here [in *La Géométrie*], and hence no fusion of positive and negative numbers to the unified area of integers" (Schubring 2005, 49). This statement assumes, and in fact it is the core idea of (Schubring 2005), that a kind of separation of numbers from line segments is a condition sine qua non for the concept of negative numbers. We argue, on the contrary, that Descartes developed a unified framework for numbers and line segments (quantities), and within it he introduced negative elements.

At the end of the nineteenth century mathematicians sought to formalize the Greek concept, as they believed, of magnitude. In these studies, the term magnitude was a synonymous with quantity (English *magnitude* and *quantity,* French *grandeur* and *quantitaté,* and German *Grösse* and *Quantität*). From the mathematical perspective, all these tries shared the common core: magnitude is an Archimedean, totally ordered semigroup.[33] When Hilbert (Hilbert 1900, 2015) introduced a definition of ordered field and axioms for real numbers, that research program faded away. One may view Descartes' *La Géeoméetrie* as its starting point.

8 Ancient Greek Numerical Practice Based on Geometry

The practice of basing numerical analysis on geometrical results mentioned above dates back to ancient Greek mathematics. Archimedes' *Measurement of a circle,* derivation of the size of chords in Ptolemy's *Almagest,* or Heron's *Metrica* provide model examples of that tradition. In the next section, we will analyze proposition I.8 of *Metrica* to illustrate in what sense a numerical analysis was based on geometry and how Descartes' field of line segments enabled to combine that kind of numerical practice and geometrical studies. To this end we need to cover basics of Euclid's arithmetic.[34]

8.1 Euclidean Arithmetic and Beyond

Euclid's *Elements* books VII to IX provide foundations for the ancient Greek number theory. Number (ἀριθμός) means the modern integer greater than 1. Zero does not occur in the *Elements,* while 1 is the so-called monad (μονάς). The first three definitions of book VII clarify the relation between monads and numbers. They read:

"A unit (μονάς) is (that) according to which each existing (thing) is said (to be) one. And a number (is) a multitude (πλῆθος) composed of units. A number is part of a number, the lesser of the greater [n < m], when it measures the greater."

We can formalize definition VII.2 as follows:

$$n = \underbrace{1 + \ldots + 1}_{n\ times}.$$

The term *n measures m* means that for some $k \in \mathbb{N}$ the equality holds

$$m = kn = \underbrace{n + \ldots + n}_{k\ times}.$$

[33] See (Błaszczyk and Fila 2020), § 6.

[34] Another way of studying arithmetic was developed by Diophantus (Christianidis and Oaks 2021). It is rather unrelated to the issues discussed here.

By multiplication of numbers as introduced by definition VII.15, we formalize

$$n \cdot m = \underbrace{m + \ldots + m}_{n \text{ times}}.$$

It can be shown[35] that Euclid's numbers form a structure $(A, +, \cdot, <)$, where addition and relation *greater than* are primitive notions characterized by axioms E1-E3, as formalized in Sect. 2 above.

Next to these common operations, in definition VII.20 Euclid introduces a proportion of numbers:

"Numbers are proportional when the first is the same multiple, or the same part, or the same parts, of the second that the third (is) of the fourth."

We formalize it by

$$m : n :: k : l \Leftrightarrow \exists p(pn = m, pl = k) \vee \exists p(pm = n, pk = l) \vee$$
$$\vee(\exists u_1, u_2)\,(\exists p, q)(m = pu_1, n = qu_1, k = pu_2, l = qu_2).$$

Contrary to the theory of magnitudes, the proportion of numbers does not refer to the relation *greater than.* There is also no counterpart of definition V.7 in Euclid's theory of numbers. Finally, axiom E5 does not hold in the domain of numbers.

In book X of the *Elements,* numbers and magnitudes, in particular their proportions, are studied within the same but unspecified framework.

In the late sixteenth and throughout the seventeenth century, there were tries to find a common ground for definitions V.5 and VII.20. Palmieri (Palmieri 2001) studies in great detail that kind of research carried out by Clavius and Galileo. They both found definition V.5 problematic as it refereed to infinitely many cases – $(\forall m, n)\,(\ldots)$ in our formalization.[36] For Clavius, it was a problem, say, in the foundations of mathematics.[37] For Galileo it also concerned his studies in natural

[35] See P. Błaszczyk, M. Fila, *On Bolzano and Greek concept of continuity* in this *Handbook.*

[36] It is worth noticing that Clavius' reading of Euclid's definition V.5 (see Palmieri 2001, 561) perfectly fits our formalization, as presented in x 2 above, except quantifiers. Note also that when Euclid applies def. V.5 in proposition VI.1, then, instead of $(\forall m, n)\,(\ldots)$, he simply takes $m = n = 3$. Indeed, it is his standard mode of operation whenever, from the modern perspective, the universal quantifier is involved.

[37] Calvius, for example, in a comment to the definition of ratio (Euclid 2007, def. V.3), portrays numbers on a par with lines and other magnitudes. Then he seeks to prove an equivalence of definitions V.5 and VII.20. Here is a sample of his argumentation: if $D \cdot E = C \cdot F$, then, by VII.19, it gives the proportion of numbers, $D/C :: F/E$, and by inversion, the desired proportion $C/D :: E/F$ obtains; see (Palmieri 2001), 558, 567, respectively. Thus, although we stated above that the multiplication of numbers is commutative, it was not that obvious in the ancient and early modern arithmetical practice.

Here is a sample of his argumentation: if $D \cdot E = C \cdot F$, then, by VII.19, it gives the proportion of numbers, $D/C :: F/E$, and by inversion, the desired proportion $C/D :: E/F$ obtains; see (Palmieri 2001), 558, 567, respectively. Thus, although we stated above that the multiplication of numbers is commutative, it was not that obvious in the ancient and early modern arithmetical practice.

science, where he often referred to the very definition of proportion.[38] Whatever motivations, no one even tried to question results of book V. More importantly, although there were other attempts to redefine the ancient proportion in the seventeenth century, none of them yielded new propositions.[39]

In general, Descartes arithmetic turned out to be more successful foundation of mathematics and natural science than a theory of proportion. And the reason is simple: It enables to bring together within a structure of an ordered field both geometrical and arithmetical knowledge. We can appreciate it by comparing Heron's proof of the formula for the area of a triangle and its modern counterparts.

8.2 Heron's Formula for the Area of a Triangle

Heron's *Metrica* (Schöone 1903) is a study on measurement of plane figures and solids. It applies a unit square (μονάς) to approximate areas of triangles, polygons, ellipses, and circles. In calculations, lines are considered multiples of a unit line (μονάς), and they are represented by natural numbers. Then, for example, through the Pythagorean theorem, Heron determines the square of a hypotenuse, and, when it is needed, approximates its square root.

In propositions I.5–6, and I.8, Heron determines numerical area (έμβαδόν) of a triangle when its sides (in fact, lengths measured by integers) are given. In I.5–6, he applies Euclid's version of the law of cosines, namely II.12–13. Regarding an acute angle triangle, Euclid's II.13 brings in the following equation (see Fig. 13):

$$AC^2, 2BC.BD = AB^2, BC^2,$$

where AC^2 stands for the square on AC etc., and $BC.BD$ – for the rectangle contained by BD and BC.[40]

Now, Heron substitutes numbers for names of line segments, namely $AB := 13$, $BC = 14$, $AC = 15$ (of μονάδων), and as a result obtains $BD = 5$. In the numerical version, however, the above formula turns into the following one

$$AC \cdot AC + 2BC \cdot BD = AB \cdot AB + BC \cdot BC.$$

In the next step, by the *Elements* I.47, Heron determines $AD = 12$, then $BC \cdot AD = 168$.

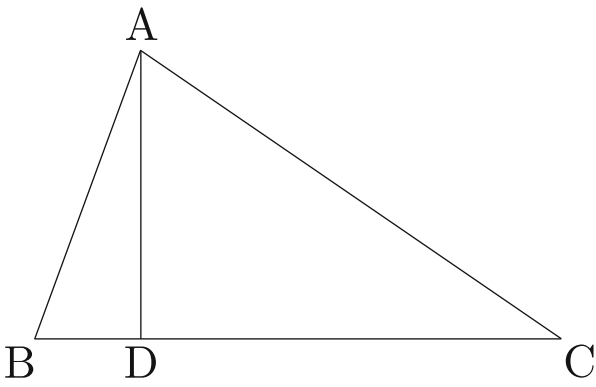
Finally, he finds that the triangle ABC is 84 of μονάδων.

[38] Galileo sought to bypass the reference to infinitely many case in definition V.5 by founding proportion on definition V.7, namely (in our stylization) $A/B :: C/D$ iff it is not the case that $A/B < C/D$, and $A/B > C/D$; see (Palmieri 2001), 576.

[39] See (Giusti 1993).

[40] For a detailed analysis of Euclid's book II, see (Błaszczyk 2020).

Fig. 13 Heron, Metrica, I.5



In I.7 – through a reference to X.54, as Taisbak (Taisbak 2014) observes – Heron states a numerical identity to the effect $\sqrt{AB^2 \cdot CD^2} = AB \cdot CD$. This proposition is not accompanying by any diagram.[41]

In I.8, Heron derives a formula for the area of a triangle without determining its height when its three sides are given.

The thesis of the proposition reads: "Let the triangle *ABC* be given, and let each AB, *BC*, *CA* be given. To find its area (ἐμβαδόν)."

At first, Heron assumes that sides of a triangle ΔABC are 7, 8, and 9 (see Fig. 14). Then he calculates $s = \frac{7+8+9}{2}$, and finds that $s(s–7)(s-8)(s-9) = 720$. By taking the squares of some guessed numbers (greater or smaller than 720), he arrives at the conclusion that $\sqrt{720}$ is about (ἔγγιστα) $26\frac{1}{2} + \frac{1}{3}$.

Referring to I.5, and I.6, Heron observes that instead of finding the height of the triangle, he can provide a "general method" of determining the area of a triangle. The proof starts with an observation that $\Delta ABC = GC.EH$, where $GC = \frac{1}{2} \times (AB + AC + BC)$, and EH is a radius of a circle inscribed in the triangle ΔABC – in our notation, s and r, respectively. To simplify the notation further, let Δ stand for ΔABC.[42] Thus, $\Delta = rs$. Then comes the crucial statements:

[1]. that between CG, EH is the side (πλευρά) of that on CG and that on EH, thus
[2]. (Artin and Schreier 1926) the area (ἐμβαδόν) of the triangle ABC multiplied (γενόμενον) by itself is equal to the square on CG by that on EH.[43]

[41] Note that in Descartes' system, $\sqrt{a^2} = a$ no matter if a is a number or a line segment. See formula (1), in Sect. 3 above.

[42] For the same reasons we apply fractions rather than symbol of proportion; moreover, symbol s^2 stands for a square on s, and rs – for a rectangle between r, s. We also add references to the *Elements* to underline a reliance of this proof on the *Elements* V.

[43] The term γενόμενον is applied on regular basis in the *Elements*, books VII, IX, i.e., in arithmetic context. On the other hand, the term πλευρά is applied both in geometrical and arithmetical contexts.

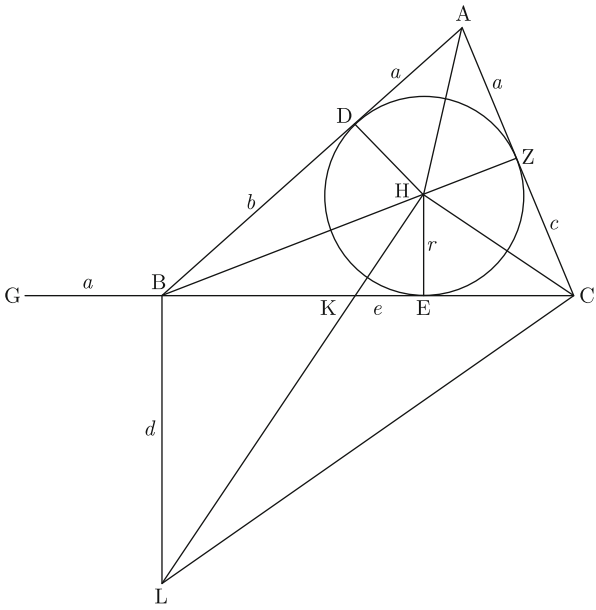


Fig. 14 Heron, *Metrica,* I.8 (modified: lower case letters added)

We can formalize them as follows[44]

$$[1]\, sr = \sqrt{s^2 r^2}, \quad [2]\, \Delta\Delta = s^2 r^2.$$

The next step brings in basic geometrical insights. Heron constructs right angle triangles ΔLHC and ΔLBC, and observes that since they share the hypotenuse LC, the quadrilateral $BHCL$ is inscribed in a circle. Thus, via *Elements,* III.22, $\angle BLC$, $\angle BHC = \pi$. It follows from the construction that $\angle DHA$, $\angle BHC = \pi$. Therefore, $\angle BLC = \angle DHA$. It entails the following proportions[45]:

$$\begin{aligned} \Delta BLC \sim \Delta DHA \quad &\rightarrow \quad \frac{b+c}{d} = \frac{a}{r} \\ &\overrightarrow{V.16} \quad \frac{b+c}{a} = \frac{d}{r} \end{aligned}$$

Further geometrical analysis yields another proportions:

[44] Shöne (Schöone 1903) and Thomas (Thomas 1957) instead of a translation of this passage provide formulas (in Greek letters): $GC \times EH = \sqrt{GC^2 \times EH^2}$, $ABC^2 = GC^2 \times EH^2$. Acerbi and Vitrac translate it literally: "mais le [rectangle contenu] par ΓΘ, EH est un [côté] du carré sur ΓΘ par celui sur EH; il en résultera donc quel'aire du triangle ABG multipliée par elle-même est égale au [carré] sur GJ par celui sur EH" ([Acerbi and Vitrac 2014], 166).

[45] We assume, $BE = BD = b$, $EC = EZ = c$, $DA = AZ = a$, $KE = e$, $BL = d$, etc., see Fig. 14.

$$\begin{array}{lll} \Delta BKL \sim \Delta EKH & \rightarrow & \dfrac{d}{r} = \dfrac{b-e}{e} \\ \dfrac{b+c}{a} = \dfrac{b-e}{e} & \overrightarrow{V.18} & \dfrac{a+b+c}{a} = \dfrac{b}{e} \\ \dfrac{s}{a} = \dfrac{b}{e} & \overrightarrow{V.I.1} & \dfrac{s^2}{sa} = \dfrac{bc}{ec} \\ HE \perp KC, \angle KHC = \pi/2 & \overrightarrow{V.I.8,16} & ec = r^2 \\ & \rightarrow & s^2r^2 = sabc. \end{array}$$

Finally, to show that the resulting formula depends only on given sides, Heron observes that

$$a = s - (b - c), b = s - (a + c), c = s - (a + b).$$

The proof of I.8 is followed by another numerical example in which Heron shows that the area of a triangle with sides 12, 13, 14 is 84.

The numerical context in Heron's propositions in obvious. It is also obvious that to get these results Heron has to root his arguments in geometry. Yet finally, when geometry is helpless, he adopts a numerical interpretation. Thus, while the proportion $\frac{s^2}{sa} = \frac{bc}{ec}$ has a geometrical sens, the formula $s^2r^2 = sabc$, or $\Delta = \sqrt{sabc}$, can be interpreted only numerically. However, in Descartes' framework, formulas s^2r^2 or $\sqrt{sabc}$, whether s,r,a,b,c are numbers, or line segments are just as meaningful.

For comparison sake, let us go back to Fig. 13 and adopt a modern notation, namely $AD = h$, $BC = a$, $AC = b$, $AB = c$, $BD = d$. Contrary to Heron's subtle arguments, modern proofs apply just basic laws, yet in algebraic stylization, namely, the area of the triangle equals to $\frac{1}{2}ab \sin \angle A$, or $\frac{1}{2}ah$, the cosine law takes the form $b^2 = a^2 + c^2 - 2\,ac \cos \angle B$, and the Pythagorean theorem is interpreted as $h^2 = c^2 - d^2$, or $\sin^2 \angle B + \cos^2 \angle B = 1$. Then, since the desired result is known, these formulas are processed according to the rules of an ordered field without any concern of their geometric meaning.[46] Although, a, b, etc., stand for real numbers, in the context of Euclidean geometry, we do not exploit the full power of the so-called completeness axiom – it just guarantees the existence of square roots.

In sum, modern proofs of the Heron's formula can be also derived in Descartes's field of line segments. In fact, they mimic the way Descartes processed formulas in *La Géométrie*, despite trigonometric identities, which in that context represent Euclid's propositions in a disguise.

9 Euclidean Versus Descartean Plane

Euclid developed geometry in the so-called synthetic fashion. Descartes adopted completely new method and his system is governed by technical issues rather than axiomatic meditations. In the nineteenth century, the name *analytic geometry* was

[46] Variety of proofs can be easily found in the Internet, see for example http://cut-the-knot.org/

coined for that way of studying algebraic curves. In the above sections, we compared basic techniques of Euclid's and Descartes' geometry, that is the theory of proportions and the arithmetic of line segments. In this section, we adopt a more abstract approach and aim to correlate these techniques with a characteristic of Euclidean and Descartean plane.[47] To this end, we reconsider Euclid's and Descartes' geometry on the so-called Cartesian plane $\mathbb{F} \times \mathbb{F}$ over an ordered field $\mathcal{F} = (\mathbb{F}, +, \cdot, 0, 1, <)$. Within that framework, the characteristic of a plane could be reduced to a property of the field $\mathcal{F}$.

Thus, a plane $\mathbb{F} \times \mathbb{F}$ is Euclidean *iff* the field $\mathcal{F}$ is closed under the square root operation. A plane characterized by this condition satisfies both circle-circle and circle-line intersection property.[48] Thus, that sole algebraic condition guarantees that straightedge and compass constructions are realizable on the plane $\mathbb{F} \times \mathbb{F}$. That is how we identify Euclid's geometry as developed in books I-IV. Regarding that part of the *Elements,* the field $\mathcal{F}$ could be Archimedean or non-Archimedean. Since Euclid's theory of similar figures requires proportions founded on axioms E1-E5, the study of that part of the *Elements* requires the field $\mathcal{F}$ to be Archimedean.[49]

Descartes' field of line segments is explicitly closed under the square root operation. But he added yet another condition, namely:

"Now to treat all the curves which I mean to introduce here, only one additional assumption is necessary, namely, two or more lines can be moved, one upon the other, determining by their intersections other curves" (Descartes 1637a, 316).

What does he mean in this sentence? Going back to Fig. 9, when we put a parabola in place of the line CNK, the resulting curve will be the so-called cubic parabola.[50] When in place of that parabola we put the cubic parabola, we get another line of a higher degree. Descartes believed that the class of curves generated by iterations of that process enables one to construct roots of a polynomial of any degree. Hence, we can identify his *additional assumption* as the curve-line intersection property, where curve means an algebraic curve, i.e., a curve defined by a polynomial equation $\psi(x, y) = 0$. We can simplify this geometrical criteria to the condition of the existence of roots of polynomials of odd degree, since it provides a neat property of the field $\mathcal{F}$.

Interestingly, when, in book III of *Geometry,* Descartes develops the first ever theory of polynomials, he observes that n-th degree polynomial $\varphi(x)$ has n roots:

"Every equation can have as many distinct roots (values of the unknown quantity) as the number of dimensions of the unknown quantity in the equation" (Descartes 1637a, 372). A few pages away, he is even more precise, as he writes.

"Neither the true nor the false roots are always real (*reelles*); sometimes they are imaginary (*imaginery*), that is, while we can always conceive of as many roots for

[47] Since the term *Cartesian plane* has the well-established meaning in mathematics, in this section, we choose a bit awkward term *Descartean plane* to emphasize the historical context.

[48] See (Hartshorne 2000), 144.

[49] There are of course Archimedean fields closed under the square root operation other then (i.e., *smaller*) the field of real numbers, as well as non-Archimedean fields closed under the square root operation.

[50] See (Descartes 1637a), 337.

each equation as I have already assigned, yet there is not always a definite quantity corresponding to each root so conceived of. Thus, while we may conceive of the equation $x^3 - 6xx + 13x - 10 = 0$ as having three roots, yet there is only one real root, 2, while the other two, however we may increase, diminish, or multiply them in accordance with the rules just laid down, remain always imaginary" (Descartes 1637a, 380).

We can factorize the discussed polynomial as follows:

$$x^3 - 6x^2 + 13x - 10 = (x-2)\left(x^2 - 4x + 5\right) = (x-2)(x-2-i)(x-2-i).$$

Next to the above remarks, there are other really insightful comments concerning polynomials.[51]

On this basis, we say that the plane $\mathbb{F} \times \mathbb{F}$ is Descartean *iff* the field $\mathcal{F}$ is real closed, i.e., (1) is closed under the square root operation, and (2) every polynomial of odd degree has a root.

A real closed field could be, again, Archimedean or non-Archimedean.[52]

Thus, the assumption that not only circles but other curves intersect straight lines, results in different characteristic of Euclidean and Descartean plane.[53] All in all, Descartes' field of line segments $(\mathbb{F}, +, \cdot, 0, 1, <)$ is a real closed field.

Viewed from the foundational perspective, the Archimedean axiom is a necessary condition of Descartes' arithmetic, as explained in Sects. 2.1 and 4. Yet, neither Descartes nor his followers, specifically Newton and Euler, paid attention to that issue. They just processed formulas in an ordered field.[54]

10 Summary

Our final conclusion that in *La Geométrie* line segments form a real closed field is based both on Descartes explicit statements and implicit rules revealed by studying the way he transforms proportions, algebraic terms, and polynomials, that is, on his mathematical practice. Thus, Descartes explicitly introduces operations on line segments: multiplication, division, and the square root. Yet, his transformation of Euclidean proportion into the equality of line segments, as encoded in the rule (ED), is implicit. Definitions (2) emphasize the dependence on his algebra of Euclidean proportions. From that stage on, the algebra of line segments develops by its own

[51] To name the most spectacular: (1) the number of positive roots is at most the number of sign changes in the sequence of polynomial's coefficients, (2) roots of a polynomial are bounded by a number determined by polynomial's coefficients.

[52] There are of course Archimedean real closed fields other than (i.e., *smaller*) the field of real numbers.

[53] Cf. (Panza 2011)

[54] Geometry over a real closed field is complete and decidable, while over a field closed under the square root is undecidable (see Ziegler 1976). One may interpret these meta-mathematical results as reaffirming optimism concerning new method voiced by Descartes at the end of book I. Yet, it looks a bit different when we consider his actual methods. Our skepticism, in the first place, concerns feasibility of his algorithm for finding roots of polynomials.

logic. While transforming algebraic terms, Descartes implicitly applies laws of an ordered field exemplified in the formulas (1). In book III, he explicitly introduces zero and negative line segments. His theory of polynomials, specifically factorization of polynomials, as well as rules (3) and (4), guarantees mathematical sense for these concepts. Finally, his explicit declaration concerning intersections of algebraic curves allowed us to characterize the structure of line segments as a real closed field.

Acknowledgments I am grateful to John T. Baldwin, Kazimierz Mrówka, Marco Panza, and Benjamin Wilck for helpful comments.

References

Acerbi F, Vitrac B (2014) Héron d'Alexandrie. Metrica, Fabrio Serra, Pisa
Artin E, Schreier O (1926) Algebraische Konstruction reeller Körper. Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität 5:85–99
Beckmann F (1967) Neue Geschichte zum 5. Buch Euklids. Archive for the History of Exact Sciences IV:1–144
Błaszczyk P, Mrówka K, Petiurenko A (2020) Decoding Book II of the *Elements.* Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia 12:39–88.
Błaszczyk P (2018) From Euclid's elements to the methodology of mathematics. Two ways of viewing mathematical theory. Annales Universitatis Paedagogicae Cracoviensis Studia ad Didacticam Mathematicae Pertinentia 10:5–15
Błaszczyk P, Fila M (2020) Cantor on infinitesimals. Historical and modern perspective. Bulletin of the Section of Logic 49(2):149–179
Błaszczyk P, Mrówka K (2013) Euklides, Elementy. Ksiçgi V-VI. Tlumaczenie i komentarz. Copernicus Center Press, Kraków
Błaszczyk P, Petiurenko A (2019) Euclid's theory of proportion revised. Annales Universitatis Paedagogicae Cracoviensis Studia ad Didacticam Mathematicae Pertinentia 11:37–61
Błaszczyk P, Petiurenko A (2020) Modes of diagrammatic reasoning in Euclid's Elements. In: Pietarinen A, Chapman P, de Smet LB, Giardino V, Corter J, Linker S (eds) Diagrammatic Representation and Inference – 11th International Conference, Diagrams 2020, Proceedings. Lecture Notes in Artificial Intelligence (LNAI) 12169. Springer, pp 50–65
Bos HJM (2001) Redefining geometrical exactness. Springer, New York
Clavius C (1589) Euclidis Elementorum. Libri XV, Roma
Cajori F (2007) A history of mathematical notations, vol 1. Cosimo Classics, New York
Crippa D (2017) Descartes on the unification of arithmetic. Algebra and Geometry Via the Theory of Proportions Revista Portuguesa De Filosofia 73(3/4):1239–1258
Christianidis J, Oaks J (2021) The Arithmetica of Diophantus. A Complete Translation and Commentary. Routledge, Abingdon, Oxfordshire
Descartes R (1637a) La Géométrie. In [17], 290–413
Descartes R (1637b) Discours de la Méthode pour bien conduire sa raison, et chercher la vérié dans les sciences, plus la Dioptrique, les Météores et la Géométrie qui sont des essais de cette Méthode. Jan Maire, Lejda
Euclid: Euclid's Elements of Geometry translated by R. Fitzpatrick (2007.); http://farside.ph.utexas.edu/Books/Euclid/Elements.pdf
Giusti E (1993) Euclides Reformatus. Bollati Boringhieri, Torino
Hartshorne R (2000) Geometry: Euclid and beyond. Springer, New York
Heiberg J (1883–1888) Euclidis Elementa. Teubner, Lipsiae

Hilbert D (2015) Grundlagen der Géométrie. Festschrift Zur Feier Der Enthüllung Des Gauss-Weber-Denkmals in Göttingen. Teubner, Leipzig (1899), 1–92. In: Volkert K (ed) David Hilbert, *Grundalgen der Géométrie* (Festschrift 1899). Springer, Berlin
Hilbert D (1900) Uber den Zahlbegriff. Jahresbericht der Deutschen Mathematiker-Vereinigung 8: 180–184
Malet A (2006) Renaissance notions of number and magnitude. Hist Math 33:63–81
Mancosu P (1996) Descartes' Géométrie. In: Mancosu P (ed) Philosophy of mathematics and mathematical practice in the seventeenth century. Oxford University Press, Oxford, pp 66–91
Mueller I (2006) Philosophy of mathematics and deductive structure in Euclid's elements. Dover, New York
Newton I (1720) Universal arithmetic. Translated from the Latin by the late Raphson. Printed for J. Senex [etc.], London
Oaks JA (2018) François Viéte's revolution in algebra. Archive for History of Exact Science 72: 245–302
Palmieri P (2001) The obscurity of the equimultiples. Clavius' and Galileo's foundational studies of Euclid's theory of proportions. Archive for the History of Exact Science 55:555–597
Pappus: Pappi Alexandrini mathematicae collections a Federico Commandino Urbinate in latinum conversae at commentariis illustratae. Pesaro (1588)
Pappus (1877) In: Hultsch F (ed) Pappi Alexandrini Coollectionis. Apud Waidmannos, Berolini
Schöone H (1903) Herons von Alexandria, Vermessungslehre und Dioptra. Opera quae supersunt omnia, Tuebner, Leipzig
Schubring G (2005) Conflicts between generalization, rigor, and intuition. Springer, New York
Smith ED, Lathan M (2007) Geometry of rene descartes. Cosimo Classics, New York
Taisbak C (1980) An Archimedean prooof of Heron's formula for the area of a triangle. Reconstructed Centaurus 24:110–116
Taisbak C (2003) ΔΕΔΟΜΕΝΑ. Euclid's data. Museum Tusculum Press, Copenhagen
Thomas I (1957) Selections illustrating the history of Greek mathematics, vol II. Harvard University Press, Cambridge, MA
Wallis J (1685) Treatise of algebra. John Playford, London
Ziegler M (1976) Einige unentscheidbare körpertheorien. Enseignement Mathmétique 28:269–280

Leonhard Euler Between Mathematics and Natural Philosophy: An Introduction to Natural Science Anleitung zur Naturlehre

Danilo Capecchi

Contents

1 Introduction 270
2 Philosophy of Nature 272
2.1 Anleitung zur Naturlehre 273
2.2 Lettres à une princesse d'Allemagne 277
2.3 Matter and Its Properties 279
2.4 The Origin of Forces 285
2.5 Concepts of Time and Space 289
3 Mechanics and Mathematics 291
3.1 Theoria Motus Corporum Solidorum Seu Rigidorum: The Motion of Mass Points 293
3.2 Measurement of Forces 303
4 The Apparent Motion and the Observer 304
5 Dynamics of Rigid Bodies 307
6 Quotations 314
7 Conclusions 318
8 Cross-References 318
References 319

Abstract

Until recently Euler was seen essentially as a mathematician than a physicist (modern meaning). After all, more than 60% of his work deals with pure mathematics, and even those whose object is mechanics and astronomy contain many sections that can be classified as mathematics. Today Euler's role as a physicist, especially as a scholar of mechanics, is re-evaluated. One of the most prolific scientific writers, Euler treated all the themes of mathematics and physics, from astronomy to optics, from electricity to magnetism, and from hydraulics to mechanics to music, leaving an indelible mark in all sectors. He also wrote many

D. Capecchi (✉)
Sapienza University of Rome, Rome, Italy
e-mail: danilo.capecchi@uniroma1.it

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_19

pages of philosophy of nature, still of interest today because written with the sober language of the mathematician. As a matter of fact, it is impossible to distinguish the various facets of the polymath Euler. The present chapter focuses on mechanics, with particular attention to the epistemological, metaphysical, and natural philosophy aspects. The texts of Euler largely collected in his *Opera omnia*, still in progress, are analyzed avoiding as much as possible the use of glasses colored by modern concepts of epistemology. What is considered important has been to render his ideas clear to a modern reader. A particular attention has been devoted to the paper *Anleitung zur Naturlehre* of 1750, published only posthumously in 1862.

Keywords

History of classical mechanics · Equations of motion · Observer · Space and time · Origin of forces

1 Introduction

Leonhard Euler (1707–1783), born in Basel, the first child of Paul Euler and Margaretha (née Brucker), spent his beginning youth not in Basel but in the nearby Swiss countryside Riehen. (Over time the Euler family name has had a variety of spellings, with Öwler and Äweler in the fourteenth and fifteenth centuries and Öwbler, Ewbler, and Ouwler in the sixteenth.) His parents were his first teachers. Margaretha is thought to have instructed him in beginning reading. His father gave an elementary education in mathematics. Like his teacher Jakob Bernoulli, Paul Euler taught his young son mathematics not as an isolated discipline but as underlying all natural knowledge, interrelated with other fields.

Euler, who required instructional preparation completed, was sent by his parents to Basel, perhaps as early as the age of eight. By the second decade of the 1700s, Basel was no longer in its golden period; nonetheless the presence of Jakob and Johann Bernoulli made it a center for scientific and mathematical research.

In Euler's day schoolboys were sometimes beaten or knocked about until they bled. At other times the teacher might find himself dragged by the hair by an enraged father making an unannounced visit to the class. There is no reason to suppose that Euler's intelligence saved him from routine indignities and brutalities.

In 1715 Paul Euler hired as a private tutor Johannes Burckhardt (1691–1743), a young theologian with a tolerable background in mathematics. At the time Burckhardt was supporting Johann Bernoulli in arguments with Brook Taylor and other members of the Royal society of London over which was superior, Leibniz's differential calculus or Newton's.

After completing his gymnasium education in 1720, Leonhard registered at the University of Basel for courses in the philosophical faculty, which covered fields of learning outside recognized professions. It was the equivalent of the modern secondary school. The philosophical faculty imparted a general education before a student chose a specialty for a higher degree. Through hard work and an astonishing

memory, Euler mastered all of his subjects. During his first 2 years, he was enrolled in Johann Bernoulli's class for beginners in geometry as well as practical and theoretical arithmetic.

Euler spent 2 years earning his *prima laurea*, roughly equivalent to a bachelor's degree, receiving it in 1722 at the age of 15 after presenting the *ad lectiones publicas*, his undergraduate thesis, *De temperantia*, *extolling moderation*. In conversations and intellectual debates, Euler began to display to a wider public his substantial learning and his command of Latin. While he advanced rapidly in lower level mathematics courses, Euler did not however quickly gain the attention of his teacher Johann Bernoulli.

In the autumn of 1723, Euler passed the examination of the philosophical faculty for the Master of Arts degree and officially received it on 8 June 1724 at the age of 17. At the graduation session on that day, Euler gave a public lecture in Latin comparing the natural philosophy of René Descartes with that of Isaac Newton and indicating the consequences of each. It was probably in consultation with Burckhardt and Johann Bernoulli that he had chosen this important and timely topic.

Euler's increasing attention toward mathematics and natural philosophy did not please his father, who obliged him to register in the theology faculty in 1723 in preparation for taking holy orders. He felt fortunate to continue Saturday meetings with the stern Johann Bernoulli. As Leibniz was dead and Newton old and less active, Bernoulli was since 1707 the leading mathematical preceptor in all of Europe. While likely concerned about employment for his genial son, Paul Euler accepted a Johann Bernoulli's request for Euler's shift out of theology.

Euler's friend and classmate Johann II, the youngest Bernoulli son, had tried gain him access to a possible tutorial with the father. But Johann Bernoulli was busy, and flatly refused to give Euler private lessons. Instead he advised him to start reading some more difficult mathematical books and work through them as diligently as he could, and if he came across some obstacle or difficulty, he could visit him every Saturday afternoon.

All suggests that Euler examined such classics as the second edition of Copernicus's *De revolutionibus orbium coelestium* from 1566 as well as Kepler's *Astronomia nova* of 1610 and Galileo's *Dialogo sopra i due massimi sistemi del mondo* published in 1632. Euler's master's lecture shows that he was studying Descartes's *Principia philosophiae* and *La géométrie*. He possibly also read Rohault's Cartesian masterful *Traité de physique*, a major physics text of the late seventeenth century, which appeared in 1671.

The interests of Euler's teacher make it probable that he read a range of works relating to the new calculus and its applications, beginning with the *Analyse des infiniment petits pour l'intelligence des lignes courbes* published in 1696 in Paris, attributed to de L'Hôpital. Euler possibly examined two texts of Varignon, a correspondent, disciple, and close friend of Johann Bernoulli: the *Projet d'une nouvelle méchanique*, published in 1687, and *Nouvelles conjectures sur la pesanteur* of 1690. Varignon's two-volume *Nouvelle mécanique ou statique* and *Éclaircissemens sur l'analyse des infiniment petits*, both from 1725, may also have been available. Euler must have studied Jakob Bernoulli's articles on the theory of infinite series,

published from 1682 to 1704 and reprinted in 1713; his *Ars conjectandi* on probability, with a preface by Nikolaus I, published posthumously in 1713; and Jakob Hermann's *Phoronomia, sive de viribus et motibus corporum solidorum et fluidorum* of 1716. Other books to which Euler referred to are John Wallis's *Arithmetica infinitorum* (1656) and Brook Taylor's *Methodus incrementorum directa et inversa* (1715). No confirming evidence exists on whether Euler saw Johann Bernoulli's pioneering articles for the Académie des sciences of Paris in 1718 on what would become the calculus of variations. (Most of the bibliographic information is drawn from Calinger (2016).)

Euler was called to the Academia scientiarum imperialis petropolitanae of St. Petersburg by his friend Daniel Bernoulli in 1727. He remained there until 1741 to reach the Académie royales des sciences et belles lettres de Berlin. Here he had some problems with Frederick II that finally pushed him back to St. Petersburg in 1766 where he remained until his death in 1783, at the age of 78 and almost completely blind. One of the most prolific writers, among the greatest mathematicians of all time, Euler treated all the themes of physics, with an approach, however, more as a "mathematician" than an experimental physicist, from astronomy to optics, from electricity to magnetism, and from hydraulics to mechanics to music, leaving an indelible mark in all sectors. He also wrote many pages of philosophy of nature, still of interest today because written with the sober language of the mathematician. His writings, more than 20 books and pamphlets and about 800 papers, are collected in his *Opera omnia* (Euler 1911–2018). It remains to complete the correspondence, written in several languages. Most are in Latin and French and some in German and Russian (Euler also had competence in Italian, Spanish, Chinese, and Japanese). Despite his strong religiousness, Euler was an illuminist philosopher-scientist, and his faith in reason transpires in all his writings.

2 Philosophy of Nature

Since entered the University of Basel, Euler, as discussed just above, had been deeply interested in the studies of theology and natural philosophy, and only 17 he gave a lecture for his master's degree comparing the natural philosophies of Descartes and Newton.

Certainly he was not a professional philosopher, but he reflected at length on the classical themes of philosophy of nature. He was not interested in discussions of a metaphysical character, and even for that, he was generally hostile to the approach to the natural philosophy and mechanics of Leibniz's school, represented in his time by his colleague Christian Wolf (1679–1754). Studying Euler's philosophy is challenging not so much to understand his works on mechanics or physics (modern sense) but rather to see how natural philosophy was being transformed into the hands of mathematicians. It is indeed his mathematics that influenced his philosophy and not vice versa; in particular his research on Calculus influenced his conception of matter and space.

Euler's ideas about philosophy of nature are scattered everywhere. The natural reference is however to the *Anleitung zur naturlehre* (Euler 1862) of the 1750s, referred in the following as the *Anleitung*, and *Lettres à une princesse d'Allemagne* (Euler 1770–1774) of 1760s, referred in the following as *Lettres*. The *Anleitung* did not spread greatly, even because of his late publication. The *Letters* had instead a different fate. Michael Faraday (1791–1867), for instance, read the letters (Calinger 2016, p. 565, note 35) but not the *Anleitung* which was probably read by Bernhard Riemann (1826–1866) (Speiser 2010, p. 46).

Euler's philosophy of nature was generally well received, praised by Voltaire, for instance. According to Alexandre Koyré, the *Lettres* may be included among prominent Newtonian popularizations (Koyré 1965, p. 18). But also by Leibniz, even though here and there, Euler criticized his philosophy of nature. Euler had no particular reason to stand on one side or the other in the philosophical debates of the period though he had a profound estimate of Newton referring to him as the *Summus Neutonus*; not being a professional philosopher, he could limit himself to accepting those ideas that seemed closer to his sensitivity as a mathematician without looking deeply into their consistency.

It should be cited however a strong criticism by two scholars well connected with Euler, d'Alembert, and Lagrange. They judged him severely. Lagrange wrote to d'Alembert, "Our friend is a great analyst but quite poor a philosopher" (Lagrange 1867–1892, Vol. 13, p. 135). (Letter of Lagrange to d'Alembert, 16th June 1769.) D'Alembert replied indignant: "It is incredible that such a great genius as him on geometry and analysis is in metaphysics so inferior to the smallest schoolboy, not to say so flat and so absurd, and it is the case to say: *Non omnia eidem Dii dedere*" (Lagrange 1867–1892, Vol. 13, p. 148). (Letter of d'Alembert to Lagrange, 7th August 1769.)

2.1 Anleitung zur Naturlehre

The *Anleitung* has been received incomplete. The end of Chap. 5 and the beginning of Chap. 6 are missing (which probably occupied about seven pages). There are reasons to think that, even with these pages, the work could not be not complete and we can therefore ask why Euler did not finished it (Speiser 2010). In its current form, the book consists of 21 chapters that can be grouped: Chaps. 1–5, the role of natural philosophy and the main properties of matter; Chaps. 7–9, general introduction to the principles of mechanics; Chaps. 10–11, apparent motion, general rules of motions; Chaps. 12–18, the property of matter; Chaps. 15–18, possible forms of matter; Chap. 19, gravity; and Chaps. 20–21, principles of hydrostatics and hydrodynamics.

The opening of the *Anleitung* suggests a traditional treatise on natural philosophy or natural science, even of Aristotelian mold. According to Euler, natural science is a science that aims to explain the causes of changes that occur on material bodies. Wherever there is such a science, he continued, it is very incomplete, since it is able to state with certainty the causes of only very few changes. Moreover, changes are restricted to those occurring on inorganic or material bodies, thus distinguishing

natural science from the science of the mind, which aims to explain mental changes. Nothing is said about living beings (Euler 1862, Vol. 2, p. 449). All changes involving material bodies must arise from the essence and from the properties of the bodies themselves. What is common to all material bodies without exception is called a property of the bodies, and therefore all things not sharing this property are excluded from the domain of material bodies. The general properties of material bodies are those shared without exception. The essence of material bodies is a property that is not only shared by all material bodies but which is such that all things having this property must of necessity be considered to be material bodies. The best way for natural science to proceed is first to investigate the general properties, that is to say the essence of material bodies, and subsequently to subject all particular types of bodies to a similar examination (Euler 1862, Vol. 2, pp. 450–453).

Following the reading of the *Anleitung*, it is understood however that the causes and the properties of which Euler spook are those useful for a mathematical treatment of the philosophy of the nature, adopting a mechanistic approach in the broad sense; they are the extension (that is, the position and or a configuration), the mobility (velocity and acceleration), the persistence or inertia (mass), and the impenetrability (force, because forces are due to impenetrability). Euler wanted to found a coherent system, able to describe all the inorganic nature, starting only from some principles of mechanics and a few other hypotheses. Even Newton, another "mathematician" who wondered about the relationship between the philosophy of nature and mathematics, in some of his *Queries* (28–31) had sketched some ideas in this direction, but they remained only queries, that is, questions.

The first general property of material bodies is the extension, and anything that has no extension cannot be regarded as a material body. We are not only convinced by our experience that all material bodies that we know possess extension, but our concept incorporates extension in such a way that we can exclude all things without extension from the category of material bodies. It follows then that whatever can be said of extension per se can without exception also be said of material bodies. Everything with extension is divisible, and divisibility can be continued ad infinitum; therefore all material bodies are infinitely divisible (Euler 1862, Vol. 2, pp. 453–455). However Euler did not enter the question of the actual corpuscular (atoms) or continuum constitution of matter. What he could say is that one knows from experience that the actual subdivision of many material bodies can be carried out to an astonishing degree and that our tools and senses are too blunt to permit this subdivision to be carried even further. However he was not discussing what can actually be done but rather the merely the possibility of taking the subdivision even further.

For liquid matter Euler however denied the existence of atoms.

> A liquid matter cannot be formed from a number of small particles that are solid and hard, for whatever the shape and arrangement of the particles might be, for it is not possible that a pressure that acts at one location will propagate in all directions with equal force. (Euler 1862, Vol. 2, p. 526; Translation by Hirsch E.) (A1)

In the following a verbatim exposition of Euler argument is referred to. If one imagines these particles in the first place to have the shape of a cube and imagines

them to be arranged in an orderly manner on one another, then one sees easily that if the uppermost is pressed down by a force, the lowest will press on the ground with equal force, but sideways no force will be exerted; therefore if many such rows fill a vessel and a force presses on them from above, then the bottom of the vessel will experience an equal force, but the sides will experience none at all. If the particles were lying together in a disordered manner, then the pressure could also be propagated sideways, but it would not be the same in all directions. But what can be said of cubic particles is equally valid for all other shapes that involve corners, for which reason most natural scientists ascribe to the particles a spherical shape; but it is easy to show that from these, if they are assumed to be solid and hard, the main property of liquids cannot be obtained. It is only necessary to imagine a heap of spheres, arranged as is usual for cannon balls, to see that if they are pressed from above, they exert no force sideways, or that at least such a force would not be the same in all directions. In addition spheres that fill a vessel cannot lie together in such a regular manner that a great dissimilarity in their arrangement would not develop, that also would prevent a uniform pressure. If one assumes the spheres to be in uniform motion, then the pressure can change but cannot be maintained the same in all directions; and one would have to regard as extremely rare a case where the pressure was the same, remembering that this constitutes the essence of a liquid. One need only to consider that if three spheres lie in a straight line, the middle one can always be pushed sideways, whatever forces may be acting on the outer ones. On the various types of bodies (Euler 1862, Vol. 2, p. 526).

An important concept that Euler considered worthy to be exposed in his foundational text is what today is called potential energy or possibly work, to which he referred to as the *effectiveness* and to which he also assigned the name *effort des forces sollicitantes* (Euler 1757b, p. 287). (It must be said that Daniel Bernoulli in his work of 1750, *Remarques sur le principe de la conservation des forces vives pris dans un sens général*, introduced something like the effectiveness (Bernoulli 1750, p. 359).) The effectiveness of a force is the integral quantity that is found if one multiplies the force with the differential of the distance to which it pushes the body and then integrates. This concept is the more of the utmost importance, because (1) the sum of the effectiveness of all forces $\int Pdx + \int Qdy + \int Rdz$ always has the same magnitude – given by the increase of living force – if one had assumed three different planes, that is, in modern term, it is invariant with the change of coordinates. (2) The whole theory of equilibrium is based on it. For it can be shown that equilibrium cannot occur, if the sum of the effectivenesses is not a minimum or occasionally a maximum: "This marvelous theorem was first derived by the world renowned President de Maupertuis and is closely connected with the other general principle of frugality. From this we see at the very least that the effectiveness has a major influence on all motions that can be produced by forces, and that *it deserves to be given a special name* [emphasis added]" (Euler 1862, Vol. 2, p. 493; Translation by Hirsch E.). It also must be remarked however that Euler seems not to be conscious that he had to limit the nature of forces in his definition; only for what are now called conservative forces – for which the integrals appearing in the definition of effectiveness are independent of the path – Euler considerations has any meaning.

One more concept Euler maintained worth to be underlined is that of pressure into a fluid with its modern meaning, in English *stress* that is internal force per unit of surface. Until then pressure was simply a force of contact, the greater, the greater the surface. Euler introduced this new concept, maintaining ambiguously the same name *pressure* both for the global force and for the force per unit of surface. Because the pressure (traditional meaning), said Euler, depends on the magnitude of the area on which it acts, it is best if one represents it by a magnitude, the pressure (new meaning), which, when multiplied by an area, expresses the magnitude of the force that acts on the area, and the constancy of the pressure rests on the fact that this magnitude is constant. If we express by $aa \times k$ the force p, which acts on the base area $S = aa$, then the pressure on any area ff is:

$$\frac{ff}{aa} k \times aa = k \times ff$$

Therefore when k remains the same, then the pressure on the liquid matter is also the same (Euler 1862, Vol. 2, pp. 523–524). To picture this completely, Euler regarded k as a height and expressed the force acting on the area ff by the volume of a cylinder of base area ff and height k, since the volume of such a cylinder is $ff \times k$. This picture is very convenient, according to Euler, because forces are best represented by weights; for this one selects a uniform matter and says that the force, by which the area ff is pressed, is just the weight of such matter that fills the volume of the cylinder of volume $ff \times k$. Therefore such a height k gives us a clear understanding of the force that presses on the inner wall of the vessel filled with liquid matter, and that at the same time presses on all the parts within. For the larger or smaller this height is, the larger or smaller, in the same proportion, will be the force of the pressure (Euler 1862, Vol. 2, p. 524). A similar statement can also been found in the *Principes generaux the l'état d'équilibre des fluides* (Euler 1757a, pp. 5–7). Here he introduced the symbol p, still used today though limited to a pressure of external loads, instead of k to indicate stress. Euler also used his concept of stress for solids bodies (Capecchi 2001, p. 27).

It may seem strange that in a treatise dealing with general aspects, Euler had also found ample space to exquisitely technical aspects concerning the laws of hydrostatics and hydrodynamics (Chaps. 20–21). It must be kept in mind however that for Euler fluids were not only the ordinary ones, for example, water and air, but also the aether (see next sections) was a fluid and as such subjected to the laws of hydraulics. David Speiser, with a bit of whiggism, suggests that the advantage of Euler over Newton, in order to provide a unified theory of the various branches of physics, is the concept of the partial derivative equations, the mathematical basis of any field theory, a new tool, created by Euler himself and d'Alembert – and others – to formulate the mechanics of fluids. And it is thanks to this new mathematical tool that Euler was the first to imagine a unified field theory, an idea, and a hope that are at the center of the ambitions of physicists (Speiser 2010, p. 45).

In fact, according to Euler, all physical phenomena, with the possible exclusion of gravity (see below), can be reduced to the interactions of four scalar fields (the

densities and pressures of gross matter and aether) and two vector fields (the velocities of matter and aether), interconnected and governed by partial derivative equations. Seen as an attempt, Euler's way of proceeding with hydrodynamics is not very different from that used in modern physics to unify field theories, despite the present more in-depth knowledge of the structure of matter and the huge amount of accumulated empirical data. Even today, the basic differential equations of a field of physics, whose theory is known, are also considered as fundamental for other sectors, in which a theory is still lacking. In the case of Euler, the basic differential equations are those of hydrodynamics, and the radical hypothesis is that of an unique aether (Speiser 2010, p. 46). Despite the considerations of Speiser, which seem convincing, in the *Anititung* there is however no mention of the aethereal fluid, neither for what concerns electricity nor for what concerns magnetism. These topics that will be treated with plenty in the *Letters*.

2.2 Lettres à une princesse d'Allemagne

At the time of the publication of the *Letters*, Euler stood at the peak of his career. He was the director of the mathematics class (division) of the St. Petersburg Academy and a well-known mathematician. Although he sent the individual letters between 1760 and 1762, he had derived them from articles dating back to the 1720s. Thus, notwithstanding the didactic character, they offer an outline of his scientific ideas and the modifications occurring in them.

The letters, 234 in number (a thousand pages), originated from lessons delivered to the princess Charlotte Ludovica Luisa, a relative of Frederik II of Prussia, but were conceived more generally for young students. They met with prodigious success; by the turn of the nineteenth century, they had been translated from the original French into eight other languages: Russian, German, Dutch, Swedish, Italian, English, Spanish, and Danish. The numbering is different from edition to edition, here that of 1770–1774 edition is considered (Euler 1770–1774).

Euler's letters essentially concern philosophy of nature considered from different points of view, that of the traditional philosopher, with considerations of *physica generalis*, and that of the philosophers emerging at the time. It is not easy to group them by themes; it can be said that the first letters (about 140) have a more general character; the others are more technical and deal with current topics in physical research, among them at least 17 letters are devoted to electricity and nineteenth to magnetism.

Among the first letters, there are five dedicated to music (4–8). We are in a period in which the musical theory was undergoing somehow a revolution with Jean-Philippe Rameau (1683–1764) among the protagonists. Euler, only 24, had written an interesting paper on music, the *Tentamen novae theoriae musicae* of 1730–1731 (Euler 1739b), which was criticized by Rameau. The two "musicians" exchanged some letters about consonances in the 1750s (Calinger 2016, pp. 362–363).

Letters (17–44) follow that deal with optics. The first four of the theory of propagation of light with particular reference to Newton; only a nod is made at

Descartes; the others face problems of geometrical optics, of brightness and color of bodies, and also of the structure of the human eye. Euler compared the "emanation" (projectile) theory attributed to Newton with his vibrational theory, reported in particular in the *Nova theoria lucis* of 1744 (Euler 1746b). He did so without mentioning Huygens who had proposed a theory in which light spread in a medium as pulses. (About the difference between the conceptions of Euler and Newton and for a justification of the absence of the name of Huygens in Euler's writings of optics, see Hakfoort (1995).) According to Euler, to explain the nature of light and color, it must be admitted celestial spaces filled with a subtle matter, he called aether, a fluid matter like air but much thinner, nevertheless, with much greater elasticity (using a modern term, stiffer than air). And as in the air, the sound is propagated, so the light propagates in the aether (Euler 1770–1774, Vol. 1, Letter 19).

In the letters from 45 to 79, there are general considerations on mechanics. They speak of gravity in a general sense, of the law of universal gravitation, to conclude with the explanation of gravity that is associated with the elasticity of the aether, without however going into details. It is here that one sees that a mature Euler is not fully satisfied of his mechanical explanation of gravity. Indeed, in the letters 46 and 54, for instance, he wrote that "philosophers have warmly disputed, whether there actually exists a power which acts in an invisible manner upon bodies; or whether it be an internal quality inherent in the very nature of the bodies, and, like a natural instinct, constraining them to descend" (Euler 1770–1774, Vol. 1, Letter 46, p. 195), "On this question philosophers are divided. Same are of opinion that this phenomenon is analogous to impulsion; others maintain, with Newton, and the English in general, that is it consists in attraction" (Euler 1770–1774, Vol. 1, Letter 54, p. 230). Thus declaring his incapacity of a definitive choice.

Letters from 80 to 87 concern aspects of natural philosophy, bodies, and spirits. In the letter 76 Euler criticized Wolf's conceptions of dynamics. He did this by referring substantially to what he had already written in his *Gedancken von den Elementen der Cörper* of 1746 (Euler 1746b) (see below).

Letters 88 to 114 address issues of ethic, psychology, and logic. Particularly in the letters 103–105, the diagrams of Euler-Venn are used to explain some logical relations. It is true that the idea of representing graphically logical relationships was not completely new; this had happened in some treatises of the eighteenth century. But it was to Euler that the mathematician John Venn (1834–1923) referred to in his studies on logic more or less a century later. Some letters about epistemology follow, from which the influence of Descartes and Locke and perhaps even of Condillac transpire – without explicit reference to them – and letters of metaphysics in which Euler repeated the criticism toward Leibniz's monad system. From letter 133 onward, physical (modern sense) problems are addressed. In particular letters from 138 to 154 deal with electricity and letters from 169 to 186 with magnetism.

Apart from a summary of a paper by Aepinus (Euler 1761), the only left written works on the theory of electricity are on the *Letters*. This notwithstanding he considered electricity a very interesting field to be explored:

> The subject which I am now going to recommend to your attention almost terrifies me. The variety it presents is immense, and the enumeration of facts serves rather to confound than to

> inform. The subject I mean is electricity, which for some time past has become an object of such importance in physics that everyone is supposed to be acquainted with its effects. (Euler 1770–1774, Letter 138, Vol. 3, p. 277; Translation in Speiser (2008)) (A2)

Euler was primarily interested in the electric field, not in the charge (he was among those who rejected the idea that there were two types of electricity). For him electricity is due to the presence of one electric fluid, the aether which in normal circumstances fills the pores which travers all bodies in all directions. The electric qualities of a body depend on how easy or how difficult the pores are to open. In conductors (modern term) – metals and such – the aether can easily flow in and out; in this case, the pores are always open, so to speak. With other materials a certain strength is required to open the pores through friction: these are the insulators (modern term).

To magnetism instead Euler dedicated some papers and letters, see, for instance, Radelet de Grave and Speiser (2004) and Euler (1748a, b). Magnetism also is associated to a fluid. Originally – in the *Anleitung*, for instance – in Euler's opinion, the same aether was responsible for gravity, magnetism, electricity, and light. In the *Lettres*, in particular in Letter 176, instead he was of the opinion that the aether associated to magnetic forces was a special kind of subtle matter and not simply the subtlest part of a single all-pervading aether (Euler 1770–1774, Letter 176, Vol. 3, pp. 118–119). He seems never to have had any doubt that the light aether is responsible for the electric phenomena as well.

2.3 Matter and Its Properties

Euler dealt at length with bodies, matter, and related properties here and there in his scientific works. However, there are some papers in which he focused specifically on the subject. Among them: *Dissertatio de igne, in qua ejus natura et proprietates explicantur* written in 1737, *Gedancken von den Elementen der Körper* of 1746, *Recherches physiques sur la nature des moindres parties de la matiere* of 1746, *Réflexions sur l'espace et le tems* of 1748, *Recherches sur l'origines des force* of 1750, *Lettres à une princesse d'Allemagne sur divers sujets de physique & de philosophie* of 1760–1761, and *Anleitung zur Naturlehre* of 1750s. (Euler published anonymously *Gedancken von den Elementen der Körper*, an anti-Wolffian work. It was immediately recognized as bearing his signature however.)

According to Euler, bodies are characterized by the following general properties: extension, mobility, inertia, and impenetrability. Among them the fundamental is impenetrability which as such should be considered the very essential property of the bodies.

> Whatever is impenetrable belongs to the category of bodies, and therefore the essence of bodies is their impenetrability, on which therefore all the other properties must be founded (Euler 1862, Vol. 2, p. 472). (A3)

A body has extension in common with space and mobility with moving images projected on the wall, both of which however do not possess the property of impenetrability.

According to Euler, from impenetrability, extension and mobility follow. His argumentation on these points has more a rhetorical than demonstrative value but is equally interesting. From the very convincing thesis (a) no-extension → no-impenetrability, from the rule of classical logic, it follows (b) impenetrability → extension. More difficult is to argue (c) impenetrability → mobility; indeed a thing can be at rest though impenetrable; thus impenetrability is not necessary for actual motion; what can be said is that impenetrability is not in contrast with the possibility of motion, that is, mobility, but this makes weak the implication (c).

Inertia, called persistence (Standhaftigkeit) in the *Anleitung*, is also associated by Euler with impenetrability being linked to mobility "because when a thing is mobile, it should also have persistence, since otherwise any change would occur without sufficient reason" (Euler 1862, Vol. 2, p. 472), and mobility implies impenetrability. But the derivation (d) impenetrability → persistence has more problems than derivations discussed previously. Euler's strategy is to prove the implication (e) mobility → persistence that with the implication (c) would give (d). The derivation of the implication (e) implies the principle of inertia that Euler, like d'Alembert, based on the principle of sufficient reason, which has been the object of criticism by many scholars. Because implication (d) would derive from two weak implications (c) and (e), it would consequently be two times weak. For a different view on this point, see Gaukroger (1982, pp. 141–142).

According to (Gaukroger 1982), the idea that impenetrability constitutes the essence of body depends upon two Euler's claims: it is unique and necessary to a body, and it is irreducible, in the sense that conceiving a body as impenetrable is a primitive clear notion. Both Boscovich in the *Theoria philosophiae naturalis* of 1758 and Kant in the *Metaphysische Anfangsgründe der Naturwissenschaft* of 1786 thought impenetrability of bodies an obscure idea requiring elucidation in terms of a more basic notion that of repulsive force. For them, repulsive force was the physically primitive notion not requiring further elucidation, just as attractive force was physically primitive (Gaukroger 1982, p. 140). Considering that Euler was deriving force from impenetrability (see below), his is exactly the opposite path from that of Boscovich and Kant.

Bodies are portions of matter; thus Euler devoted a lot of attention to the properties of matter. The *Recherches physiques sur la nature des moindres parties de la matiere*, a relatively early work, began with the following question: "It is an important question in physics and metaphysics to establish whether or not the smallest part of matter are similar to each other" (Euler 1746c, p. 289). Euler tried to give a physical answer to this metaphysical question coming to the conclusion that all the smallest parts of bodies, referred to as the *smallest molecules* of matter, are made of the same stuff. In order to speak about the smallest molecules, it is necessary to distinguish between the properly said matter (groben Materie), *gross matter*, characterizing the inertia of bodies to which only the term molecules is referred to, from another type of matter, which is usually not recognized as such, called *subtle matter*, interposed among molecules. Notice that notwithstanding Euler used the term molecules, he said nothing about a corpuscular conception of matter, that is, the molecules may be not corpuscles.

Following these premises, repeated some years later in the *Anleitung*, Euler argued that there are only two kinds of matter, a subtle matter and a gross matter. The different apparent nature and density of bodies depend on a different combination of subtle matter and gross matter, with the subtle matter that fills the pores of the gross matter that may be more or less numerous and great. The molecules of gross matter have no pores, and even if they cannot be divided in act, being extended, they can be divided with the imagination. The division of matter remained a theme dear to Euler, who maintained, if not an ambiguity, at least a lack of explanation. For liquids, with fairly simple reasoning, as already shown in the previous section, he had "succeeded" in proving that there are no indivisible finite molecules; for solids Euler escaped the problem of the actual existence of molecules; what he said concerned the extension which is always indefinitely divisible.

The thesis that there are only two kinds of matter is carried out by Euler with both physical and metaphysical arguments. Assuming a metaphysical point of view, he invoked a principle of economics, which however is not so rigid because at least two kinds are allowed. While for the gross matter Euler arrived at a more or less convincing argument on the fact that it is only of one kind, for the subtle one, he hypothesized that there may be more than one, perhaps only two. For what the uniqueness of gross matter is concerned, after making considerations based on daily experience that lead to believe that there is a single type of (gross) matter and that the differences found among the various bodies depend on the nature and quantity of the pores in it present, Euler passed to the demonstration of this affirmation, on the basis of two laws of empirical nature. The first law concerns the fact that the force of gravity varies with the inverse of the square of the distance, as Newton has shown (Euler's words), and thus the weight of bodies decreases with the square of their distance from the center of the earth. The second empirical law concerns the fact that bodies of different weight fall with the same temporal law (in vacuum).

From the law of the inverse of square, Euler proved that the density and thus the pressure of the aether toward the center of the earth which causes gravity (see below) must decrease with inverse proportion of the distance. Thus, if x is the distance from the center of the earth, the pressure of the aether should vary as $h - A/x$, being A a constant of proportionality and h the pressure of the aether when at rest, as it is for $x = \infty$. From this expression of the pressure, it is easy for Euler to prove that the weight of a body is proportional to the volume of the gross matter it contains. Then, from the independence of the temporal law of fall for bodies of different weight, and from the law of mechanics (force and mass are proportional), it easy to prove that weight and mass are proportional. Which implies that the mass is proportional to volume, and because mass for Euler is the same as quantity of gross matter, it means that the quantity of gross matter of body is proportional to its volume, and thus the density of gross matter is constant, and thus the gross matter is of one kind only (Euler 1862, pp. 542–545).

A first draft of the hypothesis of the existence of a subtle matter or aether can be found in the *Dissertatio de igne, in qua ejus natura et proprietates explicantur* that Euler wrote in 1737 for the award of 1738 proposed by the Académie des sciences of Paris, which was assigned to him (Euler 1752a). To explain the phenomenon of

combustion, Euler claimed that the flame arises from the sudden explosion of molecules of "materiae subtilis igneae compressae" and to illustrate his conjecture used a corpuscular model: "Let's imagine a great quantity of glass spheres, full of air strongly compressed, and this mass of spheres is the material we want. Suppose that only a sufficient force intervenes to break a single sphere; and it is clear that both the impetus of the air and the projection of the fragments of glass will produce a similar effect in the nearby spheres, and then from these in all the others, until all are broken, emitting with immense roar the air that they enclosed" (Euler 1752c, p. 10).

Euler went on to argue that the properties of fire, due to a specific igneous matter, are not reducible to the ordinary laws of mechanics, like the principle of inertia or the proportionality of action and reaction. The heat given off by the fire and the rapid spread of the flame in all directions violate these principles and must be explained in another way. It is necessary to resort to a further postulate: the existence of an aether, a substance distinct from the igneous particles. The flame does not instantly disperse in the air due to the resistance of this aether, much more elastic and rarefied of the igneous matter, which surrounds and holds the flame in an unstable equilibrium. The continuous shocks that are created at the limit between the aether and flame give place to vibrations that generate light and transmit it in a straight line, like it occurs for the waves in the air (Euler 1752c, p. 19). The postulate of the ether was certainly not new in the eighteenth century. Many variants of it were formulated that referred above all to the conjectures that Newton had advanced, albeit in a temporary problematic form. For a discussion of the various hypotheses discussed in the period 1740–1750 by physicists, see Cantor and Hodge (1981).

The properties of subtle matter had been discussed more in depth in the *Recherches physiques sur la nature des moindres parties de la matiere* and in the *Gedancken von den Elementen der Körper*, to be taken again in the *Anleitung*. In the second part of the former of the two papers, Euler first denied the possibility of the existence of vacuum with metaphysical considerations, scarcely convincing indeed while he did not mention a physical more convincing (for us) motivation, namely, that the presence of vacuum would not justify the oscillatory theory of light proposed in his *Nova theoria lucis*, where propagation is conceived, in analogy to that of sound, as elastic waves propagating in the medium. With regard to the nature of the subtle matter, he limited to assert that it is very different from the gross matter; it is not gross matter made of very tiny corpuscles separated by vacuum but rather a continuous substance or, with a technical modern word, a continuum.

> There will be thus at least two main kinds of matter; one [the gross matter] which gives the fabric of sensible bodies, whose the particles have all an unchangeable degree of density, which is even greater than the apparent density of gold; the other [the subtle matter] kind of matter will be that of which the subtle fluid which cause the gravity is composed, and we call the aether. (The density does not vary in time, that is, gross matter is incompressible (Euler 1862, Vol. 2, p. 510).) (Euler 1746c, p. 300) (A4)

In the *Anleitung*, Euler spook at length about his continuous *aether* and did not feel embarrassed to introduce a concept which was alien to the dominating mechanistic

philosophy where the aether had usually a corpuscular nature. Euler's aether fills the whole space, leaving no emptiness. Air consists essentially of aether with small corpuscles of gross matter. While the gross matter is incompressible, the aether is compressible:

> Gross matter is therefore not capable of any change other than in the appearance of its shape, which, if appropriate forces are available, can be changed in arbitrary ways.
>
> [...] It does not appear to be true that subtle matter also has always and everywhere the same density, such that it could through no force be driven into a smaller space. Instead an important difference between gross and subtle matter seems to be that the latter can be compressed. (Euler 1862, Vol. 2, pp. 511–512; Translation adapted from Hirsch E.) (A5)

The aether although has the property to expand or shrink and behaves like an elastic medium and has a certain density that is proper to it. However, in no case it is possible to reduce a portion of aether to a point, that is, to effectively annihilate it (Euler 1862, Vol. 2, p. 513). Euler clearly stated that the compressibility of subtle matter does not imply the violation of its property of impenetrability.

Euler did not propose a convincing, at least for a modern, explanation of the force/pressure caused by the elasticity of aether which is not deducible from any of the typical properties of matter, extension, inertia, impenetrability, and mobility and therefore should be treated as a primitive (essential?) concept. Forces on gross matter are explained in Euler's mechanics by means of impenetrability and inertia. For instance, suppose an elastic body in motion that impacts with a rigid surface, it receives a force and rebounds because of the impenetrability of the surface and the resistance to change its motion from the inertia of the ball. The pressure of the aether against a rigid surface cannot be explained in the same way because the inertia of the aether is assumed to be negligible: "because experience shows that the celestial bodies do not suffer effects with their impact with the aether that fills the universe" (Euler 1862, Vol. 2, p. 542). Certainly there is an analogy with gases; but in this case, one (Euler himself) can resort to a mechanistic explanation; this was what Daniel Bernoulli did in his *Hydrodynamica* of 1738: the pressure of a gas is determined by the impact of the particles on the walls of the vessel that contains it (Bernoulli 1738, pp. 200–203). But for Euler the aether is not formed by particles.

Euler also doubted that there is on kind only of subtle matter: "Whether there are several kinds of this subtle matter, some of which are denser than others, we shall not be discussing here, but if indeed there are several kinds, we shall refer to them collectively as subtle matter. As long as the explanation of what occurs in nature does not require several such kinds, it would be bold and against the rules of a sound science of nature if, merely following our imagination we were to increase the number of kinds of subtle matter" (Euler 1862, p. 509).

According to him, the aether that is found in the universe is compressed far beyond its natural state and consequently exerts a great elastic force on the bodies formed by the gross matter (Euler 1862, Vol. 2, p. 517), able to explain the resistance of the materials to rupture, their elasticity, and the force of gravity.

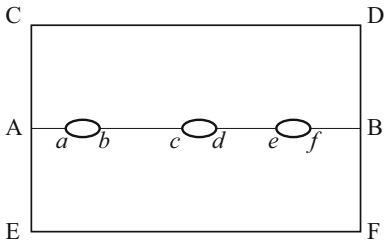


Fig. 1 Strength of bodies (Euler 1862, Vol. 2, Table XX, Fig. 235)

The resistance of solid bodies is associated with the hydrostatic pressure that the subtle compressed matter exerts on the various particles of ordinary gross matter, an explanation similar to that suggested by Galileo with the use of horror vacui. When two smooth grains of gross matter come into contact with each other, the elastic force of the aether makes them adhere strongly, and the breaking strength is maximum. If, on the other hand, the contact between the grains is not complete, then the breaking strength will be the minor the minor the contact surface.

To show this more clearly, let the two bodies of Fig. 1, ABCD and ABEF, be joined at the surface AB such that between them the cavities *ab*, *cd*, and *ef* are filled with the aether. The body ABCD will then be forced against the other body by the aether that presses on the planes CD and EF; on the other hand, it will be pushed back by the aether in the cavities *ab*, *cd*, and *ef*. Therefore the force with which the body ABCD is pressed against the other body ABEF is the result of the two constraining forces.

The lake of resistance in fluid is explained by the fact that here gross matter is so suffused by the aether that the particles have nowhere immediate contact. To see how this could happen, imagine that every particle of gross matter is surrounded by subtle matter, and the particles consequently never approach each other so closely that there would not remain some subtle matter between them. If every particle were at rest, such a mixing would be hard to understand; but if the subtle matter is in motion so that it continuously flows between the gross particles, then in this way immediate contact can be impeded (Euler 1862, Vol. 2, p. 535).

The elasticity is explained by admitting that inside bodies there are cavities (pores) filled with subtle matter that do not communicate with the outside. When a body is deformed, either compressed into a smaller space or expanded into a larger one, then there will be change in the size of its pores, some being expanded, but others being compressed.

> To explain this more clearly let us assume in the first place that a body is compressed into a smaller space. Since gross matter itself is incompressible, this cannot happen other than by making the pores smaller. In that case the apparent density of the body must increase, because the whole matter of which the body consists, or at least the gross matter, since the subtle matter can be neglected here, must have been brought into a smaller space. Let a^3 be that part of the body that is filled with gross matter, and e^3 the remaining part that only contains subtle matter, the sum of all pores taken together, then $a^3 + e^3$ is the size of the body, a is its mass and $a^3/(a^3 + e^3)$ is its density. But a^3 cannot be changed, therefore when the body is brought into a smaller space only e^3 is diminished. [. . .] If after the change in shape

> of a body the closed pores are neither larger nor smaller, then the body retains this changed shape. But if the closed pores become wider or narrower, then a force will develop in the body tending to restore the previous shape. (Euler 1862, Vol. 2, pp. 537–538; Translation by Hirsch E.) (A6)

The elasticity of commonly experienced bodies is thus explained by means of the elasticity of the aether, which is clearly a vicious circle. Thus what Euler actually explained was not the elasticity itself but rather why bodies made by an incompressible matter exhibit elasticity and why the elasticity differs from body to body. The difference in elasticity lies in the diversity of the dimensions and arrangement of the pores.

The account above referred about elasticity explains the behavior of the various elastic bodies, for example, the loss of elasticity as a result of heating. Indeed, the subtle matter contained in the pores is set into motion by heat, opening up access to previously closed pores; at the end the matter has less and larger pores than before. Therefore even though before heating there were large forces, these can vanish after heating.

The weight of the bodies, that is, gravity, is explained by admitting that the density, and therefore also the pressure of the aether surrounding the earth, increases by increasing the distance from its center. Euler is very precise in his explanation. Gravity should be associated to the gradient of pressure in the aether around the earth that determines the movement of the ethereal fluid downward with the consequent downward drag of the gross matter, determining the heaviness of the bodies (Euler 1862, Vol. 2, p. 542).

The section devoted to the heaviness of bodies ends with the consideration that even if one does not know why the pressure decreases in proximity to the planets, to stay with this doubt is always better than not knowing anything as when one says that the heaviness is due to attraction.

> Although we have to stop here and hardly can hope ever to find the cause of the diminution of the elastic force of the aether, it is easier to resign to this than to merely maintain that all bodies are by their nature endowed with a force to attract each other. For since one can not even form an understandable concept of this attraction, one can by way of contrast at least understand how it is possible that the elastic force of a liquid matter is reduced, and one also understands that this can occur in a way that is in accordance with the laws of Nature. (Euler 1862, Vol. 2, p. 547; Translation by Hirsch E.) (A7)

2.4 The Origin of Forces

Force in the eighteenth century was the fundamental magnitude of mechanics and was introduced in various ways, each with its own problems. In his youthful treatise of 1736, the *Mechanica sive de motu analytice exposita*, Euler introduced the concept of force in a classical and rather generic way. The force [potentia] is "an action on a free body that either leads to the motion of the body at rest, or changes the motion of that body" (Euler 1736, p. 78). If in this treatise Euler payed no care to

explain the nature of force, he did in subsequent works. It is worth noting that Euler had two Latin terms to indicate force: *potentia* and *vis*. To potentia he gave the technical meaning of cause of the variation of the motion of a body; to vis a generic meaning, either technical or common, for example, he can speak of *vis gravitatis* as an example of technical use, as well as of vis inertiae and the vis of impenetrability, as an example of common use.

Many years later, when he wrote the *Theoria motus corporum solidorum seu rigidorum*, with a much clearer understanding of mechanics, Euler proposed the definition:

> Definition 12. What induces to change the absolute state of bodies is named a *force* [vis]; it should be due to external causes, since the body will remain in its own state due to internal causes (Euler 1765c, p. 44). (A8)

Is there any change in the status of force (in the ontology); from a being has it became a name and instead of a potentia a vis? The question mark is justified also because in the *Anleitung*, preceding of some years the *Theoria*, Euler in a succinct proof of the laws of motion had written that force is measured by the effect it produces: Here he wrote: "A force twice as big must in the same time produce twice as big an effect, because just in view of that do we consider it twice as big" (Euler 1862, Vol. 2, p. 476). A similar position is also reiterated a few pages after Definition 12 in the *Theoria*, as will be commented in a subsequent section. In any case the above definition hides some ambiguities, always present in the scholars of the eighteenth century when they deal with force.

In the following in the attempt to decipher the complex argumentations on the concept of force, I will assume that *force* is for Euler a primitive concept, which in particular is reduced to the anthropomorphic idea of pressure. What Euler tried to explain is not what is force or pressure but what is its origin. The fact then that the Definition 12 defines force as the name given to the external cause of the change of motion does not regard ontology but simply serves to delimit the broader concept that force has in the natural philosophy of the period, somehow a hypostatization of any cause, also internal.

In the *Recherches sur l'origine des forces* of 1750, he had paused on the role of impenetrability only to explain forces. Euler observed that despite the inertia or the general property of bodies by virtue of which each body tends to preserve its state, we see that bodies that fall under our observation continually change their state. So these bodies are subject to forces that must necessarily be external (Euler 1752d, p. 423).

To illustrate the role of impenetrability to generate forces Euler considered two bodies A and B. A is at rest, while B is moving toward A. After the impact, in order to avoid its own penetration, A acts to change the state of B, and in turn B acts on A. In other words, A applies a force to B (due the impenetrability of A, which is an external cause) and B to A. Euler concluded that these forces see their source in the impenetrability only (Euler 1752d, p. 427). That is, the forces due to impenetrability are passive forces (modern term); they are not determined neither for their quantity

nor for their direction; they only are "obliged" to act when the impenetrability of a body is threatened; otherwise they do not produce any effect.

It is true that a greater action would prevent penetration as well, but it is more natural to assume a minimal value; that is forces have at all times value and direction strictly necessary to prevent penetration, so that if a minimum force is sufficient for the purpose only it acts (Euler 1752d, pp. 430–431). In this criterion of "least action," Euler took inspiration from Maupertuis, and in it, one can find a justification of the principle of least action. To Maupertuis he made explicit reference in the letter 78 of the *Lettres* (Euler 1775, Vol. 1, p. 304).

Euler argued that the forces that two bodies exchange are equal and opposite to each other. For Euler this equality of forces, which is commonly known as the principle of action and reaction, is a necessary consequence of the penetration. So, for him, it is a rational and not an empirical law.

> This equality of forces, hence the great principle of equality between action and reaction, is a necessary consequence of the nature of penetration. For, if it would be possible for the A body to penetrate the body B, body A would be equally as much penetrated by the body B; therefore, since the damage that these bodies that penetrate, is equal, these two bodies must also employ equal forces to resist penetration. So, as much as the body B is solicited by the body A, that will be solicited equally by this, both deploying exactly as much force as necessary to prevent penetration. But these two bodies acting one on the other by any force, will be in the same state as if they were compressed together by the same force (Euler 1752d, p. 434). (A9)

Euler was also able to analyze the evolution of the force due to impenetrability in the case of impact. Differently from many scholars of the period and in agreement with Johann Bernoulli (and Leibniz), he assumed that impact is not an instantaneous process: "The instantaneous impact would not be in accordance with the always respected law of nature, for which nothing can occur instantaneously, and as for a leap. According to this law a so great change, as sometimes is that occurring in the impact of bodies cannot occur without corresponding to a some interval of time" (Euler 1746a, p. 36).

In the hypothesis that the impacting bodies are linear elastic Euler furnished an explicit expression of the force of impact during the process of compression following the impact. If P is such a force, it is given by the expression $P = F/kz$, where F is the force necessary to get a penetration k, while z is the current value of the penetration. For reasons of space and also because it is not relevant for the present text, I do not repeat the mathematical steps involved in the proof. I limit myself to say that the approach and the result are essentially the same Euler had obtained in his youthful work *De communicatione motus in collisione corporum* of 1731 (Euler 1738). The difference is that now Euler had a clear concept of mass, whereas in 1731 he had not. But the equations are the same, demonstrating how mathematics may be more powerful than physics.

In subsequent works Euler associated force not only with impenetrability but with inertia also. In fact if there were no inertia, the state of bodies could be

changed, and in principle it could be argued that as impenetrability causes a change in state, it also determines forces. But it would only be a meaningless word game. Imagine, for example, stopping a moving ball by opposing it a hand; if inertia would have no role in the explication of force, one could be free to assume the ball without inertia. In this case the ball would be rejected or stopped, but the hand would not notice anything; which is contrary to what happen in the real world. To obtain an effect, it is necessary that the ball is equipped with the ability to resist changes of motion, that is, that it has not zero mass (inertia). In the *Theoria* Euler limited himself to say that forces originate from impact of bodies because of their impenetrability. He wondered if all the forces have this origin. Without excluding the existence of other forces, he was content to say that a very important class of forces has this origin.

Here's how Euler describes the role of inertia: "The cause of those forces by which the state of a body is changed may be agreed to lie not in inertia alone but in inertia coupled with impenetrability. Indeed, seeing that only bodies can be said to be impenetrable, and since bodies are necessarily endowed with inertia, impenetrability as such involves inertia, so that impenetrability alone is rightly considered the source of all forces by which the state of bodies is changed. It will therefore be proper to consider this property more exactly as being the origin of all forces" (Euler 1765c, p. 46). He thus can still pretend that only impenetrability counts because inertia is given when impenetrability is given.

The account of the force of impact between the two bodies A and B considered above may be completed as follows: when B comes into contact with A, it experiences a force which we would normally term A's force of resistance to change of state. Note that the force is not in any sense in A: what is in A is its inertia, which is not a force (according to Euler's definition) because it only maintains A's state. But this internal principle is experienced by B as a force. There is, therefore, an external force acting on B which is not internal to A. Nor does it act at a distance because it is a prior condition of there actually being a force that A and B be impenetrable and that they be in contact (Gaukroger 1982, p. 147).

Euler underlined in several occasions that the forces of impact are external to the bodies and that inertia is not a force in the sense that it is not responsible for the change in the motion of a body; it has only the role of preventing it. In his work of 1746 *Gedancken von den Elementen der Körper* (Euler 1746b), Euler criticized Leibniz's, or rather Wolf's, philosophy of forces and monads. Euler's main criticism concerned Wolf's thesis that the monads, which are the ultimate components of reality, are endowed with an internal force that intends to continually change their state. Euler believed that there was indeed an internal strength, which however tended to maintain the state and not to change it and concluded: "One must therefore stipulate that two particular well differentiated classes of things exist in the world, to one of which belong the corporeal things, whose essence consists in the force [ability] steadily to maintain their state. The other however comprises the souls and ghosts, which are endowed with a force to change their state" (Euler 1746b, p. 15).

2.5 Concepts of Time and Space

Euler discussed the concepts of space in his more technical works such as the *Mechanica sive de motu analytice exposita* and *Theoria motus corporum solidorum seu rigidorum* and in some more philosophical ones such as the *Lettres* and *Anleitung*. But the text in which it did so in a systematic way is *Réflexions sur l'espace et le tems* of 1748 (Euler 1750). Euler's discussion is interesting because it shows how a mathematical-oriented mind addressed some issues of philosophy of nature and metaphysics; he had perhaps less philosophical culture than a professor of philosophy, but he knew much more directly the topic he is dealing with. Euler's paper did not escape to Immanuel Kant (1724–1804) who read it 15 years after its publication. Kant observed that the mathematical considerations of motion and space furnished many data to guide the metaphysical speculation in the track of the truth, by thus considering mathematics and physics as a tool to avoid the void speculations of the philosophers of his time (Cassirer 1953, p. 351).

Euler defended his thesis of absolute space against the theses of "metaphysicians" who believe that only the concept of relative space makes sense. The metaphysicians in question are naturally the followers of Leibniz and Wolf, not explicitly named however. That he referred to them is evident when he stated "I strongly doubt [...] that the equality of the spaces should be judged by the number of monads that fill it" (Euler 1750, p. 331).

He addressed his thesis of absolute space with metaphysical and physical reasons. The metaphysical reason consists in affirming that a fertile concept in physics cannot be empty: "One should instead assert that both the absolute space and the absolute time, such as mathematicians look at, are real things which submit outside our imagination, because it would be absurd to sustain that pure imaginary objects could be assumed as foundation of the real principles of mechanics" (Euler 1750, p. 326). The physical reason is that only by conceiving absolute space can one explain the inertia of bodies.

Euler discussed inertia by referring it to a body at rest. He affirmed that to explain the difficulty to put it in motion from rest, its place (position) must be thought of as absolute – although in fact it is not possible to individuate any absolute system he added. Euler argued with a simple example that inertia cannot depend on the presence of nearby bodies to which refer motion considering a body that floats on still water, remaining at rest in its turn. If the water begins to flow, an observer, rigidly linked with it, sees the body move without any force being applied to it. Or if he sees the body remaining at rest with respect to the water, a careful examination shows him that this is due to the effect of water dragging. So Euler can conclude: "I strongly doubt that metaphysicians dare to sustain that bodies maintain their position with respect to other bodies thanks to their inertia, because it could be easy to show the falsehood of this explanation because of the consideration I discussed on bodies close to their neighbor" (Euler 1750, p. 328).

Euler then dealt with the case in which the position of a body is referred to fixed stars. Being them at rest – a possibility that is not certain for Euler (Euler 1765c, p. 38))

– the absolute space of the mathematicians would coincide with that of the metaphysicians. But, he wondered, how the fixed stars so far can determine inertia? This is impossible according to Euler.

This thesis, rejected by Euler, is endorsed by some modern scientists, who also refer to the reflections on the subject by the epistemologist Ernst Mach.

> As Ernst Mach has pointed out, it cannot be a coincidence that the fixed stars appear indeed fixed relative to inertial frames, and hence that it is reasonable to consider inertia as a force exerted on local objects by the totality of the objects in the entire universe. Thus, Newton's law may best be interpreted as a consequence of the basic axiom that the sum of the forces, including the inertial force, acting on a particle should be zero and of the constitutive law of inertia, which states that this inertial force should be given by $-ma$, where a is the acceleration relative to an inertial frame (Noll 2004, p. 9).

In essence Euler concluded that the preservation of the state (motion or rest) of bodies is explained only by conceiving place according to the criterion of mathematicians in an absolute space and not in relation to other bodies. And nobody can say that the principle of inertia is based only on something that exists in our head.

The reality of space is confirmed by the principle of preservation of uniform motion in a certain direction. If space was a relative concept, what sense would it have to speak of direction? It must be a direction of an absolute space that can be conceived in a natural way by abstraction.

Euler ended his considerations on space by reiterating the metaphysical argument that one cannot say that a principle accepted as true in physics by nearly all scholars (the principle of inertia), is founded on a thing that exists in our imagination only (the absolute space), and from this it must be concluded that the mathematical idea of absolute space cannot be in any way imaginary, but that there is something of real in the world which corresponds to this idea (Euler 1750, p. 329).

According to Ernst Cassirer, the *Réflexions sur l'espace et le tems* set up in fact not only a program for the construction of mechanics but a general program for the epistemology of the natural sciences. It sought to define the concept of truth of mathematical physics and contrasted it with the concept of truth of the metaphysicians. The considerations of Euler rested entirely on the foundations on which Newton had erected the classical system of mechanics. His concepts of absolute space (and absolute time, see below) were revealed not only as the necessary fundamental concepts of mathematical-physical knowledge of nature but as true physical realities. To deny these realities on philosophical grounds means to deprive the fundamental laws of dynamics – above all the law of inertia – of any real physical significance. In such an alternative, the outcome cannot be questioned: the philosopher must withdraw his suspicions concerning the "possibility" of an absolute space and an absolute time as soon as the reality of both can be shown to be an immediate consequence of the validity of the fundamental laws of motion (Cassirer 1953, p. 351).

Once established the reality of space, Euler went on to examine time. Beginning with the observation that the ideas of space and time always go together (Euler 1750, p. 331). Arguing on the reality of time is however more complicated. Euler

distinguished between the idea of time and time itself. He had no objection to the idea that time is conceived as a succession of changes. But this idea is not sufficient to assert that two time intervals are the same, as one has to say in mechanics. Thus, according to Euler, time cannot be reduced to the idea of succession of changes; there must be something else: absolute time. Euler concluded by saying that he realized that his considerations could only be taken on by philosophers who are willing to give some sense of reality to time and motion, while they will not make the slightest impression to those who consider everything as relative (Euler 1750, p. 333).

3 Mechanics and Mathematics

Euler was lucky enough to be born in a period when there was to build all the mechanics and all the Calculus starting from the foundations of the seventeenth century. His precursors and heirs have been sought. He has often been seen as the successor of Newton and Leibniz. It must be said that his role in mechanics has been overshadowed by the historians of mechanics such as Mach, Dugas, Montucla, and Dühring, very careful to the fundamentals, who see him more as a mathematician than a mechanician. After all, more than 60% of his work deals with pure mathematics, and even those whose object is mechanics and astronomy contain many sections that can be classified as mathematics. Today Euler's role, even for what concerns the fundamentals of mechanics, is re-evaluated. Some see him as the founder of modern classical mechanics as well as the one who threw the germs of the theory of relativity and quantum mechanics (Suisky 2009).

Regarding the legacy left by Euler, we must distinguish between the actual one, that is, the influence that he had on his contemporaries and the influence he could have had if all of his writings had been published and understood and also the influence it could have on modern mathematicians and physicists if they were still studied. On this last point, I limit myself to note that many of Euler's scientific writings could still well understood by a modern reader and that it is still possible to identify promising lines of research now abandoned.

In any case the genesis of Euler's thought, although interesting, is not a priority of this book. The reference to predecessors and contemporaries in mechanics is considered only on specific points, the main objective being to expose and explain his ideas. In this chapter I will focus almost exclusively on the fundamentals of Euler's mechanics, largely neglecting his elaborations in the areas of rigid body dynamics, theoretical astronomy, theory of elasticity, and fluid dynamics, even though they represent the greatest contribution of Euler to mechanics.

In my analysis, I will undergo a critical examination of the current points of view, due in large part to Truesdell's studies who edited (Vols. 12 and 13) or co-edited six volumes of the collected works of Euler and wrote appreciated articles and essays on him. He was Euler's greatest advertising agent, making him an eighteenth-century hero, a genius like Newton and perhaps superior to him. The current point of view can be summarized as follows:

1. Euler translated Newton's mechanics into Leibnizian language.
2. He introduced the use of vector calculus in mechanics.
3. He defined precisely the mechanics of the mass point, starting from the uncertain Newtonian mechanics.
4. He defined precisely the mechanics of the rigid body.
5. He made fundamental contributions to the systems of deformable bodies.
6. He founded modern hydrodynamics.
7. He introduced the concept of observer.

I will not stop to question these theses, limiting myself to say that it is one of the many myths of historiography that sees the evolution of science due to the work of isolated geniuses. It is hard to imagine that one man, not even one with exceptional memory and great workmanship and intelligence like Euler, could have done so much on his own. However, I believe that it is true that Euler was able to gather and summarize the ideas of his time, and for this reason, when one does not care about the genesis of the various ideas but wants to take stock of a certain era, the study of his research must be considered as essential.

The reconstruction of the basis of Euler's mechanics is carried out here by studying his published works – articles and books – and some letters. The main text which has been considered is *Theoria motus corporum solidorum seu rigidorum* (hereinafter *Theoria*), written in 1760 and published in 1765 (Euler 1765c). That is, I will start from a text in which the foundations of mechanics have already been laid. Of course, I will also refer to the *Mechanica sive de motu analytice exposita* (herein after *Mechanica*), published in 1736, and to articles and books that show both his researches and conceptions of natural philosophy, among which the most important is the *Anleitung*.

Euler had his own clear idea of how to develop a mechanical theory able to solve the problems that were then still waiting for a solution. Already in his early treatise, the *Mechanica*, he presented the work program that engaged him throughout his life: developing all the mechanics starting from the laws of motion of the mass point, a concept that he was among the first to specify. The mass point is the model with which the finite or infinitesimal bodies are studied when there is no interest in the extended objects. The mass of the whole body is thus imagined concentrated in one point, and the laws of kinematics and dynamics are applied to it. Here is Euler's program as reported in the general scholium of Chap. 1 of the *Mechanica*:

1. In the first place, very small bodies which can be considered as points are referred to.
2. Then the approach follows for these bodies of finite magnitudes which are rigid and are not allowed to change their shape.
3. In the third case, flexible bodies will be considered.
4. Fourthly, those which allow extension and contraction.
5. Fifthly the motion of bodies, constrained by others.
6. In the sixth case, the motion of fluids is in the agenda (Euler 1736, p. 37).

Never was Euler explicit about his metaphysical conceptions about mechanics. In particular, although in practice he treated mechanics as a purely rational discipline,

widely described with the language of Calculus, nowhere made his position explicit and never equated the status of mechanics with that of mathematics. Probably this silence was not due to an oversight but to some form of uncertainty. The only explicit statement he made about the role of mathematics in mechanics was his declaration of the preference given to the analytic treatment over the geometric one. For Euler only the former allows a systematic approach to all problems, while the latter involves the search for new routes for every problem. But this is of course true not only for mechanics but for mathematics in general, and it was for this reason that in the eighteenth century in every field of mathematics and physics, geometry was replaced by algebra and calculus, bringing the discipline of geometry to a state of decadence, from which it will recover only in the nineteenth century. Here is what Euler wrote about this point in the preface of the *Mechanica*:

> Thus, I always have the same trouble, when I might chance to glance through Newton's *Principia* or Hermann's *Phoronomia*, that comes about in using these [synthetic methods], that whenever the solutions of problems seem to be sufficiently well understood by me, that yet by making only a small change, I might not be able to solve the new problem using this method. Thus I have endeavored a long time now, to use the old synthetic method to elicit the same propositions that are more readily handled by my own analytical method, and so by working with this latter method I have gained a perceptible increase in my understanding. Then in like manner also, everything regarding the writings about this science that I have pursued, is scattered everywhere, whereas I have set out my own method in a plain and well-ordered manner, and with everything arranged in a suitable order. Being engaged in this business, not only have I fallen upon many questions not to be found in previous tracts, to which I have been happy to provide solutions, but also I have increased our knowledge of the science by providing it with many unusual methods, by which it must be admitted that *both mechanics and analysis are evidently augmented more than just a little*. (Euler 1736, Preface; Translation by Bruce I.) (A10)

The last sentence of the passage quoted above, in which Euler declared that the developments of mechanics and mathematics are closely linked, is worthy of note; mathematics allows to solve problems of mechanics; mechanics suggests cues to mathematics to treat unresolved problems.

3.1 Theoria Motus Corporum Solidorum Seu Rigidorum: The Motion of Mass Points

The first part of the *Theoria*, the *Introductio* (103 pages over a total of 527, preface excluded) reports a summary of the main results on the dynamics of the mass point, largely taking up from *Mechanica*, with many clarifications and updatings.

What a modern reader notices first, in reading the *Theoria*, though to a less extent to what happens for the *Mechanica*, is the verbosity of the exposition that in many cases instead of helping confuses and makes it difficult a reading because it does not allow to distinguish what is strictly necessary from what is superfluous. The sense of frustration is increased by the fact that in what is presented as a formal and generally axiomatic exposition, there is little or no distinction between theorems and principles.

Principles are almost never declared; the main exception concerns the principle of inertia.

Here is what Euler wrote about this point in the *Lettres*, a text following the *Theoria*, albeit slightly:

> This principle [of inertia] is commonly expressed in the two following proposition: A body once at rest will remain eternally at rest, unless it be put in motion by some external or foreign cause: Secondly, A body once in motion will preserve it eternally, in the same direction, and with the same velocity; or will proceed with an uniform motion, in a straight line, unless it is disturbed by some external, or foreign cause. In these two propositions consists the foundation of the whole science of motion, called mechanics. (Euler 1775, Vol. 1, p. 309; Translation in Euler (1802)) (A11)

While in the *Mechanica* the principle of inertia is not explicitly referred to as such, in the *Theoria* it is widely introduced with two propositions, qualified as axioms (see below). At no point is it referred to as a principle what a modern would treat as such, that is, the metaphysical principle of sufficient reason that plays a fundamental role in Euler's proofs.

The way of introducing and using principles becomes less mysterious if one reflects that the eighteenth-century mathematicians understood the term principle differently from us. They referred to a principle as a proposition placed at the foundation of a theory, requiring nothing else. Then the principle can be either first, that is, primitive, evident in itself or true empirically, or it can be a second principle if it can be demonstrated by first principles. Among the principles of mechanics, there are, for example, the principle of minimum action and that of the conservation of living forces. In Euler's time the logical status of these principles was not exactly defined; it was thought that they could be proved by first principles (and Euler will in part do so), but they still were not. Here is what Euler wrote about principles:

> XVlll. Although the principles in question are new, as they are not yet known or spread by the authors who have treated of Mechanics, it is understood, however, that the foundation of these principles cannot be new, but that is absolutely necessary that these principles should be deduced from first principles, or rather axioms, over which the doctrine of motion is established (Euler 1752a, p. 194). (A12)

The term *axiom* is used by Euler in the *Introductio* only three times to extend the validity of the properties from relative to absolute motion and to introduce the concept of inertia, with the classical meaning of a self-evident proposition. The term *theorem* is used eight times; there are then 15 definitions and 19 propositions qualified as problems. These propositions could generally be formulated as theorems because they provide the solution of the problems also. Euler's operation, perhaps similar to that of Euclid, serves to make the discussion less abstract.

The *Theoria* opens with the definitions of rest and motion:

> Definition I. Just as a body at rest remains perpetually in the same place, so a body in motion continues to change its position. Clearly a body that is observed to adhere always to the same place is said to be at rest: but that body which advances by gliding from place to place in time is said to be moving (Euler 1765c, p. 1). (A13)

The primitive concepts necessary to understand this definition, those of space and time, are introduced soon after, at a discursive level into two Explications and one Scholium. Space is introduced as an absolute entity "Now we can only conceive a notion about this space itself by abstraction, by considering all the bodies to be removed, and what is left we decide to call space" (Euler 1765c, p. 1), "a concept which was argued at length by philosophers (and uselessly)."

But the idea of absolute space, admitted Euler, cannot be used in practice neither to define the position of a body nor its motion. They must be referred to nearby bodies which must be able to maintain invariable in time positions relative to one another in order to fulfill the purpose. To determine the location of a point P, it is necessary to know its distance from at least four non-coplanar points A, B, C, and D (Euler 1765c, p. 2).

After introducing the reference system or using a modern term the observer, Euler went on to define the notions of relative motion and rest. A body is at rest with respect to A, B, C, D if its distances from them do not vary, otherwise it is in motion (Defs. 2, 3) (Euler 1765c, pp. 3–5). Note that Euler did not use the term relative, but "with respect to."

In a few definitions Euler introduced the concepts of path (Def. 4) and uniform and difform motions (Def. 5) and velocity (Def. 5), limited to uniform motion. Some comments are dedicated to the definition of velocity as the relationship between two heterogeneous quantities. The difficulty, inherent in the geometric concept of proportion and magnitude that requires comparison between uniform magnitudes, is solved by introducing units of measurement for space and time; in this way it is possible to define (the numerical value of) the velocity as the ratio between the numerical values that represent the measurements of space and time. Def. 7 concerns the direction of motion in the rectilinear and curvilinear cases, assumed as that straight line in which the motion is occurring or, if it is along a curve, as the tangent to the curve.

After these definitions Problem 1 follows, which deals with the use of the differential calculus to evaluate the space traveled by a point moving on a straight line with a velocity v assigned as function of time. Euler assumed that for a small dt (the element of time), the passed space ds, an elementary space indeed, is traversed with constant v velocity defined by $v = ds/dt$, in which v and s are explicitly thought of as generic functions of time, without particular restrictions on their regularity. (This is referred to as Proposition 3 in the *Mechanica*. "As indeed in geometry the elements of curved lines are considered to be the elements not-uniform motion is resolved into an infinite number of uniform motions. This is not the case, but it can be ignored without error. In either case the truth of the proposition is therefore apparent" (Euler 1736, p. 12; Translation by Bruce I.).)

The idea of considering constant any magnitude (in the present case the velocity) on condition of referring to a very small spatial or temporal intervals is what gives calculus a great heuristic power. If the velocity of the body v is given at individual instants, then the distances s passed in time t can be obtained with the help of the relation $ds = vdt$, the integral of which gives $s = \int vdt$. Similarly if the speed v corresponding to a distance s is known, then the time t, in which the distance s is passed, is given by the differential equation $dt = ds/v$, which gives $t = \int ds/v$.

More than one historian concedes that this approach to motion contains elements of novelty. Meanwhile, velocity is defined as a derivative (modern term, actually for Euler it was a ratio of differentials) of a function $s(t)$ which is considered as a generic mathematical function and therefore an abstract concept not necessarily connected to a geometric curve. Then time is treated in the same way as any other physical quantity, and it does not have a particular role as it happens, for example, in Newton, with his idea of fluxions (Suisky 2009, p. 132).

I do not pronounce on this judgment about Euler's originality. Certainly a similar treatment was not possible before the introduction of Calculus; furthermore it was established that Euler contributed a lot to the generalization of the concept of function. If one compares the way of treating velocity in the *Theoria* with that of the *Mechanica*, written 20 years earlier, he notices greater ease due to greater confidence both with the Calculus and the concept of function. It should be kept in mind that Euler's text also had a didactic function and the maturation of the concepts referred to was not only his but mainly of the contemporary readers who did not need to be reassured about the validity of the procedure.

Problem 2 considers the motion of a mass point on a plane curve. It is reduced to the study of two monodimensional motions projecting the velocity into two directions not necessarily orthogonal, even if, added Euler, it would be better if they were such, for reasons of simplicity of calculations. For the motion in space (Problem 3), the projection is on to three axes. It should be noted that this is only a kinematic problem where the problem of the legitimacy of decomposition does not arise.

Studying motion with respect to orthogonal coordinate axes instead of intrinsic or natural coordinates is considered a fundamental turning point in Euler's mechanics, a choice that appears first in his work *Recherches sur le mouvement des corps célestes en général* of 1747 (Euler 1749). The use of intrinsic coordinates, the standard approach at the turn of the eighteenth century, had two drawbacks. On the one hand, it required skill to find the right frame of reference. On the other hand, the approach was too difficult for bodies moving into a three-dimensional space. To be carried out consistently, it needed concepts of differential geometry such as curvature, torsion, and so on, fully developed only in the nineteenth century. The use of Cartesian coordinates does not present these problems. Although the equations may be complicated, they can be written in a standardized way. Even in this case, however, it must be said that Euler was not the first to project the equations of motion on two or three axes. According to Lagrange the first was Colin Maclaurin (1698–1746) in his *A treatise on fluxions* of 1742 (Lagrange 1811, p. 243).

> If the curve FM (Fig. 319) be described by powers directed in any manner whatsoever, and the force at any point M, resulting from the composition of these powers, act in the direction MK, and be measured by MK; let MK be resolved into the force MO in the direction of the ordinates MP ($=y$), and the force OK parallel to the base AP ($=x$); then, the time being supposed to flow uniformly, or the velocity at M being represented by the fluxion of the curve FM, the force MO will be measured by $\ddot{y}$ and the force OK by $\ddot{x}$ (Maclaurin 1801, Vol. 1, p. 298).

Johann Bernoulli also had done it according to (Maltese 2003, p. 210; Truesdell 1960, pp. 184–236, 1968, p. 112). But only with Euler, the use became systematic; indeed it became the only one.

The use of decomposing the motion into two or three directions can be seen as a step toward modern vector algebra. A same magnitude, a geometric vector, can give rise to different pairs (or triads if in space) of components as the coordinate system varies. The choice of the coordinate system can be arbitrary. To the decomposition of the motion in several directions, Euler gave the name *resolution*. The motion is said to be resolved, provided that the small interval traversed in the element of time is considered as the diagonal of a parallelogram or parallelepiped (Euler 1765c, Def. 8, p. 21). Euler also considered other types of coordinates and therefore of decomposition. In Problems 5 and 6, the polar coordinates are introduced in the plane and in the space, respectively.

In Chap. 1 of the *Introductio*, Euler carried out only kinematic considerations. In the following chapters, he went on to discuss the causes of motion. They are classified internal and external:

1. Internal causes. Responsible for the reason either for rest or motion of a body, with the exclusion of all external causes. They are able to contribute anything to change the state of motion or rest.

2. External causes. Responsible of the change of the state of motion and rest of a body.

To explain the nature and characteristics of internal causes, Euler devoted the entire Chap. 2. To achieve his purpose, he is forced to introduce the then controversial concept of absolute motion. In essence Euler said that, if there is the absolute space, of anybody one can say if it is at rest or in motion with respect to the absolute space. To these types of situation, he referred to as absolute motion or rest. This obvious fact can be expressed through an axiom:

> Axiom 1. Every body, even without being relative to other bodies, either remains at rest or moves, that is, it is either at absolute rest or in absolute motion (Euler 1765c, p. 30). (A14)

Two more axioms concern absolute motion; they express the principle of inertia, a term which Euler has not yet introduced.

> Axiom 2. A body, which is absolutely at rest, if subjected to no external actions, will persist indefinitely in the state of rest.
>
> Axiom 3. A body, which is absolutely in motion, if subjected to no external actions, continues to move uniformly along the same direction (Euler 1765c, pp. 32, 33). (A15)

These axioms, which Euler called the principles of internal motion, are justified in a simple way, basing on the principle of sufficient reason. For the state of rest, Euler argued, for instance, that because all the external causes of motion have been withdrawn, no reason is present that a body should begin to move in one direction rather than in any other: "This truth depends on the principle of sufficient reason" (Euler 1765c, p. 32).

A somewhat more sophisticated reasoning concerns the permanence in the state of uniform motion. No change in the direction can indeed occur, since there is no reason it should be deflected from that, in one rather than in all the other directions; "clearly it surely maintains the same direction, for the principle of sufficient reason." About speed [the modulus of velocity] it can be said that "unless it always remains the same, either it increases or decreases, of which neither absurdity can be said; for if it is either being increased or decreased, it must happen to follow a certain law; but what this law may be cannot be conceived in any way, since nothing surely will be agreed upon" (Euler 1765c, p. 34). The reader is asked to reflect on the validity of these demonstrations.

After the introduction of the three axioms, the first theorem (Theorem 1), which asserts that the axioms valid for absolute motion also apply to relative motion, provided that it refers to a body (to an observer) that is at (absolute) rest or in uniform rectilinear motion (Euler 1765c, p. 37), appears very simple in reality.

In Definition 11 the internal cause of motion is given a name, *inertia* (the persistence of the *Anleitung*):

> That quality of bodies, the reason for persisting in the same state present within a body itself, is called inertia, and also sometimes the force of inertia (Euler 1765c, p. 35, translation from Bruce I). (A16)

with the concept of state (absolute) introduced as follows: "While a body is either absolutely at rest or moving uniformly in a direction, it is said to persist in the same state" (Euler 1765c, p. 35).

Inertia indicates that property of bodies, whereby being at rest, means that they will continue to be at rest, therefore as if they oppose to motion; but since bodies set up in motion themselves equally oppose all to be changed, either on account of the speed or direction; this name seems a good choice. Sometime, said Euler, it is called *force of inertia*, because the body is resistant to change the state; but because often force is defined as the (external) cause which is changing the state of the body, force of inertia is not acceptable with this meaning – though Euler occasionally did. Whereby, as confusion should arise, it is better to omit the name force and refer to it by the simpler name of inertia (Euler 1765c, p. 36).

Chapter 3 is devoted to the external cause of motion; to it also is given the name *force*, with Definition 12. Though it has already been introduced before, I rewrite it for the sake of clarity:

> What induces to change the absolute state of bodies is named a *force* [vis]; it could be due to external causes, since the body will remain in its own state due to internal causes. (Euler 1765c, p. 44) (A17)

After the introduction of external and internal causes of motion, Euler could go on to demonstrate what is now known as the Newtonian equation of motion, for the one-dimensional case. In essence Euler considered as a necessary truth, that is, rational, the law that Newton considered contingent that is empirical. The demonstration takes place into two steps.

Theorem 2a. The small space [$d\omega$], through which a given body at rest is advanced in the assigned small interval of time dt by different forces, is proportional to the forces. (Euler 1765c, p. 55) (A18)

Theorem 3. If equal forces act on unequal bodies at rest, the effect [$d\omega$] produced in the same small time intervals [dt] is inversely proportional to the inertias of the bodies. (Euler 1765c, p. 56, translation from Bruce I) (A19)

The proof of Theorem 2a is very simple, based on the assumption of the additivity of the effect of different forces. Basically, said Euler, if a small body, a corpuscle, is pulled forward by a force equal to p in the short time interval dt through the small space equal to $d\omega$ and if another force equal to p is acting on the same body along the same direction, the body progresses through another equal small interval equal to $d\omega$. Thus this corpuscle acted on by a force equal to $2p$ in the same interval of time dt is pulled through the small distance equal to $2d\omega$. Similarly if n forces – or equivalently a force np – act on the corpuscle at rest for the same time interval dt, they move the body through the interval equal to $nd\omega$.

The demonstration is not convincing however. Euler assumed indeed that the effect of the sum of two forces (two causes) is the sum of their effects, that is, he is assuming a principle of superposition and this is not granted. (The principle of superposition presupposes for its validity that the two cause do not interact to give a cause different from their sum.) And even if this is conceded, the proof would be only a trivial theorem of arithmetics (at Euler's time) for which if two quantities increase of the same amount they are proportional. Moreover the demonstration follows a reasoning different from that carried out in the *Anleitung*. Here (*Theoria*), in the proof of Theorem 2a, force is considered as measured a priori; in the sense that it is independent of the formulation of dynamics, there (*Anleitung*) force was measured from effect, and the proof of the theorem becomes a simple truism.

Euler's writing is substantially contemporary to one of Daniel Bernoulli, the *Examen principiorum mechanicae, et demonstrationes geometricae de compositione et resolutione virium* of 1726 (Bernoulli 1728), in which the latter posed the problem of the logic status of the law that links acceleration to force, assuming force as defined a priori concluding that the law is empirical and not rational. Bernoulli said that the part of mechanics that deals with the balance of forces can be deduced from the principle of composition of forces, as Varignon has shown. If another principle is added to this principle, namely, that according to which the increases in velocity are proportional to the increments of time multiplied by the force, mechanics is completed, relative to the motion. Galileo used this principle. Bernoulli believed that the principle of composition of forces is a necessary truth, while that of Galileo is a contingent truth: "Nature could have made the increases in velocity in the bodies proportional to the increments of time multiplied by any function of the pressure, so that said t the time, p the pressure and v the velocity, it was not $dv = pdt$, but for example $dv = p^2dt$ or $dv = p^3dt$" (Bernoulli 1728, pp. 126–127).

The proof of Theorem 3 is analogous. Basically it is said that if one joins two bodies of equal mass under a given force, the effect is halved. With some details, following Euler's reasoning: Let consider a corpuscle having an inertia A, which at rest is moved by a force equal to p for the short interval of time dt through the small

space $d\omega$; if another corpuscle B equal to A is acted on by a force also equal to p along the same direction, it moves in the same dt by the same small space $d\omega$. If the two corpuscles are joined together into one resulting in a body with an inertia $2A$ (it is an implicit assumption by Euler), it acted by the force equal to $2p$, in the time equal dt moves still of the small space $d\omega$. Thus a force $2p$ on the body of inertia $2A$ produces the displacement $d\omega$. Similarly a force np applied to a body n A produces the same displacement $d\omega$. Thus for the Theorem 2a, the force p on a body n A produces a displacement $d\omega/n$.

Note that inertia is not equivalent to the Newtonian quantity of matter; it is rather a property of a body whose measurement is defined in an operational way as the constant of proportionality between force and displacement $d\omega$. Euler can thus introduce the concept of mass in the following way, with Definition 15:

> The *mass* or the *quantity of matter* of a body is the name given to the amount of the *inertia* which is present in that body, by which just as it tries to continue in its own state so it tries to resist all changes (Euler 1765c, p. 57). (A20)

where also the term quantity of matter is redefined by means of the concept of mass, inverting the usual approach where the mass is defined by the quantity of matter. The property of additivity is implicitly attributed to the mass as clear from the proof of Theorem 3. That is, if one joins two bodies with mass m_1 and m_2, he gets a body with mass $m_1 + m_2$.

The operational definition of mass in the eighteenth century used by Euler is not common; almost all mathematicians treat mass and inertia as proportional to the quantity of matter, taking for granted the meaning of quantity of matter (for instance, the number of atoms). The operational definition of mass will be resumed in the nineteenth century by Mach in his *Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt* written at the end of the nineteenth century (Mach 1919).

In the *Mechanica* Euler had given a different formulation of mass, and it is after a theorem – and not a definition – (Proposition 17) that the amount of inertia (vis inertia) is proportional to the amount of matter or mass (Euler 1736, p. 57). The amount of matter, however, was defined in a vague way, like the set of points (atoms?) that make up a body. Only that not all the points have the same mass. More precisely the points can be taken having the same mass when the same force exerts an equal effect on them (that is they have the same inertia). Thus the demonstration of the theorem on the equality between inertia and mass of the *Mechanica* is the result of a vicious circle of the type: the amount of matter is proportional to the amount of inertia, so the amount of inertia is proportional to the amount of matter.

At this point, by putting together Theorem 2a and Theorem 3 and using the definition of mass, Euler could enunciate Theorem 4, according to which:

> Theorem 4. If corpuscles at rest with masses in an unequal ratio, are acted on by some singular forces, the small intervals through which they are thrust forwards in the same short time intervals will be composed in the direct ratio of the forces and the inverse of the masses. (Euler 1765c, p. 59; Translation adapted from Bruce I.) (A21)

That is, for a force p and a mass A, it can be written $d\omega \propto p/A$.

In Problem 9, Euler started to replace the elementary displacement $d\omega$ measured from the rest with the variation dv of velocity. Coming to the relation (with his symbols):

$$dv = \frac{\lambda p dt}{A}$$

where λ is a proportionality constant necessary to move from a proportion to a formula, which is the natural way to work with algebra and Calculus. (Incidentally if A is the weight, as Euler always assumes, then λ coincides with the acceleration of gravity (modern meaning).) Euler basically said that the increment $d\omega$, for a given interval of time dt, is proportional to the element of velocity dv, and therefore for Theorem 4, the above relation follows.

A modern reader sees in the previous formula the second law of motion of Newtonian mechanics in the monodimensional case. To discuss whether Euler's was the first formulation of the equation of motion – it was not indeed – is not relevant here; certainly it was one of the first times the second law of motion was written in a very clear and unequivocal form. The only thing missing is the meaning of the constant λ; but this concerns only the choice of units of measurement and Euler made a choice later.

What leaves a little "surprised," especially if reading the *Theoria* immediately after the *Mechanica*, is the fact that dv is placed directly proportional to the time interval dt; Theorem 4 authorizes only to say that dv is proportional to p and inversely proportional to A. The proportionality between dv and dt had been treated as a theorem in the *Mechanica* (Proposition 15), attributing its discovery to Galileo, who actually saw it as a plausible hypothesis to be verified experimentally, or if not it, its consequence, that is, the law according to which spaces vary with the square of times.

> 135. Galileo was the first to use this theorem in the investigation of falling heavy bodies. Indeed he did not give a demonstration of it, however because of the strong agreement with phenomena, nobody doubted it anymore. (Euler 1736, p. 53) (A22)

Euler's demonstration is very simple but circular. This is how it works: consider n infinitesimal time intervals $d\tau$. The increase in velocity du is the same in each interval of time $d\tau$, assumed equal to each other, because by hypothesis the force, that is, the cause of motion is the same. So after n equal intervals $d\tau$, the velocity becomes $dv = ndu$, from which it follows that dv is proportional to $dt = nd\tau$ (Euler 1736, p. 51). The circularity is to admit the constancy of the increase in velocity over time, which is the same as saying that the increase in velocity is proportional to time.

The reasons for omitting the proof of proportionality of dv with dt in the *Theoria* are not clear, at least to me. It is possible that everything derives from Euler's, and others, studies on finite differences, in fashion in the early eighteenth century. These studies had shown that the variation of a function y of x, in a certain small interval dx,

is necessarily proportional to dx itself, that is, $\Delta y \sim \psi dx$, width ψ a proportionality constant, generally dependent on x. Euler assumed it natural to consider the variation of velocity dv that in a given interval of time, once force and mass are fixed, cannot but vary linearly with the infinitesimal interval of time dt, and therefore $dv = \psi dt$. The fact that ψ can only vary with p and A (and not also on the time t) depends on the physical nature of the problem that tells us that the variations of velocity in a certain interval are always the same (ψ= const.), or it varies as $\psi = \psi(p, A)$, but the case $\psi = \psi(t)$ is not discussed.

The motion of a body moving along a plane curve (Problem 13) following a force acting on the plane is treated by resolving the force into two (orthogonal) components and then studying two separate one-dimensional motions. It should be noted that unlike the case presented in the Problem 2, which was purely kinematic in nature, the idea of resolution or projection tends not to show the underlying physical principle, the parallelogram rule, for which the projections of forces are forces themselves that act independently of each other (without interaction). This is one of the cases in which the power of the mathematical instrument hides the physical nature of the problem. Indeed the use of Cartesian coordinates has deeper advantages than that of mere simplicity; the addition of vectors located at different points is so natural as to become customary at once, because it is reduced to the ordinary algebraic summation, giving for granted that this addiction has a physical meaning. Somehow Euler was aware of the problem and said that a force is traced back to three forces P, Q and R acting in the (orthogonal) directions x, y and z, with the "static resolution criterion" (Euler 1765c, p. 77).

The equations of motion eventually take the form (Euler 1765c, p. 85):

$$ddx = \frac{2gPdt^2}{A}; \quad ddy = \frac{2gQdt^2}{A}; \quad ddz = \frac{2gRdt^2}{A}$$

in which A is the mass, while the symbol g denotes the height through which a heavy body drops in the time of a second, in a specified region of the globe, in case mass and weight are assumed to have the same numerical value: "Let the mass of the corpuscle be equal to A, which clearly indicates as well the weight, if the corpuscle is situated in a region of the earth chosen to evaluate an absolute measure" (Euler 1765c, p. 77). (Note that g is one half the value of the acceleration of gravity G, this follows from the relation $s = 1/2Gt^2$: By assuming $t = 1$ it follows $g = 1/2G$.)

In these equations the constant λ has been replaced by $2g$. That is, a coefficient that had only the role to transform a proportion in an algebraic equation is given a mechanical meaning, the space traveled in one second. Still in his work of 1749, Euler had assumed $\lambda = 1/2$, which, in modern terms, corresponds to assume an acceleration of gravity $G = 1/2$ and thus $g = 1/4$. The choice had derived from the desire of giving a simple expression to the speed of fall from the height h, which is simply given by $v^2 = h$, instead of $v^2 = 2Gh$. No inconsistency is at play; it is just a different choice of units of measurement.

To note that the equations of motion Euler wrote are in accord with the rule of calculus of the eighteenth century, where the concept of differential was fundamental.

Using modern notation, Euler normally wrote the equation of motion as $mdv = fdt$, and not $ma = f$, where the symbols are the usual. However, he also used the acceleration a in a technical way, as the ratios dv/dt or dds/dt^2. This is what he wrote when integrating the equation of motion: [In dynamics] the effect should be measured by the acceleration or the change in the speed, that is impressed on the body in a given time: this is proportional to that force divided by the mass of the body. [...] dv is equal to the product of the acceleration and the element of distance travelled (Euler 1736, p. 84).

3.2 Measurement of Forces

The explanation of the origin of forces given in Euler's writings on natural philosophy, in the *Researches sur l'origine des forces* in particular, does not provide any criterion for their measurement. The force due to impenetrability is not measurable in itself as it exists only at the moment of impact. What, at least in principle, is measurable is its effect, which, for Euler according to the Newtonian approach, is expressed by the variation in velocity in an assigned small interval of time, in the case of the impact much less than the duration of the contact between the bodies that collide with each other.

In the writings specific on mechanics, such as the *Mechanica* and the *Theoria*, the problem is discussed a little more. In the first treatise, Euler showed no particular difficulty to introduce the measurement of forces. It could be obtained by the rules of statics, a discipline at the time considered well founded. These rules indeed allow to measure a force as function of a sample force, for instance, a weight. It is enough to evaluate how many unities of the sample force should be summed (the additivity of forces is granted, for instance, by the rule of parallelogram) to reach the equilibrium when they act in the opposite direction with respect to the force to be measured. This means that the measure of a force is a priori and independent of dynamics that is independent of the dynamical effect it produces.

Euler is more careful in the *Theoria*; here he distinguished between a static and a dynamic measure; of course when two forces measured with the two approaches give the same numerical value for Euler, the forces are equal under all the aspect, that is, there are not dynamic and static forces but only different ways to measure them; the dynamic way is used only when static measurements are not possible, as, for instance, it occurs for gravitational forces of astronomy.

Euler had established his equations of motion, Theorems 2a and 3, referring to the increment of the displacement of a body due a force starting from the rest, because only in this case he had an independent criterion to measure force, the static one. In the case of motion Euler said that nothing is known concerning the measurement of forces and the way to measure them is left to us: "Since in statics, from which we draw the measurement of forces, the bodies to which the measurements are applied, may be considered in a state of rest, and thus nothing is defined concerning the measurement of these forces when they act on the body in motion" (Euler 1765c, p. 53).

In dynamics the only way left for the measure of forces should be searched in the measure of the effects, that is, they can measured a posteriori only: "Therefore the

magnitude of these forces is determined not by the impenetrability, which clearly cannot be quantified, but from the change of the state which must be effected lest the body penetrate each other" (Euler 1765c, p. 50).

Euler, however, did not dwell at length on the problem of measuring forces and did not relate the dynamic and a posteriori measure of force, based on motions, with the static and a priori measure, based on the equilibrium. In such a way he avoided to make clear if the foundation of his mechanics is dynamic or static. Euler left comments on the dynamic measure of force only in some Explicationes of the *Theoria*, as he would not to compromise himself with strong declarations. Here he proposed the convention that if a force acting on a body in motion causes a displacement shift σ in the direction in which it acts, that force is assumed to be equal to the force that would cause the same displacement σ from the rest, evaluated, for instance, according to Theorem 2a (that is, if $\sigma = kp$ is the law derived from Theorem 2a and the shift of displacement measured for a body in motion is $\overline{\sigma}$, the force associated to this variation of motion is given by $\overline{p} = \overline{\sigma}/k$).

> 144. For the forces, then, by which bodies already in motion are acted upon, we set up this ground of measuring, so that we shall judge these equal to those which would have executed the same effect on the same bodies at rest in the same time. This ground, however, does not require proving, because it rests upon a definition and thus it was open to us to establish it. For if for any motion the small space $s\sigma$ should be equal to small space $S\sigma$, through which the same small body at rest is brought forward in that same little time by force p, we also call those forces equal. (Euler 1765c, p. 53; Translation in Pourciau (2016)) (A23)

Of course if it is admitted to measuring force through its effect, the Newtonian law of motion could be considered as a simple definition, and the concept of force could become superfluous. Only a weak ontological substratum would remain which allows to attribute some reality to force; but it is a link that can be broken without substantial changes in the formal development of mechanics. This link will be broken, shortly after the publication of the *Mechanica*, for example, by d'Alembert.

One could say that Euler preferred pragmatism to rigor. His alleged apodictic foundation of mechanics based on principles, self-evident for him, as those of impenetrability and inertia of bodies, works only if one remains at a superficial level, but above all it is not able to explain all the mechanics, for example, the link between statics and dynamics. Of course one could also criticize the evidence of the ideas of impenetrability and inertia or the lawfulness of the use of the principle of sufficient reason, but these criticisms should apply to the whole period in which Euler lived.

4 The Apparent Motion and the Observer

Although Euler believed in the existence of an absolute space, he nevertheless believed that motion, as a matter of fact, could only be studied in a relative space, defined by a certain conventional reference frame. It will then be the task of the physicist to establish what relationships exist between the relative space chosen to

study motion and the absolute space. This is, for instance, what Euler wrote in his *Mechanica*, and in other occasions:

> Because of the immense nature of space and of its unbounded nature [...] we are unable to form a fixed idea of this. Thus, in place of this immense space and of the boundaries of this, we are accustomed to defining a finite space and the limits within which bodies can move, from which we can indicate the states of motion and of rest of bodies. Thus, we are accustomed to say that a body that keeps the same situation with respect to its boundaries is at rest, and truly that which changes with respect to the same, to be in a state of motion. (Euler 1736, Vol. 1, p. 2; Translation by Bruce I.) (A24)

In the *Anleitung* Euler posed himself the problem to see what happens by changing the reference frame, in particular by considering two reference frames in motion with respect to each other, a problem which had ancient origins. Galileo founded his mechanics referring to thought experiments imagining to be in the cabin of a moving ship. Huygens also set out on a uniform motion to study the problem of collision. Euler is part of this tradition, also followed by some mathematicians at the beginning of the eighteenth century, but as usual, he specified and enlarged it. A fairly detailed analysis of his modus operandi can be found in (Maltese 2000; Bertoloni 1993).

Already in one of his works of 1739 in which he addressed the problem of the influence of the speed of light propagation in astronomical observations, Euler faced the purely kinematic problem to correct the observational data to take into account the fact that the observers (spectatores in Latin) based on the earth mobile with respect to the fixed stars (Euler 1739a).

In Chap. 10 of the *Anleitung*, Euler systematically introduced the concept of observer, to which he referred to with the German term Zuschauer, which is properly translated as spectator, basically a fixed or mobile coordinate frame (without clock). There are some important results that clarify what today is known as the Galilean principle of relativity. Thanks to the use of Cartesian coordinates and calculus, the mathematical passages that Euler had to face appears very simple to a modern.

The first proposition states:

> If the observer moves at constant speed along a straight line, and estimates directions correctly, then all bodies that are either at rest or are moving at constant speed in a straight line will appear to him to remain in the same state. (Euler 1862, Vol. 2, p. 497; Translation by Hirsch E.)

Notice that it is not specified very clearly the observer is considered to be moving with respect to the absolute space. The second proposition basically states that the equations of motion are the same for an absolute system and for a system that moves uniformly with respect to it:

> If the observer moves uniformly in a straight line, and if he judges directions correctly, that is according to parallel running lines, then the maintenance of the apparent movement requires the same forces as the true movement, however much the apparent movement may differ from the true movement. (Euler 1862, Vol. 2, p. 497; Translation by Hirsch E.) (A25)

In the two propositions, the statement asserting that the direction is judged correctly means in modern term that the observer is moving without rotation.

In the following I refer with some details Euler's proof. Let's assume three coordinate axes in the absolute space, OA, OB, and OC and call the components of the motion (velocities) along the axes u, v, w. The motion of the observer is defined by velocities α, β, γ that by assumption are constant. To the observer, the motion of the body as seen from the axis OA will appear the smaller, the faster his own motions. The apparent motion in the direction of the axes OA, OB, and OC will be $\underline{u} = u - \alpha, \underline{v} = v - \beta, \underline{w} = w - \gamma$, respectively. From the equations of motion that Euler had already developed, it follows that for the maintenance of the true (absolute) motion, one requires three forces, $MP = P,\ MQ = Q,\ MR = R$, so that:

$$P = \frac{Mdu}{ndt}; \quad Q = \frac{Mdv}{ndt}; \quad R = \frac{Mdw}{ndt}$$

where M is the mass and n a normalizing factor. Now replace u, v and w, by $\underline{u}, \underline{v}, \underline{w}$, in these equations. Since α, β and γ are constants, the differentials remain unchanged, so that it is apparent that for the maintenance of the apparent motion the same forces P, Q and R are required as for the true motion.

In a third proposition, perhaps the newest one, Euler considered an observer which moves in a translational way (without turning) unevenly. In this case, he concluded that fictitious forces must be considered alongside real forces.

> But if the observer does not move uniformly in a straight line, *but does estimate directions correctly* [emphasis added], then to maintain the apparent motion of all bodies, in addition to the forces that are actually acting on the bodies, further forces are required that will at every instant and in every body produce the change that takes place at the location of the observer, but acting in the opposite direction. (Euler 1862, Vol. 2, p. 498; Translation by Hirsch E.) (A26)

In this case also the proof is very simple. However arbitrarily the motion of the observer may change, with reference to the three assumed axes; it can always be represented by the three motions (velocities) α, β and γ by taking these quantities to be variable. Now if the true movement requires the forces:

$$P = \frac{Mdu}{ndt}; \quad Q = \frac{Mdv}{ndt}; \quad Q = \frac{Mdw}{ndt}$$

By replacing u, v, w by $u - \alpha$, $v - \beta$, $w - \gamma$, the maintenance of the apparent motion will require the following three forces:

$$\text{Kraft} \quad MP = \frac{Mdu}{ndt} - \frac{Md\alpha}{ndt} = P - \frac{Md\alpha}{ndt}$$
$$\text{Kraft} \quad MQ = \frac{Mdv}{ndt} - \frac{Md\beta}{ndt} = Q - \frac{Md\beta}{ndt}$$
$$\text{Kraft} \quad MR = \frac{Mdw}{ndt} - \frac{Md\gamma}{ndt} = R - \frac{Md\alpha}{ndt}$$

Therefore, apart from the forces P, Q, R that actually act on each body, three additional forces are needed that in each body produce the same change that occurs at the location of the observer, but in the opposite direction.

The above proposition of Euler finds its applications on several occasions. In the *Theoria*, for example, he finally applied it to astronomy, in order to eliminate the non-inertial motion of the observer's frame based on the earth (Euler 1765c, pp. 120–121). In his important work on water wheels, Euler studied the action of a jet of water coming out of a rotating pipe (Euler 1756). To determine this action, the observer is located in a frame rigidly linked with the pipe (Maltese 2000, p. 342; Bertoloni 1993, pp. 322–326). An approach that will be followed by the nineteenth-century engineers involved the study of turbines. In a more general study of 1755, Euler arrived to formulate relations that have analogies, a part from some mistakes in calculation, with the famous relations that Gaspard-Gustave de Coriolis (1792–1843) found almost a century later (Bertoloni 1993, p. 326; Maltese 2000, pp. 342–342), with reference to a problem similar to that faced by him, that is a rotating wheel.

5 Dynamics of Rigid Bodies

The problem of the motion of rigid bodies, particularly of their rotation, has been the subject of careful studies since the middle of the seventeenth century. It was then presented as the problem of the composite pendulum and its reduction to a simple pendulum. Descartes, Roberval, and Huygens had dealt with the problem (Capecchi 2014) and then the Bernoullis. In 1745 Daniel Bernoulli wrote to his friend Euler at the end of 1745 of the problem as a very difficult and almost impossible to solve for anyone (Truesdell 1968, p. 257).

Euler succeeded in the enterprise shortly after Bernoulli's letter with his famous paper *Decouverte d'un nouveau principe de mecanique* (Euler 1752b), presented at the Académie royale de science et belles lettres de Berlin in 1750. The locution "noveau principe" referred to in the title of the paper at a first reading appears to coincide with those that today go under the name of equations of motion of the mass point in a system of three coordinated axes, which Euler had already presented in another of his works, the *Recherches sur le mouvement des corps célestes en général* of 1749, without referring to it as a new principle. Historians have often questioned in which sense this principle was new. Generally, in the past, it was considered that the title adopted by Euler was improper; after all it was not Newton who had written the equations of motion? And in any case Euler himself had rewritten them a year before. More recently, however, a great originality has been sustained, thus passing from one extreme to another (Truesdell 1968; Maltese 2001).

There are some novelties with respect to the paper of 1749. The mass to which "Newton's equations" refer are not necessarily of isolated bodies; they may also be of infinitesimal bodies, that is, the elementary mass dM of the corpuscles which are – true or imaginary – part of an extended body, for the moment considered as rigid. And the forces acting on these corpuscles are not only the external forces acting on

the whole body but also the constraint reactions that come from the surrounding corpuscles to maintain the constraint of rigidity. So the novelty of Euler would be of method, in the sense that the motion of an extended body is studied using the same laws that apply to a mass point, thus reducing mechanics to a single principle, as Euler had always declared he wanted to do, and of merit because in these equations a broader concept of force is used.

The historians who praise Euler's role, however, are silent on the fact that d'Alembert in 1749, a year before Euler presented his work, wrote *Recherches sur la precession des equinoxes* that although starting from different principles of mechanics presented an approach that was equivalent to that of Euler from a mathematical point of view. It equally based on infinitesimal elements of the extended body, the earth, by calculating the necessary force to keep it in motion, as a product of the infinitesimal masses for their acceleration due to gravitation forces, according to relations derived from the study of the kinematics of rigid bodies and by imposing the equivalence of the (static) moments of these forces with those of the external forces. Among other things, with his principles, d'Alembert was not obliged to call for the constraint reactions, which somehow caused embarrassment to Euler (see below). However, regardless of whether Euler's work was original or not, it was certainly fundamental for the evolution of his thought and also of the mechanics of eighteenth century.

Here is as Euler presented his new principle: Let consider an infinitely small body, whose mass M (dM) is assumed to be contained in a single point. Suppose that that body of mass M has received an arbitrary motion along the direction x and that it is acted on by arbitrary forces P, we will have:

$$2Mddx = Pdt^2$$

with the symbols having the usual meaning. "And it is this formula alone, which contains all the principles of mechanics" (Euler 1752b, p. 195). A similar reasoning can be repeated for the direction y and z and forces Q, R arriving to the equations (Euler 1752b, p. 196):

$$2Mddx = Pdt^2; \quad 2Mddy = Rdt^2; \quad 2Mddx = Rdt^2$$

which are exactly the same as those reported in the paper of 1749.

> It is thus from this same great principle that it will be necessary to derive all the rules which we need in order to determine the motion of a solid body whose axis of rotation does not remain immobile. For this purpose, it will be necessary to consider not only all the elements of the body, but also their mutual connection, in virtue of which all the elements maintain among themselves the same order and the same distances. For the motion of the whole body is composed of the motions of all its elements, and the motion of each of these must follow the principle which I have just explained, given that each element participates in the forces [*f*] which act on the body, and, in addition, *is affected by certain forces [r] which prevent it from losing its connection with the others* [emphasis added]. But before we can determine the effect of the forces to which the elements are subjected, we must consider in general the motion to which such a body is susceptible. (Euler 1752b, p. 197; Translated by Langton SG.) (A27)

The motion of a rigid body subject to generic forces is decomposed into the motion of the center of gravity and the motion around it, that is, a rotation. For the first motion, Euler merely stated that it obeys the laws of motion valid for the mass point by considering the mass of the whole body concentrated in its center of gravity and subjected to a force equal to the resultant of all the forces acting on the body.

To study the rotating motion, Euler considered a frame fixed with the origin in the center of gravity of the body and three non-rotating axes AO (z), BO (y), and CO (x) and studied the motion, displacement, velocity, and acceleration, of a mass corpuscle dm, which occupies the position x, y, z, when the body rotates around an axis passing through O. Rotation is defined by the components λ, μ, and ν of the angular velocity vector with respect to the axes x, y, z.

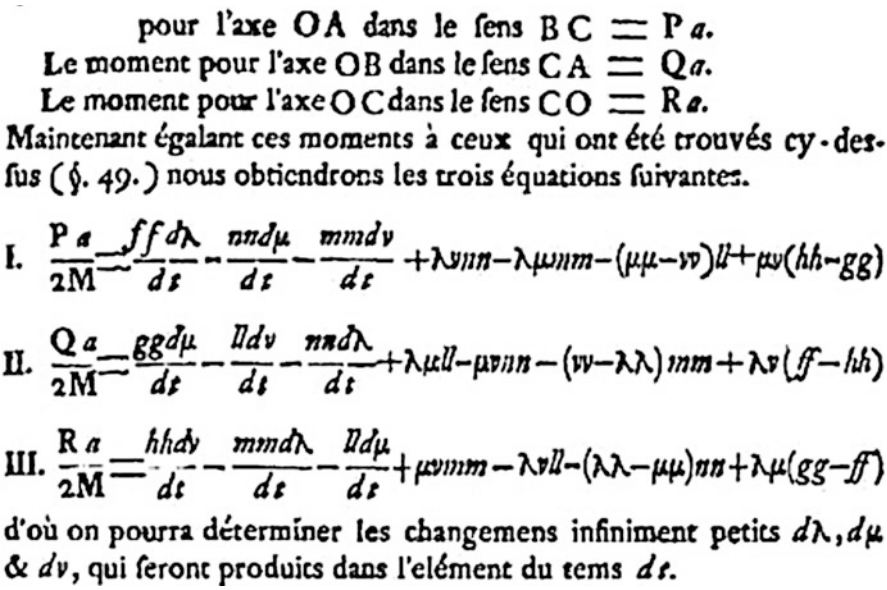
Once found the expression of ddx, ddy, ddz, starting by the kinematic relations of x, y, z and the angles of rotation, he equated the resultant moment of the masses multiplied by acceleration (ma) to the resultant moment of the forces ($f+r$) acting on the body with respect to the fixed axes. (In modern term, given the equation of motion $ma = f + r$, Euler made equal the moments of the two members of the equality.) In the analysis of the forces, Euler believed he could ignore the constraint forces (r). For which "it is to be remarked that the internal forces mutually cancel one another, so that the continuation of the motion requires only the external forces, to the extent that those forces do not mutually cancel" (Euler 1752b, p. 206). It is not clear here whether Euler refers to the principle of action and reaction or, more simply, to the shared axiom of mechanics that a body cannot modify its motion due to internal forces. Truesdell thinks that Euler drew his idea from Daniel Bernoulli (Truesdell 1968, pp. 252–256).

The result of Euler's calculations is shown in Fig. 2. In the Eqs. I, II, and III, M is the mass of the whole body; ff, gg, and hh are in the order the axial moments of inertia (Euler term) with respect to the axes z, y, x, while ll, mm, and nn are the mixed moments of inertia with the pairs of axes xy, xz, yz, respectively, and λ, μ, ν are the components of the vector of angular rotation according to the axes z, y, x. Pa, Qa, Ra are the (static) moments of the external forces with respect to the axes z, y, x.

Euler referred to the Eqs. I, II, and III as "three formulas which contain the new principles of Mechanics which we need in order to determine the motion of solid bodies." They are quite complex, especially because the moments of inertia are referred to fixed axes, thus: "it is necessary to calculate anew for each instant the values ll, mm, nn, ff, gg, hh, because they will vary as a result of the change in the position of the body with respect to the three [fixed] axes" (Euler 1736, p. 214).

Euler took up the subject of the rigid body motion in later works, the results of which are summarized in his *Theoria* of 1760 where there are no substantial new elements compared to what Euler had written in the previous 10 years. Besides writing the equations of motion of the rigid body, Euler also dealt with complex problems, including gyroscopic motions.

A superficial reading of this "applicative" part makes it clear the perspectives that calculus has opened to the mathematical investigation of motion. It also makes manifest the theoretical and practical difficulties that arise in the solution of the differential equations of motion which requires the development of new fields of mathematics.



pour l'axe OA dans le ſens BC = P a.
Le moment pour l'axe OB dans le ſens CA = Q a.
Le moment pour l'axe OC dans le ſens CO = R a.
Maintenant égalant ces moments à ceux qui ont été trouvés cy-deſſus (§. 49.) nous obtiendrons les trois équations ſuivantes.

$$\text{I.}\quad \frac{P\,a}{2M} = \frac{ff\,d\lambda}{dt} - \frac{nn\,d\mu}{dt} - \frac{mm\,d\nu}{dt} + \lambda\nu nn - \lambda\mu mm - (\mu\mu - \nu\nu)ll + \mu\nu(hh - gg)$$

$$\text{II.}\quad \frac{Q\,a}{2M} = \frac{gg\,d\mu}{dt} - \frac{ll\,d\nu}{dt} - \frac{nn\,d\lambda}{dt} + \lambda\mu ll - \mu\nu nn - (\nu\nu - \lambda\lambda)mm + \lambda\nu(ff - hh)$$

$$\text{III.}\quad \frac{R\,a}{2M} = \frac{hh\,d\nu}{dt} - \frac{mm\,d\lambda}{dt} - \frac{ll\,d\mu}{dt} + \mu\nu mm - \lambda\nu ll - (\lambda\lambda - \mu\mu)nn + \lambda\mu(gg - ff)$$

d'où on pourra déterminer les changemens infiniment petits $d\lambda$, $d\mu$ & $d\nu$, qui ſeront produits dans l'elément du tems dt.

Fig. 2 Motion of a rigid bodies about a movable axis (Euler 1752b, p. 213)

To confirm what was already clear to Euler since 1736, that mechanics and mathematics proceed together, because a part of mechanics has now become mathematics itself.

However, the way in which Euler reached the equations of motion is also interesting from a foundational point of view. He did so by making explicit what was implicit in his article of 1750. He started from the idea of *elementary (disturbing) forces*, that is, those forces that applied to each element of the body separately may produce the same change in their state as likewise actually found (Euler 1765c, p. 117). The elementary forces, if applied in the opposite direction to the body, reduce it to rest. This is, as shown in the next chapter, what goes under the name of d'Alembert principle (introduced in 1743). Euler did not mention d'Alembert, and with a little reason because he had already used this approach in his previous works, but without giving him the relief that he is giving it here. To distinguish the *elementary disturbing forces* from the "ordinary forces," he will refer to the latter as the *actual forces*. Euler referred to the principle of d'Alembert as a "well known circumstances which can be deduced from the principles of metaphysics, always fully effective according to reason. But this principle should be correctly understood; and indeed it is usually proposed generally in a vague manner, so that its conclusions cannot be reached without any risk" (Euler 1765c, pp. 119–120).

> Moreover here the actual forces acting bear the causes in turn, that we designate by the letter *V*; then the effect is twofold: on the one hand, by which the motion of the body is affected, in place of this the elementary forces are to be assumed immediately effecting a change in the motion, which forces likewise we denote by the letter *T*. Now the other effect consists in trying to change the structure of the body, in place of this the forces affecting the bonds must be taken, which we denote by the letter *S*. Therefore since from the cause *V* there is produced

> the effect equal to $T + S$, it must be considered that $V = T + S$, thus on collecting terms together, $S = V - T$, in short as we have found above. (Euler 1765c, pp. 119–120; Translation by Bruce I.) (A)

The equations of motion to translation and rotation of a rigid body are obtained by equating the resultant and (static) moment of the elementary and external forces, according to the rules of statics.

Euler devoted a particular care to constraint forces, that is, the forces, to which the constraints of body must resist, which can be obtained as the resultant of the forces actually disturbing the body and the elementary forces applied in the opposite direction. The actual forces with the elementary forces taken away must give the forces affecting the constraints. How a body resists to these forces depends on its structure and on the way the parts of the body are connected together, which refers to physics rather than to mechanics. However, Euler said we only assume bodies, of which we consider the motion, to be provided with a sufficient level of stiffness, so that no change in the shapes be apparent. Of course it is sufficiently plausible that the connection of the parts is not so strong to not give way a little bit, although this could escape our senses. If it is true, then clearly no body is able to be taken as rigid, unless those where in general there are no forces trying to disturb the bond, since even by the smallest forces a certain change in the shape is being produced. "Now whatever such rigid bodies, assumed here, are present in the world or not, this question does not touch on the present discussion, since in all disciplines it is not needed that the objects are in existence to be contemplated" (Euler 1765c, p. 120).

Studies carried out by Euler on the motion of a rigid body in the period 1750–1760 are very well summarized in (Maltese 2000, 2003). In 1751 Euler employed axes rigidly linked to the body (in the following *body axes*) and noticed the existence of axes of free rotation (Euler 1767). After János András Segner (1704–1777) had discovered in 1755 the existence of at least three of such axes named principal axes of inertia, Euler explored their properties in *Du mouvement de rotation des corps solides autour d'un axe variable* (Euler 1765b) and arrived in 1758 to the three equations of motion of a rigid body, which are named after him in the paper *Du mouvement de rotation des corps solides autour d'un axe variable* of 1758 (Euler 1765a) and which are much simpler that those reported in Fig. 2.

They are in the form (Euler 1765a, p. 165) (Note that the two papers of 1751 and 1758 have very similar title and sometimes are confused. This has occurred also in the Euler archive.):

$$P = Maa\frac{d \propto \cos}{2gdt} + M(cc - bb)\frac{\propto\propto \cos\beta\cos\gamma}{2g}$$

$$Q = Mbb\frac{d \propto \cos\beta}{2gdt} + M(aa - cc)\frac{\propto\propto \cos\gamma\cos\alpha}{2g}$$

$$R = Mcc\frac{d \propto \cos\gamma}{2gdt} + M(bb - aa)\frac{\propto\propto \cos\alpha\cos\beta}{2g}$$

where $\propto$ is the modulus of the angular velocity; α, β, γ are the angles the axis of rotation makes with the body axes a, b, c, the principal axes of inertia passing from

the center of gravity; *aa*, *bb*, *cc* are the moment of inertia with respect to the principal axes; *P*, *Q*, *R* are the component of the (static) moment of the external forces with respect to the body axes; *M* is the total mass of the body (actually the weight); and *g* is one half of the acceleration of gravity. These equations are much simpler than those referred to in Fig. 2.

After 1760, Euler came back to the study of the motion of rigid bodies. In a work of 1775, *Nova methodus motum corporum rigidorum determinandi* (Euler 1776), Euler now 70 years old and with impaired sight rewrote the equations of motion of the rigid body in the form shown in Fig. 3, where *x*, *y*, *z* are the coordinates of the mass point *dM* with respect to a fixed coordinate system; *P*, *Q*, *R* and *S*, *T*, *U* are, respectively, the components with respect to *x*, *y*, *z* of the resultant of forces and (static) moments with respect to the center of gravity of external forces; *i* is the acceleration of gravity. The integral must be extended to the body of which the motion is studied.

Truesdell (Truesdell 1968) gives much importance to the Eqs. I–VI of Fig. 3. In his opinion:

1. It is the first time that the equations of motion appear altogether, rotation and translation.
2. They have general validity, that is, they apply to all types of bodies, even those that are not rigid.
3. The equations apply not only to the whole body but also to each of its finite or infinitesimal portions.

Eodem modo pro axe I B obtinebimus hanc aequationem:

$$\int x\,d\,M\left(\frac{d\,d\,z}{d\,t^2}\right)-\int z\,d\,M\left(\frac{d\,d\,x}{d\,t^2}\right)=i\,T.$$

Tertia vero aequatio erit pro axe I C

$$\int y\,d\,M\left(\frac{d\,d\,x}{d\,t^2}\right)-\int z\,d\,M\left(\frac{d\,d\,y}{d\,t^2}\right)=i\,U.$$

§. 29. Hac igitur ratione ſex nacti ſumus aequationes, quas hic coniunctim conſpectui exponamus

I. $\int d\,M\left(\frac{d\,d\,x}{d\,t^2}\right)=i\,P$ IV. $\int z\,d\,M\left(\frac{d\,d\,y}{d\,t^2}\right)-\int y\,d\,M\left(\frac{d\,d\,z}{d\,t^2}\right)=i\,S$

II. $\int d\,M\left(\frac{d\,d\,y}{d\,t^2}\right)=i\,Q$ V. $\int x\,d\,M\left(\frac{d\,d\,z}{d\,t^2}\right)-\int z\,d\,M\left(\frac{d\,d\,x}{d\,t^2}\right)=i\,T$

III. $\int d\,M\left(\frac{d\,d\,z}{d\,t^2}\right)=i\,R$ VI. $\int y\,d\,M\left(\frac{d\,d\,x}{d\,t^2}\right)-\int z\,d\,M\left(\frac{d\,d\,y}{d\,t^2}\right)=i\,U.$

In

Fig. 3 Equations of motion of a rigid body (Euler 1776, p. 224)

4. The Equations of moments IV, V, and VI are treated as if they were independent of the Eqs. I, I, and III.

Truesdell's theses deserve respect, considering the long study he has dedicated to Euler, but they are ideologized, in the sense that Truesdell having made Euler his champion tends to attribute him the authorship of the principles used in modern continuum mechanics.

I will comment on Truesdell's thesis point by point.

1. It is true that the equations of equilibrium to translation and those to rotation appear together for the first time in Euler's texts and therefore also in the history of mechanics. But this is not particularly significant. Euler had in fact already written since 1750 equations similar to those presented in the work of 1775. Any compiler by reading Euler's texts would effortlessly summarize his mechanics with the equations of Fig. 3.
2. It is true that the equations of Fig. 3 have general validity. They represent the equilibrium equations of mechanics regardless of the nature of the bodies to which they refer, hence independently of what today goes under the name of constitutive relationship, that is, the fact that the body be rigid, or perfectly flexible as a chain or elastic (the *elastica*, for instance), or fluid body. For example, in Euler's text of 1775, it can be assumed that Euler had consciously defined the constraint of rigidity starting from coordinate transformation expressed by the relations:

$$x = f + FX + F'Y + F''Z$$

$$y = g + GX + G'Y + G''Z$$

$$x = h + HX + H'Y + H''Z$$

where x, y, z are the coordinates of the element dm, in the fixed coordinate system; X, Y, Z are the coordinates of the same element with respect to the body axes, whose axes which, for instance, coincide with the barycentric principal axes of inertia; f, g, h are the coordinates of the center of gravity, in the fixed coordinate system. $F, F', F'', G, G', G'', H, H', H''$ are the cosines directors of the body axes with respect to the fixed coordinate system. Substituting these relations in the equations of Fig. 3, one obtains the equations of motion of the rigid body. But it is not at all certain that Euler was aware of the generality of his results. At least this does not result from the reading of the work of 1775, in which explicit reference is made to the rigid body only.
3. In the work of 1775, the integrals are extended to the whole body only.
4. Reading the work, it seems that Euler proceeded in the same way as in 1750; that is, he started from the forces on the mass element dm and then evaluated the resultant for translation and rotation. However, compared to 1750 or even 1760, he forgot the constraint reactions and their role. It is obviously difficult that this

was a forgetfulness. It remains to be seen whether Euler ignored constraint reactions for simplicity of treatment, considering that he intended essentially to present equations of the motion of the rigid body in a different form from what has been done in previous works. Or instead he chooses of adopting two independent principles instead of just one. This however would completely transform the axiomatic of Euler and would void his ideal of unification and rationalization. A discipline based on a single principle demonstrated a priori, and therefore purely rational, would pass to a physical-mathematical discipline based on two principles, of which no explanation can be given, which confer hypothetical nature to the theory, leaving its verification to the agreement with the experimental results, perhaps too modern for Euler.

According to Truesdell his thesis would also be confirmed by the approach followed by Euler in the study of deformable continua. An in-depth study of the study in this direction is not of interest for the following writing; moreover, most likely it would not lead to a conclusive result.

6 Quotations

A1 Eine flüssige Materie muss zu allererst diese Eigenschaft haben, dass ihre Theilchen nicht aneinander befestigt sind, so dass ein jegliches Teilchen ohne einigen Widerstand von den übrigen abgesondert und in Bewegung gesetzt werden kann

A2 La matière, sur laquelle je voudrois à présent entretenir V.A. me fait presque peur. La variété en est surprenante, & le dénombrement des faits sert plutôt a nous éblouir qu'à nous éclairer. C'est de l'Electricité dont je parle, & qui depuis quelque tems en devenüe un article si important dans la Physique, qu'il n'est presque plus permis à personne d'en ignorer les effets.

A3 Alles was undurchdringlich ist, gehört in das Geschlecht der Körper, und daher besteht das Wesen der Körper in der Undurchdringlichkeit, in welcher folglich alle, übrigen Eigenschaften ihren Grund haben müssen.

A4 II y aura donc deux especes de matiere, l'une qui fournit l'étoffe à tous les corps sensibles et dont toutes les particules ont la même densité, qui est très considerable et qui surpasse meme de plusieurs fois celle de l'or; l'autre espece de matiere sera celle dont ce fluide subtil qui cause la gravite est compose, et que nous nommons l'ether.

A5 Die grobe Materie ist also an sick selbst keiner anderen Veränderung fähig als in Ansehung ihrer Figur, welche, wenn hinlängliche Kräfte vorhanden, auf alle mögliche Arten verändert werden kann.

[. . .] Dass die subtile Materie auch allezeit und allenthalben eine, beständige Dichtigkeit haben sollte, dergestalt, dass dieselbe durch keine Kräfte in einen kleineren Raum getrieben werden kannte, scheint der Wahrheit nicht gemäss zu sein. Vielmehr mochte auch hierin ein

Hauptunterschied zwischen der groben und subtilen Materie bestehen, dass sick diese zusammendrücken liesse.

A6 Um dieses deutlicher darzuthun, so wollen wir erstlich setzen, ein Körper werde in einen kleinern Raum zusammengepresst. Weil sich nun die grobe Materie fur sich nicht zusammenpressen lasst, so kann dieses nicht anders geschehen, als wenn die Poren kleiner gemacht werden. In diesem Falle muss demnach die scheinbare Dichtigkeit des Körpers wachsen, weil die ganze Materie, woraus der Körper besteht, oder zum wenigsten die grobe, da die subtile in Ansehung derselben für nichts zu achten, in einen kleinern Raum gebracht worden. Es sei a^3 der Theil des vom Körper eingenommenen Raumes, welcher mit grober Materie angefüllt ist, e^3 aber der ülbrige Theil, so nur subtile Materie in sich enthalt, oder die Summe von allen Poren zusammen genommen, so wird $a^3 + e^3$ die Grosse des Körpers, as seine Masse und $a^3/(a^3 + e^3)$ seine Dichtigkeit aüsdrucken. Nun aber kann a^3 nicht verändert werden, daher, wenn der Körper in einen kleinern Raum gebracht wird, so wird nur e^3 verringert oder [...] Wenn nach geschehener Veränderung der Figur eines Körpers die verschlossenen Poren weder grosser noch kleiner werden, so behalt der Körper diese veränderte Figur. Wenn aber die verschlossenen Poren weiter oder enger werden, so wird sick in dem Körper eine Kraft äussern, sick wieder in seine vorige Figurherzustellen.

A7 Ungeachtet wir aber hier stehn bleiben müssen und kaum hoffen können, jemals die wahre Ursache dieser Verminderung der elastischen Kraft des Aethers zu ergründen, so kann man sich doch damit leichter begnügen; als wenn man blosserdings vorgiebt, alle Körper seien von Natur mit einer Kraft begabt, einander anzuziehen. Denn da man sich von diesem Anziehn nicht einmal einen verständlichen Begriff mach en kann, so kann man im Gegentheil zum wenigsten überhaupt einsehn, wie es möglich sei, dass die elastische Kraft einer flüssigen Materie vermindert werde, und man begreift auch, dass dieses auf eine den Gesetzen der Natur gemässe Art geschehen könne.

A8 Qiucquid statum corporum absolutum mutare valet, id vis vocatur; quae ergo, cum corpus ob causas internas in statu suo esset permansurum, pro causa externa est habenda.

A9 XXXL. Cette égalité des forces, d'ou depend le grand principe de l'égalité entre l'action & reaction, est une suite necessaire de la nature de la penetration. Car, s'il étoit possible que le corps A pénétrat le corps B, le corps A seroit précisément autant péenétre par le corps B; donc, puisque le danger que ces corps se pénètrent, est égal de part & d'autre, il faut aussi que ces deux corps employent des forces égales pour resister a la penetration. Ainsi, autant que le corps B est sollicité par le corps A, précisément autant sera celui-cy sollicité par celui-là, l'un & l'autre déployant exactement autant de force qu'il faut pour prevenir Ia penetration. Or ces deux corps agissant l'un sur l'autre par une force quelconque, se trouveront dans le même état que s'ils étoient comprimes ensemble par Ia même force.

A10 Idem omnino mihi,cum Neutoni Principia et Hermann Phoronomiam perlustrare coepissem, usu venit, ut, quamvis plurium problematum

solutiones satis percepisse mihi viderere, tamen parum tantum discrepantia problemata resolvere non potuerim. Illo igitur iam tempore, quantum potui, conatus sum ex synthetica illa methodo elicere easdemque propositiones ad meam utilitatem analytice pertractare, quo negotio insigne cognitionis meae augmentum percepi. Simili deinde modo alia quoque passim dispersa ad hanc scientiam spectantia scripta sum persecutus, quae omnia ad meum usum methodo plana et aequabili exposui atque in ordinem idoneum digessi. Hoc in negotio occupatus non solum in plurimas antea nondum tractatas incidi quaestiones, quas feliciter solutas dedi:sed etiam complures peculiares methodos sum adeptus, quibus tam mechanica quam ipsa analysis non parum augmenti accepisse videantur.

A11 Or on énonce communément ce principe par deux propositions, dont l'une porte, qu'un corps étant une fois en repos demeure éternellement en repos, à moins qu'il ne soit mis en mouvement par quelque cause externe ou étrangere. L'autre proposition porte, qu'un corps étant une fois en mouvement, conservera toujours éternellement ce mouvement avec la même direction & la même vitesse, ou bien sera porté d'un mouvement uniforme suivant une ligne droite, à moins qu'il ne soit troublé par quelque cause externe ou étrangere.

A12 Quoique les principe dont il s'agit ici soient nouveaux, entant qu'ils ne sont pas encore connus ou étalés par les Auteurs, qui on traité la Mécanique, on comprend néanmoins, que le fondement de ces principes ne saurait être nouveau, mais qu'il est absolument nécessaire, que ces principes soient déduits des première principes, ou plutôt des axiomes, sur le quels toute la doctrine du mouvement est établie.

A13 Definitio I. Quemadmodum Quies est perpetua in eodem loco permanentia, ita Motus est continua loci mutatio. Corpus scilicet, quod semper in eodem loco haerere observatur, quiescere dicetur:quod autem labente tempore in alia atque alia loca succedit, id moveri dicitur.

A14 Axioma I. Omne corpus, etiam sine respectu ad alia corpora, vel quiescit vel movetur, hoc est vel absolute quiescit vel absolute movetur.

A15 Axioma 2. Corpus, quod absolute quiescit, si nulli externae actioni fuerit subiectum, perpetuo in quiete perseverabit.

Axioma 3. Corpus, quod absolute movetur, si nulli externae actioni subiiciatur, secundum eandem directionem motu aequabili progredi perget.

A16 Definitio 11. Proprietas illa corporum, quae rationem perseverationis in eodem statu in se continet, inertia appellatur, quandoque etiam vis inertiae.

A17 Definitio 12. Qiucquid statum corporum absolutum mutare valet, id vis vocatur; quae ergo, cum corpus ob causas internas in statu suo esset permansurum, pro causa externa est habenda.

A18 Theorema 2a. Spatiola, per quae idem corpusuclum quiescens eodem tempusculo dt a diveris viribus promovetur, sunt ipsis viribus proportionalia.

A19 Theorema 3. Si aequales vires corpuscula inaequalia quiescentia sollicitent, effectus eodem tempusculo producti erunt reciproce inertiae corpusculorum proportionales.

A20 Definitio 15. Massa corporis vel quantitas materiae vocatur quantitas inertiae, quae in eo corpore inest, qua tam in statu suo perseverare quam omni mutationi reluctari conatur.

A21 Theorema 4. Si corpuscula ratione massae inaequalia quiescant atque a viribus quibuscunque singula sollicitentur, erunt spatiola, per quae eodem tempusculo protrudentur, in ratione composita ex directa virium in inversa massarum.

A22 Theoremate hoc ex solutione problematis invento primus est usus GALILAEUS ad motum gravium delabentium investigandum. Eius quidem demonstrationem non dedit, sed tamen propter insignem eius cum phaenomenis congruentiam de eo amplis dubitari noluit.

A23 Pro viribus ergo, quibus corpora iam mota sollicitantur, hanc dimetiendi rationem stabilimus, ut eas aequales iudicemus iis, quae in iisdem corporibus quiescentibus eodem tempore eundem effectum essent praestaturae. Haec autem ratio non indiget probatione, quia definitioni innititur nobisque adhuc liberum fuerat eam constituere. Si enim promotu quocunque spatiolas $s\sigma$ aequalia fuerint spatiola $S\sigma$, per quod idem corpusculum quiescens tempusculo eodem profertur a vi p, huic etiam illas vires aequales appellamus

A24 Quoniam autem immensi illius spatii cuiusque terminorum, [...] nullam nobis certam formare possimus ideam; loco huius immensi spatii eiusque terminorum considerare solemus spatium finitum, limitesque corporeos, ex quibus de corporum motu et quiete indicamus. Sic dicere solemus, corpus, quod respect horum limitum situm eundem conservat, quiescere, id vero, quod situm eodem respectu mutat, moveri.

A25 Wenn der Zuschauer gleichgeschwind in einer graden Linie fortrückt und die Gegenden richtig, das ist, nach gleichlaufenden Linien schätzet, so werden zur Unterhaltung der scheinbaren Bewegung, wie sehr dieselbe auch von der wahren unterschieden sein mag, eben diejenigen Kräfte erfordert, als zur Unterhaltung der wahren Bewegung.

A26 Wenn aber der Zuschauer sich nicht gleichförmig in einer graden Linie bewegt, dennoch aber die Gegenden richtig schätzet, so werden, um die scheinbare Bewegung aller Körper zu bewerkstelligen, noch ausser den Kräften, welche wirklich auf dieselben wirken, solche Kräfte erfordert, welche in einem jeden Körper alle Augenblicke eben die Veränderung hervorbringen, welche in dem Orte des Zuschauers vorgeht, aber nach einer umgekehrten Richtung.

A27 XXIV. C'est done de ce même grand principe, d'où il faudra dériver les regles, donc no us avons besoin pour determiner le mouvement d'un corps solide, lorsque l'axe de rotation ne demeure pas immobile. Pour cet effet il faudra considerer non seulement tous les élémens du corps, mais assi leur liaison mutuelle, en vertu de laquelle tous les élémens conservent entr'eux le meme ordre & les mêmes distances. Car le mouvement du corps entier est composé des mouvemens de tous ses élémens, & le mouvement de chacun doit suivre le principe, que je viens d'expliquer, entant que chaque élément

participe des forces qui agissent fur le corps, & qu'il est outre cela sollicité par de certaines forces, qui l'empêchent, qu'il n'abandonne la connéxion avec les autres. Or avant que de déterminer cet effet des forces, auxquelles les élémens sont assujettis, il faut considerer en general le mouvement, donc un eel corps en susceptible.

A28 Hic autem vires actu sollicitantes vicem causae gerunt, quam littera V designemus; deinde effectus est duplex: alter, quo motus corporis afficitur, cuius loco assumi debent vires elementares mutationem motus immediate efficientes, quas vires simul littera T denotemus. Alter vero effectus in conatu structuram corporis turbandi consistit, cuius loco sumi debent vires compagem afficientes, quas littera S notemus. Cum igitur a causa V producatur effectus $= T + S$, censeri debet $V = T + S$, unde colligitur $S = V - T$, prorsus uti invenimus. Verum in tanta rerum metaphysicarum caligine malim demonstrationem allatam adhibere ad principium metaphysicum illustrandum.

7 Conclusions

The justification by the scholars of the eighteenth century of the foundations of mechanics required substantial involvement in metaphysics and epistemology to introduce fundamental notions: the nature of space, time, force, constitution and properties of bodies, nature of motion. This effort was pursued by Euler not with the classical and organic approach of canonical philosophy but with the pragmatism of a mathematical philosopher. Euler's role in mechanics and philosophy has been overshadowed by the historians such as Mach, Dugas, Montucla, Dühring, very careful to the fundaments, who see him more as a mathematician than a mechanician or a philosopher. After all, more than 60% of his work deals with pure mathematics, and even those whose object is mechanics and astronomy contain many sections that can be classified as mathematics. Today Euler's role, even for what concerns the fundaments of mechanics, is reevaluated. Euler was lucky enough to be born in a period when all the mechanics and all the calculus were to build starting from the foundations of the seventeenth century. The section has showed that he used his luckiness however as much as fruitful it was possible.

8 Cross-References

► Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice

References

Bernoulli D (1728) Examen principiorum mechanicae, et demonstrationes geometricae de compositione et resolutione virium. Commentarii Academiae Scientiarum Petropolitanae 1:126–142

Bernoulli D (1738) Hydrodynamica, sive de viribus et motibus fluidorum commentarii. Dulsseker, Strasbourgh

Bernoulli D (1750) Remarques sur le principe de la conservation des forces vives pris dans un sens général. Mémoires de l'Académie Royale des Sciences et Belles Lettres de Berlin 4:356–364

Bertoloni M (1993) The emergence of reference frames and the transformation of mechanics in the enlightenment. Hist Stud Phys Sci 23:301–335

Calinger R (2016) Leonhard Euler. Mathematical genius in the enlightenment. Princeton University Press, Princeton

Cantor G, Hodge M (eds) (1981) Conceptions of ether. Studies in the history of the ether theories. Cambridge University Press, Cambridge

Capecchi D (2001) La tensione secondo Cauchy. Hevelius, Benevento

Capecchi D (2014) Attempts by Descartes and Roberval to evaluate the centre of oscillation of compound pendulums. Early Sci Med 19(3):211–235

Cassirer E (1953) Substance and function and Einstein's theory of relativity. Dover, New York

Euler L (1736) Mechanica sive motus scientia analytice exposita (2 vols). Academiae Scientiarum, Saint Petersburg

Euler L (1738) De communicatione motus in collisione corporum. Commentarii Academiae Scientiarum Petropolitanae 5:159–168

Euler L (1739a) Explicatio phaenomenorum quae a motu successivo lucis oriuntur. Commentarii Academiae Scientiarum Petropolitanae 11:150–193

Euler L (1739b) Tentamen novae theoriae musicae ex certissismis harmoniae principiis dilucide expositae. Academiae Scientiarum, Saint Petersburg

Euler L (1746a) De la force de percussion et de sa veritable measure. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 1:21–53

Euler L (1746b) Gedancken von den Elementen der Körper. In: Various (ed) Leonhardi Euleri Opera omnia (in progress), III, vol 2, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp 347–366

Euler L (1746c) Recherches physiques sur la nature des moindres parties de la matiere. Opuscula Varii Argumenti 1:287–300

Euler L (1748a) De observatione inclinationis magneticae. In: Leonhardi Euleri Opera omnia (in progress), III, vol 10, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp 110–137

Euler L (1748b) Dissertatio de magnete. In: Various (ed) Leonhardi Euleri Opera omnia (in progress), III, vol 10, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp 138–179

Euler L (1749) Recherches sur le mouvement des corps célestes en général. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 3:93–143

Euler L (1750) Reflexions sur l'espace et le tems. In: Various (ed) Leonhardi Euleri Opera omnia (in progress),III, vol 2, Teubneri GB; [then]Fussli Turici; [then]Birkhäuser, Basel, pp 376–383

Euler L (1752a) Avertissement au sujet des recherches sur la precession des equinoxes. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 6:412

Euler L (1752b) Decouverte d'un nouveau principe de mecanique. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 3:185–217

Euler L (1752c) Dissertatio de igne, in qua ejus natura et proprietates explicantur. In: Recueil de piéces qui ont remporté le prix de l'Académie royale des sciences. Contenant les pieces depuis 1738 jusqu'en 1740, vol 4, Imprimerie Royale, Paris, pp 5–21

Euler L (1752d) Recherches sur l'origine des forces. In: Various (ed) Leonhardi Euleri Opera omnia (in progress), II, vol 5, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp 109–131

Euler L (1756) Théorie plus complete des machines qui sont mises en mouvement par la reaction dell'eau. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 10:227–295

Euler L (1757a) Principes generaux de l'état d'équilibre des fluides. In: Leonhardi Euleri Opera omnia (in progress), II, vol 12, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp 2–53
Euler L (1757b) Principes generaux du mouvement des fluides. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 11:274–315
Euler L (1761) Comment to F.U.Th. Aepinus' Descriptio ac explicatio novorum quorundam experimentorum electricorum. In: Leonhardi Euleri Opera omnia (in progress), III, vol 10, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp CXXI–CXXIV
Euler L (1765a) Du mouvement de rotation des corps solides autour d'un axe variable. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 14:154–193
Euler L (1765b) Recherches sur la connoissance mecanique des corps. Mémoires de l'Académie des Sciences et Belles Lettres de Berlin 14:131–153
Euler L (1765c) Theoria motus corporum solidorum seu rigidorum ex primis nostrae cognitionis principiis stabilita et ad omnes qui in huiusmodi corpora cadere possunt accommodata, presented 1765. Röse, Rostck
Euler L (1767) Du mouvement d'un corps solide quelconque lorsqu'il tourne. In: Various (ed) Leonhardi Euleri Opera omnia (in progress), II, vol 8, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp 313–356
Euler L (1770–1774) Lettres à une princesse d'Allemagne sur divers sujets de physique & de philosophie (3 vols). Steidel & Compagne, Mietau, Leipzig
Euler L (1775) Lettres à une princesse d'Allemagne sur divers sujets de physique & de philosophie (3 vols). Societé Typographique, Berne
Euler L (1776) Nova methodus motum corporum rigidorum degerminandi. Novi Commentarii academiae scientiarum Petropolitanae 20:208–238
Euler L (1802) Letters of Euler on different subjects in physics and philosophy addressed to a German princess, 2 vols. Translated into English by Hunter H. Murray et al, London
Euler L (1862) Anleitung zur Naturlehre. In: Fuss P, Fuss N (eds) Opera postuma (2 vols), III, vol 1. Eggers & Socios, Saint Petersburg, pp 449–560
Euler L (1911–2018) Leonhardi Euleri Opera omnia (in progress). Teubneri GB; [then] Fussli Turici; [then] Birkhaüser, Basel
Gaukroger S (1982) The metaphysics of impenetrability: Euler's conception of force. Br J Hist Sci 15(2):132–154
Hakfoort C (1995) Optics in the age of Euler. Cambridge University Press, Cambridge
Koyré A (1965) Newtonian studies. Champan & Hall, London
Lagrange JL (1811) Mécanique analytique. In: Serret JA, [Darboux G] (1867–1892) (ed) Oeuvres de Lagrange (14 vols), vol 11, Gauthier-Villars, Paris
Lagrange JL (1867–1892) Ouvres de Lagrange (14 vols). Serret JA [Darboux G] (eds). Gauthier-Villars, Paris
Mach E (1919) The science of mechanics: a critical and historical account of its development. Translated into English by McCormack TJ. Open Court, Chicago
Maclaurin C (1801) A treatise on fluxions (2 vols). Baynes et al, London
Maltese G (2000) On the relativity of motion in Leonhard Euler's science. Arch Hist Exact Sci 54 (4):319–348
Maltese G (2001) Da "F=ma" alle leggi cardinali del moto: Sviluppo della tradizione newtoniana nella meccanica del 700. Hoepli, Milan
Maltese G (2003) The ancients' inferno: the slow and tortuous development of 'Newtonian' principles of motion in the eighteenth century. In: Essays on the history of mechanics in memory of Clifford Ambrose TruesdelI and Edoardo Benvenuto. Springer, Basel, pp 199–222
Noll W (2004) Five contribution to natural philosophy. Carnegie Mellon University, Pittsburgh
Pourciau B (2016) Instantaneous impulse and continuous force: the foundations of Newton's *Principia*. In: Iliffe E, Smith GE (eds) The Cambridge companion to Newton. Cambridge University Press, Cambridge, pp 93–186

Radelet de Grave P, Speiser D (2004) Introduction to volume III/10 of Leonhardi Euleri Opera omnia, electricity, heat and magnetism. In: Leonhardi Euleri Opera omnia (in progress), III, vol 10, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel, pp XIII–CXXIV

Speiser D (2008) Euler, the principle of relativity and the fundamentals of classical mechanics. In: Fuss P, Fuss N (eds) Discovering the principles of mechanics (1600–1800), Birkhäuser, Basel, pp 179–224

Speiser D (2010) Il fisico Leonhard Euler. Il Club dei Volterriani 10:31–56

Suisky D (2009) Euler as physicist. Springer, Berlin

Truesdell CA (1960) The rational mechanics of flexible or elastic bodies. In: Leonhardi Euleri Opera omnia (1911–, in progress), 2, vol 1, part 2, Teubneri GB; [then] Fussli Turici; [then] Birkhäuser, Basel

Truesdell CA (1968) Essays in the history of mechanics. Springer, New York

Algebraic Versus Geometric Thought and Expression in the Early Calculus

Viktor Blåsjö

Contents

1 Introduction 324
2 Absence of Trigonometric Functions in Early Calculus 324
3 Introduction of Trigonometric Functions by Euler 329
4 Logarithmic Functions 330
5 Dimensional Homogeneity 331
6 Did Barrow Prove the Fundamental Theorem of Calculus? 333
7 Conclusion 337
8 Cross-References 338
References 338

Abstract

The language of the early calculus was much more geometrical than the analytic and algebraic style that was pioneered by Euler and still dominates today. For instance, functions such as $\sin(x)$ and $\log(x)$ were largely absent from the early calculus, with geometric paraphrases used in their place. From a modern standpoint, one may be inclined to assume that the eventual triumph of the more analytic perspective was a straightforward case of progress, and that the geometric aspects of the early calculus were a historical artifact ultimately hampering this development. Interestingly, however, in private notes, the pioneers of the calculus showed a readiness to disregard traditionalism and operate freely in a more proto-modern style than they allowed themselves in their publications. This suggests that the adherence to the geometrical mode in published works was a deliberate choice selected with full awareness of the analytic alternative. Indeed, the geometrical paradigm was no mere blind conservatism or lip service to classical foundations; rather, it arguably had genuine merits, for example, as an intuition-boosting heuristic strategy.

V. Blåsjö
Universiteit Utrecht, Utrecht, The Netherlands
e-mail: V.N.E.Blasjo@uu.nl

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_14

This aspect of the early calculus can serve as a case study that illuminates the relation between official expression and informal thought in mathematics more generally. For one thing, it complicates the common historiographic assumption that fidelity to historical thought is best achieved by following the original text's mode of expression as closely as possible.

Keywords

Leibniz · Newton · Barrow · Euler · History of infinitesimal calculus · Fundamental theorem of calculus · Dimensional homogeneity · Anachronism · Historiography

1 Introduction

When studying a historical mathematical text, should we attempt to interpret it in terms of later mathematical concepts? Rampant anachronism can be a mark of insensitivity and a closed mind; it can drown out subtleties and rob us of one of the greatest rewards of studying history in the first place – that of expanding our own mindset and perspective. But banning all anachronistic analysis as illegitimate risks making historical scholarship conceptually impoverished and pathologically preoccupied with incidental details. The question is ultimately one that not only the historian but also the philosopher of mathematical practice must grapple with: Can a mathematician's written word be taken as constitutive of their underlying thought?

The early development of the infinitesimal calculus in the late seventeenth century provides a rich store of case studies for these questions. In addition to interesting differences between the original and modern forms of the calculus, we are fortunate that extensive personal manuscripts by the pioneers of the calculus have been preserved (and in many cases recently published for the first time), helping us further triangulate the relations between thought and expression, tradition and innovation.

2 Absence of Trigonometric Functions in Early Calculus

The early practitioners of the calculus virtually never used any trigonometric expressions such as $\sin(x)$ in their calculus. Neither Newton nor Leibniz ever wrote $\sin(x)$ in any calculus formula in any of their published works. They and many others did write expressions like this in purely geometric contexts, such as referring to sine tables. But they never treated $\sin(x)$ or any other trigonometric expression as a function in a calculus context. Unlike us, they did not see the differentiation and anti-differentiation of such expressions as among the most basic and commonly used calculation rules. They never described curves or solutions to differential equations in terms of such expressions.

How can this be, when any modern calculus textbook is packed with those kinds of uses of trigonometric functions? How can one have a functioning calculus without this tool? In fact, the early practitioners of the calculus were not in any way handicapped by this choice. Let us look at a few examples of how they managed without trigonometric functions. Let us consider some examples of problems where the treatment in a modern calculus textbook seems most essentially based on trigonometric functions, and see how Newton and Leibniz managed to treat these problems just as well, if not better, in other terms.

Simple harmonic motion – such as the bobbing motion of a weight on a spring – is a prototype example of a physical phenomenon described by a sine curve. The harmonic motion of a spring follows from Hooke's law that force is proportional to extension, $F \propto x$. Combined with Newton's law $F = ma = m\ddot{x}$, this means that the differential equation for the motion of the weight on the spring is $x = -k\ddot{x}$. Today, we would express the solutions to this equation in terms of sine and cosine functions. Yet in the seventeenth century people managed without it.

Newton dealt with this problem. Let's see how he did it. Newton considered a more general situation: motion in a force field where the force is proportional to distance from the origin, and directed toward the origin. This is analogous to a gravitational force field, but with a different force law. For any given starting position, a body can be launched into a circular orbit if it is given just the right sideways velocity. If the speed is increased or diminished by any amount, this would cause a deviation from the circular path (Fig. 1). Thus a specific orbital speed is naturally and intrinsically associated with this force field and this starting position.

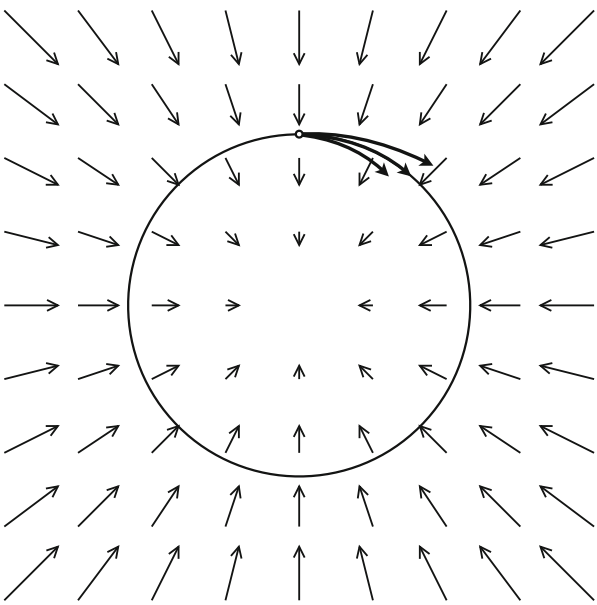
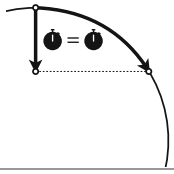
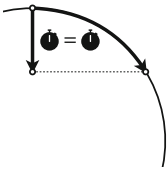
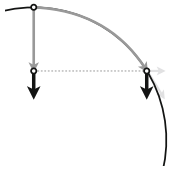


Fig. 1 Uniqueness of circular orbit through a given point in a $\mathbf{F} = -k\mathbf{r}$ force field

Table 1 Behavior of a body being dropped from $(x, y) = (0, R)$ in a $\mathbf{F} = -k\mathbf{r}$ force field

	Position at given time	Time to reach given position	Speed at given position
In terms of formulas	$y(t) = R\cos\left(\sqrt{k}t\right)$	$t(y) = \frac{\arccos\left(\frac{y}{R}\right)}{\sqrt{k}}$	$\dot{y}(y) = -\sqrt{k}R\sin\left(\arccos\left(\frac{y}{R}\right)\right)$
In terms of the associated circular-orbit motion	y-position of the orbital motion at that time 	Time in which orbital motion reaches that y 	y-direction speed of orbital motion at that y 

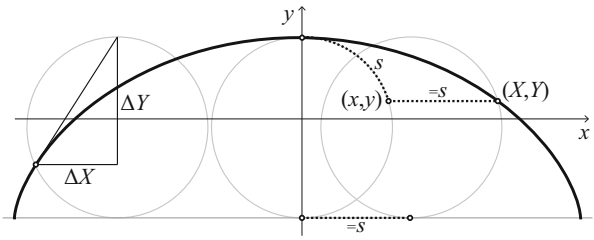
By taking this speed as a reference, one can describe harmonic motion in this field in a natural and precise way without the need for formulas.

This is precisely what Newton does in the *Principia*, Book I, Prop. 38, which is about the simple harmonic motion of a dropped object in a $F \propto r$ force field. He expresses everything in terms of the associated circular-orbit motion. Table 1 puts this approach side by side with the modern approach using formulas. Newton's way has a number of advantages. For instance, the equivalence of the three different results shown is immediately evident from the geometric descriptions but takes some algebra to derive from the formulas. Furthermore, Newton's mode of expression uses language intrinsic to the system itself: the naturally associated orbital motion. The formula approach, on the other hand, becomes ugly and opaque precisely because it uses extrinsic frames of reference. The constants k and R depend on our choice of units of force and position, and the trigonometric functions assume a unit radius and hence have to be scaled in various ways. Those are conventions that are external to the specific scenario at hand, yet they dominate the formulas. This is why the formulas are opaque and hide the simple dynamics of the system: the formulas are primarily focused on accommodating the system to a fixed external reference frame rather than on describing the system in the clearest and most intuitive terms.

Another example where modern calculus textbooks make trigonometric functions seem indispensable is the cycloid. "The only convenient way of representing a cycloid is by means of parametric equations," one standard textbook proclaims (Simmons 1996, 592). In the discussion that follows, the expressions sin and cos occur 40 times in the space of two pages, making the cycloid a good candidate for the most trig-intensive example in the standard calculus repertoire.

Leibniz did not think this was "the only convenient way of representing a cycloid." In fact, the very opposite is the case: Leibniz explicitly and repeatedly used the cycloid as his prime example of how "perfectly" the calculus could represent curves, but his representation is very different from the "only convenient" one of modern textbooks. In his very first paper on the integral calculus, he stresses that his "equation expresses the relation between the ordinate y and the abscissa x

Fig. 2 The cycloid



perfectly" and "from it all the properties of the cycloid can be derived" (Blåsjö 2017, 65).

Leibniz's "perfect" way of expressing the cycloid involves no trigonometry and no parametric equations. It is based instead on a simple geometrical property of the cycloid shown in Fig. 2: The three dashed lengths are all equal, since the arc s measures both the amount of rotation needed to bring the tracing point from the top position down to (x, y), and the amount of arc that has been in contact with the ground during this rotation, that is, the distance traveled, or the horizontal displacement $X - x$. Therefore an equation for the cycloid is $X = x + s$. This equation can be expanded, if desired, by expressing x using the Pythagorean Theorem and s as an arc-length integral. Assuming that the generating circle has unit radius, this gives

$$X(y) = x + s = \sqrt{1 - y^2} + \int_y^1 \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy = \sqrt{1 - y^2} + \int_y^1 \frac{dy}{\sqrt{1 - y^2}}$$

This is essentially Leibniz's "perfect" equation for the cycloid, except he centers his coordinate system on our $y = \ -\ 1$, so he gets the equivalent $\sqrt{2y - y^2} + \int dy/\sqrt{2y - y^2}$ instead.

Leibniz does not spell out how "all the properties of the cycloid can be derived" from this equation, but this can indeed be done. Let us consider how such an approach can lead to the same results that are derived using trigonometric expressions in modern textbooks such as Simmons (1996).

What is the area A of the cycloid, in terms of the area a of the generating circle? Integrating the equation $X = x + s$, we get

$$\frac{A}{2} = \int_{-1}^{1} X dy = \int_{-1}^{1} x dy + \int_{-1}^{1} s dy = \frac{a}{2} + [ys]_{y=-1}^{y=1} - \int_{-1}^{1} y \frac{ds}{dy} dy = \frac{3a}{2} + \int_0^{\pi} y ds$$
$$= \frac{3a}{2} + [x]_{s=0}^{s=\pi} = \frac{3a}{2}$$

so the area of the cycloid is three times the area of the generating circle, $A = 3a$. We see that Leibniz's equation for the cycloid indeed lends itself very well to calculation. The steps of this calculation do not correspond to the standard modern solution and

arguably compare favorably to it. A key ingredient of the modern proof is the use of a trigonometric addition formula to express $\cos^2(\theta)$ in terms of $\cos(2\theta)$. The Leibnizian proof does not need this technical machinery. The modern approach ultimately reduces the integral to the standard antiderivative $\int \cos(\theta)d\theta = \sin(\theta)$. The Leibnizian approach does the same thing in the step $\int yds = x$, which is nothing but a different way of saying $\int \cos(\theta)d\theta = \sin(\theta)$. Thus, the Leibnizian calculus contains the exact equivalent of our standard trigonometric derivatives and antiderivatives, but expresses them in other terms.

What is the tangent of the cycloid at any given point? In fact, the tangent passes through the top point of the rolling circle. This too is derived with quite elaborate trigonometric machinery in Simmons (1996) and other standard textbooks. From Leibniz's equation, we can get this result without any trigonometry. For we find that

$$\begin{aligned}\frac{d}{dy}X &= \frac{d}{dy}(x+s) = \frac{d}{dy}\left(\sqrt{1-y^2} + \int_y^1 \frac{dy}{\sqrt{1-y^2}}\right) \\ &= \frac{-y}{\sqrt{1-y^2}} - \frac{1}{\sqrt{1-y^2}} = \frac{-(1+y)}{\sqrt{1-y^2}} = \frac{-(1+y)\sqrt{1-y^2}}{1-y^2} = \frac{-\sqrt{1-y^2}}{1-y} = \frac{-x}{1-y}\end{aligned}$$

which indeed corresponds to $\frac{\Delta X}{\Delta Y}$ in the triangle in Fig. 2, as claimed. So, in this case too, the Leibnizian calculus handles the problem quite elegantly without any need for trigonometric functions. (In modern terms the differentiation of the integral $\int_y^1 dy/\sqrt{1-y^2}$ is nothing but the derivative of the arccosine, but re-expressing it in such terms adds nothing of value or substance.)

These examples show that although the calculus of Leibniz and Newton is different from ours, it should not be assumed inferior. Modern calculus books leave us with the impression that "the only convenient way of representing a cycloid is by means of parametric equations" involving trigonometric functions, and so on for various other problems such as harmonic oscillation. Seeing, then, that Newton and Leibniz used other methods, one may be inclined to assume that their methods are inherently clumsier, and that these historical mathematicians would have recognized as much on the spot if we explained it to them; the only reason they didn't take this step must have been that they were conceptually limited by old ways of thinking, one naturally assumes. But the above examples suggest that the sense of superiority of the anachronistic modern perspective is hubristic. There are in fact ways in which the older approach can not only readily match what the modern approach can do, but even has a number of outright advantages over it. This suggests that Newton and Leibniz may very well have adhered to their style of calculus as a matter of conscious choice, with full awareness of the possibility of a more formula- or function-based approach.

In fact, there is even one intriguing bit of direct textual support for this interpretation. Although trigonometric formulas are completely absent from all calculus publications for several decades from the inception of the calculus onwards, there is one obscure early manuscript in which Newton does precisely what the hypothesis of

a conceptual limitation postulated that he could not conceive: namely, he *does* use trigonometric formulas in calculus calculations, very much in the style of Euler or a modern textbook. In this manuscript from around 1680, Newton repeatedly writes "fl Cos A" – the fluxion of the cosine of the variable angle A – in the course of one particular problem (MS Add.3963.7, 55r; MP.IV.459). Despite seemingly never working with such expressions before or since, Newton handles them effortlessly and incorporates them in his calculations without ado. It seems that Newton opted for this mode of expression in this particular problem because of its complexity; the more geometric phrasing he normally used (including in the other problems in this very manuscript) would be very verbose in this case, which would probably impede the clarity and overview of the calculations. Thus, Newton seems to have treated "fl Cos A" as a trivial shorthand. This shows that there was certainly no conceptual obstacle that held Newton back from using such expressions more extensively. Clearly, he realized that calculus could be built on such expressions, but consciously opted against it.

3 Introduction of Trigonometric Functions by Euler

Trigonometric functions were eventually introduced into the standard calculus repertoire by Euler. As late as 1754, this was still referred to as the "new calculus of sines" (Katz 1987, 316; Euler *Opera* 1.14.543). The context that led Euler (1739) to introduce this "new calculus" was the differential equation

$$2a\ddot{s} + \frac{s}{b} + \frac{a}{g}\sin\left(\frac{t}{a}\right) = 0,$$

a periodically forced harmonic oscillator, whose solution is

$$s(t) = D\cos\left(\frac{t}{\sqrt{2ab}}\right) + C\sin\left(\frac{t}{\sqrt{2ab}}\right) - \frac{a^2 b\sin\left(\frac{t}{a}\right)}{g(a-2b)},$$

unless $2b = a$, in which case

$$s(t) = D\cos\left(\frac{t}{a}\right) + C\sin\left(\frac{t}{a}\right) + \frac{at}{4g}\cos\left(\frac{t}{a}\right).$$

Thus, unlike the simple harmonic oscillator that had long been handled by geometric paraphrase, the solutions of this differential equation involve combinations of trigonometric functions with different periods and trigonometric components multiplied by a nonconstant function (Fig. 3). The complexity of these solutions makes it very difficult to replace them with a geometrical description or to do without formulas of this type. As Euler put it, "there appear ... motions so diverse and astonishing that one is unable altogether to foresee until the calculation is finished" (Katz 1987, 318). Just as Newton had been led to use cosine formulas more

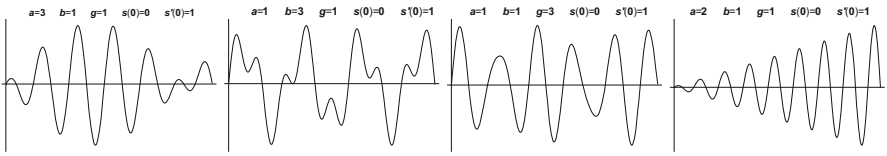


Fig. 3 Examples of solutions to the periodically forced harmonic oscillator considered by Euler (1739)

than half a century earlier in a particularly convoluted problem, so also to Euler the need to introduce such functions arises due to the complexity of the problem. In this way, it is possible to regard Euler's step not as a change in outlook, but merely as a consequence of the same priorities that were present even in Newton's time: the same outlook, consistently applied, leads to different styles in different contexts.

4 Logarithmic Functions

Just as trigonometric functions were absent from the early calculus, so was the logarithm function $\log(x)$. Despite explicitly dealing with the solution to the differential equation $\frac{dy}{dx} = \frac{a}{x}$ and referring to logarithmic curves by name in this connection, early calculus works by, for example, Leibniz (1684) and Johann Bernoulli (1692) refrain from using any kind of expression such as $\log(x)$ in a calculus setting. This could be interpreted as a shortcoming that hampered these works (Jahnke 2003, 111; Kowalewski 1914, 169). Interestingly, however, manuscript evidence shows that the reluctance to treat $\log(x)$ as one of the fundamental functions of the calculus cannot be attributed to an inability to conceive of the possibility of doing so. For Leibniz demonstrably did conceive of it years before, when he wrote "Log y" in a calculus formula in a 1675 manuscript (AA.VII.5.325). Clearly, then, his decision not to write any formulas involving $\log(x)$ even when dealing with problems involving logarithms must be considered a conscious choice rather than unreflective adherence to older modes of expression.

Despite these notable early absences, the logarithm function entered official calculus discourse much sooner than trigonometric functions. The first published use of a logarithm expression in a calculus formula appears to be Leibniz (1694, 369), who writes "log. x." Despite never having used this notation before in print, Leibniz simply starts using it *in media res* without any indication that he has expanded the core repertoire of calculus functions, and with no qualms about taking its properties (such as "log. $x = \int dx : x$") for granted as if they were common knowledge. This makes sense if he regarded "log. x" as a trivial shorthand for what was already well known in other (primarily verbal) terms.

Explicit uses of the logarithm function in calculus formulas did not catch on very quickly. Leibniz himself only used it passingly on one other occasion in print (Leibniz 1695, 314). Manfredi (1707), in what is effectively the first published integral calculus textbook, used the logarithm function quite extensively. Manfredi's

notation is "lx" for log(x). Johann Bernoulli also used this notation on a number of occasions, yet as late as 1716 he still felt the need to explain that "by *lx* I understand the logarithm of that x" (Bernoulli 1716, 228). Thus, even two decades after its first appearance in print, Johann Bernoulli evidently considered the logarithm as a calculus function still not fully established as an elementary part of the standard vocabulary and notation of the calculus.

To conclude, trigonometric and logarithmic functions are seen as an indispensable part of calculus today, but historically the calculus was up and running for a decade before the logarithm function was used in a calculus context, and half a century before trigonometric functions followed suit. This was not for lack of occasion to use such functions: on the contrary, plenty of situations in which we would use these functions were encountered, and often explicitly recognized as concerning trigonometric or logarithmic relations. But this was expressed in verbal and geometric prose, rather than by explicit formulas.

From a modern point of view, this may strike us as potentially a sign of a cognitive limitation: perhaps these early practitioners of the calculus were inhibited by an older geometrical and prosaic paradigm of mathematical thought that prevented them from embracing the power of a more modern formal and purely analytic approach. If we knew only the published works of these mathematicians, then the historical record could be construed as fitting this hypothesis. After all, the calculus community relatively soon decided to favor the analytic style rather than the original approach.

Nevertheless, a strong case can be made that this anachronistically tempting hypothesis misses the mark. Manuscript evidence shows that the analytic approach – in the form of explicit use of trigonometric and logarithmic expressions in calculus formulas – was in fact considered by the creators of the calculus at a very early date, even though they generally opted against this approach in their published works. Indeed, one could argue that the approach they did take was the one best suited for their purposes. The strength of the modern analytic approach over the older style lies especially in handling complicated relationships with many components and parameters. So instead of seeing the transition from geometric to analytic expression as one from ignorance to enlightenment, one can interpret it as the natural outcome of sound preferences that remained consistent throughout: the analytic approach was introduced precisely when the occasion called for it – that is, when the efficiency and generality of formula-crunching outweighed the intuitive appeal of geometric expression.

5 Dimensional Homogeneity

Requiring equations to be dimensionally homogenous was another geometrically motivated choice persistent in the early calculus but at odds with later analytic practice (Bos 1974, 7). Thus a "length" such as y cannot equal an "area" such as x^2, for instance, so one would rather write $ay = x^2$ to get a dimensionally balanced equation. This goes hand in hand with an emphasis on geometric interpretation

rather than analytic formulas. Thus, for example, Johann Bernoulli, in his calculus lectures of 1692, writes the differential equation for the exponential curve as $ydx = ady$, and when he separates the variables to $dx = ady : y$ he then proceeds, for the sake of geometrical interpretation, to explicitly multiply by a to get $adx = aady : y$, which he then interprets visually as an equality of areas (Johann Bernoulli *Opera* III.421; Blåsjö 2017, 84). Bernoulli describes the resulting curve verbally and geometrically as a "Logarithmica," but has no symbolic notation for this at this time. By contrast, Euler, in his calculus textbook, happily works with dimensionally unbalanced equations such as $dy = dx/x$ (Euler 1755, § 180) and treats them purely analytically.

But again it would be a mistake to think of the geometric conservatism of the early calculus as a limitation of thought. Just as Leibniz privately used analytic notation for logarithm functions, so also in early manuscripts he was much more lax about working with equations violating dimensional homogeneity than in his later published works (Bos 1974, 12; Leibniz AA VII.5.xxxvi). A 1694 letter by Johann Bernoulli (Leibniz AA III.6.167) is explicit about what was probably widespread practice: working with a differential equation, he finds that a certain expression "$=dy$" but then immediately adds: or for the sake of dimensional homogeneity "$=bdy$" where b is a new constant introduced solely to balance the equation dimensionally. That is to say, Bernoulli evidently had no problems working in an analytic mode that ignored formal geometric convention when expedient, only to then translate the end result into official form as a veritable afterthought. One finds the same type of reasoning in Leibniz's letters as well (Leibniz AA III.5.188).

This suggests that the convention of dimensional homogeneity was consciously adopted for the positive insights it could bring, without prejudice to other modes of reasoning. Indeed, dimensional analysis is a useful heuristic tool to this day (Pólya 1957, 202–205). Leibniz explicitly points out its usefulness as a check on calculations (AA VII.5.292). It was also very useful for purposes of geometric interpretation (the two examples from Bernoulli just mentioned are instances of this), which was often sensibly pursued in the early calculus as we have seen.

A related respect in which informal private practice was more flexible than formal expression was with regard to curve plotting. Impressively, accurate figures are quite common in the early calculus, but accompanying text is often couched in formal language quite aloof from the concrete, hands-on perspective that must have gone into producing the figures. Again, private manuscripts reveal a greater flexibility of thought than official discourse. Leibniz's description of the catenary $y = (e^x + e^{-x})/2$, for example, is in its published form rather stiltedly formulated in terms of the classical language of ratios. Privately, however, he operated with a more free-wheeling numerical approach, including using a decimal representation of e that he never revealed in print (Raugh and Probst 2019).

A similar example concerns how Jakob Bernoulli corrected his mistaken belief that the radius of curvature of any curve must be infinite at an inflection point. On intuitive grounds, Leibniz and Bernoulli had believed this to be a general principle. Indeed, it holds for simple inflection points, but there are exceptions. For instance, $x^3 = y^5$ has an inflection point at the origin, yet the radius of curvature there is zero.

This is because there are really multiple singularities in one at that point. Bernoulli deals with this by resolving the singularity. Instead of $x^3 = y^5$, consider $x^3 = y^5 - b^2y^3$ for small b. This pulls the singularities apart. Now, the inflection point at the origin is simple again and the general rule holds: infinite radius of curvature. Bernoulli published this finding, but it is interesting to compare his manuscript notes with the published account (Bernoulli 1999, 151–154, 255–258). First of all, it is striking that Bernoulli writes his equations in dimensionally homogenous form ($aax^3 = y^5$ and $aax^3 = y^5 - bby^3$) in his published article, but uses unbalanced equations ($x^3 = y^5$ and $x^3 = y^5 - bby^3$) in his private notes. Furthermore, in his notes Bernoulli used detailed numerical calculations to plot the curve and explore its behavior. A table of x and y values, computed to two decimal places, occurs right next to the figure. This is suppressed or deemed unworthy of inclusion in the published version, where the figure appears without such supporting calculations and the formal presentation gives the impression that the conclusions were reached by abstract reasoning in eminently classical form. The look behind the scenes, at the numerical curve plotting and visual checks in Bernoulli's private notebook, reminds us once again that official published expression can give a misleading impression of the underlying thought process.

6 Did Barrow Prove the Fundamental Theorem of Calculus?

Isaac Barrow (1670) proved certain geometric theorems that could be interpreted as equivalents of the fundamental theorem of calculus. Barrow was writing before the introduction of the concepts and symbolism of derivatives and integrals. He speaks in purely geometrical terms, such as areas and tangents of curves. Should he nevertheless be regarded as having had an insight effectively equivalent to the fundamental theorem of calculus? Some have answered yes (Child 1916, vii, 31; Heath 1917, 133; Nauenberg 2014, 343), and seem to have judged that the direct translatability of Barrow's theorems into this modern form is in and of itself a compelling reason to accept this conclusion. At the other extreme, some emphasize precisely the geometric form and absence of symbolism as a conclusive reason to deny Barrow this insight (Bos 1980, 64–65; Wagner 2001; Sonar 2018, 55). It is advisable to steer clear of both of these extremes when trying to understand the meaning and significance of historical texts. Though these two interpretations are opposites, they share an excessive emphasis on the statement of theorem and proof in isolation. A sounder approach is to investigate how these theorems functioned in Barrow's thought as a whole, and how he saw them fitting into a bigger picture. Historians who have taken such a perspective have tended to came down against the claim that Barrow's theorems are equivalent to the fundamental theorem of calculus (Whiteside 1961, 367–368; Mahoney 1990, 236; Katz 2009, 539).

When taking this balanced approach, the historian can use modernized notation with benefit, as a tool for clarification and analysis, without falling into either of the opposite traps of the "mathematician's" anachronistic naiveté and the "historian's" knee-jerk rejection of all anachronistic devices. Let us take Barrow's relation to the

fundamental theorem of calculus as a case study illustrating this methodology. Thus, we shall first translate Barrow's results into modernized form, while remaining agnostic as to whether this is a faithful representation of his thought or not. This will enable us to see what the global organization of Barrow's text says about how he viewed the theorems in question.

The fundamental theorem of calculus has two parts:

$$\frac{d}{dt}\int_a^t y(x)dx = y(t) \tag{FTC1}$$

$$\int_a^b y'(x)dx = y(b) - y(a) \tag{FTC2}$$

By choosing the coordinate system so that $y(0) = 0$ (which corresponds to Barrow's geometrical treatment), and using the notation $Y(x) = \int_0^x y(t)dt$, we can write this as

$$\frac{d}{dx}Y = y \tag{FTC1}$$

$$\int_0^x y'(t)dt = y(x) \tag{FTC2}$$

What we would interpret as results about derivatives, Barrow expresses in terms of subtangents and subnormals (i.e., the distance from the point where the tangent or normal cuts the axis to the point on the axis perpendicularly below the point on the curve in question; Fig. 4). In modernized language, the subtangent is $\sigma(y) = y/y'$ and the subnormal is $\eta(y) = yy'$.

The purported equivalent of FTC1 is Barrow's X.11. In our modernized form, it says $\sigma(Y) = Y/y$. Using the expression for the subtangent in terms of derivatives, we see that this is indeed equivalent to $Y' = y$, or FTC1. The question is: is this theorem, in Barrow's mind, tied to the specific geometrical configuration, or is it a more structural and general result that it would be natural to call upon in any situation when we seek the rate of change of any anti-derivative or accumulation function? Barrow's own text – and in particular his treatment of a polar version of X.11 – gives us strong reason to favor the former rather than the latter interpretation, as follows.

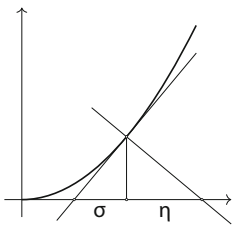


Fig. 4 The subtangent σ and subnormal η of a given curve

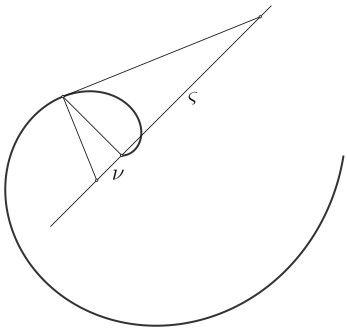


Fig. 5 The polar subtangent ς and polar subnormal ν

Throughout his treatise, Barrow interleaves results of the above form (in effect, assuming perpendicular rectilinear axes) with analogous results for curves defined in polar terms. In modernized form, $R(\theta) = \int_0^\theta (r(\vartheta))^2/2d\vartheta$ is the area-counting-function of the polar curve $r(\theta)$ (that is to say, $R(\theta)$ equals the area swept out by $r(\theta)$ from 0 to θ). For polar curves, the subnormal and subtangent are defined not as segments of the x-axis, but as segments of the line perpendicular to the radial line through the origin point (Fig. 5). The polar subtangent is $\varsigma(r) = r^2/r'$. The polar subnormal is $\nu(r) = r'$.

Barrow's Proposition X.13 is a polar analog of X.11. It says: For the tangent at the point where $r(\theta)$ and $R(\theta)$ intersect, $\varsigma(R) = 2$ (equivalently: $R' = R^2/2$). (Formally, the theorem only enables us to find the tangent at one particular point, but this is not an essential restriction. To find the tangent at another point on the curve $R(\theta)$, say where $R = kr$, we can let $\overline{r} = r/k$, and apply the theorem to the point of intersection of $\overline{r}(\theta)$ and $R(\theta)$, giving $\varsigma(R) = 2/k^2$.) Barrow proves X.13 from first principles, using the geometry of the polar differential triangle of $R(\theta)$. He presents X.13 as a "theorem of the same kind" as X.11, but he makes no use of the latter in the proof of X.13.

In modern terms, we could instead reason as follows. We know that $\varsigma(R) = R^2/R'$ for any polar curve from the geometry of the polar subtangent. In this case, $R(\theta)$ is the area-counting function of $r(\theta)$, which means that $R(\theta) = \int_0^\theta (r(\vartheta))^2/2d\vartheta$. We can easily find its derivative using FTC1. It is $R'(\theta) = r(\theta)^2/2$. So the polar subtangent is $\varsigma(R) = 2R^2/r(\theta)^2$. In particular, when $R = r$, this simplifies to $\varsigma(R) = 2$, which is Barrow's theorem X.13. This is a much shorter and easier proof than the elaborate deduction from first principles used by Barrow. The fact that Barrow did not use X.11 to prove X.13 in this way is thus quite compelling evidence that he lacked precisely that way of thinking about X.11 that warrants the FTC1 the epithet "fundamental."

Let us now turn to FTC2. In Lecture XI, Barrow presents "some Theorems ... relating to the Mensuration of Magnitudes by Tangents or Perpendiculars to Curves." That is to say, this lecture explains how to find areas in terms of tangents – or integrals in terms of derivatives. Indeed, we find a geometric equivalent of FTC2 in Barrow's XI.19. As in the case of FTC1, Barrow's statement and proof of

Table 2 Results proved by Barrow (1670) in modernized form

XI.1	$\int \eta dx = y^2/2$
XI.2	$\int \eta y dx = y^3/3$
XI.3	$\int \eta y^2 dx = y^4/4$
	$\int \eta y^3 dx = y^5/5$
XI.4	$\int Y y dx = Y^2/2$
XI.5	$\int \sqrt{Y} y dx = 2\sqrt{Y^3}/3$
XI.7	$\int \eta + x dx = \frac{1}{2}\sqrt{x^2 + y^2}^2$
XI.10	$\int \sigma dy = Y$
XI.12	$\int \sigma y dy = \int y^2 dx$
XI.13	$\int \sigma y^2 dy = \int y^3 dx$
XI.16	$\int \sigma^2 dy = \int \sigma y dx$
XI.17	$\int \sigma^3 dy = \int \sigma^2 y dx$
XI.19	$\int y' dx = y$
XI.20	$\int (y')^2 dx = \int y' dy$
XI.21	$\int (y')^3 dx = \int (y')^2 dy$

Theorem XI.19 could quite reasonably be construed as corresponding directly to FTC2. Unless one takes a fundamentalist zero-tolerance hardline against any interpretation that transcends surface expression, there is nothing in Barrow's formulation of XI.19 that explicitly indicates any essential conceptual difference between it and FTC2. In other words, Barrow's XI.19 *could* be thought of as a full-fledged FTC2. But the more interesting question is: *did* Barrow think of it this way?

To answer this question, we must not focus on XI.19 itself, but rather look at the context in which this theorem occurs. Much as in the case of FTC1, a contextual perspective strongly suggests that Barrow did not think of XI.19 as fundamental or a conceptual cornerstone of the entire chapter. Rather, he seems to have viewed it as one particular geometrical result among many of a similar kind. Excluding polar curves, the area results that Barrow presents in Lecture XI are shown in Table 2.

In modern terms, it is trivial to verify these results by substituting the respective derivative expressions for η and σ and using FTC2 as well as basic "Leibnizian algebra" with differentials such as $\frac{dy}{dx}dx = dy$.

Today, we think of FTC2 as much more generic and prototypical than the other theorems in Barrow's Lecture XI. But in the geometrical language used by Barrow, it does not have an evidently exceptional status. XI.19 is not any more generic than XI.1 or XI.10 or even XI.7, which has a very direct geometrical meaning since $\eta + x$ is the distance from origin to end of subnormal and $\sqrt{x^2 + y^2}$ is the distance from origin to the relevant point (x, y) on the graph of $y(x)$.

Barrow proves these results by direct area considerations and not by verifying by differentiation that the right-hand sides are antiderivatives of the integrands, as we would in light of FTC2. He never uses XI.19 for such a purpose. There is no indication that he conceived of the possibility of taking XI.19 to be the centerpiece of the theory

in this sense (or any other sense for that matter). Admittedly, XI.19 and its variations occur at the end of this sequence of results, which might suggest that it is the key result that crowns the entire development. Conceivably, Barrow might have perceived the fundamental nature of this result, yet opted to include and prove the previous theorems independently, perhaps for pedagogical, illustrative, or reference purposes. This possibility can be ruled out, however, by some revealing remarks later in the work.

In Appendix III to Lecture XII, Barrow offers a certain Theorem IV which he praises as "the most Fertile of all the Propositions foregoing. The greater Part of them being either contained in it, or deduced easily from it." The theorem in question may be considered a general result on change of variables. In slightly modernized terms, it states that if $y(x)$, $g(x)$, $h(y)$ are functions such that $\frac{\eta(y)}{y(x)} = \frac{h(y)}{g(x)}$ for all x (which in modern terms is equivalent to $g = y'h$) then $\int g dx = \int h dy$. Barrow notes the following special cases. If $h = 1$, we get XI.19. If $h = y$, we get XI.1. If $y = g$, we get XI.10. From this Barrow concludes: "I cannot but accuse my Foresight, for not having first laid down this Theorem . . . and then having deduced the rest from it, . . . which I observe may be done." This very strongly suggests that Barrow did not see the possibility of using his FTC2-equivalent XI.19 to unify the entire sequence of results in Lecture XI. He expressly admits that he would have loved to give such a unified treatment if he had seen a way of doing so. Only later did he realize that this was in fact possible, and even then he still did not see XI.19 as fundamental but rather saw the unification as coming from another theorem altogether that does not correspond to the modern view of the centrality of the FTC.

Thus, in the case of Barrow's purported equivalents of both FTC1 and FTC2, Barrow misses major opportunities to use them to do the kind of work that FTC1 and FTC2 would do for us. This suggests that Barrow's theorems should not be considered equivalent to FTC1 and FTC2. The mere fact that Barrow used geometrical language does not in and of itself entail this conclusion. On the contrary, the early pioneers of the calculus often used geometrical language similar to that of Barrow, yet showed no signs of being constrained in their thinking when the occasion called for formal, nongeometrical applications of calculus principles abstractly in ways that are not visualizable in terms of tangents and areas (e.g., Engelsman 1984, Chap. 2). Hence a similar surface form of expression can be in one case indicative of a genuine conceptual limitation and in another case not. Rather than focusing on the surface form in which an idea is expressed, attention to its functional role in the author's broader argument is a better indicator of which aspects of the form of expression are incidental or essential to the underlying thought. Reconstructions in modernized mathematical terms can be a useful tool for such purposes.

7 Conclusion

The early calculus used geometrical language and shunned the analytic style that came to predominate later. Many historians have tended to view this as a major conceptual divide (e.g., Speiser 2008, 108, 130). More mathematically inclined

scholars, however, have been inclined to push back against this perspective and perceive more continuity in the underlying ideas than the surface form of expression would suggest (e.g., Fraser 2020, 198). The examples from Newton, Leibniz, and Bernoulli analyzed above weigh in favor of the mathematical interpretation. With regard to Barrow's purported geometrical version of the fundamental theorem of calculus, on the other hand, we have a reverse state of affairs. Mathematically inclined readers have been those most ready to accept the equivalence, while the historians' perception of conceptual discontinuity is borne out by closer analysis.

In the former case, the mathematician's intuition was best supported by the historian's methods: The mathematician's sense that the early practitioners of the calculus did not lack the ability to reason in a manner functionally equivalent to the modern analytic approach is borne out by textual evidence that they indeed explicitly used more analytic approaches in private manuscripts that in many cases have only recently been published through meticulous efforts of specialized historians. Conversely, in the Barrow case, the historian's intuition is arguably best supported by the mathematician's methods: A thoroughly modernistic reconstruction of Barrow's reasoning, far from being tantamount to a simplistic anachronistic fallacy, makes a compelling case against accepting the equivalence of Barrow's theorems and their modern counterparts. Efforts to understand past mathematical thought are well served by drawing on the strengths of each these diverse perspectives and approaches.

8 Cross-References

▶ Descartes' Transformation of Greek Notion of Proportionality
▶ Heuristics and Mathematical Practice

References

Barrow I (1670) Lectiones geometricae: in quibus (praesertim) generalia curvarum linearum symptomata declarantur, London. Reprinted in William Whewell (ed) The mathematical works of Isaac Barrow. Cambridge University Press, 1860, 155–316. English translations in Barrow (1735) and Child (1916)

Barrow I (1735) Geometrical lectures: explaining the generation, nature and properties of curve lines. Stephen Austen, London

Bernoulli J (1999) Die Werke von Jakob Bernoulli, Bd. 5. Differentialgeometrie, edited by David Speiser, André Weil, & Martin Mattmüller. Birkhäuser, Basel

Bernoulli J (1692) Lectiones Mathematicae de Methodo Integralium, aliisque. In: Bernoulli J (ed) Opera III, pp 385–558. Partial German translation in Kowalewski (1914)

Bernoulli J (1716) Problema: data serie linearum per rectae in eadem Linea constantis variationem prodeunte invenire aliam seriem linearum, quarum quaevis priores omnes ad angulos rectos secabit. Acta Eruditorum:226–230

Blåsjö V (2017) Transcendental curves in the Leibnizian calculus. Elsevier, Oxford, United Kingdom

Bos HJM (1974) Differentials, higher-order differentials and the derivative in the Leibnizian calculus. Arch Hist Exact Sci 14(1):1–90
Bos HJM (1980) Newton, Leibniz and the Leibnizian tradition, chapter 2. In: Grattan-Guinness I (ed) From the Calculus to set theory, 1630–1910. Princeton University Press, Princeton
Child JM (1916) The geometrical lectures of Isaac Barrow. Open Court Publishing, Chicago and London
Engelsman SB (1984) Families of curves and the origins of partial differentiation, Elsevier North-Holland mathematics studies, vol 93
Euler L (1739) De novo genere oscillationum. Commentarii academiae scientiarum Petropolitanae 11:128–149. (presented 1739). *Opera Omnia*, Series 2, vol 10, 78–97. E126, 1750
Euler L (1755) Institutiones calculi differentialis, Saint Petersburg
Fraser C (2020) Review of Shank (2018). J Mod Hist 92(1):197–198
Heath TL (1917) Review of Child (1916). Math Gaz 9(130):131–134
Jahnke HN (2003) A history of analysis. American Mathematical Society. Providence, Rhode Island, USA
Katz VJ (1987) The Calculus of the trigonometric functions. Hist Math 14(4):311–324
Katz VJ (2009) A history of mathematics: an introduction, 3rd edn. Addison-Wesley
Kowalewski G (1914) Die erste Integralrechnung: Eine Auswahl aus Johann Bernoullis Mathematischen Vorlesungen über die Methode der Integrale und anderes, Ostwald's Klassiker der exakten Wissenschaften 194. Engelmann, Leipzig & Berlin
Leibniz GW (1684) Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractas, nec irrationales quantitates moratur, & singulare pro illis calculi genus. Acta Eruditorum:467–473
Leibniz GW (1694) Constructio propria problematis de Curva Isochrona Paracentrica. Acta Eruditorum:364–375
Leibniz GW (1695) Responsio ad nonnullas difficultates, a Dn. Bernardo Niewentiit circa methodum differentialem seu infinitesimalem motas. Acta Eruditorum:310–316
Leibniz GW (AA) (1923–present ongoing) Sämtliche Schriften und Briefe, Gottfried-Wilhelm-Leibniz-Gesellschaft, Academie Verlag, leibnizedition.de
Mahoney MS (1990) Barrow's mathematics: between ancients and moderns. In: Feingold M (ed) Before Newton: the life and times of Isaac Barrow. Cambridge University Press, Cambridge
Manfredi G (1707) De constructione aequationum differentialium primi gradus, Bologna
Nauenberg M (2014) Barrow, Leibniz and the geometrical proof of the fundamental theorem of the Calculus. Ann Sci 71(3):335–354
Newton I (MP) (1967–1981) The mathematical papers of Isaac Newton, 8 vols, edited by D. T. Whiteside. Cambridge University Press, Cambridge, pp 1967–1981
Pólya G (1957) How to solve it, 2nd edn. Princeton University Press, Princeton
Raugh M, Probst S (2019) The Leibniz catenary and approximation of e – an analysis of his unpublished calculations. Hist Math 49:1–19
Simmons GF (1996) Calculus with analytic geometry, 2nd edn. McGraw-Hill, New York, USA
Sonar T (2018) The history of the priority dispute between Newton and Leibniz. Birkhäuser, Basel, Switzerland
Speiser D (2008) Discovering the principles of mechanics 1600–1800, edited by Kim Williams & Sandro Caparrini. Birkhäuser
Wagner J (2001) Barrow's fundamental theorem. Coll Math J 32(1):58–59
Whiteside DT (1961) Patterns of mathematical thought in the later seventeenth century. Arch Hist Exact Sci 1(3):179–388

Evolution of Leibniz's Thought in the Matter of Fictions and Infinitesimals

Monica Ugaglia and Mikhail Katz

Contents

1 Introduction 342
2 The Quest for Actuality 344
2.1 Actuality and the Ultimate Elements of Physics 344
2.2 Syncategorematicity and Potentiality 345
3 Leibniz's Trajectory in Five Phases 349
3.1 An Overview 349
3.2 Five Phases 352
4 Conclusion 381
References 382

Abstract

In this chapter, we offer a reconstruction of the evolution of Leibniz's thought concerning the problem of the infinite divisibility of bodies, the tension between actuality, unassignability, and syncategorematicity, and the closely related question of the possibility of infinitesimal quantities, both in physics and in mathematics.

Some scholars have argued that syncategorematicity is a mature acquisition, to which Leibniz resorts to solve the question of his infinitesimals – namely the idea that infinitesimals are just signs for Archimedean exhaustions, and their unassignability is a nominalist maneuver. On the contrary, we show that syncategorematicity, as a traditional idea of classical scholasticism, is a feature of young Leibniz's thinking, from which he moves away in order to solve the same problem, as he gains mathematical knowledge.

M. Ugaglia
Montecassiano, Italy
e-mail: monica.ugaglia@gmail.com

M. Katz (✉)
Department of Mathematics, Bar Ilan University, Ramat Gan, Israel
e-mail: katzmik@math.biu.ac.il

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_149

We have divided Leibniz's path toward his mature view of infinitesimals into five phases, which are especially significant for reconstructing the entire evolution. In our reconstruction, an important role is played by Leibniz's text *De Quadratura Arithmetica*. Based on this and other texts, we dispute the thesis that fictionality coincides with syncategorematicity, and that unassignability can be bypassed. (In this chapter, we employ "syncategorematic" as a shorthand for "eventually identifiable with a procedure of exhaustion" and, as a consequence, "involving only assignable quantities." We also identify syncategorematicity with potentiality, as suggested by some part of the scholastics, and in Sect. 2.2, we show that the two characterizations are equivalent, provided that "potentiality" is intended in the correct way: namely as in the unending iterative procedures of Greek mathematics.) On the contrary, we maintain that unassignability, as incompatible with the principle of harmony, is the ultimate reason for the fictionality of infinitesimals.

Keywords

Infinite divisibility · Infinitesimals · Infinita terminata · Potential infinity · Syncategoremata · Fictionality · Unassignability · Leibnizian metaphysics

1 Introduction

Leibniz's thinking about infinitely small quantities underwent a complex evolution, both from the mathematical and from the epistemological point of view.

In his maturity, Leibniz was aware of the difficulties that the notion of an infinitely small quantity raises, but his writings suggest that his epistemological awareness has matured as a result of his mathematical proficiency. While the young Leibniz applied almost naively in physics the new mathematical discoveries of his time, of which he had only a superficial knowledge, in his maturity, he develops a more complex view of the relationship between mathematics and physics (and metaphysics), and becomes more cautious both in the application and in the definition of mathematical notions.

In particular, Leibniz eventually comes to master the notion of unassignability and the consequences it involves. This leads him to a cogent treatment of infinitesimals in mathematics – infinitesimals are noncontradictory notions – and to a rejection in physics – they are incompatible with the principle of harmony. In other words, this leads Leibniz to his mature view of infinitesimals as accidental impossibilities, or "well-founded fictions."

We have divided Leibniz's path toward this position into five phases, which can be roughly summarized as follows.

The young Leibniz shares with many contemporary philosophers both a scholastic background and the need of correcting it in the wake of the modern mechanistic philosophy. At this stage, Leibniz's thought does not require infinitely small quantities, for both in mathematics and in physics, he adopts a traditional syncategorematic approach: both extension and matter are continuous, which means that they are

potentially infinitely divisible, and every obtained part is assignable. Take a continuous quantity – for instance, a segment – you can divide it, and then divide it, and then divide it. . . and so on, indefinitely. Since the procedure has no end, no ultimate constituents can be reached, in either mathematics or physics, and since the quantity obtained at every step of the procedure is smaller than the one obtained at the previous step, but is nonetheless finite, no unassignable infinitesimal quantities can be reached, either (phase I: 1668–1669).

Quite early, for metaphysical reasons and in opposition to Descartes, the question of the ultimate constituents of physical bodies becomes a matter of primary importance for Leibniz. However, the solution he offers before his stay in Paris – marking the start of Leibniz's mathematical training and his acquaintance with the contemporary debate – is a very problematic one. At the very beginning (phase II: 1670–1671), Leibniz simply redefines matter as a continuous stuff infinitely divided – and not only divisible – into parts. Both actuality and unity of these parts are due to their *conatus*, or unextended amounts of motion. The geometrical counterpart of this physical situation is a hybrid in which the infinite divisibility of the continuum has to coexist with the actuality of unextended "fat points," which Leibniz identifies with Cavalieri's indivisibles. In this phase, unassignability is accepted, without any further elaboration.

In the first years of the Parisian stay (phase III: 1672–1674), a slightly modified version of this solution is devised, which seems to solve the problem of actuality in physics but still raises difficulties in mathematics.[1] The core idea is to put discreteness before continuity: physical bodies do not come from an infinitely divided continuous stuff but from the aggregation of discrete basic elements. As in phase II, the ultimate elements are actually infinite and "glued" together by their *conatus* of motion. In mathematics too, the situation is similar, with the only difference that now points are extended, although unassignable. Apparently, in this phase, to keep together extended but unassignable quantities and assignable ones is not a problem for Leibniz. Unassignable quantities are not called fictions, nor referred to using cognate verbal forms.

Only towards the end of his stay in Paris, Leibniz becomes aware of the existence of a problem, as the notes he wrote in phase IV (1675–1676) show, as well as his further writings. In particular, Leibniz rejects the possibility that physical quantities are not all comparable to each other: a world containing some portion of matter which has an unassignable ratio to another portion would be less perfect than a world where all portions are comparable.

For this reason, the solution Leibniz offers in this period to the problem of the composition of physical bodies is only apparently similar to the one described in phase III. In both cases, we have an infinitely divided material stuff and a sort of glue that keeps the parts together, but while in phase III the glue, namely the *conatus* of motion, corresponded to physical objects, in phase IV, material parts are held together by metaphysical bonds. In other words, in phase III, a geometrical description was needed, which did contain both the parts of a body, which are assignable, although infinitely divisible, and their unassignable *conatus*. In phase IV, instead, a

[1] See further in note 21.

geometrical description is needed only for the assignable physical parts, while the metaphysical "cement" is no longer constrained.

It is exactly in phase IV that Leibniz starts referring to unassignable quantities as "fictions" and links them to other "imaginary" notions, such as imaginary numbers, that are noncontradictory but do not admit any traditional geometrical representation.

Finally (phase V: 1677– ...) the separation between mathematics, physics, and metaphysics is gradually brought to completion. In particular, mathematics ceases to be the foundation of physics, and geometry ceases to be its natural description: the true essence of physics is metaphysics, and as such it is non-representable. In particular, being extraneous to the category of quantity, metaphysical objects can be thought of as unassignable in a stronger sense and without problematic consequences.

2 The Quest for Actuality

2.1 Actuality and the Ultimate Elements of Physics

In the very first years of his philosophical career, here described as phase I, Leibniz still operates in a scholastic context, where physical objects are continuous and continuity implies potentiality, by definition. But as soon as he begins to develop his own positions, the need to go beyond potentiality, grounding physical objects on something actual, becomes a key feature of Leibniz's philosophy.

The search for actuality, however, is forced to coexist with another crucial feature, namely the infinite divisibility of matter, which Leibniz inherits from scholasticism and will never abandon. The tension between these two seemingly irreconcilable positions, combined with a firm belief in keeping metaphysics out of physics, is what characterizes Leibniz's first moves towards an original theory of physical substance.

In the early 1670s, here described as phases II and III, we witness Leibniz's first unsuccessful attempts to conciliate actuality and infinite divisibility, without resorting to metaphysics. The idea was to locate actuality in some ultimate elements of matter, making it derive not from a vague metaphysical form, but from a much more concrete, and mathematically manageable, (tendency to) motion. In Leibniz's intentions, by virtue of their motion, the ultimate parts of a physical body should be able to stay connected together, as a whole, while maintaining their individuality. But neither the continuous parts of phase II nor the discrete atoms of phase III were up to the task.

At this point, Leibniz shifts his attention from material parts to their connections, namely their tendencies to motion, or *conatus*, whose inherent actuality makes them appear a better candidate to ground physical substances. Moreover, having been defined as unassignable, it seemed natural to interpret the *conatus* as the physical counterpart of mathematical indivisibles, a topic that had captured the attention of Leibniz.

From the writings of this period, however, it is not clear how, and to what extent, the elimination of any material constituent can be finally implemented. Also, the question of the coexistence of unassignable and assignable quantities is left unanswered. Probably unsatisfied by this incomplete picture, starting from the last years of his stay in Paris, here described as phase IV, Leibniz changes his approach: he

dismisses the search for infinitesimal actual elements in physical bodies and looks elsewhere to find the source of the actuality he needs.

More generally, he dismisses the search for any mathematical quantitative causes in physics, and turns to metaphysics, leaving to mathematics the (approximate) description of the phenomenal realm of nature, but not its causal explanation.

In the following subsection, we analyze in more detail the notions of syncategorematicity and potentiality, and their relationship as conceived in the scholastic logic.

2.2 Syncategorematicity and Potentiality

The distinction between categorematic and syncategorematic infinite has a grammatical origin. C*ategoremata* were the parts of speech which possess a definite meaning of their own. *Syncategoremata* were such parts as depend on *categoremata* in order to acquire a definite meaning.[2]

Although in grammar *syncategoremata* and *categoremata* formed two mutually exclusive and collectively exhaustive classes, the notion of *syncategorema* was imported to logic, with a broadened range of application. For medieval logicians, a *syncategorema* is a term that has a signification, but an indefinite one, and can be defined by the presence of other words.

In this context, a *categorema* can be employed in a "syncategorematic" way, and the distinction between categorematic and syncategorematic use of the same term becomes one of the standard ways of solving *sophismata*, namely puzzling sentences, that can be true or false depending on the interpretation of their parts.

The sentence "unlimited are limited" (*infinita sunt finita*) was a paradigmatic example of such a puzzle. The way of solving it by resorting to the opposition categorematic/syncategorematic made the distinction between categorematic and syncategorematic (use of the word) infinite a very common one.

Given the traditional Aristotelian distinction between actual and potential infinite, meant as philosophical notions, it is not surprising that the logico-grammatical level and the philosophical one became conflated, leading to talk about a categorematic and a syncategorematic infinite *tout court*. This is so, especially since the logicians themselves, introducing the grammatical distinction, referred to Aristotle's philosophical discussion of the infinite.

At least, this is the case of Peter of Spain, whose *Summulae Logicales* (Peter of Spain 2014), written in the second quarter of the thirteenth century, became the most influential introduction to logic of the late Middle Ages. Peter deals with the infinite in the last treatise of the book, devoted to distributions (*distributiones*). As we read in chapter 1:

[2]Both the distinction and the term *syncategoremata* appear in the *Institutiones grammaticae*, written at the beginning of sixth century by the Latin grammarian Priscian and well known in the Middle Ages. On syncategoremata in grammar and in logic, see in particular Kretzmann (1982).

> **[P1]** Distribution is the multiplication of a common term produced by a universal sign. When someone says 'every man,' for example, that term 'man' is distributed or confused for whatever is below it by this sign 'every,' and so in this case there is multiplication of a common term. [*Summulae Logicales* XII.1 Transl. Copenhaver][3]

After a short introduction, Peter discusses the most important distributive signs and distinguishes them as *distributive of substance*: every, nothing, whole... and *distributive of accidents* (quality and quantity).

The infinite is the last among the signs distributive of quantity. Peter starts his discussion explaining the diverse meanings of the term "infinite" in philosophy. He quotes Aristotle's *Physics* almost verbatim[4]:

> **[P2]** Next comes the name 'infinite' which is said in five ways. In one way, the infinite is said to be what cannot be crossed-over – as a word is said to be infinite as to sight because it is invisible, since it is not naturally suited to be seen. In another way, the infinite is said to be what is incompletely crossed in that it has not yet been crossed, even though it is naturally suited to be crossed, as when someone is crossing a space and has not yet come to its limit.
>
> In a third way, the infinite is said to be in regard to succession, as an augmentable number is infinite when succeeded by unity or by another number. But in a fourth way it is in regard to division, like the continuous, for everything continuous is divisible without limit. Hence, it is defined by Aristotle in the sixth book of the *Physics* as follows: "the continuum is divisible into what is always divisible."[5] And in a fifth way, the infinite is said in both those senses, by addition and by division, like time. For since time is continuous, it is divisible without limit, and in this way, it is infinite by division, and since one time comes after another time, in that way, by the addition of one time to another, it is infinite by addition. [*Summulae Logicales* XII.36 Transl. Copenhaver, with changes][6]

The first two ways of saying that something is infinite are in fact improper ways of speaking (as Aristotle himself explains), which Peter does not consider. He discusses the proper senses of the term, focusing on their crucial common feature: namely their being defined in an iterative way:

---

[3] "Distributio est multiplicatio termini communis per signum universale facta. Ut cum dicitur omnis homo, iste terminus homo distribuitur sive confunditur pro quolibet suo inferiori per hoc signum omnis, et sic est ibi multiplicatio termini communis."

[4] *Physics* III 4, 204a2–7.

[5] *Physics* VI 2, 232b24–25.

[6] "Sequitur de hoc nomine infinitum, quod dicitur quinque modis. Uno enim modo dicitur infinitum quod non potest pertransiri—ut vox dicitur infinita quoad visum quia est invisibilis, quia non est apta nata videri. Alio autem modo dicitur infinitum quod habet transitum imperfectum eoquod nondum transitum est, tamen est aptum natum transiri, ut cum aliquis transit spatium et nondum venit ad finem eius. Tertio modo dicitur infinitum secundum appositionem, ut numerus augmentabilis est infinitus appositione unitatis vel alterius numeri. Quarto autem modo secundum divisionem, ut continuum, omne enim continuum divisibile est in infinitum. Unde diffinitur ab Aristotile in sexto *Phisicorum* sic: continuum est divisibile in semper divisibilia. Quinto autem modo dicitur infinitum utroque modo, scilicet appositione et divisione, ut tempus. Cum enim sit continuum, est divisibile infinitum, et sic est infinitum divisione, et cum post unum tempus veniat aliud tempus, sic per appositionem temporis unius ad aliud est infinitum appositione."

[P3] And as to these last three significations, the definition goes this way: the infinite is that whose quantity – for items that take quantity – can always get something more, so that if, after one part of a line, another part was taken on, and after that, a third, and the end of it could never be reached, then the line would be called 'infinite'. [*Summulae Logicales* XII.36 Transl. Copenhaver, with minor changes][7]

Here the reference is to Aristotle's definition of potential infinite:

[A] the infinite is in virtue of another and another thing being taken, over and over again; and what is taken is finite, over and over again; but it is a different thing, over and over again. [*Physics* III 6, 206a27–29. Translation ours][8]

Then Peter comes to the dialectical parts of his discussion and explains how infinite can be used both for a common and for a distributive term:

[P4] But the usual claim is that the infinite is sometimes used for a common term, and then this proposition: "infinite are finite" is equipollent with this one "some infinite are finite". Sometimes, however, it is used for a distributive term, and then the former "infinite are finite" is equipollent as to distribution with this one: "the-more than-whatever are finite". And it is confirmed in this way: "the-more than-one are finite; the-more than-two are finite; the-more than-three are finite" and so on for others; "therefore, the-more than-whatever are finite".

And then this is said to produce an 'inter-laddered' distribution (or else interrupted or discontinuous) because this word 'more' in the first proposition stands for two and beyond, in the second for three and beyond, and in this way, it keeps going higher step by step, or like a ladder. And therefore, the phrase 'more than whatever' produces an inter-laddered distribution because my term 'whatever' stands for some things and my term 'more' for others, as it goes higher numerically—as has been said [*Summulae Logicales* XII.37 Transl. Copenhaver, with minor changes][9]

Finally, summarizing the solution of the sophisma, Peter maintains the equivalence between syncategorematic and distributive use of the term "infinite":

[7] "Et quoad has tres significationes ultimas diffinitur sic: infinitum est cuius quantitatem accipientibus semper est sumere aliquid extra, ut si post unam partem linee accipiatur alia, et post illam tertiam, et numquam posset attingi terminus eius, tunc linea diceretur infinita."

[8] ὅλως μὲν γὰρ οὕτως ἔστιν τὸ ἄπειρον, τῷ ἀεὶ ἄλλο καὶ ἄλλο λαμβάνεσθαι, καὶ τὸ λαμβανόμενον μὲν ἀεὶ εἶναι πεπερασμένον, ἀλλ᾽ ἀεί γε ἕτερον καὶ ἕτερον

[9] "Solet autem poni quod infinitum quandoque sumitur pro termino communi, et tunc hec propositio, 'infinita sunt finita', equipollet huic 'aliqua infinita sunt finita' Quandoque autem sumitur pro signo distributive, et tunc ista, 'infinita sunt finita', equipollet huic quoad distributionem, scilicet 'quolibet plura sunt finita'. Et probatur sic: 'Uno plura sunt finita'; 'duobus plura sunt finita'; 'tribus plura sunt finita', et sic de aliis; ergo 'quolibet plura sunt finita'. Et tunc dicitur facere interscalarem distributionem (sive interruptam vel discontinuant) quia hec dictio plura in prima propositione supponit pro duobus et deinceps, et in secunda pro tribus et deinceps, et sic semper gradatim sive scalariter ascendendo. Et ideo hec oratio quolibet plura facit interscalarem distributionem quia pro aliis supponit hoc quod dico quolibet et pro aliis hoc quod dico plura, secundum numerum ascendendo—ut dictum est."

> **[P5]** A distinction made by others is that 'infinite' can be a common term, and in this way the original proposition is false; or else it can be a syncategorematic expression that in itself indicates distribution, as has been said, and in that way, they claim that the original proposition is true. [*Summulae Logicales* XII.38 Transl. Copenhaver, with minor changes][10]

In Peter's exposition, the relation is clear, between the syncategorematic infinite, i.e., the infinite used for a distributive term, and the infinite in regard to addition, or to division, as described in **[P3]**: namely Aristotle's potential infinite, as defined in **[A]**.

But while there is a general agreement concerning the identification of syncategorematicity with distributivity, scholars are divided on the possibility of equating syncategorematic and potential infinite.[11] This depends on the interpretation of Aristotle's definition of potential infinite: once the particular kind of potentiality Aristotle ascribes to the infinite is correctly understood, as in Peter's analysis, the identification with syncategorematicity is straightforward. On the contrary, no similarity can be detected between the two notions if one tries to ascribe to the infinite the "standard" notion of potentiality: namely the one Aristotle usually deploys, and contrasts with actuality, when he speaks of objects, but explicitly prevents from using in the case of the infinite:

> We must not take potentiality here in the same way as that in which, if it is possible for this to be a statue, it actually will be a statue, and suppose that there is an infinite which will be in actuality [*Ph*. III 6, 206a18–20. Translation ours][12]

But then, what is the peculiar meaning Aristotle assigns to the notion of potentiality in the phrase "potential infinite"? The answer is in the definition itself:

[10] "Alii distinguunt eoquod infinitum potest esse terminus communis, et sic prima est falsa; vel potest esse dictio sincathegorematica importans in se distributionem, ut dictum est, et sic ponunt ipsam esse veram."

[11] It is worth noting that scholars who deny any connection between syncategorematicity and potentiality ascribe this position to Peter of Spain, unfortunately without any further specification. Uckelman (2015), p. 2368, for instance, says that: "it is clear that the syncategorematic/categorematic and the potential/actual distinctions are simply not of the same sort [. . .] in agreement with Peter of Spain, who, according to Moore, explicitly rejected any coincidence between the two distinctions (Moore 1990, p. 51)." But Moore (1990) simply states that: "Peter had explicitly denied any coincidence," giving no reference. In fact, a passage from the treatise *De exponibilibus* was quoted with the same intent by Duhem (1913) (VII p. 61. Translated in Duhem 1985, p. 50) who mistakenly attributed it to Peter. The inauthentic *De exponibilibus* is not included in De Rijk's edition of Peter's *Treatises* and maybe for this reason modern commentators who deny the relation syncategorematic/potential cannot give an exact reference for their key passage. In general, on the distinction, and on the identification syncategorematic/potential and categorematic/actual, see, for example, Kretzmann (1982). On the meaning of syncategorematicity in Leibniz, see Antognazza (2015) and Ugaglia (2022). On the iterative reading of potentiality and its connections with Greek proofs by exhaustion, see Ugaglia (2009, 2018).

[12] οὐ δεῖ δὲ τὸ δυνάμει ὂν λαμβάνειν, ὥσπερ εἰ δυνατὸν τοῦτ᾽ ἀνδριάντα εἶναι, ὡς καὶ ἔσται τοῦτ᾽ ἀνδριάς, οὕτω καὶ ἄπειρον ὃ ἔσται ἐνεργείᾳ·

> **[A]** the infinite is in virtue of (**A1**) another and another thing being taken, over and over again (*aei*); and (**A2**) what is taken is finite, over and over again (*aei*); but (**A3**) it is a different thing, over and over again (*aei*). [*Physics* III 6, 206a27–29. Translation ours][13]

As is clear, and as was clear to Peter of Spain, **[A]** is not a conventional Aristotelian definition, but an operational characterization. It does not describe an object but an action – "taking" something – and the iteration of the action, signaled by the adverb *aei*. For each thing we take (*first step*), there is another thing to take (*second step*), and another to take beyond that (*third step*), and another beyond, and so on, over and over again (*aei, next steps*).

Once the potential infinite is understood in this way, it is immediate to interpret it as "a syncategorematic expression that in itself indicates distribution" to use Peter's words.[14] In fact, as he explains in **[P4]**, if the term "infinite" is used in this sense, the sentence "infinite are finite" can be rephrased as: "the-more than-one are finite (*first step*); the-more than-two are finite (*second step*); the-more than-three are finite (*third step*) and so on for others (et sic de aliis, *next steps*)."[15]

3 Leibniz's Trajectory in Five Phases

3.1 An Overview

Leibniz's mature position concerning the ultimate composition of bodies, and the related issue of their division, is rather clear (see Sect. 3.2.5). It is part of an articulated philosophical construction, embracing mathematics, physics, and metaphysics, that can be summarized as follows.

Since a complete separation exists between the ideal realm of mathematics, the phenomenal realm of physics, and the true realm of metaphysics, a completely different solution is offered to the problem in the three contexts.

Metaphysical real entities are indivisible unextended unitary wholes, which do not admit division of any kind. However, these substances, as any other metaphysical notion, are unknowable and inaccessible per se. As unknowable, in particular, and extraneous to the category of quantity, they can be described only through

[13] See note 9 for the Greek text. The condition **A3** has the only aim of discriminating between procedures that are infinite in a proper sense and procedures that are infinite because they are periodic. In the latter case, one comes back to something already taken after a certain number of steps.

[14] See also William of Sherwood's *Syncategoremata*, where the infinite in its syncategorematic significance is explicitly related to induction (O'Donnell 54–55, translated in Kretzmann 42–43).

[15] The sophisma itself "infinite are finite" is nothing but a deliberately equivocal way of restating Aristotle's condition **A2**.

analogical speech; as inaccessible, they can be approached only through their phenomenal "expressions," namely physical bodies.[16]

Physical bodies are extended and composed of extended parts. These parts are actually infinite, but only in this sense: there is no part that is not actually subdivided into others:

> It is perfectly correct to say that there is an infinity of things, i.e., that there are always more than one can specify. But it is easy to demonstrate that there is no infinite number, nor any infinite line or other infinite quantity, if these are taken to be genuine wholes. The Scholastics were taking that view, or should have been doing so, when they allowed a 'syncategorematic' infinite, as they called it, but not a 'categorematic' one. [*Nouveaux essais sur l'entendement humain*, II. XVII. A6.6 157. Transl. Remnant-Bennet][17]

Moreover, the parts of a physical body do not form a continuous magnitude, but only a contiguous one. Here one must note that neither the notion of actual infinity employed in physics – also known in literature as "syncategorematic actual" – nor that of contiguity are mathematical notions.[18]

In fact, in mathematics, Leibniz maintains his traditional genuinely syncategorematic characterization of continuity as potentially infinite divisibility, which means that the parts into which a geometrical continuum can be divided are only potential:

> As for the first point, it follows from the very fact that a mathematical body cannot be analyzed into primary constituents that it is also not real but something mental and designates nothing but the possibility of parts, not something actual. A mathematical line, namely, is in this respect like arithmetical unity; in both cases the parts are only possible and completely indefinite. A line is no more an aggregate of the lines into which it can be cut than unity is the aggregate of the fractions into which it can be split up. [. . .]
>
> But in real things, that is, bodies, the parts are not indefinite – as they are in space, which is a mental thing – but actually specified in a fixed way according to the divisions and subdivisions which nature actually introduces through the varieties of motion. And granted that these divisions proceed to infinity, they are nevertheless the result of fixed primary constituents or real unities, though infinite in number. Accurately speaking, however, matter

[16] The notion of "expression" is repeatedly discussed throughout the correspondence with Arnauld from September 1687 [A2.2. 230–8].

[17] "A proprement parler, il est vrai qu'il y a une infinité de choses, c'est à dire qu'il y en a tousjours plus qu'on n'en peut assigner. Mais il n'y a point de nombre infini ny de ligne ou autre quantité infinie, si on les prend pour des veritables Touts, comme il est aisé de demonstrer. Les écoles ont voulu ou dû dire cela, en admettant un infini syncategorematique, comme elles parlent, et non pas l'infini categorematique."

[18] On the expression "syncategorematic actual" for denoting Leibniz's physical infinite, which is currently employed in Leibnizian studies, see Arthur (2015, 145). It has been defended as an innovative mathematical notion, for instance, in Arthur (2019), criticized on that point in Bussotti (2019). On the mathematical inconsistency of Leibniz's physical notion of infinite, see also Ugaglia (2022).

> is not composed of these constitutive unities but results from them. [*Leibniz to De Volder* June 30 1704; G II, 268. Transl. Loemker][19]

Despite the differences, both in physics and in mathematics, no infinitesimal unassignable quantities can be reached while proceeding with the division: the quantity obtained at any step, in fact, is in a given and finite ratio with the one obtained at the previous step, and with the one obtained at the next one.

But while to postulate the existence of infinitesimal quantities in physics could raise some difficulties,[20] the introduction of infinitesimal quantities in mathematics does not carry such risks and is highly useful. Although they do not follow Euclid V Definition 4 and do not admit a traditional geometrical representation, Leibniz provides rules that make his unassignable infinitesimals mathematically sound, noncontradictory objects.[21]

In other words, while geometrical curves are continuous, namely indefinitely divisible into potential parts, the problems they represent can be solved by using infinitesimal quantities. Moreover, the language of infinitesimals turns out to be useful in speaking, analogically, of some metaphysical objects.

The situation is similar to that with imaginary numbers, which at Leibniz's time had neither a geometrical representation nor a physical counterpart, but were well-defined, noncontradictory, and useful mathematical notions. And in fact, as for

[19] "Quod primum attinet, eo ipso quod corpus mathematicum non potest resolvi in prima constitutiva, id utique non esse reale colligatur, sed mentale quiddam nec aliud designans quam possibilitatem partium, non aliquid actuale. Nempe linea mathematica se habet ut unitas arithmetica, et utrobique partes non sunt nisi possibiles et prorsus indefinitae; et non magis linea est aggregatum linearum in quas secari potest, quam unitas est aggregatum fractionum in quas potest discerpi [. . .] At in realibus, nempe corporibus, partes non sunt indefinitae (ut in spatio, re mentali), sed actu assignatae certo modo, prout natura divisiones et subdivisiones actu secundum motuum varietates instituit, et licet eae divisiones procedant in infinitum, non ideo tamen minus omnia resultant ex certis primis constitutivis seu unitatibus realibus, sed numero infinitis. Accurate autem loquendo materia non componitur ex unitatibus constitutivis, sed ex iis resultat."

[20] Even if the existence of infinitesimal quantities in physics is not an absolute impossibility – namely it does not involve any contradiction – it is contrary to the harmony of things. In fact, a world containing "some portion of matter which has an unassignable ratio to another portion" would be less perfect than a world where all portions are comparable. On this point, it is interesting to read the 1698 correspondence between Leibniz and Bernoulli, vigorous proponent of the existence of infinitesimal quantities in physics. In particular, in his letter of July 5, 1698 (GM III 503–504), Bernoulli asks why only a size (*gradus*) of infinity is allowed, namely the one proportionate to our intellect, and not the infinite others (he means here infinitesimals and higher degree of infinitesimals dx, ddx. . .). In his answer of July 12, Leibniz replies that: "from the actual division it follows that in every part of matter, there is a world – so to speak – made of innumerable creatures; but we ask if there is some portion of matter which has an unassignable ratio to another portion, or if there is a line bounded (terminata) on both ends, having an infinite or an infinitely small ratio, to another line. We assume that in our Calculus, fruitfully, but it does not follow that this holds in nature. But this is the subject of another research" (GM III 516). Accidental impossibility of Leibniz's infinitesimals has been argued in Esquisabel and Raffo Quintana (2021).

[21] The rules are discussed in the *Nova Methodus pro maximis et minimis*, published in 1684 in the *Acta Eruditorum* [GM V pp. 220–226].

imaginary numbers, in his mature philosophy, Leibniz refers to infinitesimals using the language of imagination – to feign, fictions – and contrasts them with more traditional mathematical notions, referred to as ideal concepts, which are the image, although idealized, of physical situations.[22]

3.2 Five Phases

As mentioned above, Leibniz reached this final position through a complex path, which we will trace in the present section. We have focused our attention on five phases, which are especially significant for reconstructing the entire evolution. Except for phase I, where Leibniz's physical and mathematical analysis do coincide, we have treated them separately, as far as possible.

3.2.1 Phase I (1668–1669)

While in his mature philosophy Leibniz will analyze physical bodies in terms of individual metaphysical substances, he was initially confident that one should banish from physics any metaphysical explanations and replace them by purely mechanical ones. In particular, Leibniz attempted to reduce the Aristotelian triad *matter-form-change*,[23] which he considered too abstract, to the "mechanical" triad *extension-shape-motion*. Wrote Leibniz:

> The one question is whether Aristotle's abstract theories of matter, form, and change can be explained by magnitude, shape, and motion. This the Scholastics deny and the philosophical reformers affirm. The latter opinion seems to me to be not only the more true but also the more consistent with Aristotle. [*Leibniz to Thomasius*, April 1669. A2.1[2]. 25–26. Transl. Loemker with minor changes][24]

A physical body is the compound of a passive element, or matter, which Leibniz identifies with extension,[25] and an active one, or form, which he identifies with shape. Both extension and shape are dealt with in a rather simplistic way, in accordance with the fact that in this phase, Leibniz's knowledge of mathematics, which essentially means Greek geometry, is rather basic, and for the most part acquired by reading philosophical works. In particular, extension is continuous,

[22] On Leibniz's ideal notions, see Sherry and Katz (2014) and Katz and Sherry (2013).

[23] The noun *change* is our rendering of Aristotle's μεταβολή, which includes motion (κίνησις, namely local change) as well as qualitative and quantitative changes.

[24] "Hoc unum in quaestione est, an quae Aristoteles de materia, forma, et mutatione abstracte disputaverit, ea explicanda sint per magnitudinem, figuram, et motum. Id Scholastici negant, Reformatores affirmant. Reformatorum sententia mihi non solum verior, sed et Aristoteli magis consentanea videtur." Here and throughout the chapter, we have rendered "figura" with "shape," instead of Loemker's "figure."

[25] In physical bodies, extension is coupled with impenetrability, which sometimes Leibniz prefers to call antitypy (ἀντιτυπία): "To be extended, however, is nothing else but to be in space, and antitypy is the impossibility of being in the same space with another thing" [*Leibniz to Thomasius*, April 1669. A2.1[2]. 36. Transl. Loemker].

and continuity means for Leibniz potentially infinite divisibility, that is syncategorematicity, as it was for Aristotle.[26] The terminology itself of unassignability and infinitesimals is absent in this phase: "infinitely small" just means "potentially infinitely divisible."[27]

> Matter has quantity too, though it is indefinite, or interminate as the Averroists call it. For being continuous, it is not cut into parts and therefore does not actually have boundaries. [*Leibniz to Thomasius*, April 1669. A2.1[2]. 26–27. Transl. Loemker][28]
>
> Space is a primary extended being or a mathematical body, which contains nothing but three dimensions and is the universal locus of all things. Matter is a secondary extended being, or that which has, in addition to extension or mathematical body, also a physical body, that is, resistance, antitypy, solidity, the property of filling space, impenetrability[29] [*Leibniz to Thomasius*, April 1669. A2.1[2]. 34. Transl. Loemker with minor changes][30]

Motion intervenes in the generation of physical bodies. In particular, it is responsible for their shape:

> Form (in the technical sense) as educed from the potentiality of matter, is nothing more than this particular shape of the whole as generated from this particular motion of matter, and from this particular situation of parts: e.g., the form of a square ⊡ *c* is educed from the union of two triangles, Δ *a* and ∇*b*, and from the potentiality of their matter to be in reciprocal contact by motion. [*Leibniz to Thomasius*, October 1668. A2.1[2].17][31]

[26] Aristotle characterizes the continuum as "infinitely divisible" (*Physics* III 1, 200b18; I 2, 185b10–11; VI 6, 237a33; VI 8, 239a22) or "divisible into what is always (*aei*) divisible" (*Physics* VI 2, 232b24–25; see IV 12, 220a30; VI 1, 231b15–16; VI 6, 237b21; VIII 5, 257a33–34; *On the heavens* I 1, 268a6–7). Infinite divisibility is an essential property of the *continuum*, which directly follows from Aristotle's definition of continuity, as a relational notion: two parts of a whole are said to be continuous with one another when they are in contact and their ends therefore become one (συνεχῆ μὲν ὦντὰ ἔσχατα ἕν *Physics* VI 1, 231a21–22 "we defined continuous things whose ends are one"); see also *Physics* V 4, 228a29–30; V 3, 227a11–12; *Cat.* 6, 4b20–5a14. We discussed the continuum as always (*aei*) divisible in Sect. 2.2.

[27] Cf. the analysis of the infinitely small in *Physics* III.6, where Aristotle argues against the existence of indivisible lines (ἀτόμους γραμμάς 206a17–18).

[28] "Quantitatem quoque habet materia, sed interminatam, ut vocant Averroistae, seu indefinitam, dum enim continua est, in partes secta non est, ergo nec termini in ea actu dantur."

[29] Resistance, antitypy, solidity, the property of filling space, and impenetrability are synonyms and refer to the impossibility, for a physical body, of being in the same space with another thing, as Leibniz explains a few lines later.

[30] "Spatium est Ens primo-extensum, seu corpus mathematicum, quod scilicet nihil aliud continet quam tres dimensiones, estque locus ille universalis omnium rerum. Materia est ens secundo-extensum, seu quod praeter extensionem vel corpus mathematicum habet et corpus physicum, id est, resistentiam, ἀντιτυπίαν, crassitiem, spatii-repletivitatem, impenetrabilitatem."

[31] "Formam enim educi ex potentia materiae, nihil aliud est, quam ex hoc materiae motu, ex hoc partium situ, hanc totius figuram oriri: v.g. ex unione duorum triangulorum, Δ *a* et ∇*b*, et ex potentia materiae eorum ad contactum mutuum per motum, educitur forma quadrati ⊡ *c.*" In this case, the resulting square does not keep trace of the way in which it has been created. The same example of the square arising from two triangles appears in a 1676 note – see footnote 71 – but here Leibniz resorts to metaphysics in order to obtain a form which keeps trace of its constituents.

The world is a plenum, made of continuous bodies, actually divided from each other:

> This makes it clear that, if mass were created discontinuous or separated by emptiness in the beginning, there would at once be certain concrete forms of matter. But, if it is continuous in the beginning, forms must necessarily arise through motion. For division comes from motion, the bounding of parts comes from division, their shapes come from this bounding, and forms from shapes; therefore, forms come from motion [. . .] From this it is clear that every arrangement into a form is motion, and the vexatus problem of the origin of forms is answered.[32] [*Leibniz to Thomasius*, April 1669 A2.1[2].26–27. Transl. Loemker][33]

Phase I in Summary

Physics: physical objects are continuous; continuous = in(de)finitely divisible; parts are only potential and finite. Both the approach and the terminology employed are syncategorematic.

Mathematics: mathematical objects are continuous; continuous = in(de)finitely divisible; parts are only potential and finite. Both the approach and the terminology employed are syncategorematic.

3.2.2 Phase II (1670–1671)

Continuity as conceived by Aristotle, and as employed by Leibniz in phase I, means potentiality of the parts, and potentiality means indefiniteness. However, indefiniteness, in Leibniz's opinion, has more to do with Descartes' thoughts that with reality; see item (2) below. Phase II represents Leibniz's first attempt to overcome this difficulty without abandoning his initial purpose of leaving aside any metaphysics. However, Leibniz's poor knowledge of mathematics will prove to be a hindrance to effective implementation of the project.[34]

Phase II in Physics

The project was to give some sort of individuality to the parts of a physical body, still treating them as continuous. For doing so, Leibniz broadens the notion of continuity

[32] As Leibniz himself wrote some years later: "Tempus erat quo credebam, omnia Motuum Phaenomena principiis pure Geometricis explicari posse, nulli Metaphysicis propositionibus assumptis" [There was a time when I believed that all the phenomena of motion could be explained on purely geometrical principles, assuming no metaphysical propositions. *Principia mechanica ex metaphysicis dependere* 1678/81 A6.4.1976, Transl. Ariew-Garber].

[33] "Ex his patet siquidem ab initio massa discontinua seu vacuitatibus interrupta creata sit, formas aliquas statim materiae concreatas esse; sin vero ab initio continua est, necesse est, ut formae oriantur per motum (nam de annihilatione certarum partium ad vacuitates in materia procurandas, quia supra naturam est, non loquor) quia a motu divisio, a divisione termini partium, a terminis partium figurae earum, a figura formae, ergo a motu formae. Ex quo patet, omnem dispositionem ad formam esse motum, patet quoque solutio vexatae de origine formarum controversiae."

[34] On Leibniz's knowledge of mathematics before his stay in Paris, see, for instance, Hofmann (1974), pp. 1–11.

and affirms – without any concern for the mathematical consistency of the statement – that a continuous body is not merely infinitely divisible, but it is infinitely divided into parts, and these parts are actual:

> (1) *There are actually parts in the continuum*
> (2) *these are actually infinite*, for Descartes's indefinite is not in the thing but in the thinker [*Theoria motus abstracti* 1671, A6.2. 264. Transl. Arthur][35]

Being actual, the parts of a physical body are organized in a precise way. Therefore, in this phase, form does not depend only on the external shape of the body, as in phase I, but also on the disposition of its internal parts:

> For Aristotle too, substantial form (with the exception of mind) is not an absolute entity, but only a *logos*, a ratio, a proportion, a number, a structure of internal parts. [*Hypothesis Phisica Nova* 1671. A6.2. 247][36]

But how can these parts stay together, forming a continuous body? To solve the problem, Leibniz resorts to motion: being in motion, parts experience an endeavor (*conatus*)[37] to penetrate each other, and this causes their extremities to become one. But the coincidence of the extremities of the parts is the condition Aristotle requires in order to have a continuous, unitary whole (see note 27):

> I should think that the *conatus* of the parts toward each other, or the motion through which they press upon each other, would itself suffice to explain the cohesion of bodies. For bodies which press upon each other are in a *conatus* to penetrate each other. The *conatus* is the

[35] "Dantur actu partes in continuo [. . .] eaeque infinitae actu, indefinitum enim Cartesii non in re est, sed in cogitante." We find the same assumption in Leibniz's *Hypothesis Phisica Nova*, dated shortly before and also known as *Theory of concrete motion*: "Quae res ob divisibilitatem cujuslibet continui in partes quantumvis parvas in infinitum non est difficilis explicatu" [*Hypothesis Phisica Nova* 1671. A6.2 224]. Here physical bodies are thought as made of parts, or *bullae* (bubbles), in a more or less solid state, depending on the circumstances: "Principio autem ex fluidi aestuatione et fusione per lucem seu calorem ortae sunt bullae innumerabiles ac magnitudine crassitieque variantes. Nam quoties subtilia perrumpere per densa conantur, et est quod obsistat, formantur densa in cavas quasdam bullas, motumque partium internum, ac proinde consistentiam seu cohaesionem (per nostram *de motu Theoriam*, theor. 17) nanciscuntur." [*Hypothesis Phisica Nova* 1671. A6.2 226].

[36] "Certe formas substantiales (demta mente) etiam Aristoteli non esse ens absolutum, sed tantum λόγον, rationem, proportionem, ἀριθμόν, structuram partium intimam."

[37] Leibniz borrows the notion of *conatus* from Hobbes' endeavor, which has a very similar definition, and a very similar role in physics. However, while Leibniz's *conatus* are unextended, Hobbes' endeavors are finite quantities, although negligible, like his points: "I define endeavour to be Motion made in less Space and Time then can be given; that is, less than can be determined or assigned by Exposition or Number; that is, Motion made through the length of a Point, and in an Instant or Point of Time" [*De Corpore* 3.15.2]. For Hobbes, a point is not "that which has no quantity" but "that whose quantity is not at all considered" [ibid.]. On the role of Hobbes as the mediator between the young Leibniz and mathematics, and particularly Cavalieri's method, see, for instance, Hofmann (1974).

> beginning (*initio*); the penetration is the union. But when bodies begin to unite, their beginnings or extremities (*termini*), are one. Bodies whose extremities are one, or τὰ ἔσχατα ἕν, are according to Aristotle's definition not only contiguous but continuous, and truly one body, movable in one motion [*Leibniz to Hobbes*, July 1670. A2.1[2]. 92. Transl. Loemker with minor changes][38]
>
> If one body strives against another, this is the beginning of mutual penetration or union or that the boundaries of the two are one, as Aristotle defines a continuum [*Leibniz to Arnauld*, November 1671. A2.1[2]. 279. Transl. Loemker][39]

In his more applied writings, Leibniz describes this situation in concrete terms. The behavior of the parts of physical bodies, which he calls "bubbles," is explained in analogy with what happens in a turbulent fluid: as the bubbles which are continuously produced and destroyed in a fluid, also the parts of a body are elastic and constantly move, breaking and recombining with each other.

Phase II in Mathematics

As in phase I, Leibniz does not draw a clear distinction between physics and its mathematical description, but while in phase I continuity as potentially infinite divisibility had a clear meaning both in physics and in mathematics – a body is made of continuous parts in the same way as a segment is made of segments – the composition of bodies as conceived by Leibniz in phase II is more difficult to translate in mathematical terms.

The main problem is the mathematical *status* of the actually infinite parts and their relationship with unassignable endeavors.

Later in the *Theory of abstract motion*, we learn that Leibniz rejects *minima* (i.e., points as constituent parts of the continuum in the modern sense) and *maxima*, where a maximum is an unbounded infinite line (the kind we have in modern Euclidean geometry). The reason for both is a violation of the part-whole principle that would ensue via Galileo's paradox:

> (3) *There is no minimum in space or body*, that is, nothing which has no magnitude or part. For such a thing has no situation, since whatever is situated somewhere can be touched by several things simultaneously that are not touching each other, and would thus have several faces; nor can a minimum be supposed without it following that the whole has as many minima as the part, which implies a contradiction. [*Theoria motus abstracti* 1671, A6.2. 264. Transl. Arthur][40]

[38] "Ego crediderim ad cohaesionem corporum efficiendam sufficere partium conatum ad se invicem, seu motum quo una aliam premit. Quia quae se premunt sunt in conatu penetrationis. Conatus est initium, penetratio unio. Sunt ergo in initio unionis. Quae autem sunt in initio unionis, eorum initia vel termini sunt unum. Quorum Termini sunt unum seu τὰἔσχαταἕν, ea etiam Aristotele definitore non jam contigua tantum, sed continua sunt, et vere unum corpus, uno motu mobile."

[39] "Si corpus conetur in corpus, esse ambo in initio penetrationis seu unionis, seu extrema eorum unum esse ut continuum definit Aristoteles [ὦν] τὰ ἔσχατα ἕν."

[40] "Nullum est minimum in spatio aut corpore, seu cujus magnitudo vel pars sit nulla: talis enim rei nec situs ullus est, cum quicquid alicubi situm est, simul a pluribus se non tangentibus tangi possit, ac proinde plures habeat facies; sed nec poni Minimum potest, quin sequatur tot esse totius, quot partis minima, quod implicat."

On the other hand, he acknowledges the existence of indivisibles – namely the mathematical counterpart of endeavors – and argues as follows:

> (4) *There are indivisibles or unextended things*, otherwise neither the beginning nor the end of a motion or of a body is intelligible. This is the demonstration: any space, body, motion, and time has a beginning and an end. Let that whose beginning is sought be represented by the line *ab*, whose midpoint is *c*, and let the midpoint of *ac* be *d*, that of *ad* be *e*, and so on. Let the beginning be sought to the left, on *a*'s side. I say that *ac* is not the beginning, since *dc* can be taken away from it without destroying the beginning; nor is *ad*, since ed can be taken away, and so on. Therefore, nothing is a beginning from which something on the right can be taken away. But that from which nothing having extension can be taken away is unextended. Therefore, the beginning of a body, space, motion, or time (namely, a point, an endeavour, or an instant) is either nothing, which is absurd, or is unextended, which was to be demonstrated [*Theoria motus abstracti* A6.2. 264. Transl. Arthur with minor changes][41]

Endeavors are unextended amounts of motion, and indivisibles, or "fat points" (our term) are unextended portions of bounded lines.

Moreover, unlike Euclidean or Aristotelian points, which are "unextended beginnings" of lines but do not enter in their composition, Leibnizian "fat points" do make lines (and "fat lines" do make surfaces. . .).

As a consequence, even if they are smaller than any assignable quantity, "fat points" can be compared with each other. In Leibniz's words, they are different in "magnitude": bigger in a bigger curve, smaller in a smaller one:

> The cone is made by as many circles parallel to the basis as points of the height [*Nugae pueriles* 1670. A7.4. 57–58][42]
>
> Whence the unassignable arc of a bigger circle is greater than that of a smaller one [*Theoria motus abstracti* A6.2. 267. Transl. Arthur][43]
>
> An *arc* smaller than any that can be given is still greater than its chord, although this is also smaller than can be expressed, i.e., consists in a point. But that being so, you will say, an *infinitangular polygon* will not be equal to a circle: I reply, it is not of an equal magnitude, even if it be of an equal extension: for the difference is smaller than can be expressed by any number [*Theoria motus abstracti* A6.2. 267. Transl. Arthur][44]

[41] "Dantur indivisibilia seu inextensa, alioquin nec initium nec finis motus cor porisve intelligi potest. Demonstratio haec est: datur initium finisque spatii, corporis, motus, temporis alicujus: esto illud cujus initium quaeritur, expositum linea *ab,* cujus punctum medium *c,* et medium inter a et c sit *d,* et inter *a* et *d* sit *e,* et ita porro: quaeratur initium sinistrorsum, in latere a. Ajo ac non esse initium quia ei adimi potest *dc* salvo initio; nec *ad,* quia *ed* adimi potest, et ita porro; nihil ergo initium est, cui aliquid dextrorsum adimi potest. Cui nihil extensionis adimi potest, inextensum est; initium ergo corporis, spatii, motus, temporis (punctum nimirum, conatus, instans) aut nullum, quod absurdum, aut inextensum est, quod erat demonstrandum." Note that from the same premises, and with a similar proof, Aristotle reached the opposite conclusion: a motion has neither a beginning nor an end (Physics VI.5; cf. VI 1 where Aristotle's proves the impossibility of indivisibles).

[42] "Constabit conus ex tot circulis basis parallelis quot sunt puncta altitudinis."

[43] "Unde arcus inassignabilis circuli majoris major est quam minoris"

[44] "Arcus minor quam qui dari potest, utique chorda sua major est, quanquam haec quoque sit minor quam quae exponi potest, seu consistat in puncto. At ita, inquies, Polygonum infinitangulum non erit circulo aequale: respondeo, non esse aequalis magnitudinis, etsi sit aequalis extensionis: differentia enim minor est quam ut ullo numero exprimi possit"

> A point can be bigger than another, but in a smaller ratio (*in ratione minore*), than the one that can be exhibited, or infinite with respect to any sensible quantity [*Leibniz to Arnauld*, November 1671. A2.1[2]. 278][45]

The main problem with this picture is the necessity of distinguishing between "having extension" and "having magnitude," a problem that Leibniz thinks he has solved by employing the locution "having magnitude" as synonymous with "having parts," where parts are not necessarily "external to each other." In fact, Leibniz says, parts can be indistant, and this is the basis of Cavalieri's method too:

> (5) *point is not that which has no part*, nor that whose part is not considered; but that which *has no extension*, i.e., whose parts are indistant, whose magnitude is inconsiderable, unassignable, is smaller than can be expressed by a ratio to another sensible magnitude unless the ratio is infinite, smaller than any ratio that can be given. But this is the basis of the *Cavalierian Method*, whereby its truth is evidently demonstrated, inasmuch as one considers certain rudiments, so to speak, or beginnings, of lines and figures smaller than any that can be given. [*Theoria motus abstracti* A6.2. 265. Transl. Arthur][46]
>
> The nature of points is undeniably admirable: indeed, although a point is not divisible into parts external to each other, however it is divisible into parts which before were not external to each other, or into parts that before were penetrating each other. [*Leibniz to Oldenbourgh*, September 1670. A2.1[2]. 103][47]

Summarizing, Leibniz tries to reconcile the actuality of the infinite parts into which physical bodies are divided – and not only divisible – and the potentiality of matter. As material parts, in fact, their magnitude can decrease ad infinitum (no *minima*).[48]

But the conflation between actuality and potentiality is not the only flaw in Leibniz's construction. Furthermore, the physical notion of *conatus*, that is of an endeavor of motion, creates problems when one tries to translate it into mathematics.

Two physical endeavors can have the same extension (in fact, they have no extension) but a different "intension": for instance, when one begins a faster motion

[45] "Esse punctum puncto majus, sed in ratione minore, quam quae exponi potest, seu ad sensibilem quamcumque infinita."

[46] "Punctum non est, cujus pars nulla est, nec cujus pars non consideratur; sed cujus extensio nulla est, seu cujus partes sunt indistantes, cujus magnitudo est inconsiderabilis, inassignabilis, minor quam quae ratione, nisi infinita, ad aliam sensibilem exponi possit, minor quam quae dari potest: atque hoc est fundamentum Methodi Cavalierianae, quo ejus veritas evidenter demonstratur, ut cogitentur quaedam ut sic dicam rudimenta seu initia linearum figurarumque qualibet dabili minora"

[47] "Admirabilis profecto est natura punctorum; quanquam enim punctum non sit divisibile in partes positas extra partes, est tamen divisibile in partes antea non positas extra partes, seu in partes antea se penetrantes"

[48] A different interpretation has been propounded in Arthur (2009, pp. 17–18), which identifies actually infinite parts with points: in this case, the whole physical construction could be traced back to Cavalieri's theory, and Leibniz's construction would appear more consistent. In principle, this would be an appealing interpretation, the only problem being that there is no basis for it in Leibniz's text. In particular, he never identifies the physical notions of part and point, and repeatedly refers to points as extremities of parts.

than the other. But what does it mean that two unextended "fat points" do have a different "magnitude"? Leibniz's explanation in term of indistant inner parts, which can penetrate each other, does not have a clear meaning in mathematics. However, this analysis is never elaborated by Leibniz, who merely claims that with the identification of his "fat points" with indivisibles he has demonstrated Cavalieri's method, too.[49]

Phase II in Summary

Physics: physical objects are actually infinitely divided into continuous material parts (no *minima*) whose extremities coincide. Extremities are held together by unextended endeavors of motion.

Mathematics: material extension is represented by a continuum, as in phase I. Endeavors of motion are represented by unextended "fat points," unassignable and indivisible, which have parts but parts that are indistant.

3.2.3 Phase III (1672–1674)

Phase III, which coincides with the first years of Leibniz's stay in Paris, marks the beginning of his mathematical career. It is in Paris that Leibniz undertakes serious training in mathematics and enters in relation with Huygens, and then with other leading mathematicians, who will become his correspondents.

In phase II, Leibniz's knowledge of the contemporary methods for solving the problems related to infinity was rather poor. He realized the importance of Cavalieri's indivisibles, and more generally of unassignable infinitely small quantities, but he lacked the technical tools necessary in order to transform a philosophical intuition into a mathematical investigation.

The idea of a point with indistant parts, for instance, while appealing and maybe effective in a dialectical argument, was very hard to disambiguate and to translate into a sound mathematical notion. Analogously, the coexistence of actuality and potentiality implied in Leibniz's analysis of physical bodies would lead to serious problems, if mathematically analyzed.

But in phase II, Leibniz did not seem aware of the existence of such difficulties, which become clear only in subsequent years, as his mathematical knowledge grew. For this reason, even if in terms of effective results phase III is merely a transitional phase, it is a crucial step in our analysis. Even though Leibniz did not publish anything, the large amount of handwritten notes he left, often inconclusive and vitiated by inconsistencies, show nevertheless the awareness of the impasse which phase II was leading to, and the attempt to avoid it.

[49] On the ambiguous relation between indivisibles and motion, see Cavalieri's reaction to Guldin's critique (*Exercitationes geometricae sex*, Exercitatio III, cap. VII). Cf. Cavalieri (1966) tr. Radice 808–810.

Phase III in Physics

The first novelty, which Leibniz introduces in physics in order to keep actuality and potentiality separate, consists in reversing the arrow of the analysis: if, by dividing a continuum, one can only reach potentially divisible parts, the only way to obtain a matter made of genuine actual ultimate parts is to start from these parts.

Both in phases I and II, the basic stuff – from which physical bodies was derived – was a continuous entity: once continuity was broken, in virtue of motions, parts were created, which entered into the composition of physical bodies. In phase III, by contrast, the ultimate element is a plurality of basic material entities, discrete from the outset, spherical, which aggregate in order to give rise to physical bodies. Leibniz calls them "atoms" or *bullae*, or *terrellae*:

> Sensible bodies are aggregates of countless (innumerabilium) terrellas [*Propositiones quaedam physicae*, 1672. A6.3.32][50]

By introducing discrete elements, which have a proper individuality and self-sufficiency, Leibniz solves any ambiguity concerning their *status*: the ultimate parts are actual and indivisible.

The problem is that these atoms are meant to be part of bodies which are continuous: how is this possible? Leibniz solves the problem in one direction, showing how a continuous stuff can be obtained starting from discrete elements. As in phase II, in fact, he resorts to motion for welding together the material atoms which form a body: due to their endeavors to motion, the discontinuities between different atoms disappear:

> If a body is forced (*conetur*) into the place of another, these two bodies are continuous [*De consistentia corporum* 1672/3. A6.3.95][51]

Moreover, due to their smallness, even the empty spaces which exist between the spherical atoms become negligible and do not interrupt continuity:

> It is necessary that matter is uninterrupted, and it completely fills the space, or at least it leaves some unnoticeable (*insensibiles*) empty islands, without any breaking of continuity. [*Demonstratio substantiarum incorporearum*, 1672. A6.3. 78][52]

[50] "Corpora sensibilia sunt aggregata innumerabilium terrellarum." A *terrella* is a little spherical loadstone, as described in Gilbert, *De Magnete* 1600.

[51] "Si corpus unum conetur in alterius locum, ea duo corpora continua sunt."

[52] Leibniz will come back on the status of *vacua intermixta*, which he calls also *vacua metaphysica* and their incomparable magnitude some years later, in phase IV: "[. . .] because spheres, however many they may be, do not fill a whole. However, all emptinesses collected into one have no greater ratio to any assignable space than the angle of contact has to a straight line" [*De summa rerum* 1675/6. A6.3. 526. Transl. Parkinson]. "A metaphysical vacuum is an empty place, no matter how small, provided that it is genuine and real. A physical plenum is consistent with a metaphysical vacuum which is unassignable" [*De summa rerum* 1675/6. A6.3. 473. Transl. Parkinson]. Note that in this period and in this context, the term "metaphysical" still retains a negative connotation, keeping track of the "empty words" imputed to scholasticism.

But once physical bodies are continuous, they must be also infinitely divisible:

> This sounds paradoxical – I must admit – but is no less necessary than the actual division of the continuum, in infinite parts (which follow from this exact proof, consistently with Cartesian teachings) and other marvels of the internal nature [*Propositiones quaedam physicae*, 1672. A6.3.57][53]

And then the question arises, how can individuality and infinite divisibility coexist?

Reading Leibniz's notes, it seems that the only possible answer is that they cannot, and that the only way to obtain a consistent description of physical bodies is to eliminate any (atom of) matter:

> From this we understand that matter and motion, or *conatus*, are the same [. . .] And getting back to our issue, it is clear from these facts that matter does not consist of extension, but of motion, and a lot of matter is nothing more than a lot of motion. [*Propositiones quaedam physicae*, 1672. A6.3 56][54]
>
> Hence it follows that there is no matter in body distinct from motion, since it would necessarily contain indivisibles. [*De minimo et maximo* 1672/3. A6.3. 100][55]

In other words, motion can contain indivisibles, matter cannot. And here two difficulties arise: is a motion made of unassignable indivisibles continuous? And if it is, in what sense? Moreover, if bodies are nothing but motion, what happened to the atoms of matter? Leibniz's writings of this period do not contain an answer, and in fact, even if the starting point was different, the physical theory of continuity Leibniz builds in phase III ends up with the same inconsistencies as the one he built in phase II.

Phase III in Mathematics

In mathematics, the situation is less problematic, although some doubt remains concerning the composition of the continuum. Apparently, Leibniz's mathematical continua can be read as entirely composed of "fat points," since now these points are no longer seen as unextended, but as extended infinitely small lines. As their ancestors, these new points are unassignable, i.e., infinitely smaller than any assignable line:

> Continuous space is made of parts smaller than anything determinable by us [*Demonstratio substantiarum incorporearum*, 1672. A6.3.81]

---

[53] "Res est fateor paradoxa sed non minus necessaria, quam divisio continui actualis, in partes infinitas (quae ex hac ipsa demonstratione confessione Cartesii sequitur) aliaque naturae interioris admiranda"

[54] "Unde intelligi potest, materiam et motum vel conatum idem esse [. . .]. Patet ergo ex his, ut viam redeam, materiam non consistere in extensione, sed in motu, et nihil aliud esse mulatm materiam, quam multum motum."

[55] "Hinc sequitur nullam esse in corpore materiam a motu distinctam, ea enim necessario indivisibilia contineret."

> In a continuum there are some infinitely small things, or infinitely smaller than any perceptible one [*De minimo et maximo* 1672/3. A6.3.98].
>
> A point can be infinitely smaller than another [*De minimo et maximo* 1672/3. A6.3.99][56]

In this period "infinitely small," "unassignable" and "point" are employed as synonyms:

> In short: each of these scalene triangles is a point, because their base and height are infinitely small lines. For this reason, all together, arranged on a single straight line obtained from the extended evolute (in fact, the triangles are as many as the points of the chord) do not make but a line [*Varia de cycloeide* 1673 A7.4 p. 74][57]

Here indivisibles are defined as synonymous with infinitely small:

> Indivisibles are to be defined as infinitely small, or as having infinite ratio, or difference, with a sensible quantity [A7.4 p. 265][58]

During these years, the term "infinitesimal" makes its appearance, but it is always employed in the phrase "infinitesimal (part) of" and denotes a particular case of infinitely small line: the infinitesimal part of a quantity is one of the infinitely many equal parts into which the quantity has been divided[59]:

> We suppose that the abscissa AE is divided into equal infinitely many equal parts as EF, FG [. . .] And the straight line EF, or FG will be infinitely small, and each one will be an infinitesimal of the abscissa AE [*Methodus nova investigandi tangentes linearum curvarum*. . . 1673 A7.4 p. 657][60]
>
> And this is the proof: suppose that MNE is the characteristic triangle of the figure, that is the one whose height MN is an infinitesimal of the height of the figure [*Trigonometria inassignabilium* 1673 A7.4 p. 497][61]

[56] "Spatium Continuum componitur ex partibus qualibet a nobis determinabili minoribus." "Sunt aliqua in continuo infinite parva, seu infinities minora, quovis sensibili dato." "Punctum unum alio potest esse infinities minus."

[57] "Idem brevius: quodlibet ex his triangulis scalenis punctum est, cum eius basis et altitudo sint lineae infinite parvae. Ergo omnia simul, in unam rectam ex evoluta extensa conflatam disposita (tot enim sunt scalena, quot in ea chordae, seu puncta) non nisi lineam facient." Note that not only infinitely small lines are called points but also infinitely small figures (cf. in particular *Plagula* 16_4 A7.4 pp. 304–331).

[58] "*Indivisibilia* definienda sunt infinite parva, seu quorum ratio ad quantitatem sensibilem (vel differentia) infinita est."

[59] It is interesting to note that a similar position, namely the idea that one can reach infinitesimals at an infinitieth stage of getting smaller and smaller, will be maintained by Bernoulli, and repeatedly criticized by the mature Leibniz, in their 1698 correspondence (GM III 529–554, see in particular pp. 539, 541 and 551).

[60] "Intelligatur abscissa AE dividi in partes aequales infinitas, quales sunt EF. FG. [. . .] Eritque recta EF. vel GH. infinite parva, eademque erit infinitesima rectae AE. abscissae."

[61] "Demonstratio haec est: intelligatur triangulum figurae characteristicum esse MNE. cuius scilicet altitudo MN est infinitesima altitudinis figurae." Cf. A7.4 p. 432; 493; 494; 497; 663.

Infinitesimals are often employed as "unities":

> And we suppose EE infinitesimal, that is the unity for the construction = 1 [*Varia circa functiones*... 1673 A7.4 p. 755][62]
>
> I'll grant you that, but I maintain that this 1 must be denoted in this way 1r. In fact it is not 1, or a certain number, but is something indivisible, which in this construction works as unit, that is to say the infinitesimal part of a certain line, divided in thought into infinite equal parts [*De paraboloeidum et hyperboloeidum quadratura II* 1673 A7.4 p. 607][63]

As in phase II, the notion of unassignable quantities does not seem particularly to worry Leibniz, but the new writings bear witness to a deepening in his analysis. On one hand, the vague notion of indistant parts disappears; on the other hand, we encounter the first mentions of imagination: a curve is divided "in thought" (*cogitatione*) into infinitesimal pieces.

Moreover, a hiatus is being created between what we perceive in physics, namely a continuous matter, made of assignable parts, and what is behind the appearance, that is motion and its unassignable *conatus*. The apparent and the deeper level seems to have two different mathematical counterparts.

3.2.4 Phase IV (1675–1676)

In phase IV, we witness the reemergence of metaphysical explanations, which Leibniz had strongly criticized in phase I, but with a new awareness, for now metaphysics does not supplant mathematics, but will complement it. The idea is that mathematics provides a description of natural phenomena, which although approximate and idealized makes physical problems manageable and as predictable as possible. But mathematics does not give the true laws of physics, which are instead provided by metaphysics. As Leibniz recalls decades later:

> Mais quand je cherchay les derniers raisons du Mechanisme et les loix mêmes du mouvement, je fus tout surpris de voir qu'il etoit impossible de les trouver dans les Mathematiques, et qu'il falloit retourner à la Metaphysique. C'est ce qui me ramena aux Entelechies, et du materiel au formel. [*Leibniz to Remond*, 1714. GIII 606][64]

Phase IV in Physics

Concerning the problem of the composition of physical bodies, Leibniz resorts to metaphysics in order to get the actuality he needs. Although physical bodies are infinitely divisible – namely they are capable of all resolutions into parts – the

[62] "Et posita EE infinitesima seu unitate constructionis = 1."

[63] "Ego fateor, sed aio illud 1 notandum esse hoc modo 1r. neque enim esse 1, seu numerum illum, sed esse indivisibile aliquod, quod in praesenti constructione unitatis personam sustinet, sive quod est infinitesima pars lineae cuiusdam, in partes aequales infinitas cogitatione divisae."

[64] "But when I looked for the ultimate reasons for mechanism, and even for the laws of motion, I was greatly surprised to see that they could not be found in mathematics but that I should have to return to metaphysics. This led me back to entelechies, and from the material to the formal" [Transl. Loemker].

potentiality of these parts is overcome by attributing to parts not only a motion but a soul, or mind[65]:

> And so one must examine rigorously whether there follows the perfect division of a liquid into metaphysical points, or only into mathematical points. For mathematical points could be called *indivisibiles* of Cavalieri, even if they are not metaphysical, i.e., are not minima. But if it could be shown that a liquid is divided to a greater or lesser extent, it would follow that a liquid is not resolved in[to] indivisibiles. But one could defend the composition of a liquid out of perfect points, even if it is never wholly resolved into them, since it is capable of all resolutions, and will cease when its cement – namely, mind and motion – ceases. [*De Summa Rerum* 1676. A6.3. 474. Transl. Parkinson][66]
>
> And a single indivisible body – or ἄτομον – arises, of whatever magnitude, every time it has a soul. [*Metaphysical notes* 1676. A6.3. 393]

In this way, the material parts of a physical body can be said to be unities in a proper sense – namely indivisible actual wholes – and simultaneously infinite.

The situation is apparently similar to the one described in phases II and III: we have an infinitely divided material stuff, and "something" that keeps the parts together. The crucial difference is that while in phases II and III both the material parts and the "cement," that is their *conatus* of motion, were physical entities, in phase IV, the bond is not physical but metaphysical, due to the fact that parts have minds, or souls.

So, while in phase III a geometrical representation of both the infinitely divisible assignable parts and their unassignable *conatus* was needed, because both belonged to physics, in phase IV, a geometrical description is needed only for the infinitely divisible physical parts. As metaphysical notions, indeed, minds reside on another level.

In particular, they can be conceived as unassignable and unextended without any need to resort to a definition of "magnitude" different from "extension." In phase IV, in fact, being unextended is something more than a quantitative property: minds are unextended, because as any metaphysical entity, they are extraneous to the category of extension, and more generally to the category of quantity. As such, they are also unassignable.

[65] In this period, Leibniz uses the Latin "mens" (mind) to mean every kind of mind-like substances. In later writings, he distinguishes between minds (*mentes*), souls (*animae*), and "bare" monads. Monads are capable of perceptions; souls are capable of perception and consciousness; minds are capable of rational perceptions and self-consciousness.

[66] "Adeoque examinandum rigorose, an perfecta divisio liquidi in puncta metaphysica, an vero tantum in puncta mathematica sequatur. Nam puncta mathematica possent appellari indivisibilia Cavaleriana, etsi non sint metaphysica seu minima. Quod si ostendi posset plus minusve dividi liquidum, sequetur non resolvi liquidum in indivisibilia. Posset tamen defendi liquidum componi ex punctis perfectis, etsi nunquam in illa prorsus resolvatur, vel ideo quia omnium est resolutionum capax, et caemento cessante, mente scilicet et motu, cessabit."

If at the very beginning the separation between minds and their expression is not so clear,[67] as Leibniz himself admits in later writings, in his mature theory, any conflation of metaphysical and material elements of a body disappears:

> Many years ago, when my philosophy was not yet sufficiently mature, I located souls in points [. . .] But after further reflection, I discovered that in this way we are not only led into innumerable difficulties, but that there is also here a certain, as it were, category mistake (μετάβασιν εἰς ἄλλο γένος). Those things that pertain to extension should not be attributed to souls, and their unity and multitude should not be taken from the category of quantity, but rather from the category of substance.[68] [*Leibniz to Des Bosses* April 30 1709; G II, 372. Transl. Look-Rutherford. See Leibniz (2007)]

Leibniz's reference to a category mistake is significant, because it points at the very heart of the problem: the conflation of substantial and quantitative analysis of physical bodies.

As we mentioned, in phase I, Leibniz thought that he could avoid any difficulty of this kind by banishing metaphysics from physics: no substantial immaterial forms are needed to define physical objects, which are completely identified by their geometrical shape, and mechanical implications.[69] Physical objects are continuous, divisible into potentially infinite assignable extended parts, where the adjective *infinite* is taken in its quantitative meaning.

The main problem with this geometric/mechanical interpretation was the indefiniteness of the parts into which bodies can be analyzed, but of which they cannot be "made," at least in a proper sense. In phases II and III, Leibniz tried to solve the problem by hypothesizing actual ultimate elements, which held together by motion make a continuum, but have a proper unity. However, his move raised two new difficulties: the mathematical inconsistency of the model, which should bring together the actuality of the parts and their quantitative infinity, and the impossibility of distinguishing a unitary body from a simple aggregate of parts, or atoms, without resorting to some kind of form.

Leibniz's minds are a first attempt to solve both problems from a new perspective. On one hand, in fact, minds give unity to the ultimate constituents of bodies, and keeping track of how the body has been created, make it distinguishable from other bodies with the same shape:

> We say that the effect involves its cause; that is, in such a way that whoever understands some effect perfectly will also arrive at the knowledge of its cause. For it is necessary that there is some connection between a complete cause and the effect. But on the other hand there is this obstacle: that different causes can produce an effect that is perfectly the same.

[67] Scholars' interpretations are discordant on this point. On the coexistence of matter and souls, see, for instance, Garber (2009a). The thesis of a wholesale separation from the start is maintained in Mercer (2001).

[68] The relation between simple substances and souls is discussed in detail in Garber (2009a, 2016).

[69] In phase I, the only metaphysical notion Leibniz allowed was a supreme mind, which is the ultimate cause of motion, and hence of interaction between bodies.

> For example, whether two parallelograms or two triangles are put together in the appropriate way (as is evident here) the same square will, as is clear, always be produced. Neither of these can be distinguished from the other in any way, not even by the wisest being. So, given a square of this kind, it will be in the power of no one – not even the wisest being – to discover its cause, since the problem is not determinate. The effect, therefore, seems not to involve its cause. So if we are certain, from some other source, that the effect does involve its cause, then it is necessary that the method of production must always be discernible in squares that have been produced. And so it is impossible that two squares of this kind should be perfectly similar; for they will consist of matter, but that matter will have a mind, and the mind will retain the effect of its former state. [*De summa rerum* 1676. A6.3. 490-1. Transl. Parkinson p. 51][70]

On the other hand, as metaphysical objects, minds are nonquantitative notions, and the problem of their mathematical representation simply disappears. In fact, on one level, we have the appearance of physical objects, that is the extended material bodies we perceive, and we can approximately represent using the quantitative objects of traditional geometry: a physical body will be idealized as a continuous, in(de)finitely divisible, geometric solid, every part of which is comparable with every other.

On a totally separate level, we have the metaphysical foundation of the same physical object, which we do not directly perceive, and cannot represent using traditional geometry:

> Le corps est un aggrégé de substances, et n'est pas une substance à proprement parler. ll faut par conséquent que partout dans le corps il se trouve des substances indivisibles, ingénérables et incorruptibles, ayant quelque chose de répondant aux ames. Que toutes ces substances ont tousjours esté et seront tousjours unies à des corps organiques, diversement transformables. Que chacune de ces substances contient dans sa nature *legem continuationis seriei suarum operationum*, et tout ce qui luy est arrivé et arrivera. [*Leibniz à Arnauld 23 March 1690*, A2.2. 311–12]

("The body is an aggregate of substances, and is not a substance strictly speaking. It is therefore necessary that everywhere in the body there be substances that are

[70] "Dicimus effectum involvere causam suam; id est ita ut qui perfecte intelligat effectum aliquem, etiam ad causae eius cognitionem perveniat. Utique enim necessaria quaedam inter causam integram et effectum connexio est. Sed contra obstat, quod diversae causae producere possunt eundem perfecte effectum, exempli causa, sive duo parallelogramma, sive duo [triangula], uti oportet, et hic apparet, componantur, semper idem plane quadratum prodibit; quorum alterum ab altero nullo poterit discerni modo; ne a sapientissimo quidem; ita ut in nullius ne sapientissimi quidem potestate futurum sit, ex dato quadrato ejusmodi invenire eius causam, quoniam problema non est determinatum. Videtur ergo effectus non involvere suam causam. Quare si aliunde nobis certum sit, effectum involvere suam causam, necesse est in productis quadratis modum productionis semper discerni posse Adeoque impossibile esse, ut duo quadrata eiusmodi sint perfecte similia, quia ex materia constabunt, ea autem mentem habebit, et mens retinebit effectum status prioris." As noted by Garber in commenting this quotation: "This would seem to lead directly to the later view that individual substances, complete beings, require a soul as the seat of their complete individual concept: there must be form in bodies in order for them to contain the marks and traces of their previous states" [Garber 2009b, p. 72].

indivisible, non-generatable and incorruptible, possessing something corresponding to souls. And that all these substances have always been and will always be attached to organic bodies, variously transformable. And that each of these substances contain in its nature *legem continuationis seriei suarum operationum*, and all that has happened to it and will happen to it.")

Summarizing, in phase IV, the transition occurs towards Leibniz's definitive theory of physical bodies, which are nothing but the phenomenal "expression" of the aggregation of infinite unextended metaphysical simple substances. Now, as we will show more in detail in our analysis of phase V, the way in which physical bodies express metaphysical substances is only an aspect of a more general situation in which the whole physics is grounded on metaphysics, rather than on mathematics, as it was before.

Phase IV in Mathematics

It is in mathematics that this separation has the most important consequences. In fact, once mathematics is no longer the foundation of physics, but at most provides an (approximate) description of it, it can be extended to include concepts which do not have any physical counterpart. In particular, lacking a physical model, these notions lack a geometrical representation too: since they cannot be shown in the phenomenal realm, they can only be imagined, or feigned.

In phases II and III, Leibniz's physical bodies were made of matter *and* motion: infinite assignable portions of a continuous material stuff held together and made actual by their unassignable endeavors of motion. A geometrical representation that would account consistently for both aspects turned out to be problematic, having to make the actuality of the unassignables coexist with the infinite divisibility of the continuum in assignable parts.

Gradually, Leibniz abandoned the attempt to keep them together: physical bodies have a natural representation as continuous geometrical extensions, while unassignable infinitely small quantities become imaginary notions, or fictions, and we will conclude the analysis of phase IV following the path of Leibniz's infinitely small quantities from traditional geometry towards fictionality.

The crucial step of this path is the awareness of the consequences of unassignability, namely the non-Archimedean nature of infinitely small quantities. In particular, being in no ratio to finite magnitudes, when compared with them infinitely small quantities can be ignored, without leading to any contradiction.[71]

We argue that the syncategorematic reading represents a step in Leibniz's path toward his mature conception of infinitesimals, but not the final result. This is evident from the reading of the notes he wrote in the years 1675/1676, and in particular of *De Quadratura Arithmetica* (*DQA*), Leibniz's longest mathematical work.

In DQA, three methods for the quadrature of a curve are employed: the traditional indirect Archimedean method of exhaustion, Cavalieri's method of indivisibles, and Leibniz's infinitesimal techniques. In his seminal study of *DQA*, Eberhard Knobloch

[71] The only requirement is that in deleting, and more generally in manipulating infinitely small quantities, one follows certain well-defined rules; see note 22.

considered as substantially equivalent Leibniz's and Cavalieri's method and argued that indirect Archimedean methods are furnished so as to justify these speedy but risky techniques. Leibniz himself warns the reader concerning the risks inherent in the method of indivisibles:

> J'ai pris plaisir à poursuivre ce raisonnement car il offre un exemple des précautions à prendre lorsqu'on raisonne sur l'infini dans la méthode des indivisibles. [*DQA* 1676 Prop. XXII *scholium* A 7.6 583, translated by Parmentier p. 177][72]

("I took pleasure in pursuing this reasoning because it offers an example of the precautions to be taken when reasoning about infinity in the method of indivisibles.")

We argue that infinitesimal methods are rigorous, while according to Leibniz, Archimedean methods must be handled with care. In fact, if we have to apply an exhaustion-like procedure to an infinite area, we must be careful, because some absurd results could appear.

With the introduction of *infinita terminata*, and the rules for computing with them (see footnote 83 below), Leibniz provides a tool for escaping such difficulties,

> et montre qu'on ne peut pas toujours s'empresser d'affirmer que la totalité d'un espace infini possède une certaine propriété, sous prétexte que celle-ci est toujours vérifiée par un nombre fini de parties. [*DQA* 1676 Prop. XXII *scholium* A 7.6 583, translated by Parmentier p. 177][73]

("and shows that one cannot always hasten to affirm that the totality of an infinite space possesses a certain property, under the pretext that this property is always verified by a finite number of parts.")

The DQA is sometimes used for supporting the claim that Leibniz's mature infinitesimals are inherently syncategorematic, so that the occurrences of the term "infinitesimal" are merely signs for traditional exhaustion procedures or else foreshadow "epsilontic" ones.[74] However, an attentive reading of DQA shows the opposite.

[72] "Libenter hanc contemplationem persecutus sum, quia specimen exhibet cautionis circa ratiocinia de infinito, et methodum indivisibilium"

[73] "ostenditque non semper ex partium finitarum perpetuo abscissarum proprietate quadam ad totius infiniti spatii proprietatem posse prosiliri"

[74] The idea of "syncategorematic infinitesimals" – a locution that never appears in Leibniz known writings but has been introduced in Arthur 2013 in order to explain Leibniz's mature conception of infinitesimals – comes from the idea of a syncategorematic infinite: just as a syncategorematic infinite refers to the procedure of indefinitely increasing a finite quantity, which remains in any case finite (see Sect. 2.2), syncategorematic infinitesimals refer to the procedure of indefinitely decreasing a finite quantity, which remains in any case finite. In other words, infinitesimals are a shorthand for Archimedean exhaustion-style procedures. It is clear that the definition itself of syncategorematic infinitesimals depends on the assumption that for Leibniz, the counterpart of infinitesimals are unbounded infinities. Based on Leibniz's own texts, we reject this assumption. Similarly, we reject the idea that Leibniz's infinitesimals foreshadow the current traditional theory of limits, usually referred as "epsilontic." On the epsilontic interpretation of Leibniz's infinitesimal calculus and its history, see in particular Breger (2008), pp. 187–189: "The connection between infinitesimals and what we now call epsilontics was *obvious enough* for 17th-century mathematicians" (p. 187).

In fact, it can be divided in two parts, containing respectively a traditional syncategorematic and a new infinitesimal approach.[75]

The first part of DQA (Prop I–X) does not contain anything new concerning infinitely small and infinite quantities. All one needs in order to understand propositions I–X is the syncategorematic stuff typical of Greek geometry: we take a finite quantity and then we divide (or we multiply) it indefinitely. We can go on indefinitely due to the properties of the continuum: at every step, we have a finite quantity, which at every step is smaller (bigger), without end. Both the notion of the asymptote (a line that can be made bigger and bigger without meeting the curve) and that of negligible error (a quantity that can be made lesser and lesser without becoming zero) belong to this traditional approach. Cavalieri's method of indivisibles is thought by Leibniz to fit into this scheme: a curvilinear area is filled with rectangles with infinitely small basis; since one can show that, diminishing the basis, the remaining part can be made lesser than any given quantity, the equivalence between the "sum" of infinite rectangles and the curvilinear area is demonstrated (in fact this is closer to Torricelli's method than Cavalieri's).

In most cases, the traditional method works perfectly, but there are instances where, as Leibniz explicitly says, it is not safe. In particular, it becomes "slippery" (lubrica) when one has to consider an infinite area, for instance, the one contained between a curve and its asymptote. In this case, one can reach a contradiction, and in Proposition XXII, he shows a typical case: an infinite area is proved to be equal to the same area minus a finite part[76]:

> On aboutira de la même manière, pour les autres hyperboloïdes, à un tout égal au tiers ou au quart d'une de ses parties. Voici une élégante manière de montrer combien sont périlleux les raisonnements sur l'infini lorsqu'ils ne sont pas pour ainsi dire guidés par le fil d'une démonstration. [*DQA* 1676 Prop. XXII *scholium* A 7.6 583, translated by Parmentier pp. 177–179][77]

("We will arrive in the same fashion, for the other hyperboloids, at a whole equal to a third or a quarter of one of its parts. Here is an elegant way of showing how perilous

[75] In fact, Leibniz introduces two novelties in DQA: (1) he introduces some "fictions" i.e., lines which are infinite and bounded. As we read in Prop. XI: "dans la mesure où elle s'appuie sur une fiction et aboutit sans difficulté dès qu'a été admise une ligne à la fois bornée (terminata) et infinie" [DQA, Prop XI, p. 99]. But also (2) he generalizes the method of indivisibles (see Prop. VI and Parmentier's footnotes, especially n. 2 p. 63). And in fact he calls Cavalieri's method "la methode des indivisibles ordinaire"). Here the novelty consists in the introduction of what Parmentier calls "rectangles intermédiaires, dont les arêtes ne correspondent pas aux points de la courbe" (DQA note 2 p. 63).

[76] "Eodem modo in aliis colligetur totum partis suae trientem aut quartam partem esse. Eleganti argumento quam lubrica sit ratiocinatio circa infinita, nisi demonstrationis filo regatur" (Translated by Parmentier). Historians often claim that it is Leibniz's method with infinitesimals that is not safe and therefore requires justification by means of an exhaustion like procedure (on this interpretative line, see Rabouin and Arthur (2020)). In fact here Leibniz is claiming on the contrary that only his direct method is safe, which does not involve unbounded infinities. There is currently a lively debate concerning the Leibnizian calculus; see e.g., Bair et al. (2021) and references cited therein.

[77] "Eodem modo in aliis colligetur totum partis suae trientem aut quartam partem esse. Eleganti argumento quam lubrica sit ratiocinatio circa infinita, nisi demonstrationis filo regatur." Here Leibniz is clearly referring to arguments using infinite wholes, rather than arguments using infinitesimals.

reasoning about infinity is when it is not, so to speak, guided by the thread of a demonstration.")

It is worth noting that the contradiction follows from having assumed that the quantities involved in the proof are always finite, at any step of the procedure. But this is exactly the crucial property of syncategorematic procedures and of the unbounded infinities they involve.[78]

By contrast, starting from Prop. XI, Leibniz lays the foundation of a new method, distinct from the traditional one, and which clearly has the purpose of avoiding ambiguity and contradictions related to the use of unbounded infinities[79]:

> Bien plus, à supposer que les questions dont je traite soient susceptible de démonstrations directes, je n'hésiterais pas à avancer qu'on n'en pourrait donner aucune qui n'admette ces quantités fictives, infiniment petites ou infinies, reportez-vous plus haut au scholie de la proposition VII[80]

("Furthermore, assuming that the questions I am treating are susceptible of direct proofs, I would not hesitate to assert that one could not give any [proofs] which do not exploit these fictitious quantities, infinitely small or infinite, see above at Scholium of Proposition VII.")

> [. . .] Le présent petit traité en fournira un exemple propre à convaincre de la fécondité de la méthode; ce que j'ai dit dans le scolie de la proposition 22 fournira d'autre part un exemple particulier de raisonnement sûr, là où la méthode de Cavalieri prise au sens strict n'est pas fiable [*DQA* 1676 Prop. XXIII *scholium* A 7.6 585-586, translated by Parmentier].[81]

[78] Although this demonstration has been used to argue the inherently contradictory nature of Leibniz's *infinita terminata* (see, for instance, Arthur (2013), pp. 556–558, where an early version of the proof is used), it actually proves the contradictory nature of *infinita interminata* rather than of *infinita terminata*.

[79] Significantly, Knobloch admits that contradictions are due to the use of, specifically, *un*bounded infinity: "Leibniz underlines that the method of infinite quantities is dangerous. We have to apply the thread of an indirect proof if we have in mind to rely on it. Why? Because the infinite is characterized in a privative way. In reality, there is no last point, no end, we have only relied on a fiction in order to admit it in mathematics." (Knobloch 1994, p. 270). But the fiction is not the slippery unbounded infinite: it is the reliable bounded infinite.

[80] Although they are not defined until Prop. XI, fictive infinite quantities are mentioned in the *Scholium* to Prop. VII. This fact seems to have led some readers to conflate the two methods. But in Prop. VII, Leibniz does not use any fiction: he merely notes fact that he has been forced to use a *reductio ad absurdum* because he does not know any direct method. In fact, he fears that no direct solutions are possible, unless one agrees to use fictive quantities, namely infinitesimals and *infinita terminata*: "j'ai même des raisons de craindre qu'on ne puisse pas le faire d'une manière naturelle et sans faire intervenir des quantités fictives, je veux dire infinies ou infiniment petites" (imo ratione habeo, cur verear ut id fieri possit per naturam rerum sine quantitatibus fictitiis, infinitis scilicet vel infinite parvis assumptis *DQA* VII scholium, translated by Parmentier, pp. 69–71).

[81] "Imo si quidem possibile est directas de his rebus exhiberi demonstrationes, ausim asserere, non posse eas dari, nisi his quantitatibus fictitiis, infinite parvis, aut infinitis, admissis. Adde supra prop. 7. schol. [. . .] Cujus specimen totus hic libellus erit si quis methodi fructum quaerit; securitatis autem excemplum peculiare dabunt quae diximus *schol. ad prop 22* ubi Cavalerii methodus crude sumta infida est." What may have misled some readers is Leibniz's reference to the method of indivisibles. They take this to be a reference to infinitesimals, whereas we argue that Leibniz is referring to exhaustion-type arguments.

("This little treatise will provide a suitable example to convince of the fecundity of the method; what I said in the Scolium of Proposition 22 will, on the other hand, furnish a particular example of reliable reasoning, where Cavalieri's method taken in the strict sense is not reliable.")

Here Leibniz furnishes an example where the traditional method is not reliable, whereas the introduction of Leibniz's fictions furnishes a more reliable method.

The main ingredients of Leibniz's method are bounded infinities (*infinita terminata*). Here the infinitely small quantities (μ) are fractions whose denominators are infinitely large but bounded quantities (μ)λ, both called fictions. It does not matter – Leibniz says – if these quantities are *in rerum natura*. Of course, for using them safely, one needs to provide the appropriate rules, indeed provided by Leibniz.[82]

These quantities (bounded infinities and infinitesimals) in the DQA are treated as different from what Leibniz calls indivisibles. In particular, Leibniz says that:

> Au demeurant il semblera à certains paradoxal de parler de Lignes infinies qui soient cependant bornées, je dois donc faire remarquer qu'il existe une grande différence entre indivisible et infiniment petit, et de la même manière une grande différence entre infini et non borné [*DQA* 1676 **Prop. XI]**[83]

("At any rate, it will seem paradoxical, to some, to speak of infinite Lines which are nevertheless bounded, I must therefore point out that there is a great difference between indivisible and infinitely small, and in the same way a great difference between [bounded] infinite and unbounded.")

Here Leibniz emphasizes the difference between indivisibles and infinitesimals, which supports our interpretation. Leibniz also wrote:

> La Géométrie des indivisibles est trompeuse si on n'y adjoint quelques précisions au sujet des infiniment petits; utiliser des points véritablement indivisibles comporte un certain risque et il faut au contraire employer des lignes qui, pour être infiniment petites, n'en sont pas moins des lignes et sont, par conséquent, divisibles. [*DQA* 1676 **Prop. XI]**

("The Geometry of the Indivisibles is misleading if one does not add to it some clarifications concerning the infinitely small; using truly indivisible points carries a certain risk and on the contrary it is necessary to use lines which, while being infinitely small, are nonetheless lines and are, therefore, divisible.")

[82] In particular, having constructed the infinite rectangle with basis (μ) and length (μ)λ, Leibniz proves that, depending on the properties of the curve, it can be: (a) finite, (b) infinite, or (c) infinitely small. The proof of (a) is made extending to the infinite rectangle some properties which hold for finite areas. The proofs of (b) and (c) are made using the rectangle xy where x is the "first ordinate" and y is the "first abscissa." Summarizing, Leibniz shows that if x: 1 = m: (μ)λ with (μ)λ infinitum terminatum, then x is infinitely small. if (μ): 1 = m: y with (μ) infinitely small, then y is infinite (terminatum).

[83] Note that the contrast is a chiastic one: indivisibles and unbounded go together, and infinitesimal and infinite (i.e., bounded infinite) go together, as is clear from the subsequent discussion.

This clearly establishes that when Leibniz is talking about indivisibles, he is not talking about infinitesimals.

In sum, in *De Quadratura Arithmetica*, Leibniz presents the infinitesimal method as more reliable than the Archimedean one, because the latter sometimes relies on infinite wholes which are contradictory. In better-known passages, however, he justifies his infinitesimal method by an appeal to Archimedean exhaustion. It is possible that the erroneous idea that in *De Quadratura Arithmetica* Leibniz speaks of unreliability of his method, rather than of other methods involving unbounded infinities, comes from these later passages.

In phase IV, one is far from the unproblematic approach of phases II and III, when Leibniz was convinced that, although unassignable, infinitely small quantities could simply be handled via the method of exhaustion. Now he feels the need to justify their legitimacy.

As sketched in our analysis of the *De Quadratura Arithmetica,* Leibniz's strategy consists of two steps. On one hand, he shields his unassignable infinitely small quantities from the accusation of not complying with the part-whole axiom, and in this way, he renders them consistent from a mathematical point of view. On the other hand, he introduces for them a peculiar *status*, that of "fictions," which holds together their mathematical consistency and their lacking of any physical counterpart.

Concerning the part-whole axiom, Leibniz specifies that only unbounded infinities (*infinita interminata*) contradict it, and for this reason are absolutely impossible.[84] By contrast, bounded infinities imply no contradiction, and for this reason are only accidentally impossible.[85]

Concerning the *status* of his unassignable infinitely small quantities, Leibniz assimilates them to other physically problematic but useful, well-defined, and well-known notions such as imaginary numbers, which appear in mathematics but cannot be exhibited or geometrically constructed, as we have shown at the end of phase III and as reaffirmed by Leibniz in phase IV in the following terms:

> I call *imaginary constructions* the constructions that we can feign (fingere), but we can show (exhibere) only with your mind. They are nevertheless possible [. . .]: we can imagine for instance a quantity that grows – or is shortened or is transformed or is moved – following a

[84] We know that already in 1672, Leibniz declared infinite wholes to be contradictory: "Hence it follows either that in the infinite the whole is not greater than the part, which is the opinion of Galileo and Gregory of St. Vincent, and which I cannot accept; or that infinity itself is nothing, i.e., that it is not one and not a whole." [*Notes on Galileo's Two New Sciences* 1672. A VI. 3, p. 168; translated in Arthur 2001, p. 9].

[85] "Therefore if the essence of a thing can be conceived, provided that it is conceived clearly and distinctly (e.g., a species of animal with an uneven number of feet, also a species of immortal beast), then it must already be held to be possible, and its contrary will not be necessary, even if its existence may be contrary to the harmony of things and the existence of the divine intelligence, and consequently it never will actually exist, but it will remain per accidens impossible. Hence all those who call impossible (absolutely, i.e., per se) whatever neither was nor is nor will be are mistaken." (*Confessio Philosophi* A VI 3, 128, transl. Sleigh).

> given law, to which apparently no "ratio" can be associated. And it can happen that one can find a physical or even a geometrical realisation (effectio) of a construction that at the beginning seemed imaginary. [*De descriptione geometrica curvarum transscendentium* 1675 A7.7 516][86]
>
> One must discuss if $\sqrt{-1}$ is nothing or if it contains a truth. In fact, even if it cannot be done, it can be thought, not in itself but using characters and analogy (I call this way of thinking a blind one). In fact, as incommensurables are such that their powers are commensurable, in the same way imaginary are such that their powers are real, and impossibles are such that their squares are possibles, like $\sqrt{-1}$, whose square is -1, even if we assume that in nature there is no quantity corresponding to that: in fact, it is enough that its character is useful, because together with other ones it can express something real. [*Imaginariae usus ad comparationem Circuli et Hyperbolae* 1675 A7.7 561][87]

Like imaginary numbers, unassignable infinitely small quantities are intrinsically not geometrically representable: how could one represent a non-Archimedean quantity in a continuous framework such as that of Greek geometry? Furthermore, they cannot be viewed as the idealization of some physical objects, which would not respect the harmony of things. Summarizing: since unassignable infinitely small quantities do not contradict the part-whole axiom, they are possible (more precisely, not absolutely impossible) notions, but since they cannot be exhibited – either in geometry or in physics – they are fictions. The same position will be reiterated by Leibniz in subsequent years:

> And some extractions of roots are such that roots are surd and they do not exist in natura rerum, and we call them imaginary, and the problem is impossible, as when our analysis shows that the requested point must be exhibited by the intersection of a specific circle and a specific straight line, in which case it may happen that this circle by no means reaches this line, and then the intersection is imaginary [. . .] There is a big difference between imaginary quantities, or those impossible by accident, and absolutely impossible ones, which involve a contradiction: e.g., when it is found that solving a problem requires that 3 be equal to 4, which is absurd. But imaginary quantities, or [quantities] impossible by accident, namely quantities that cannot be exhibited for lack of a sufficient condition, which is required for having an intersection, can be compared with infinite and infinitely small quantities, which are generated in the same way [. . .] And it is true that calculus necessary leads to them, and people who are not sufficiently expert in such matters get entangled and think they have reached an absurdity (*absurdum*). Experts know instead that this apparent impossibility [apparentem illam impossibilitatem] only means that a parallel line is traced instead of a

[86] "*Constructiones imaginarias* voco, quas fingere possumus, exhibere non nisi mente non possumus. Sunt tamen possibiles [. . .] tales sunt si imaginemur magnitudinem aliquam certa lege crescere aut contrahi aut transformari, aut etiam moveri, cujus reapse praescribendae nulla appareat ratio. Et fieri poterit aliquando ut constructionis quae initio imaginaria apparuit tandem physica aut etiam Geometrica effectio reperiatur." The passage was discussed also in Katz et al. (2021).

[87] "An autem quantitas $\sqrt{-1}$. sit omnino nihil, an vero contineat nescio quid discutiendum diligentius. Etsi enim non possit effici, potest tamen quodammodo intelligi, non seipsa sed ope characteris et analogiae, exemplo illius cogitationis quam caecam appello. Et vero quemadmodum sunt incommensurabiles quae sunt potentia commensurabiles, ita sunt imaginariae quarum potentiae sunt reales; seu impossibiles quarum quadrata sunt possibilia, ut $\sqrt{-1}$, cujus quadratum -1 etsi autem ponatur nihil omnino esse in natura tali quantitati respondens; sufficit tamen ejus characterem esse utilem, quia cum aliis junctus realia exprimit."

> straight line making the required angle, and this parallelism is the required angle, or quasi-angle [*Elementa nova matheseos universalis* 1683 A 6.4 520–521][88]

The idea is that traditionally Greek geometry was considered the natural idealization of physics, and Greek geometry did coincide with the whole mathematics.[89] Subsequently, mathematical notions were found which did not belong to Greek geometry – roughly speaking, Greek geometry became a subset of a larger mathematics – but the idea survived that the idealization of physics has to be Greek geometry. The elements belonging to the complement of Greek geometry, having no direct physical counterpart, are then accepted, but they are regarded as fictions, or imaginary notions.

3.2.5 Phase V (1677–. . .)

Phase V leads Leibniz to his ultimate theory of the composition of bodies, a theory in which the separation among mathematics, physics, and metaphysics is complete. In particular, Leibniz eliminates any residual ambiguity by providing a metaphysical foundation to any feature of a physical body. Not only the actuality and the unity of the parts but their physical extension too is now explained in term of metaphysical notions.

Basically, Leibniz's idea is to extend to physical composite substances his analysis of simple metaphysical substances. Simple metaphysical substances – which the mature Leibniz also calls monads – can be analyzed into an active force (*vis activa*), or substantial form, and a passive force, or prime matter.[90] In an analogous way, in phase V, Leibniz expresses both extended matter and substantial form of physical bodies

[88] "Et quaedam extractiones tales sunt, ut radices illae surdae nec in natura rerum extent, tunc dicuntur imaginariae, et problema est impossibile, ut cum analysis ostendit punctum quaesitum debere exhiberi per intersectionem certi circuli et certae rectae, ubi fieri potest ut ille circulus ad illam rectam nullo modo perveniat, et tunc intersectio erit imaginaria [. . .]. Multum autem interest inter quantitates imaginarias, seu impossibiles per accidens, et impossibiles absolute quae involvunt contradictionem, ut cum invenitur ad problema solvendum opus esse, ut fit 3 aequ. 4 quod est absurdum. Imaginariae vero seu per accidens impossibiles, quae scilicet non possunt exhiberi ob defectum sufficientis constitutionis ad intersectionem necessariae, possunt comparari cum Quantitatibus infinitis et infinite parvis, quae eodem modo oriuntur [. . .] quod adeo verum est ut saepe calculus ad eas necessario ducat, ubi harum rerum nondum satis periti mire torquentur et in absurdum se incidisse putant. Intelligentes vero sciunt apparentem illam impossibilitatem tantum significare, ut loco rectae angulum quaesitum facientis ducatur parallela; hunc parallelismum esse angulum illum seu quasi angulum quaesitum."

[89] Still in Leibniz's days, algebraic notions were viewed as a sort of symbolic "translation" of geometrical quantities, and a "rigorous" proof meant a geometrical proof. On this point, see Breger (2008), §3. Cf. Parmentier (1995), pp. 26–27.

[90] While prime matter, as pure passivity, cannot be further specified, Leibniz describes the active power, or *entelecheia*, of a simple substance in terms of a law, the *law of the series*, which expresses the inner states of the substance itself (see GP III, 66; 356; GP IV, 512). To explain the way in which the law of the series rules the sequence of the inner states, Leibniz uses the analogy of a mathematical sequence of numbers: "L'essence des substances consiste dans la force primitive d'agir, ou dans la loy de la suite des changemens, comme la nature de la series dans les nombres" (*Excerpta ex notis meis marginalibus ad Fucherii responsionem primam in Malebranchium critica*, 1676, A6.3. 326). On the relation between the law of the series and simple substances in general, see, for instance, Whipple (2010) and the bibliography quoted there.

in terms of forces: a primitive passive force, corresponding to matter – the "*primary matter* in the schools, if correctly interpreted"[91] – and a primitive active force, corresponding to substantial form. As for Aristotle, form and matter, together, constitute a substance, for Leibniz, primitive forces, active and passive, do the same:

> Primitive active force, which Aristotle calls first entelechy and one commonly calls the form of a substance, is another natural principle which, together with matter or [primitive] passive force, completes a corporeal substance. This substance, of course, is one *per se*, and not a mere aggregate of many substances, for there is a great difference between an animal, for example, and a flock. [*On Body and Force* (1702), AG 252][92]

As Aristotelian primary matter and form, primitive forces are causes in a wide sense, for they are responsible for the whole being of a physical substance, and consequently for its behavior, and its observable properties:

> Composite substance does not formally consist in monads and their subordination, for then it would be a mere aggregate or a being per accidens. Rather, it consists in primitive active and passive force, from which arise the qualities and the actions and passions of the composite which are discovered by the senses, if they are assumed to be more than phenomena. [*Leibniz to Des Bosses* (1716). See Leibniz (2007), AG 203–4][93]

In particular, mechanical properties of physical substances – namely their internal coherence and their reciprocal interactions, and therefore the local motions behind them – are due to peculiar properties of these primitive forces, which Leibniz calls derivative forces:

> By derivative force, namely, that by which bodies actually act on one another or are acted upon by one another, I understand, in this context, only that which is connected to motion (local motion, of course), and which, in turn, tends further to produce local motion. For we acknowledge that all other material phenomena can be explained by local motion. [*Specimen Dynamicum* (1695), AG 120][94]

As for metaphysical notions in general, we have direct access neither to primitive forces nor to derivative ones: at most, they can be partially understood by means of

[91] *Specimen Dynamicum* [GM VI, 237].

[92] "Vis activa primitiva quae Aristoteli dicitur ἐντελέχεια ἡ πρώτη, vulgo forma substantiae, est alterum naturale principium quod cum materia seu vi passiva [primitiva] substantiam corpoream absolvit, quae scilicet unum per se est, non nudum aggregatum plurium substantiarum, multum enim interest verbi gratia inter animal et gregem" (GP IV, 395).

[93] "Substantia composita non consistit formaliter in monadibuset earum subordinatione, ita enim merum foret aggregatum, seu ens per accidens, sed consistit in vi activa, et passiva primitiva, ex quibus oriuntur qualitates et actiones passionesque compositi, quae sensibus deprehenduntur, si plus quam phaenomena esse ponantur" (GP II, 517–518)

[94] "Vim ergo derivativam, qua scilicet corpora actu in se invicem agunt aut a se invicem patiuntur, hoc loco non aliam intelligimus, quam quae motui (locali scilicet) cohaeret, et vicissim ad motum localem porro producendum tendit. Nam per motum localem caetera phaenomena materialia explicari posse agnoscimus" (GM VI, 237).

analogical speech. In particular, if we read primitive forces, active and passive, as substances, then derivative forces, active and passive, must be read as their faculties (*facultés*), or ways of being (*façons d'être*).[95] If otherwise we read the primitive force in analogy with the law of the series (of the states of the substance; see note 91), then derivative forces can be related to the single terms of the series:

> Derivative force is itself the present state when it tends toward or preinvolves a following state, as every present is pregnant with the future. But that which persists, insofar as it involves all cases, contains primitive force, so that primitive force is, as it were [*velut*], the law of the series, while derivative force is, as it were, a determination which designates some term in the series. [*Leibniz to De Volder* (1704). Transl. Loemker][96]

At this point, Leibniz differentiates his derivative active force from scholastics' faculties, which are traditionally interpreted as purely passive powers (*dynameis*). Derivative active force, instead, is endowed with an active tendency, which naturally tends towards its fulfillment:

> Active force differs from the mere power familiar to the Schools, for the active power or faculty of the Scholastics is nothing but a close possibility of acting, which needs an external excitation or a stimulus, as it were, to be transferred into action. Active force, in contrast, contains a certain act or entelechy and is thus midway between the faculty of acting and the act itself and involves a *conatus*. It is thus carried into action by itself and needs no help but only the removal of an impediment. [*De primae philosophiae emendatione, et de notione substantiae*, 1694. Transl. Loemker][97]

The explanation is resumed at the beginning of the *Specimen Dynamicum*, where Leibniz insists on the notion of *conatus*, now called *nisus* and defined in a scholastic way as a disposition, or appetite, which completes the force:

> Elsewhere we urged that in corporeal things there is something over and above extension, in fact, something prior to extension, namely, that force of nature implanted (*inditam*) everywhere by the Creator. This force does not consist in a simple faculty, with which the schools

[95] "Les Puissances primitives constituent les substances mêmes, et les puissances derivatives, ou si vous voulés, les facultés, ne sont que des façons d'estre, qu'il faut deriver des substances" (*Nouveaux Essais*, IV.iii.6. A VI, 6, 379).

[96] "Vis autem derivativa est ipse status praesens dum tendit ad sequentem seu sequentem praeinvolvit, uti omne praesens gravidum est futuro. Sed ipsum persistens, quatenus involvit casus omnes, primitivam vim habet, ut vis primitiva sit velut lex seriei, vis derivativa velut determinatio quae terminum aliquem in serie designat" (GP II, 262).

[97] "Differt enim vis activa a potentia nuda vulgo scholis cognita, quod potentia activa Scholasticorum, seu facultas, nihil aliud est quam propinqua agendi possibilitas, quae tamen aliena excitatione et velut stimulo indiget, ut in actu transferatur. Sed vis activa actum quendam sive ἐντελέχειαν continet, atque inter facultatem agendi actionemque ipsam media est, et conatun involvit; atque ita per se ipsam in operationem fertur; nec auxiliis indiget, sed sola sublatione impedimenti" (GP IV 469).

> have been content, but is further endowed with *conatus* or *nisus*, attaining its full effect unless it is impeded by a contrary conatus. [*Specimen Dynamicus* 1695, AG 118][98]

Summarizing, in phase V, Leibniz reads extended physical bodies as the mere phenomenal "expression" of the aggregation of infinite unextended metaphysical simple substances. Here a simple substance simply means the union of an active force, or substantial form, and a passive force, or prime matter. On a purely metaphysical level, then, physical bodies too can be analyzed in terms of primitive forces, active and passive. In their turn, primitive forces are meant to act in virtue of certain faculties, which Leibniz calls derivative forces, active and passive. Finally, derivative forces "produce" all the observable physical effects. For instance, all manifestations of local motion come from derivative active forces, via their *conatus*, or *nisus*.

The actuality of derivative forces manifests itself in observable but phenomenal local motions:

> For, strictly speaking, motion (and likewise time) never really exists, since the whole never exists, inasmuch as it lacks coexistent parts. And furthermore, there is nothing real in motion but a momentary something which must consist in a force striving *[nitente]* toward change. Whatever there is in corporeal nature over and above the object of geometry or extension reduces to this. [**AG 119**]

In a similar way, the unity and self-sufficiency of metaphysical substances, namely of primitive forces, has its phenomenal expression in the physical extension of bodies, every part of which has a proper individuality.[99] In other words, due to its monadic origin, physical matter is not a continuous stuff, like geometrical extension is, but a contiguous one, every part being characterized and individuated by a different motion:

> As for the first point, it follows from the very fact that a mathematical body cannot be analyzed into primary constituents that it is also not real but something mental and designates nothing but the possibility of parts, not something actual. A mathematical line, namely, is in this respect like arithmetical unity; in both cases the parts are only possible and completely indefinite. A line is no more an aggregate of the lines into which it can be cut than unity is the aggregate of the fractions into which it can be split up. [. . .]
>
> But in real things, that is, bodies, the parts are not indefinite – as they are in space, which is a mental thing – but actually specified in a fixed way according to the divisions and

[98] "In rebus corporeis esse aliquid praeter extensionem, imo extensione prius, alibi admonuimus, nempe ipsam vim naturae ubique ab Authore inditam, quae non in simplici facultate consistit, qua Scholae contentae fuisse videntur, sed praeterea conatu sive nisu instruitur, effectum pleno habituro, nisi contrario conatu impediatur" (GM VI, 235).

[99] In Leibniz's words, "every part of matter is actually divided into other parts" [*Leibniz to Arnauld*, November 28-December 8, 1686. A2.2. 122. Transl. Ariew-Garber]; "In truth, matter is not a continuous thing but a discrete one, actually divided at infinity" [*Leibniz to De Volder*, October 11 1705; G II, 278. (A2.4. 303600). Transl. Loemker].

> subdivisions which nature actually introduces through the varieties of motion. And granted that these divisions proceed to infinity, they are nevertheless the result of fixed primary constituents or real unities, though infinite in number. Accurately speaking, however, matter is not composed of these constitutive unities but results from them. [*Leibniz to De Volder* June 30 1704; G II, 268. Transl. Loemker][100]
>
> In truth, matter is not a continuous thing but a discrete one, actually divided *ad infinitum*. [*Leibniz to De Volder*, October 11, 1705; G II, 278][101] And certainly every extended thing is divisible, in such a way that parts can be assigned in it, but in matter they are actually assigned, while in extension they are only potential, as in number. [*Ad Schedam Amaxariam*][102]

The crucial novelty, which solves the last remaining difficulties in phase IV, is that now all the ultimate "components" of physical bodies – namely primitive and derivative forces, as well as their tendencies – are metaphysical entities. And as metaphysical entities, the question of their geometrical representation simply disappears:

> Therefore, I concluded from this that, because we cannot derive all truths concerning corporeal things from logical and geometrical axioms alone, that is, from large and small, whole and part, shape and position, and because we must appeal to other axioms pertaining to cause and effect, action and passion, in terms of which we can explain the order of things, we must admit something metaphysical, something perceptible by the mind alone over and above that which is purely mathematical and subject to the imagination, and we must add to material mass [massa] a certain superior and, so to speak, formal principle. Whether we call

[100] "Quod primum attinet, eo ipso quod corpus mathematicum non potest resolvi in prima constitutiva, id utique non esse reale colligatur, sed mentale quiddam nec aliud designans quam possibilitatem partium, non aliquid actuale. Nempe linea mathematica se habet ut unitas arithmetica, et utrobique partes non sunt nisi possibiles et prorsus indefinitae; et non magis linea est aggregatum linearum in quas secari potest, quam unitas est aggregatum fractionum in quas potest discerpi [. . .] At in realibus, nempe corporibus, partes non sunt indefinitae (ut in spatio, re mentali), sed actu assignatae certo modo, prout natura divisiones et subdivisiones actu secundum motuum varietates instituit, et licet eae divisiones procedant in infinitum, non ideo tamen minus omnia resultant ex certis primis constitutivis seu unitatibus realibus, sed numero infinitis. Accurate autem loquendo materia non componitur ex unitatibus constitutivis, sed ex iis resultat."

[101] "Revera materia non continuum sed discretum est actu in infinitum divisum."

[102] "Et sane omne extensum divisibile est, ut partes in eo assignari possint, actuque sint assignatae in materia, potentiales vero sint in ipsa extensione, ut in numero." A similar statement is reiterated a few pages later: "Est enim ut saepe dixi dispositio compossibilium phaenomenorum, Geometriaque est possibilitatum ipsarum tractatio. Et ideo extensio ipsa possibilis sive continua non habet partes determinatas, non magis quam Unitas numerica; quod secus est de physicis, seu existentibus ubi divisiones sunt actu factae. Quemadmodum non datur minima fractio in numeris, ita nec minimum in Geometria, punctumque est extremum non pars. Sed in rebus ipsis dantur minima simplices nempe substantiae." (*Ad Schedam Hamaxariam*, ad. 16, in LH IV, 3, 5c, Bl. 1–2. Transcription by Massimo Mugnai and Heinrich Schepers).

> this principle form or entelechy or force does not matter, as long as we remember that it can only be explained through the notion of forces. [*Specimen Dynamicum* (1695), AG 125][103]

What has to be dealt with using mathematics is only the phenomenal "expression" of metaphysical forces, namely physical bodies and their interactions. As usual, dealing with physics, one must be aware that mathematical descriptions are idealizations, useful but necessarily approximate. For instance, Leibniz defines matter as a contiguous stuff but, as we mentioned in phase IV, contiguity is not a mathematical property. The better mathematical description of a contiguous physical body is still a continuous mathematical solid, where continuity still means, as in phase I, potentially infinite divisibility, and indefiniteness of parts.[104]

More generally, this holds for each physical notion involving extension – whether matter, time, or motion – which therefore admit a traditional geometric representation, and can be dealt with using the traditional tools of Greek geometry, and in particular the method of exhaustion, without any need for infinitesimal quantities. Of course, since one can solve traditional geometrical problems using infinitesimal calculus, one can solve cinematic problems using x, t, and dx, dt. But while x and t have a direct counterpart in a measurable, and geometrically representable extensions, dx and dt have no direct physical counterparts. In other words, even if space and time are only potentially infinitely divisible (without ever attaining infinitesimals), we can fictively introduce infinitesimals of space, or of time.

---

[103] "Hinc igitur, praeter pure mathematica et imaginatione subjecta, collegi quaedam metaphysica solaque mente perceptibilia esse admittenda, et massae materiali principium quoddam superius, et ut sic dicam formale addendum, quandoquidem omnes veritates rerum corporearum ex solis axiomatibus logisticis et geometricis, nempe de magno te parvo, toto et parte, figura et situ, colligi non possint, sed alia causa et effectu, actioneque et passione accedere debeant, quibus ordinis rerum rationes salventur. Id principium Formam, an ἐντελέχειαν, an Vim appellemus, non refert, modo meminerimus per solam virium notionem intelligibiliter explicari." (GM VI, 241–242). Cf. the quotations in footnote 64 and 66 For analogous statements, see, for instance, what Leibniz's wrote some years before, in his *Discourses on Metaphysics* (1686): "Mais la force, or cause prochaine de ces changemens est quelque chose de plus reel, et il y a assez de fondement pour l'attribuer, à un corp plus qu'à l'autre; aussi n'est ce que par là qu'on peut connoistre à qui le mouvement appartient d'avantage. Or cette force est qualque chose de different de la grandeur, de la figure et du mouvement, et on peut juger par là que tout ce qui est conçu dans les corps ne consiste pas uniquement dans l'étendue et dans ses modifications, comme nos modernes se le persuadent. Ainsi nous sommes encor obligés de rétablir quelques estres ou formes, qu'il ont bannies. Et il paroist de plus en plus quoyque tous les phenomenes particuliers de la nature se puissent expliquer mathematiquement ou mechaniquement par ceux qui les entendent, qui neantmoins les principes generaux de la nature corporelle et de la mecanique même sont plustot metaphysiques que Geometriques, et appartiennent plustost à quelques formes ou natures indivisibles comme causes des apparences qu'à la masse corporelle et étendue" (GP IV, 444. English translation in AG 51–52).

[104] "A continuum is a whole whose parts are indefinite, and so is the space if we abstract everything that is in it." [*Divisio terminorum*. . .1683 A6.4. 565]. The relation between continuity and indefiniteness of parts is at the basis of Aristotle's discussion of the potentially infinite division of the continuum, in *Physics* III.4–8 and VI.1, a discussion which Leibniz continued to appreciate far beyond his "aristotelian" phase I: "Aristote a fort bien expliqué le plein et la division du continu contre les atomistes" [*Remarques sur la doctrine Cartesienne* 1689. A6.4. 2047].

As we argued at the end of phase IV, this is finally the sense in which we interpret Leibniz's idea of fictionality: a mathematical notion is a fiction when it lacks a traditional geometrical representation and therefore has no physical counterpart.[105]

We conclude our analysis by observing that phase V offers a new perspective on the usefulness of fictions like infinitesimals, which Leibniz employs in an analogical way to speak of otherwise inexpressible metaphysical notions. For instance, in the *Specimen Dynamicus* (GM VI, 238), a *nisus* is described as an elementary, infinitely small solicitation, from the infinite repetition of which *impetus* arises. An analogous relation is described between the dead derivative force (*vis mortua*), to which the elementary solicitation belongs, and the living derivative force (*vis viva*).

Our analysis of Leibniz's approach to force is compatible with the analysis by Rutherford, who wrote:

> In the end, I propose, it is Leibniz's view that, in and of itself, force is not a mathematically representable property. To understand the sense in which force is real, we must turn to a different theoretical framework altogether, that of metaphysics, wherein force is represented as a modification of a substantial power, or principle of change. (Rutherford 2008, p. 257)

The connection with infinitesimal calculus together with the necessity of taking it as an analogy is acknowledged by Leibniz himself and made explicit in a 1698 letter to De Volder:

> By analogy with geometry, or my analysis, solicitations are as dx, speeds as x, and [living] forces are as xx, or $\int xdx$.[106]

And this is in accordance with Leibniz's belief that metaphysical notions elude our human faculties, and the way in which we can speak of them necessarily involves reasoning by *analogy*:

[105] In fact the introduction of infinitesimals, as well as other fictional mathematical notions, contributed to the release of algebra from geometry: "Leibniz avait adopté une nouvelle manière de pratiquer la géometrie. Le calcul entre ses mains n'était plus la traduction des figures et des problèmes mais devenait le vrai point de départ" (Parmentier 1995, p. 104). On this point, see also Sherry and Katz (2014): "In this respect Leibniz recognized a hidden power of algebra, realizing that it was more than a short hand for what could be expressed geometrically. As impressed as Leibniz was by imaginary expression for real roots, he was more impressed by the power of algebra, and more generally, symbolic thinking, to unify disparate domains" (p. 170). Even though he starts from different premises – that infinitesimals are compendia for Archimedean procedures and foreshadow the epsilontic approach to the notion of limit – an analogous conclusion is reached by Breger 2008 (pp. 195–198).

[106] "Ut ita secundum analogiam Geometriae seu analysis nostrae solicitationes sint ut dx, celeritates ut x, vires ut xx seu ut $\int xdx$" [GP II, 156].

> On ne doit point s'étonner que je tâche d'éclaircir ces choses par des *comparaisons* prises des mathématiques pures, où tout va dans l'ordre, et où il y a moyen de les démêler par une méditation exacte, qui nous fait jouir, pour ainsi dire, de la vue des idées. [*Théodicée*, § 242][107]
>
> ("One should not be surprised that I try to clarify these things by comparisons taken from pure mathematics, where everything goes in order, and where there is a way to disentangle them by a precise meditation, which makes us benefit, so to speak, from the view of ideas.")
>
> Je trouvay donc que leur nature consiste dans la force, et que de cela s'ensuit quelque chose d'*analogique* au sentiment et à l'appetit; et qu'ainsi il falloit les concevoir *à l'imitation* de la notion que nous avons des ames. [*Systeme nouveau* (1695)][108]
>
> ("I therefore found that their nature consists in the force, and that from this it follows something analogous with regard to feeling and appetite; and that thus it was necessary to conceive them in imitation of the notion that we have of souls.")

4 Conclusion

In this chapter, we have attempted to trace the evolution of Leibniz's thinking regarding infinitesimals, both in mathematics and in physics.

We have observed how Leibniz's increasing mathematical awareness led him to gradually abandon his initial scholastic position, centered around the idea of the potential infinite divisibility of the continuum, in favor of more informed reevaluations, that involve the notion of unassignability.

We have divided Leibniz's journey into five phases, which can be summarized as follows:

Phase I: Both in mathematics and in physics, Leibniz adopts a traditional syncategorematic approach: both extension and matter are continuous, which means that they are potentially infinitely divisible, and every obtained part is assignable. The problem arises of the actuality of the elementary constituents of physical matter.

[107] GP VI, 261–262. "je ne prends les mots dans ce sens extraordinaire que parceque je ne trouve point des termes plus propres pour m'exprimer" (*Additions à l'Explication du systeme nouveau touchant l'union de l'ame et du corps*, 1698. GP IV, 582). "L'Idée positive de cette substance simple ou Force primitive est toute trouvée, puisque elle doit toujours avoir en elle un progres reglé de perceptions, suivant l'Analogie qu'elle doit avoir avec notre ame" (Leibniz to Lady Masham 30 june 1704. GP III, 357). On the role of analogy in Leibniz's philosophy, see Rescher (2013), Chap. 7, and Pasini (2016).

[108] GP IV 479, translated in AG 139. Italics ours. Of course, it is not only the notion of force that cannot be translated in mathematical terms: think about the notion of simple substance, or that of hypercategorematic infinite, or even possible worlds etc.. . . and the mathematical analogies which, this notwithstanding, Leibniz employs to illustrate them: "Vous savés donc que lorsque les conditions d'un point qu'on demande, ne le determinent pas assés, et qu'il y en a une infinité, ils tombent tous dans ce que les Geometres appellent un lieu, et ce lieu au moins (qui est souvent une Ligne) sera determiné. Ainsi vous pouvés vous figurer une suite reglée de Mondes, qui contiendront tous et seuls le cas dont il s'agit, et en varieront les circonstances et les consequences. Mais si vous posés un cas qui ne differe du monde actuel que dans une seule chose definie et dans ses suites, un certain monde determiné vous repondra" (*Théodicée*, § 414, GP VI, 362–363).

Phase II: In an attempt to resolve the problem, Leibniz separates the treatment of physics from that of mathematics. He redefines matter as a continuous stuff infinitely divided – and not only divisible – into individual parts. Both actuality and unity of these parts are due to their conatus, or unextended amounts of motion. The geometrical counterpart of this physical situation is a hybrid in which the infinite divisibility of the continuum has to coexist with the actuality of unextended "fat points," which Leibniz identifies with Cavalieri's indivisibles. In this phase, unassignability is accepted, without any further elaboration.

Phase III: Perhaps dissatisfied with the conclusions reached in phase II, Leibniz introduces some variations. The most significant one is the inversion of perspective in physics: discreteness is considered a foundational notion in relation to continuity: physical bodies do not come from an infinitely divided continuous stuff but from the aggregation of discrete basic elements, that as in phase II are "glued" together by their conatus of motion. In mathematics, the situation is similar, with the only difference being that now points are extended, although unassignable. Apparently, in this phase, to keep together extended but unassignable quantities and assignable ones is not a problem for Leibniz. Unassignable quantities are not called fictions, nor referred to using cognate verbal forms.

Phase IV: Leibniz delves deeper into the concept of unassignability and further separates mathematical inquiry from physical inquiry. Unassignable quantities are useful in mathematics but cannot have a physical counterpart. In fact, he rejects the possibility that physical quantities are not all comparable to each other: a world containing some portion of matter which has an unassignable ratio to another portion would be less perfect than a world where all portions are comparable. Leibniz starts referring to unassignable quantities as "fictions" and links them to other "imaginary" notions, such as imaginary numbers, that are noncontradictory but do not admit any traditional geometrical representation, nor a physical counterpart. In physics, metaphysical explanations reappear alongside the mathematical ones, and in some cases, as a replacement for them.

Phase V: The separation between mathematics, physics, and metaphysics is gradually brought to completion. In particular, mathematics ceases to be the foundation of physics, and geometry ceases to be its natural description: the true essence of physics is metaphysics, and as such it is non representable. In particular, being extraneous to the category of quantity, metaphysical objects can be thought of as unassignable in a stronger sense, and without problematic consequences.

We are far away from the reassuring "syncategorematic" vision of the beginning: mathematical awareness had led Leibniz to remarkable achievements in mathematics, but it had also compelled him to introduce greater complexity in physics.

References

Antognazza MR (2015) The hypercategorematic infinite. Leibniz Rev 25:5–30

Arthur R (Tr) (2001) The labyrinth of the continuum. Writings on the continuum problem, 1672–1686. G. W. Leibniz. Yale University Press, New Haven

Arthur R (2009) Actual infinitesimals in Leibniz's early thought. In: Kulstad M, Laerke M, Snyder D (eds) The philosophy of the Young Leibniz. Studia Leibnitiana Sonderhefte 35. Franz Steiner, Stuttgart, pp 11–28
Arthur R (2013) Leibniz's syncategorematic infinitesimals. Arch Hist Exact Sci 67(5):553–593
Arthur R (2015) Leibniz's actual infinite in relation to his analysis of matter. In: Leibniz GW (ed) interrelations between mathematics and philosophy, 137–156. Edited by Goethe NB, Beeley P and Rabouin D. Archimedes, 41, Springer, Dordrecht
Arthur R (2019) Leibniz in Cantor's paradise. In: De Risi V (ed) Leibniz and the structure of sciences. Modern perspectives on the history of logic, mathematics, epistemology. Boston studies in philosophy and history of science. Springer, Cham, pp 71–109
Bair J, Błaszczyk P, Ely R, Katz M, Kuhlemann K (2021) Procedures of Leibnizian infinitesimal calculus: an account in three modern frameworks. Br J Hist Math 36(3):170–209. https://doi.org/10.1080/26375451.2020.1851120, https://arxiv.org/abs/2011.12628
Breger H (2008) Leibniz's calculation with Compendia. In: Goldenbaum and Jesseph (2008), pp 185–198
Bussotti P (2019) Review of Arthur 2019 for Zentralblatt. See https://zbmath.org/?q=an%3A07217301
Cavalieri B (1966) Geometria degli indivisibili (ed: Lombardo Radice L). Utet, Torino
Duchesneau F (2008) Rule of continuity and infinitesimals in Leibniz's physics. In: Goldenbaum and Jesseph (2008), pp 235–253
Duhem P (1913–1956) Le Système du Monde. Hermann, Paris
Duhem P (1985) Medieval cosmology. Theories of infinity, place, time, void and the plurality of worlds. Edited and translated by Roger Ariew. The Chicago University Press, Chicago/London
Esquisabel O, Raffo Quintana F (2021) Fiction, possibility and impossibility: three kinds of mathematical fictions in Leibniz's work. Arch Hist Exact Sci 75(6):613–647. https://doi.org/10.1007/s00407-021-00277-0
Garber D (2008) Dead force, infinitesimals, and the mathematicization of nature. In: Goldenbaum and Jesseph (2008), pp 281–306
Garber D (2009a) Leibniz: body, substance, monad. Clarendon Press, Oxford
Garber D (2009b) What did Leibniz learn about body in January 1678? In: Kulstad M, Laerke M, Snyder D (eds) The philosophy of the Young Leibniz. Studia Leibnitiana Sonderhefte 35. Franz Steiner, Stuttgart, pp 67–82
Garber D (2016) Monads on my mind. In: Leibniz's metaphysics and adoption of substantial forms. Between continuity and transformation. The new synthese historical library. Springer, Dordrecht
Goldenbaum U, Jesseph D (2008) Infinitesimal differences. Controversies between Leibniz and his contemporaries. Walter de Gruyter, Berlin/New York
Hofmann JE (1974) Leibniz in Paris 1672–1676. His growth to mathematical maturity. Cambridge University Press, Cambridge
Katz M, Sherry D (2013) Leibniz's infinitesimals: their fictionality, their modern implementations, and their foes from Berkeley to Russell and beyond. Erkenntnis 78(3):571–625. https://doi.org/10.1007/s10670-012-9370-y, https://arxiv.org/abs/1205.0174
Katz M, Kuhlemann K, Sherry D, Ugaglia M (2021) Leibniz on bodies and infinities: rerum natura and mathematical fictions. Rev Symb Log 1–31. https://doi.org/10.1017/S1755020321000575
Knobloch E (1994). The infinite in Leibniz's mathematics – The historiographical method of comprehension in context. In Trends in the historiography of science, pp. 265–278. Gavroglu K, Christianidis J, Nicolaidis E (eds.), Dordrecht: Kluwer.
Kretzmann N (1982) Syncategoremata, Exponibilia, Sophismata. In: Kretzmann et al. (1982), pp 211–245
Kretzmann N, Kenny A, Pinborg J (1982) The Cambridge history of later medieval philosophy: from the rediscovery of Aristotle to the disintegration of scholasticism 1100–1600. Cambridge University Press, Cambridge
Leibniz GW (1875–90) Die philosophischen Schriften. Edited by Carl Immanuel Gerhardt. Weidmann, Berlin (cited as [GP])
Leibniz GW (1923–) Sämtliche Schriften und Briefe. Akademie, Darmstadt/Leipzig/Berlin (cited as A. series, volume, page). See https://leibnizedition.de

Leibniz GW (1949–1863) Die mathematische Schriften. Edited by Carl Immanuel Gerhardt. Asher and Halle: H.W. Schmidt, Berlin (cited as [GM])
Leibniz GW (1989a) Philosophical essays. Edited and translated by Roger Ariew and Daniel Garber. Hackett, Indianapolis (cited as [AG])
Leibniz GW (1989b) Philosophical papers and letters, 2nd edn. Synthese historical library, vol 2. Leroy E. Loemker, Editor and Translator. Kluwer Academic, Dordrecht/Boston/London (cited as [L])
Leibniz GW (2007) The Leibniz – Des Bosses correspondence. Translated, Edited, and with an Introduction by Brandon C. Look and Donald Rutherford. Yale University Press, New Haven
Mercer C (2001) Leibniz's Metaphysics: Its origins and development. Cambridge University Press, Cambridge
Moore AW (1990) The infinite. Routledge, London
Parmentier M (1995) G. W. Leibniz: la naissance du calcul différentiel, 2nd edn. Introduction, traduction et notes par M. Parmentier. Vrin, Paris
Pasini E (2016) Mathematical similes in Leibniz's *Theodicy*. In: Camposampiero MF, Geretto M, Perissinotto L (eds) Theodicy and reason logic, metaphysics, and theology in Leibniz's Essais de Théodicée (1710). Edizioni Ca' Foscari, Venice, pp 63–83
Peter of Spain (2014) Summaries of logic. Text, translation and notes by Brian P. Copenhaver, Calvin Normore and Terence Parson. Oxford University Press, Oxford
Rabouin D, Arthur R (2020) Leibniz's syncategorematic infinitesimals II: their existence, their use and their role in the justification of the differential calculus. Arch Hist Exact Sci 74:401–443
Rescher N (2013) Studies in Leibniz's cosmology, vol 13. De Gruyter, Berlin/Boston. https://doi.org/10.1515/9783110332384.155
Rutherford D (2008) Leibniz on infinitesimals and the reality of force. In: Goldenbaum and Jesseph (2008), pp 255–280
Sherry D, Katz M (2014) Infinitesimals, imaginaries, ideals, and fictions. Stud Leibnitiana 44(2): 166–192
Uckelman SL (2015) The logic of categorematic and syncategorematic infinity. Synthese 192: 2361–2377
Ugaglia M (2009) Boundlessness and iteration: some observations about the meaning of ἀεί in Aristotle. Rhizai 6(2):193–213
Ugaglia M (2018) Existence vs conceivability in Aristotle: are straight lines infinitely extendible? In: Piazza M, Pulcini G (eds) Truth, existence and explanation – FilMat 2016 studies in the philosophy of mathematics. Boston studies in the philosophy and history of science 334. Springer, Cham, pp 249–272
Ugaglia M (2022) Possibility vs iterativity: Leibniz and Aristotle on the infinite. In: Ademollo F, Amerini F, De Risi V (eds) Thinking and calculating. Essays in logic, its history and its philosophical applications in honour of Massimo Mugnai. Springer, Berlin, pp 255–270
Whipple J (2010) The structure of Leibnizian simple substances. Br J Hist Philos 18(3):379–410

René Thom: From Mathematics to Philosophy

Athanase Papadopoulos

Toute ma métaphysique sous-jacente, c'est d'essayer de transformer le conceptuel en géométrique, le logique en dynamique.
(René Thom, from a letter to Claire Lejeune, February 20 1980, quoted in (Lejeune 1978))

Contents

1 Introduction 386
2 *Vita*: From Mathematics to Philosophy 387
3 Thom, Bourbaki and Mathematical Education 396
4 Qualitative and Quantitative 400
5 Thom's Rediscovery of Aristotle 403
6 Topology from an Aristotelian Perspective 406
7 Aristotelian Biology 408
8 Philosophy of Singularities 411
9 Topology and Ethology 413
10 Linguistics and Morphology 413
11 Psychoanalysis 415
12 A Metaphysics of Analogy 417
13 Another Return to the Classics 419
14 Conclusion 422
Appendix: Jean Cavaillès and Albert Lautman 424
References 425

Abstract

In this chapter, I will discuss René Thom's approach to philosophy based on his mathematical background. At the same time, I will highlight his connection with Aristotle, his criticism of the modern view of science as a predictive process, his

A. Papadopoulos (✉)
Institut de Recherche Mathématique Avancée (IRMA) and Centre de Recherche et d'Expérimentation sur l'Acte Artistique (ITI CREAA), Université de Strasbourg and CNRS, Strasbourg Cedex, France
e-mail: athanase.papadopoulos@math.unistra.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_98

ideas on mathematical education, his position with respect to the French school of mathematics that was dominant in his time, and his relationship with the philosophical community. I will also touch upon the connections between Thom's ideas and those of Leibniz, Riemann, Freud and others.

Keywords

Reni Thom · Aristotle · Biology · Predictive science · Singularities · Morphology form · Morphogenesis · Analogy · Bourbari · Catastrophe theory · Linguistics · Cobordism · Structural stability · Hylemorphism · Homeomerous · Stratification · History of topology · Morphogenesis form · Morphology · Qualitative science

1 Introduction

René Thom belongs to a line of preeminent mathematicians who were thoroughly engaged in philosophy. After a brilliant but short career in mathematics culminating in the Fields medal, he devoted himself to philosophy, in the context of a return to the ideas of Aristotle, the philosopher of nature, of whom he considered himself a heir. Thom had a monumental program: to lay down the foundations of a new natural philosophy based on mathematics, and more particularly on geometry and topology. Natural philosophy, from his point of view, included physics, embryology, optics, hydrodynamics, biology, linguistics, ethology, semantics, psychoanalysis and other fields.

With his philosophical broad program, Thom was continuously an intriguing figure for linguists, psychoanalysts, biologists and philosophers of sciences, let alone mathematicians. He opened up with these fields a dialogue which, in the history of mathematics, is unprecedented by the width of its spectrum. Philosophers of various sciences discussed and argued with him on technical grounds, defending their ways of thinking while trying to understand his original program, which, after all, seems quite natural for a mathematician: describing nature using the fundamental ideas of topology, from the most basic notions of open and closed set, boundary, singularity of a differentiable function, to the more elaborate concepts of cobordism, structural stability, analytic continuation, etc.

Thom's philosophical program was initiated in his first published book, *Stabilité structurelle et morphogénèse* (Structural stability and morphogenesis) (Thom 1972), published in 1972, and it was continued in his *Modèles mathématiques de la morphogenèse* (Mathematical models of morphogenesis) (Thom 1974b), published in 1974 and his *Esquisse d'une sémiophysique: Physique aristotélicienne et théorie des catastrophes* (Smiophysics, a sketch; Aristotelian physics and catastrophe theory) (Thom 1990b) published in 1988. He wrote other books with a strong philosophical flavor. He also published a large number of papers on philosophy. I will quote several of them in the present chapter, and there are many more.

Thom explained at several occasions the reasons for his passage to philosophy, and I shall dwell on them in what follows. He was also thoroughly involved in the reflection on mathematical education, and he participated to the major educational debates of his time, in particular, on the introduction of the so-called modern math in schools, in France and elsewhere, and on the new university programs. His article

Les mathématiques modernes: une erreur pédagogique et philosophique? (Modern mathematics: an educational and philosophical error?) Thom (1970) was translated into several languages and published in various journals. In both domains of pre-university and university programs, he was a defender of classical geometry, against the exaggerated weight given to algebra, set theory, information science and applied mathematics, and I will also comment on this in what follows.

The plan of the rest of this chapter is the following. In Sect. 2, I report on a few important moments from Thom's life. In particular, I review some events that were foreshadows of his future career as a mathematicians, and on certain factors that led him to abandon mathematical research and devote himself to philosophy. Section 3 is intended to situate Thom with respect to the Bourbaki group who was leading mathematical research and setting the rules for mathematical education in France, starting in the 1950s. At the same time, I indicate some of Thom's ideas on education and on mathematical writing. Section 4 is an overview of Thom's belief that, contrary to a common modern dogma, science is not quantitative but always qualitative, and of the debate that this belief generated among scientists. In Sect. 5, I comment on the relation of Thom's ideas with those of Aristotle, a filiation he has declared on many occasions and which he explains in several articles. In Sect. 6, I comment on the different assertions of Thom according to which he found in the philosophical ideas of Aristotle his own ideas on form, morphogenesis, and also the first notions of topology. In Sect. 7, I explain how Thom shared Aristotle's ideas on biology as a qualitative science, his ideas on hylemorphism, on homeomerous, and on anhomeomerous parts, and how he considered his program in this field to be also in the lineage of the one of the Stagirite. Thom established relations between the mathematical notion of stratification and other topological notions with ideas and concepts that Aristotle introduced in his work on the parts of the animals. Section 8 is concerned with Thom's ideas on the role played by the mathematical theory of singularities in the general framework of prediction in science. The theory of singularities becomes, from this point of view, a philosophical topic. Then follows a short Sect. 9, which contains a few notes on Thom's ideas on morphology applied to this field. The next two Sects. 10 and 11 respectively, are concerned with Thom's work on these two fields, which he viewed as domains of applications of his ideas on morphology. Section 12 is a glimpse into Thom's ideas on analogy, again, in the tradition of Aristotle. In the last section before the conclusion of this chapter, titled *A return to the classics*, I mention some ideas of three mathematicians that were predecessors of Thom's reflections, namely, Leibniz, Riemann and Grassmann, pointing out some connections with Thom's ideas.

2 *Vita*: From Mathematics to Philosophy

René Thom, at several occasions, recounted some events from his childhood that are related to the shift that occurred later in his intellectual activity, from mathematics to philosophy. I have already commented on this in the long biography (Papadopoulos 2018a) that I devoted to him in the first chapter of the book (Papadopoulos 2018b). In this section, I will give a short *Vita* of Thom, highlighting some episodes from his life that are connected with his journey from mathematics to philosophy.

René Thom was born in 1923 in Montbéliard, a small town in the East of France, now in the Burgundy-Franche-Comté region. The city was occupied by the Germans between 1940 and 1945[1], the period in which Thom was finishing secondary school and entering higher education. His parents owned a grocery store in Montbéliard. His father had a solid education, with a good knowledge of Latin, and he wrote poetry. Thom used to say that he gained his open-mindedness from having met people with very different backgrounds at the family grocery store. His parents encouraged him and his siblings to get a higher education.

Thom had the profile of a studious French schoolchild. From what he recounted, we know that he was not particularly brilliant at school, but some of his recollections from that period are harbingers of his life as a future mathematician, and I will mention some of them.

The first reminiscence dates from age ten or eleven, when Thom had set himself the task of systematically exploring what every theorem he knew on the geometry of three-dimensional space becomes in dimension four. He writes about this in an interview published in 1989 (Nimier):

> [This] was, if I may say so, my first attempt to do something original; but it was a way for me to understand, say, how a system of two planes in $\mathbb{R}^4$ was made, etc., and I think I had arrived to a pretty good intuition at that time. I was already seeing in four-dimensional space.[2]

The second fact I mention took place when Thom was thirteen or fourteen. We read in (Porte) that he came across a book on differential calculus in the public library of his hometown and was seduced by it. But Thom's main interest was Euclidean geometry. He writes in (Thom 1991c, p. 10): "I had a certain taste for Euclidean geometry which immediately attracted me very much." On the other hand, he had no taste for algebra, which, he says, was akin to certain automatisms. He later maintained this opinion on this field, a field which he considered "not very instructive," and he kept his ardor for Euclidean geometry, "our old Euclidean geometry," as he used to call it affectionately. On the same subject, he writes in (Thom 2004):

> There was one encounter that was quite decisive, and this was with Euclidean geometry in the ninth and tenth grades, where I really developed a taste for geometric thinking and proof.[3] I had made it a point of honor to find solutions to all conceivable problems in

[1] The city of Montbéliard oscillated a few times between France and Germany. From 1042 to 1793, it was part of the Holy Roman Empire, after which it was attached to France. It was occupied by the Prussians during the Franco-Prussian War of 1870, and liberated one year later by the so-called Bourbaki army, named after Charles-Denis Bourbaki (1816–1897), a general of the army of Napoleon III and son of the Greek colonel Constantin-Denis Bourbaki. The latter was educated in France and he served in the French army between 1787 and 1827. The pseudonym Nicolas Bourbaki, which was chosen randomly, refers to the name of general Bourbaki.

[2] The translations from the French are mine. For some more subtle sentences, I have included the French original in a footnote.

[3] I have transformed "classe de troisième et de seconde" of the French educational system into their equivalent in the American one.

> elementary geometry, in the geometry of the triangle, and this is why, since that time, I have kept a preference for Euclidean geometry that my colleagues and fellow mathematicians do not forgive me.

Besides geometry, Thom was intrigued by dynamics. In the interview (Nimier), he recalls his early interest and his motivation for this subject which, a few years later, became one of his favorite topics of reflection:

> Around age seventeen, I started being interested in dynamics. I don't remember on what occasion it was, but I had handed in a paper to my mathematics teacher, where I talked about the eternal return seen from a dynamical point of view, the theories of the eternal return. . . It was the idea that one could have of a space-time, a universe in which there would be the eternal return, that is to say, where the dynamics would be periodic, but I believe that it was almost the first time that I really thought things in terms of dynamics.

During his high-school years, Thom was equally fond of literature, poetry, history and science. He especially enjoyed reading Greek and Latin classical authors. In 1939, the question arose as to the choice of the second part of the *baccalauréat*.[4] Thom hesitated between the literature curriculum (which, in France, was called the *philosophy* section) and the science one (called the *elementary mathematics* section). He opted for the second. He explained his choice in his interviews with Émile Noël (Thom 1991c, p. 9):

> We knew that this section offered more opportunities than the first one; it was perhaps an illusion, but we were convinced of that. Above all, we were in 1939, at the beginning of the war, and our parents, who had fought during the First World War, used to tell us: "try be a artilleryman, you are less exposed than in the infantry!" To be an artilleryman, you had to have studied mathematics. This element probably weighed heavily in the birth of my mathematical vocation.

Thom obtained his *baccalauréat 2ème partie* in June 1940, in a *lycée* in Besançon, the large city close to Montbéliard. He was 16.[5] The period was hard for the Thom family, as for many in Europe, because of the Second World War. Furthermore, in those years, studying at high school represented a major cost for families. Thom, as a gifted child from a modest background, had obtained a small government scholarship for his education. We read in an interview reported on by Michèle Porte (Porte) that Thom's parents were really struggling to make their financial needs meet and to put their children through education.[6]

After his *baccalauréat 2ème partie*, Thom, together with his older brother, managed to cross the Swiss border, escaping from the Germans who had occupied

[4] In the old educational French system, the "baccalauréat," that is, the national high school diploma, had two parts. The "baccalauréat 2ème partie" corresponded to the last year in high school.

[5] Finishing school at age 16 is not unusual in France for a talented child.

[6] "Mon père a tiré le diable par la queue toute sa vie. . .Mes parents se sont vraiment saignés pour me faire faire des études."

Montbéliard and the region around it. Their stay in Switzerland was hard for they were penniless, and Thom recounts that they suffered there from starvation. With the help of the French authorities, the two brothers returned to France, and they ended up in Lyon, which was still in the free zone.[7] There, with the help of a family friend, who at the same time was a remote cousin, René Thom enrolled again in a high school, since he had nothing better to do, and he presented for the second time the *baccalauréat 2ème partie*, this time in the philosophy section.

After Thom graduated anew from high school, the preparation of the competition for entering the École Normale Supérieure was a natural choice for a talented young man like him, interested in science, literature, and philosophy. In France, such a preparation is offered by some selective high schools,[8] and Thom was admitted at the *Lycée Saint-Louis*, a prestigious secondary institution in Paris.[9] The preparation lasts in principle 2 years, in classes called *classe de mathématiques supérieures* and *classe de mathématiques spéciales*. Thom spent 3 years at the *Lycée Saint-Louis*, as his first attempt to enter the École Normale was unsuccessful and he had to repeat the year of *mathématiques spéciales*.

At several occasions, Thom recounted that right after he was admitted to the École Normale, he announced to the school director that rather than studying mathematics, his desire was to study philosophy of science, precisely, "in the path of Cavaillès and Lautman." The acting director at that time was the physicist Georges Bruhat, who strongly advised Thom not to follow this path, saying that it would be more useful for him to concentrate his efforts on the preparation of the *agrégation* in mathematics, cf. (Nimier).[10]

At the end of this chapter, I have included an appendix on the two philosophers of science Jean Cavaillès and Albert Lautman.

Thom's teacher at the École Normale was Henri Cartan, who took him under his tutelage and who played later an important role for him, especially at the beginning of his career as a mathematician.

Thom finished his studies at the *École* in 1946, after having passed the competitive examination of the agrégation. On the recommendation of Cartan, he was appointed to the University of Strasbourg.[11]

Thom was expected to submit a PhD thesis in the years that followed his appointment, and his first task was to find an appropriate research subject. He chose the field of algebraic topology. He explained later that it was easier for a beginner like him to work in a new discipline such as this one, where the ground to

[7] The city of Lyon was occupied later by the Germans, in November 1942.

[8] This is again peculiar to the French system, where the preparation of the entrance exams to the so-called *grandes Écoles* is done in specialized classes provided by some high schools, after the usual secondary school program.

[9] The *Lycée Saint-Louis*, before Thom enrolled there, had among its pupils André Weil, Laurent Schwartz, Gustave Choquet and several other preeminent mathematicians.

[10] This is a French diploma allowing, in principle, to become a teacher in a high school, but many students who plan to teach at university also pass this exam.

[11] In those times, one could enter academia in France without having a doctorate.

be explored was virgin and where there were still easily accessible open problems (Thom 1992d).

Thom's work on cobordism, which was to ripen in his mind in the few years following his arrival in Strasbourg, inaugurated the implementation of the methods of differential topology in algebraic topology, so that the subject of topology of manifolds became a combination of the two, differential and algebraic topology. In fact, at the time Thom entered the field of topology, both points of view (the algebraic and the differential) were in a state of rapid expansion and were undergoing major transformations.[12] Algebraic topology and the study of differentiable functions occupied Thom during his whole mathematical career. He obtained his doctoral degree in 1951, with a dissertation titled *Espaces fibrés en sphères et carrés de Steenrod* (Thom 1952). Cartan, who was officially Thom's thesis advisor, became involved in this work after Thom had written a first version, in order to correct the writing and make it intelligible.[13] Several letters exchanged between Thom and Cartan, edited by André Haefliger in the form of an article in volume I of Thom's Mathematical Collected Works (Thom 2017), concern the finalization of the writing of this thesis. The article contains excerpts from these letters, which show that Cartan strongly committed himself during the last months before the defense, to make the thesis understandable. The general impression that emerges from this correspondence, as well as from some letters dating from the previous year concerning a paper that Thom had submitted to Cartan for publication in the *Comptes Rendus de l'Académie des Sciences* (and which was eventually published in the proceedings of a conference), is that Cartan did not consider Thom capable of writing complete proofs. The correspondence between Thom and Cartan also shows that the latter consulted Samuel Eilenberg about Thom's dissertation and that Eilenberg began by expressing serious reservations about the validity of the proofs. In particular, in a letter to Thom, Cartan writes that Eilenberg "believes that it is absolutely impossible to know whether the theorems in question are correct, and consequently he considers

[12] Thom entered topology just after a major revival of this field that started in the 1930s. Topologists tried to classify manifolds and were looking for algebraic invariants. The notions of fibre space and fibre bundle became central in algebraic topology around the year 1950, and the topology of manifolds passed important milestones in the space of very few years: In 1956, Milnor proved the existence of exotic differentiable structures on 7-dimensional spheres. In 1961, he disproved the so-called *Hauptvermutung der kombinatorischen Topologie* ("main conjecture of combinatorial topology"), a conjecture formulated in 1908 par Steinitz et Tietze, asking whether any two triangulations of homeomorphic spaces are isomorphic after subdivision. The higher-dimensional analogue of the Jordan curve theorem (the so-called Jordan–Schoenflies theorem) was obtained in 1959–1960 (works of Mazur, Morse and Brown). In 1961, Smale proved the Poincaré conjecture for dimensions ≥ 7. The same year, Zeeman obtained the corresponding result for dimension 5, and the year later, Stallings for dimension 6. Thom considers that he arrived in topology at the right time. He writes in (Thom 2004): "I think that researchers who started in mathematics after 1960 have had infinitely more merit than than people like me who came along at just the right time."

[13] In those days, the task of the thesis advisor was generally that of a guarantor at the moment of the defense. In general, suggesting the subject of the thesis was not part of his role. At any rate, in Thom's case, it was not Cartan who proposed the topic of the work. I discussed this at length in the article (Papadopoulos 2018a).

that it would be very imprudent to publish them as they are in a thesis, without a true proof." We note incidentally that among the changes that Cartan advised Thom to make to his thesis, he recommended the elimination of a complete chapter which seemed to him very unclear and almost impossible to make readable in order to defend the thesis in due time,[14] hoping that Thom could at least work properly on the rest; in fact, this was a chapter on cobordism, a notion that Thom had introduced, and which became later the ground of the work for which he was awarded the Fields medal.

From an exterior point of view, the years Thom spent in Strasbourg (1947–1963) were those of his purely mathematical activity. In his interviews *Prédire n'est pas expliquer*, talking about this period, he summarizes it as, in the first place, that of an abandonment of philosophy in favor of mathematics, and also a period ending with a passage to mathematics as a tool for the description of the real world. This is how mathematics is used in *catastrophe theory*, a theory which he started to develop under the name *model theory*, whose goal was to provide a mathematical framework to ideas from all domains of knowledge, including the physical, social and biological ones.[15] He writes (Thom 1991c, p. 24):

> From 1950 on, I abandoned these philosophical preoccupations to devote myself to mathematics. This lasted until 1956-1957. Then I went through a phase of depression: progress in mathematics was made by others. They led to theories so algebraically complicated that I could not follow them anymore. I gave up. . . . But one must still do something! So I started to look for possible applications of the mathematical theories I knew. That is how I oriented myself towards catastrophe theory.

By the end of his stay in Strasbourg, Thom came across Christopher Zeeman's article *Topology of the brain*, an article which opened for him the door to new possibilities of representing the phenomena of life by mathematical models. He writes in his introductory talk of the 1982 symposium *Logos and Catastrophe Theory* (Thom 2004):

> The idea of biological modeling, I owe it to a certain extent to Christopher Zeeman who, in 61, had published an article called Topology of the brain, where he had shown the great theoretical possibilities of mathematical modeling of the most complex biological activities. I was very impressed by this article at the time.

At the Edinburg ICM (14-21 August 1958), Thom was awarded the Fields medal for his work on cobordism. Later, he described this problem as "to know when two manifolds can be deformed into each other without having singularities at any moment in this deformation." He considered that he had partially solved this problem. Concerning the prize itself, he declared in the interview *Paraboles et*

[14] This was in the summer of 1951. Cartan had arranged an invitation for Thom to Princeton for the following fall, and the thesis had to be defended before.

[15] The sub-title of Thom's book *Stabilité structurelle et morphogénèse* is *Essai d'une théorie générale des modèles* (An attempt for a general theory of models).

catastrophes (Thom 1980, p. 23): "It was a bit of a good fortune. I seized the right moment by being able to exploit the techniques that Cartan, Ehresmann, Serre and many others had taught me." He wrote in his autobiographical article (Atiyah and Iagolnitzer 1977) that ultimately, this prize left him with a bitter taste, because the work done was too little, compared to what others did after him. "In this sense," he says, "this prize represented for me a certain fragility that the future made even more visible. The consequence, if there was one, was only the invitation to the new Institut des Hautes Études Scientifiques." Indeed, Thom was offered a research position at this newly founded institute.[16] In (Thom 1991c, p. 27), he writes:

> I left Strasbourg and came, answering the call of the founder of the IHÉS. I had more leisure time; I was less preoccupied with teaching and administrative tasks. My purely mathematical productivity seemed to be on the wane and I began to focus more on the periphery, i.e., possible applications.
>
> In addition to optics, I wondered if there were some possible applications to biology. I finally understood where these exceptional singularities came from: they are linked to the fact that the trajectories of light rays are special, because they satisfy a variational principle, Fermat's principle; they have special properties that make caustics "catch" singularities more easily than they should. It is the transfer of this idea that led me to the theory of catastrophes, to the mathematical part that leads to it.

Thom declared at several occasions that his interest in forms arose from his observation of caustics.

Thom later had another "revelation," during a visit to the Poppelsdorfer museum in Bonn, when he saw plaster cast models of the various phases of the development of a frog embryo in the process of gastrulation. In particular, he noticed that some lines in these representations were, from the mathematical point of view, singularities that were familiar to him and he realized that the whole process may be interpreted using topology and geometry. He gave an interpretation of this beautiful geometry in terms of the unfolding of a wave front in an appropriate space that projects onto the ordinary space and later, he described all the process of development of the frog as a sequence of successive unfoldings of a singularity. This was the starting point for the application of his ideas on singularity theory to embryology (Thom 1991c, p. 111).

[16] The institute was founded by Léon Motchane (1900-1990), a Russian-born immigrant who started with small jobs in Switzerland and then Germany and who later became a wealthy businessman in banking and insurances. Motchane liked mathematics and he started studying them relatively late in his life, encouraged by Paul Montel. At age 54, he obtained a doctorate, with Gustave Choquet as supervisor, and he decided to dedicate an important part of his time and fortune to the foundation of a mathematical institute. He managed to obtain funding from several large private companies. The institute was founded in 1958, and Thom was among the first professors hired. The Institute was situated in Paris, in a two-room apartment in the Thiers Foundation, located in a private mansion in the 16th arrondissement, before moving four years later to Bure-sur-Yvette, in the domain called Bois-Marie, which became its permanent location. Today, the French government is the main source of funding for the IHÉS. Motchane represented Grothendieck for his reception of the Fields medal at the Moscow ICM (1966).

Thom's stay in Strasbourg lasted from 1947 to 1963, with a few breaks for visits he made to the United States, each for several weeks, and an interruption during the academic year 1953–1954, which he spent in Grenoble.

Thom introduced the word *catastrophe* in his book *Stabilité structurelle et morphogénèse* (1972), to denote an abrupt phenomenon that appears and disappears continuously. The work quickly became the reference for the theory that became known under the name *Catastrophe theory*. Speaking of this theory, Thom recalls the role of the optical experiments he conducted in 1959-1960 in Strasbourg as an impetus. He declares in his interview (Thom 1991c, p. 26):

> I came to [catastrophe theory] in a rather natural way: it was an evolution that led me, starting from a purely mathematical problem, the one called the generic singularities of a map, to see whether this theorem had physical applications. I was then at the University of Strasbourg and I had the opportunity to do some optical experiments. A colleague, who was a physicist, lent me some instruments, a spherical mirror, a prism and a dioptre. I was able to produce some caustics, and I made them vary by changing a little bit the position of the parameters; I observed how they were deformed. That is precisely the thing that interested me. [. . .] From there I started the theory of catastrophes: I realized that there were types, variations of caustics, that I did not foresee, and that I had to explain to myself the appearance of these singularities. It took me two or three years to understand where this came from. This is a banal phenomenon, but that is where I started from.

Optical caustics are curves that appear as envelopes of a beam of light rays reflected in a curved mirror. Later, in 1991, in his lecture titled *Leaving mathematics for philosophy* (Thom 1992c), Thom recalled that his first surprise during his experiments on optics was to see that the caustics obtained using a spherical mirror or a linear diopter give rise to singularities that he had not expected to see.

In 1961, during one of his trips to the United States, Thom met Salomon Lefschetz, who was working on structural stability, a notion that was introduced in the 1930s by Aleksandr Andronov and Lev Pontryagin. Several years later, in his introductory talk of the 1982 conference *Logos et théorie des catastrophes* (Logos and catastrophe theory) (Thom 2004), Thom wrote, regarding Lefschetz: "He understood very well the crucial importance of this notion and he did a lot to develop this theory in the Western circles." Later on, Thom used this concept in a decisive way in his mathematical and philosophical work, and it became a central element of his theory of morphogenesis. In fact, the title of his first book, published in 1972, whose aim is to provide a mathematical formalism to tackle general morphogenesis problems, is precisely *Stabilité structurelle et morphogénèse* (Thom 1972). He writes in the introduction: "This book is the first systematic attempt to think in geometrical and topological terms the problems of biological regulation, as well as those posed by the structural stability of any form." Thom, whose main objective became to understand form in nature, considered that only structurally stable morphologies and phenomena can be understood. He even tried to formalize

thermodynamics using the notion of structural stability, although he realized later on that he failed, and that some weaker notion of structural stability was needed.[17]

Let me make a parenthetical remark, concerning physics regarded by philosophers from Greek Antiquity, since, as we shall recall below, Thom often referred to them. Plato considered that there cannot be any knowledge of nature, that is, that physics does not exist, since the physical world is permanently subject to change, and it is impossible to have a real knowledge of something which is perpetually changing. Aristotle, on the contrary, in his *Physics*, was puzzled by evolution, and by things that are changing. His *Physics* is precisely a physics of motion and of change. Structurally stable phenomena makes a compromise between changing and unchanging objects. A dynamical system is structurally stable if, at each moment of its evolution, there is an interval of time around this moment such that the system obtained at any given time in this interval is identical to the original system.[18]

Thom's main philosophical program, that of describing nature and language in geometric terms, is already largely initiated in his book *Stabilité structurelle et morphogénèse* (1972). He writes in the Preamble: "This is a work that wishes to place itself in the line of a deceased discipline, namely *Natural philosophy*," and the reference to Aristotle is clear. A good half of this book is devoted to questions of embryology, where Thom proposes a formulation of this field using a geometric language.

In his second book, titled *Modèles mathématiques de la morphogenèse* (Thom 1974b), published in 1974, Thom defines morphogenesis as "the creative or destructive process of forms" (Thom 1974b, p. 10). Thom declares that his research on morphogenesis is based on two sources: on the one hand, his research in differential topology on structural stability, and on the other hand, his reading of embryology treatises. He considered that the applications of the theory he proposes go far beyond embryology and even biology, to include subjects such as physical chemistry, optics, hydrodynamics and others.

At IHÉS, Thom continued publishing on mathematics, but not in a classical style, that is, not on new theorems and proofs. Most of his mathematical works were addressed to biologists, philosophers, linguists and other nonmathematicians, and some of these articles were published in philosophical journals. Thom declared later that his drift toward biology and the other sciences was due in part to the presence of Alexander Grothendieck at IHÉS. He writes on this subject, in an autobiographical text (Atiyah and Iagolnitzer 1977, p. 82):

[17] Thom wrote: "initially, the theory of structural stability seemed to me to be of such scope and generality that with it, I could hope to replace thermodynamics by geometry, to geometrize in a certain sense thermodynamics, eliminating from thermodynamical considerations all the stochastic and measurable aspects, to retain only the corresponding geometric characterization of the attractors. It is certain that the instability phenomena of the attractors that have been discovered since then show that such a hope is false or, at least, that it would be necessary to weaken considerably the notion of structural stability." (Thom 2004, p. 587) (1989).

[18] The mathematical formulation of this notion asserts the existence of a conjugacy between a system and close-enough systems so that in some sense the two systems are similar.

> [. . .] Relations with my colleague Grothendieck were less agreeable for me. His technical superiority was crushing. His seminar attracted the whole of Parisian mathematics, whereas I had nothing to offer. That made me leave the strictly mathematical world and tackle more general notions, like the theory of morphogenesis, a subject which interested me more and more and led me towards a very general form of "philosophical" biology. The director, L. Motchane, had no objection (or if he did he kept it for himself). And so I ran what was known as a "crazy" seminar, which lasted the best part of my first year at the institute. Three or four years passed, during which I returned to Aristotle and classical Greek.

3 Thom, Bourbaki and Mathematical Education

At the École Normale Supérieure, Thom received from Cartan an education in the purest Bourbaki tradition. Soon after, in Strasbourg, he was seduced by ideas of Ehresmann, another eminent member of the Bourbaki group. But later, he became no less than an acerbic critic of the spirit that reigned in that group to which he owed his mathematical training. In his interviews *Paraboles et catastrophes* (Thom 1980*, p. 23*), reporting on a Bourbaki session to which, at the request of Cartan, he participated while the was still a student at the Énormale, he writes:

> According to the tradition, I had to attend a presentation of the texts and I was asked for my opinion in case of disagreement. . . . I was tired of listening (it was a bit boring!) and sometimes I even fell asleep during the sessions. In a sense, this was not a bad thing after all: later, in fact, I became an opponent of the Bourbaki view of mathematics. . . . They were really ultra-formalists! They presented mathematics in a rigorous, ascetic manner; but in doing so, they ended up dealing only with those portions of mathematics that were appropriate to this rigorous presentation. For this reason, much of the mathematics that was practiced at the time could only escape their attention. . . . Even today there are large sections of mathematics — the calculus of variations, partial differential equations, qualitative dynamics, etc. – that are considered as not "clean" enough (or "dead"!) to be included in the Bourbaki anthology. Thus, there is a genuine incompatibility between the Bourbaki repertoire and living mathematics. Bourbaki has embalmed the mathematics that he has reduced, so to speak, to a mummy!

Rigor, systematization, and axiomatization were among the notions at the foundation of the Bourbaki style. In the article *Modern mathematics: a pedagogical and philosophical error?* (Thom 1970), Thom expresses a certain retreat with respect to formalization, and he speaks of the limits of axiomatization. He protests against the disappearance of traditional Euclidean geometry (the geometry of the triangle) from French curricula, in favor of set-theoretic and algebraic theories, where students no longer take initiatives to solve construction problems, but learn abstract notions and apply blind algebraic recipes from a pre-established scheme. He then comes to the notion of "meaning," and he raises the question of why we do mathematics. Is it a gratuitous game, a random product of our brain activities, or is it a way to comprehend the universe? At the same time, he wonders about the meaning of the word "rigor" in mathematics. His answer is: "A proof is rigorous, if it arouses, in any sufficiently educated and prepared reader, a state of evidence that leads to adhesion" (p. 230). This is not the orthodox view on rigor, where a proof is rigorous if it consists of a logical sequence of implications, starting from well-known statements.

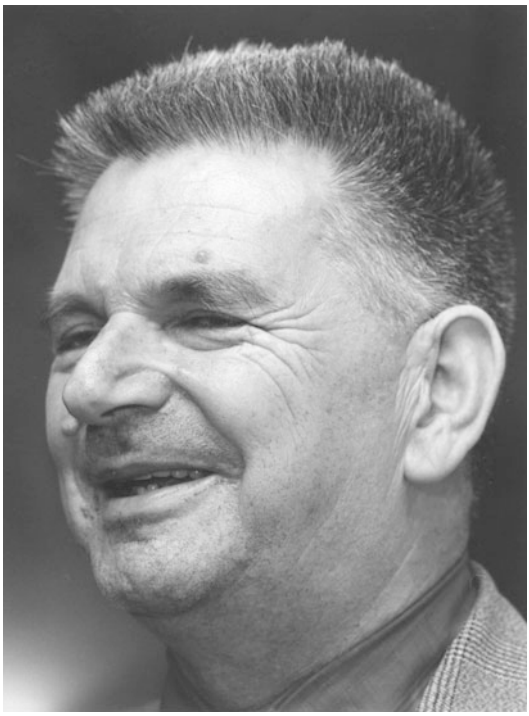


Fig. 1 René Thom in 1985. (Photo Gerd Fischer, Munich. Copyright Gerd Fischer. Source: Archives of the Mathematisches Forschungsinstitut Oberwolfach)

He recalls that the etymology of the word "theorem" is "vision," declaring that "there is no need for great axiomatic constructions or refined conceptual machineries to judge the validity of a reasoning." He writes (p. 232):

> There has been a lot of talking in recent years about the importance of axiomatization as an instrument of systematization and discovery. Yes, an instrument of systematization, certainly; but for what concerns discovery, the matter is more than doubtful. It is characteristic that, from the immense effort of systematization of Nicolas Bourbaki (which, by the way, is not a formalization, because Bourbaki uses a non-formalized metalanguage), no new theorem of any importance has emerged. And if researchers in mathematics refer to Bourbaki, they find their food much more often in the exercises—where the author has pushed the concrete material—than in the deductive body of the text. This must be said quite clearly: axiomatization is a research intended for specialists, and has no place either in secondary teaching nor at university (except, of course, for professionals who wish to specialize in the study of the foundations). This is why accusations of inconsistency addressed to Euclidean geometry are in fact irrelevant at the level of the local intuitive validity of the reasoning, which is the only level that matters.

On p. 236 of the same article, he writes:

> It is not certain that even in pure mathematics, every deduction can have a set-theoretic model. The ill-tamed paradoxes that undermine formal set theory are there to remind the mathematician of the dangers of the exaggerated use of these seemingly innocent symbols.

> Perhaps, even in mathematics, quality remains, and resists any set-theoretic reduction. The old Bourbakist hope, to see the mathematical structures emerge naturally from the hierarchy of sets, their subsets and their combinatorics, is probably only a fantasy. No one can reasonably escape the impression that the important mathematical structures (algebraic structures, topological structures) appear as data fundamentally imposed by the external world, and that their irrational diversity finds its justification only in reality.

Thom often referred to Gödel to remind us that absolute rigor in mathematics does not exist, the latter having shown that a completely formalized system of arithmetic (like a machine) is either inconsistent (leading to a contradiction), or incomplete.[19]

In a dialogue bearing the title *Métaphysique extrême* (Extreme metaphysics), Thom declares (Thom 2021, p. 86): "Formalists are people who always tell you: 'Oh, natural language is horrible, it tolerates every kind of ambiguity, it is not possible to do mathematics with it.' I have never done anything but natural language in mathematics, plus a few symbols from time to time."

In a series of lectures he gave in 1981, titled *Mathématiques essentielles* (Thom 1981), given at the Solignac Abbey, a medieval monastery in the Limousin (South of France), the main idea that Thom tried to communicate is that "what justifies the 'essential' character of a mathematical theory is its ability to provide us with a representation of reality." He declares (Thom 1981, pp. 2–3):

> The various branches that constitute mathematics are not all essential. By "real" I mean both the reality of the external world—whether it is given to us by the immediate perception of the world around us, or by a mediate construction as is scientific vision; that is to say, that the "real", in its most immediate form, is also apprehended by introspection—as an "immediate data of consciousness".

Thom reiterated this thesis in a talk he gave in 1982 at the École Normale, titled *Les réels et le calcul différentiel, ou la mathématique essentielle* (The reals and differential calculus, or essential mathematics) (Thom 1990d) and in several other writings. He was often accused of giving non-rigorous demonstrations, but these non-rigorous demonstrations were part of his style, which was that of a continuous search for meaning instead of rigor, and a refusal of any conventional way of writing mathematics.[20] In his interview with Jacques Nimier (Nimier), he expresses his doubts and hesitations, which were part of his philosophy:

> There are clean theories and dirty theories, and personally I always have more sympathy for a dirty theory. The clean theories are the theories where things are well presented, where the concepts are clearly defined, the problems more or less well defined as well. Whereas the dirty theories are the theories where one does not know very well where to go, one does not

[19] I would like to point out the interesting paper (Farmaki and Negrepontis 2021) by Farmaki and Negrepontis titled *The paradoxical nature of mathematics*, in which the authors argue that the deductive strength in Mathematics is strongly related to its paradoxical nature, and in fact, that it comes, maybe exclusively, from its proximity to the contradictory.

[20] See e.g., (Jaffe and Quinn 1993) and Thom's response in (Atiyah et al. 1994).

> know how to organize things and where the main directions are, etc. From this point of view, in fact, I have never been a Bourbakist, because Bourbaki likes clean things; whereas I think that it is necessary to get one's hands dirty, and sometimes even dirtier in mathematics.

As a passionate of topology in the sense of the theory of forms and their deformations, Thom was not attracted by algebra, nor by number theory, nor by set theory whose formalism had invaded mathematical writing. He adopted a way of thinking that included mathematics in a very broad perspective, that of motion, form and change of form. Gradually, he integrated the mathematical objects he had been working on into his new framework of thought, that of natural philosophy, where mathematics becomes the model of the physical and biological world. In his article *Les racines biologiques du symbolisme* (The biological roots of symbolism) (Thom 1978e), he returns to the opposition between algebra and geometry, this time associating topology with geometry:

> There is a certain opposition between geometry and algebra. The fundamental material of geometry, of topology, is the geometrical continuum; pure and unstructured extent is *par excellence* a "mystical" notion. Algebra, on the contrary, shows a fundamentally "diaeretic" operative attitude. Topologists are the children of the night; algebraists, on the other hand, wield the knife of rigor with perfect clarity.

Thom's judgment on algebra was shared by some other preeminent mathematicians. Michael Atiyah, referring to Goethe, considers that when a mathematician solves a question using algebra, he sells his soul. He writes in his article *Mathematics in the 20th Century* (Atiyah 2002):

> Algebra is the offer made by the devil to the mathematician. The devil says: "I will give you this powerful machine, it will answer any question you like. All you need to do is give me your soul: give up geometry and you will have this marvelous machine". [. . .] When you pass over into algebraic calculation, you essentially stop thinking; you stop thinking geometrically, you stop thinking about the meaning.

Mostly in reaction to the proponents of "modern mathematics" that imposed new programs in school and university curricula, Thom published, in 1962, an article on envelopes (Thom 1962), a classical subject which originates in the work of the seventeenth century mathematician Christiaan Huygens. The subject had been in the high school curriculum before the modern math reform, but had disappeared, certainly because of a difficulty it presented in comparison with subjects like linear algebra or analysis: in the study of envelopes, there is no general method, and one had to reason case by case. This was the kind of mathematics that Huygens preferred, and Thom followed him. Furthermore, this subject accounts for phenomena of life: Thom, again like Huygens, had observed the envelopes of light rays undergoing reflection or refraction on a surface or curve and the structure of the singularities they form. He was interested in the distinction between those that are stable or generic and those that are not, as well as in problems of classical plane geometry that the theory of envelopes leads to. Furthermore, Thom included the study of envelopes in that of

generic singularities of differentiable functions, one of his favorite subjects. In fact, in his auto-biographical article (Thom 1989b), Thom writes that he started the study of differentiable functions in order to understand envelopes.

4 Qualitative and Quantitative

From the beginning of the 1970s, Thom became involved in epistemology, in particular in the question of what is expected from science. In his books and articles, he expressed his ideas on the predictive attribute of theory, on the role of experimentation, and on analogy as a fundamental tool for reflection. From epistemology, he was gradually led to the great problems of philosophy. He turned more and more to Aristotle, with his categories and genera, expressing reservations about mathematical logic as the basis of thought, stressing the fact that it may lead to contradictions and making clear his preference to Aristotelian logic, a logic used in a natural way, and based on reality.

Thom liked Aristotle's informal and lively style. He writes in his *Esquisse d'une sémiophysique*: "The very style of Aristotle, far from the axiomatic precision expected of the logician, is that of a thinking that seeks itself, made of returns on oneself, always in struggle with its object; this testimony of a constant effort of thinking filled me with a kind of immense sympathy" (Thom 1990b, p. 13). He emphasized at several occasions that the fact that the theory he was proposing had no quantitative character was an echo of Aristotelian biology. This is the subject of his article *Qualitative and quantitative in evolutionary theory with some thoughts on Aristotelian Biology* (Thom 1996b). He used to describe Aristotle as "the apostle of the qualitative against the quantitative" (Thom 1990b, p. 172). In his 1988 article *Un rapport problématique: la rencontre théorie-expérience* (A problematic relationship: the theory-experiment encounter) (Thom 1988c), he contests the frequently repeated postulate that modern science was born at the beginning of the seventeenth century as a result of the development of the experimental method. Science, he insists, is the consequence of theoretical reflection, and no experimentation is useful without a causal analysis that sets its limits and allows for an adequate interpretation of the results.

Thom declared in 1984, in a communication to the Académie des Sciences titled *La méthode expérimentale: un mythe des épistémologues (et des savants)?* (The experimental method: A myth of epistemologists (and Scientists)?) (Thom 1985) that only two types of causal analyses are available for science: the first one, of Aristotelian origin, is that of efficient causes, based on natural language, explaining the phenomena by *ad hoc* entities, and the second one, linked to mathematics, essentially reducing the question of causality to the solution of a differential system whose solutions are completely determined by the initial solutions. In biology, he says, the first scheme is massively used to deduce precepts that have no demonstrative value. In the same article, Thom writes that the expression "experimental method" is not only a myth, but also an oxymoron: it contains a contradiction.

Indeed, he recalls that, while the word "method" refers to a direction of thinking, the word experiment originates in the latin word *experientia*, which means a trial.[21] Thus, rather than a "method," experimental science is a *pratice*, a practice, which according to Thom, comes under the heading of exploration carried out at random, if it is performed without a theory supporting it. The aim of experimentation would be, according to him, to find "that which one has not looked for," and, as such, it is illusory in many respects. He considers that approximation should not be part of science as an enterprise aiming at constituting a knowledge with universal validity. Specifically, Thom criticizes what he calls "contemporary biology," which uses the experimental method with no scruple.

A fierce debate followed Thom's lecture at the Academy, notably with the physicist Anatole Abragam, a fervent supporter of the experimental method, and who was also an academician. The text of Thom's lecture as well as Abragam's answer are published in the volume *La philosophie des sciences aujourd'hui* (Philosophy of sciences today), edited by J. Hamburger (Hamburger 1986).

Thom returned to this question several times in his writings. In his interviews (Thom 1991c, p. 130), he writes: "If one reduces science to a set of recipes that work, then, intellectually speaking, we are not better than a rat who knows that when a lever is pressed, food will fall into his bowl. The pragmatist theory of science brings us back to the situation of the rat in its cage." This is illustrated on the book cover (Thom 1991c) (Fig. 2).

In the same book (Thom 1991c, p. 79), Thom returns to a quote by Rutherford, "Qualitative is nothing but poor qualitative," which he strongly contests. He is convinced that the contrary is true. He writes: "All of topology is to be added to the chapter of examples of this conviction. How is a sphere different from a ball? This is really not quantitative. How is a circle different from a disk? It is not a question of quantity, it is a problem of quality."

One characteristic of Thom's theory of morphogenesis which will be criticized by the proponents of a theory of pure experimentation, is that it is a completely geometrical theory, "independent of the substrate of forms and of the nature of the forces that create them" (Thom 1972, p. 8). According to him, the fact that a theory is independent of the substrate makes it applicable at the same time to domains which are a priori very different from each other, like biology and linguistics. The ground space is simply an open set in a topological space, i.e., an abstract substrate, devoid of any matter. Thom writes: "This may seem difficult to admit, especially for experimentalists who are accustomed to cutting to the chase and continually struggling with a reality that resists them."

Thom said later that he ended up being a philosopher of science because of the need to respond to the criticisms directed toward catastrophe theory. He declared in an interview (Thom 1991c, p. 90):

[21] Etymologically, the word "method" consists of two words: the word is a combination of the words μετά, which means "after," or, "to follow" and ὁδός, which means a "way," or a "path."

Fig. 2 This book cover is Thom's representation of a predictive science: A rat who knows that if he presses a lever, food will fall into his bowl. (Figure © Flammarion)

I was indeed subjected to criticisms that were on the epistemological ground. I devoted myself for a while to the philosophy of science before engaging in more general philosophy [. . .]. I started somehow from the validity of the theory of catastrophes, and I came to be interested in the position of science in general, in what can be expected from it from the point of view of knowledge. It was there that I began to develop critical positions towards the so-called experimental method and the general belief in the virtues of experimentation as leading us to progress. I believe that experimentation by itself can hardly lead to progress. [. . .] But I found myself hitting a kind of wall: in a science like biology, for example, people refuse the theoretical necessity of imagination.

5 Thom's Rediscovery of Aristotle

Thom has always had a deep reverence for the philosophers of classical Greece. From the end of the 1960s on, he constantly referred to them. In his book *Structural stability and morphogenesis* (Thom 1972) which we already mentioned several times, he writes: "In fact, all the fundamental intuitions in morphogenesis are already to be found in Heraclitus; my unique contribution will be to have placed them in a geometrical and dynamical framework which one day will make them accessible to quantitative analysis."

Among the great figures of antiquity, Aristotle occupies a central place in Thom's work. On several occasions, he expressed his surprise for having found in Aristotle's writings ideas that he had previously held himself. He also shared with the Philosopher a boundless curiosity and a desire to understand all the phenomena of nature. In his book, *Semiophysics, a sketch*, which bears the subtitle *Aristotelian physics and catastrophe theory*, he writes: "It is only rather recently, almost incidentally, that I discovered Aristotle's work. I was almost immediately fascinated by this reading" (Thom 1990b, p. 12). In his *Esquisse*, published in 1988, he writes (Thom 1990b, p. 12):

> [. . .] If I add that I found in Aristotle the concept of genericity (ὡς ἐπὶ τὸ πολύ), the idea of "stratification" that we can glimpse in Aristotle the biologist in the decomposition of the organism into homeomers and anhomeomers, the idea of the decomposition of the genus into species as an image of bifurcation, we will agree that there was something to be astonished about.

Thom also discovered in Aristotle the idea of significant fact and significant form, which played an important role in his philosophy. He even found the essence of the mathematical notions of cobordism, genericity and stratification in the writings of the Stagirite, in particular in his biology. I shall talk about this in more detail in the next sections. The titles of some of Thom's articles are significant in announcing the subject at hand. Let me mention some of them, where the reference to Aristotle or Aristotelian ideas is obvious:

- *Mathématique et réalité: faut-il croire Aristote?* (Mathematics and reality: should we believe Aristotle?) (Thom 1987a, b).
- *Considérations sur la finalité: d'Aristote à Geoffroy Saint-Hilaire* (Considerations on finality: From Aristotle to Geoffroy Saint-Hilaire) (Thom 1986).
- *Les intuitions topologiques primordiales de l'aristotélisme* (The primary topological intuitions of Aristotelism) (Thom 1988b).
- Structure et fonction en biologie aristotélicienne (Structure and function of Aristotelian biology) (Thom 1990e).
- *Ésquisse d'une sémiophysique: Physique aristotélicienne et théorie des catastrophes* (Semiophysics, a sketch: Aristotelian physics and catastrophe theory) (Thom 1990b).
- *Causality and finality in theoretical biology: a possible picture* (Thom 1989a).

- *Homéomères et anhoméomères en théorie biologique d'Aristote à aujourd'hui* (Homeomerous and anhomeomerous since Aristotle until today) (Thom 1990c).
- *Matière, forme et catastrophes* (Matter, form and catastrophes) (Thom 1991a).
- *Actualité de la physique aristotélicienne* (Actuality of Aristotelian physics) (Thom 1992b).
- *Biologie aristotélicienne et topologie* (Aristotelian biology and topology) (Thom 1995).
- *Qualitative and quantitative in evolutionary theory with some thoughts on Aristotelian Biology* (Thom 1996b).
- *Histoire de la transversalité d'Aristote à la théorie des catastrophes* (History of transversality, from Aristotle to catastrophe theory) (Thom 1996a).
- *The hylemorphic schema in mathematics* (Thom 1997).
- *Comment la biologie moderne redécouvre la kinèsis d'Aristote* (How modern biology rediscovers Aristotle's *kinesis*) (Thom 1998).
- *Aristote topologue* (Aristotle as a topologist) (Thom 1999).

I will highlight several ideas that lie behind these works in the next section. Here I will only make a few comments on some papers in this list.

In the article *Actualité de la physique aristotelicienne* (Thom 1992b), Thom starts by pointing out a few principles from Aristotelian physics that can be found in an equivalent form in catastrophe theory. He dwells on the distinction that Aristotle makes between the phenomena that take place in the celestial, or "supra-lunar" world, and those of the sub-lunar one. The former is governed by God-given uniform circular motions, whereas the latter is subject to one great principle, namely, the effort of the four elements (fire, air, water and earth) which moves them toward what Aristotle calls their "natural place." It is in this distinction that Thom finds the origin of Aristotelian physics. He makes a comparison between this division and the modern classification of dynamical systems into reversible Hamiltonian ones and irreversible gradient-like dissipative ones. This leads him to present Aristotelian physics in an axiomatic way. He starts with an axiom of existence of entities (*ousiai*), continuing with their division into primary and secondary ones: The first ones include the living being, identified with a three-dimensional topological ball separated from the exterior world by a boundary (the skin), and the second ones are only "qualities" that affect the primary ones. The other axioms concern the *state* of an entity, the notion of *power*, and the *natural* transformations between entities. Entities interact with each other, and qualities are classified according to *genus*. The analogy between the major concepts of Aristotelian physics and those of catastrophe theory is outlined in terms of a few correspondences and principles:

(1) Natural place ↔ genericity;
(2) genus decomposition ↔ bifurcation;
(3) homeomerous/anhomeomerous ↔ regular and irregular point in a morphology;
(4) biological organization ↔ stratification;
(5) The continuity assumption;
(6) Form is the boundary of matter.

The fact that in Aristotle's *Physics*, form is the boundary of matter, is an idea that Thom interprets using his background as a topologist. He writes (Thom 1992b, p. 93) that the ambition of catastrophe theory is to keep this Aristotelian flexible scheme. We shall dwell on this in the next section.

I will review now some ideas of Thom on form.

Thom was, like Aristotle, a philosopher of form. Understanding form and its evolution, which Thom connected with his research in the topology of manifolds and function spaces, became an obsession for him in his philosophical works. He highlighted several passages in which his predecessor discusses form. In the *Metaphysics*, Aristotle writes that what makes the essence of an object is its *form* (Aristotle 1928) 1032b1-2, 1035b32. In the *De Anima*, he maintains that the soul is the form of a living being (Aristotle 1927) (412a11). In the *Physics*, he writes that potentially, flesh or the bone do not have yet their own nature and do not exist by nature, as long as they have not received the form of flesh and bone, that is, a definable form, that which can be stated to define the essence of the flesh or the bone (Aristotle 1930) (193b).

In the *Esquisse*, Thom talks about Aristotle's theory of hylemorphism, to which he adheres completely. This is a theory which considers that any being (object or individual) is composed in an indissociable way of matter (*hylé*, ὕλη) and of form (*morphè*, μορφή), matter being, for him, essentially something in power, a substrate which awaits to receive a form, in order to become a substance, the substance of being, or being itself. Thom notes that the word ὕλη was originally used to designate shapeless wood, and that it is to Aristotle himself that we owe its introduction in philosophy, while the latter was putting forward the notion of continuity, both for matter and for form. Thom writes on this topic (Thom 1990b, p. 12):

> Of course, I knew that the hylemorphic scheme which I use in the formalism of catastrophes had its origin in the work of the Stagirite. But I was unaware of the essential point, namely, that Aristotle had tried, in the *Physics*, to build a theory of the world based not on number, but on the continuum. He had thus realized (at least partially) the dream I have always had of developing a "Mathematics of the Continuum" which takes the continuum as its starting point, if possible, without any recourse to the intrinsic generativity of number.

Talking again about form in Aristotle, Thom writes in (Thom 1991c, p. 53):

> I see matter in an Aristotelian perspective, a kind of continuum that can acquire form. Form can be external, visible, or internal. The internal form is what one would call, from the semantic point of view, a quality. Aristotle's *materia signata* is a matter provided with quality. I think that any quality can be seen precisely, to a certain extent, as a spatial form, an extended form in an abstract space. The original matter, if I may say so, is a bit like Aristotle's raw material, it is the substrate which can receive any kind of quality, any predicate. The raw material is a kind of idealization which, very quickly, acquires qualities, forms.

Thom declared several times in his writings that he shared with Aristotle the conception of the soul as the form of the body. He comments on this in (Thom 1991c, p. 100):

> The capacity of a body to be the support of a soul is something that presents itself as a structure or as a law of a formal nature, associated precisely with form in the morphological sense, in the spatiotemporal sense of organization. The whole is obviously linked to the form of the flows running through the organism: blood, neuronal, or more generally metabolic: for me all this is form, and something comes out of it as a residual form with an organizing character: the soul. But the structure, the character (intrinsic in a certain sense) of this residual form, like the result of a formal structure, originates in this gigantic object that is the consideration of the form of all the molecular and physical movements of our organism.

6 Topology from an Aristotelian Perspective

In his books and in several articles, Thom highlighted and commented on several passages in the works of the Stagirite which support his conviction that the latter was aware of basic notions of topology. He declared in particular that a careful reading of Aristotle's *Physics* shows that the latter understood the topological distinction between a closed and an open set. The reader might try to imagine what kind of intuition can one have on the notion of open and closed set and the distinction between them if one does not have the language of set theory and topology. Thom writes in his *Esquisse* (Thom 1990b, pp. 167–168):

> A careful reading of the *Physica* leaves little doubt but that [Aristotle] had indeed perceived this difference. "It is a whole and limited; not, however, by itself, but by something other than itself" (207a24-35)[22] could hardly be interpreted except in terms of a bounded open set. In the same vein, the affirmation: "The extremities of a body and of its envelope are the same" (211b12) can be identified, if the envelope is of a negligible thickness, with the well-known axiom of general topology: "Closure of closure is closure itself" expressed by Kuratowski at the beginning of this century. This allows the Stagirite to distinguish two infinites: the great infinite that envelops everything and the small infinite that is bounded. The latter is the infinite of the continuum, able to take an infinity of divisions (into parts that are themselves continuous). Whence the definition he proposes: "The infinite has an intrinsic substrate, the sensible continuum" (208a).

Chapter 7 of Thom's *Esquisse* is titled *Perspectives in Aristotelian biology*. The first part concerns topology and bears the title *The primordial topological intuitions of Aristotelianism: Aristotle and the continuum*. It starts as follows (Thom 1990b, p. 165):

> We shall present here those intuitions which we believe sub-tend all Aristotelianism. They are ideas that are never explicitly developed by the author, but which—to my mind—are the framework of the whole architecture of his system. We come across these ideas formulated "incidentally" as they were condensed into a few small sentences that light up the whole corpus with their bright concision.

[22] Here and in the following, while I am giving the relevant passages in Aristotle, I am translating Thom's interpretation of the passages in Greek.

In his articles *Les intuitions topologiques primordiales de l'aristotélisme* (Thom 1988b) and *Matière, forme et catastrophes* (Thom 1991a), Thom returns to these matters, explaining that the modern topological distinction between an open and a closed set is also expressed in Aristotle's distinction between form and matter, and between actuality and potentiality. We mentioned this concerning his distinction between actual and potential infinity. He declares that the difference between the notion of bounded and unbounded open set in Aristotle's philosophical system lies in the fact that the former may exist as a substrate of being whereas the latter cannot (Thom 1988b, p. 396). He notes the presence of the notion of boundary in formulae such as: "Form is the boundary of matter," (Thom 1988b, p. 398) and "Actuality is the boundary of potentiality" (Thom (1988b), p. 399) and (Thom (1991a), p. 380). He points out that the paradigmatic substance for Aristotle, which is the living being, is nothing else than a ball in Euclidean space whose boundary is a sphere, that is, a closed surface without boundary. Shapeless matter is enveloped by form – *eidos* – in the same way as the boundary of a bronze statue defines its form. The boundary of a living organism is its skin, and its "interior" exists only as a potentiality.

The distinction between openness and closeness in the topological sense in Aristotle's works is also pointed out by Thom in his discussion of homeomerous and anhomeomerous parts of animals, which we review in Sect. 7 below, in the context of Aristotle's biology.

Another guiding idea in Thom's philosophical work which has an Aristotelian origin is the opposition between the discrete and the continuous. He declares, about this opposition: "For me, the fundamental aporia of mathematics is indeed in the opposition discrete-continuous. And this aporia dominates at the same time all thought" (Thom 1991c, p. 81). In this interview, he recalls Zeno's paradox supported by the story of Achilles and the tortoise and he declares:[23]

> The continuous is in a way the universal substratum of thought, and of mathematical thought in particular. But one cannot think anything effectively without having something like the discrete in this continuous unfolding of the mental processes: there are words, sentences, etc. The *logos*, the discourse, is always discrete; there are words entering in a certain succession, but discrete words. And the discrete immediately calls for the quantitative. There are points: we count them; there are words in a sentence: we can classify them quantitatively by the grammatical function they occupy in the sentence, but there is still an undeniable multiplicity.

Thom basically believed in the continuous character of the universe, of phenomena and of their substrate. Discontinuities only operate on the continuous. He was opposed to what he used to call a "modern vulgate," which comes essentially from information science and which says that everything is expressed in bits (Thom 1991c, p. 62). In the *Esquisse d'une sémiophysique* (Thom 1990b, p. 172), there is

[23] I take this opportunity to point out an extremely interesting interpretation of Zeno's paradoxes by S. Negrepontis in (2019), based on a thorough interpretation of Plato's dialogue *Parmenides*. In short, Negrepontis' explanation is that Zeno's aim in his argumentation is to show that the sensible Beings are separate from the true/intelligible Beings.

a long passage in which he discusses Aristotle's decision to leave Plato's Academy, a decision which he explains by a disagreement with the master, concerning the notion of continuous. In particular, Thom writes that Aristotle's "revolt against his master is that of the topologist against the imperialism of the arithmetician." He writes: "Aristotle does not go into much detail about the way he sees the raw material. In my opinion, it must always be reduced to a continuum, to an extent. I am a universal topologist! I have a real metaphysics of the continuum" (Thom 1991c, p. 54).

The notion of boundary is another key idea which Thom found in Aristotelian physics. He writes in (Thom 1991c, p. 110):

> For Aristotle, the form of a physical object is something like its boundary; in the abstract sense of the definition, the *eidos* is also something like a form in an abstract space, with its boundary. There is intelligible matter, which in some way is compressed by its definition. *Orismos* means definition; it is almost the same word as *oros* which means boundary. This is quite remarkable. . . . To define is to draw the borders. I think that there's something quite profound about that intuition.

In the two articles *The primary topological intuitions of Aristotelism* (Thom 1988b) and *Aristotle as a topologist* (Thom 1999), Thom pushed his ideas to the point of finding Stokes' famous formula in the work of the Stagirite. He explained in particular how he found this formula in a notion of minimal limit of an "enveloping body" contained in Aristotle's Physics (211b 11-12), a formula that he translates as: "The edge of the enveloped body and the edge of the enveloping body are together." This identification of Stokes' formula with a formula of Aristotle is echoed again by Thom in his article *Matière, forme et catastrophes* (Matter, Form and Catastrophe) (Thom 1991a, p. 381), an article published in the proceedings of a conference held at the Unesco headquarters in Paris at the occasion of the 23rd centenary of the Philosopher. If we add to this that Thom foresaw in the latter the mathematical ideas of genericity, stratification and bifurcation, we will not be surprised to see the reference to Aristotle in the subtitle of the *Esquisse* (Thom 1990b, p. 13).

The notion of cobordism, which was at the center of Thom's work in the 1950s, is also incorporated into this philosophical framework *via* the notion of boundary, which is its primary element. He states in his 1992 interview *La théorie des catastrophes* (Thom 1992d):

> The whole unity of my work revolves around the notion of boundary, because the notion of cobordism is only a generalization of it. The notion of boundary seems all the more important to me today as I am interested in Aristotelian metaphysics. For Aristotle, the boundary is a principle of individuation. The marble statue is matter in the block from which the sculptor drew it, but it is its boundary which defines its form.

7 Aristotelian Biology

Like Aristotle, Thom was the philosopher of biology. And basically, biology, from Thom's point of view, was not far from topology, nor from philosophy: "The great problem of biology," he says in a 1992 interview on catastrophe theory

(Thom 1992d), "is the relationship between the local and the global. Knowing the organism, why the organs have this or that property at this point and not another, etc., that is the big problem." He then adds:

> The problem of the relation between the local and the global is a problem that is essentially philosophical, that has to do with extent; this problem is the object of topology. Topology is essentially the study of the means to make the junction between a known local situation and a global property to be found, or vice versa: knowing a global property of space, find the local properties, around each point. There is a kind of deep methodological unity between topology and biology.

In his article *Les mathématiques et l'intelligible* (The mathematics of the intelligible) (Thom 1975b), Thom introduces what he calls mathematical thought in the context of biology and anthropology, a thought that, according to him, goes beyond the usual language in describing external phenomena with their mixture of determinism and indeterminism.

Chapter 5 of Thom's *Esquisse* is titled *The general plan of animal organization*. It is in the lineage of the zoological treatises of Aristotle, expressed in the language of modern topology. Thom writes in the introduction: "This presentation might be called an essay in transcendental anatomy; by this I mean that animal organization will be considered here only from the topologists' abstract point of view." He adds: "We shall be concerned with ideal animals, stylized images of existing animals, leaving aside all considerations of quantitative size and biochemical composition, to retain only those interorganic relations that have a topological and functional character." In the same chapter, Thom returns to Aristotle's notion of homeomerous and anhomeomerous, in relation with the stratification of an animal's organism, formulating in a modern topological language the condition for two organisms to have the same organization. Aristotle considers this notion at the beginning of the *History of Animals* (Aristotle 1910): "Regarding the parts of the animals, some are non-composite: these are those which can be divided into *homeomers*, like flesh is divided into flesh; others are composed: these are those which can be divided into *anhomeomers*; the hand, for example, cannot be divided into hands, nor the face into faces." Thom explains that a homeomerous part of an animal (ὁμοιομερής) has generally a boundary structure constituted by anhomeomerous parts, which implies that the substrate of a homeomerous part is, topologically speaking, not closed. Mathematics, philosophy and biology are intermingled in this interpretation of a living organism, following Aristotle, with formulae like: "The opposition homeomerous-anhomeomerous is a 'representation' (a homomorphic image) of the metaphysical opposition: potentiality-act. As the anhomeomerous is part of the boundary of a homeomerous of one dimension higher, we recover a case of the application of *act as boundary of the potentiality*." (Thom 1988b, p. 400).

In Thom's words, an animal organism is a three-dimensional ball O equipped with a stratification which is finite provided we neglect the details that are too fine. For example, when we consider the vascular system, we may take into account only the elements that can be seen with the naked eye: arterioles and veinlets. Seen from

this point of view, the homeomerous parts are the three-dimensional strata: blood, flesh, the inside of bones, etc.; the two-dimensional strata are the membranes: skin, mucous membrane, periosteum, intestinal wall, walls of the blood vessels, articulation surfaces, etc., the one-dimensional strata are the nerves: vessel axes, hairs, etc.; and the zero-dimensional strata are the points of junction between the one-dimensional strata or the punctual singularities: corners of the lips, ends of hairs, etc.

This leads Thom to the definition of two organisms O and O' "to have the same organisation" by the existence of a homeomorphism h: $O \to O'$ that preserves this stratification.

Thom also developed his ideas on the stratification of an animal body in the series of lectures given in 1988 at the Solignac Abbey (Thom 1981). On p. 7 of these notes, he addresses the question of when two animals have isomorphic stratifications, and he uses for that the notion of isotopy between stratified spaces: Two sets E_1, E_2 have isotopic stratifications if there exists a stratification of the product $E \times [0, 1]$ such that the canonical projection p: $E \times [0, 1] \to [0, 1]$ is of rank one on every stratum of E, with $E_1 = p^{-1}$ (1) and $E_2 = p^{-1}$ (2). He considers that this notion is implicitly used by Aristotle in his classification of the animals, insisting on the fact that the latter neglected all the quantitative differences and was only interested in the qualitative ones.

Thom also talks about stratification in Aristotle in chapter 5 of his *Esquisse*, a chapter titled *The general plan of animal organization*, in relation with the notion of homeomerous introduced by the Stagirite. He gives this idea of stratification in biology an official status in his work, by defining his *phenomenological equivalence relation*, a relation which at the same time allows us to take a new look at Aristotle's classification of the elements of a living body. He writes in his article *Structure et fonction en biologie aristotélicienne* (Thom 1990e), that this allows "the identification of two anomalies by an isotopy between stratifications, either according to genus, or refined according to species."

In relation with Aristotle's classification of the parts of the animals, and more generally of a general plan of animal organisms, Thom revived a famous controversy that took place in 1830 between Geoffroy Saint-Hilaire and Cuvier, in which the former used a purely geometrical classification principle of animal organs whereas the latter defended criteria based only on function; see (Thom 1978c), an article in which Thom proposes an approach to this dilemma using catastrophe theory.

Genericity is another idea that Thom finds in Aristotle's biology. He declares in his 1992 interview *La théorie des catastrophes* (Thom 1992d) that Aristotle considered that what matters in the study of nature is not to know everything that can happen, but everything that happens most often. In the Metaphysics, as in other texts, the Philosopher speaks at length about the difference between what happens most often and what is only an accident (συμβεβηκός). He writes in particular that all science is either about what is always or what is most often; otherwise, he says, how could one learn or teach others?[24] Thom establishes a parallel between this and the

[24] *Metaphysics* (Aristotle 1928) 1027a20.

theory of singularities: we study generic singularities, and those can be classified; their number is finite. There are important pathological singularities, but since they are not generic, the general theory does not deal with them.

In the series of interviews *To predict is not to explain*, Thom returns to the homological Stokes formula $d^2 = 0$ and to its biological interpretation in terms of Aristotelian zoology: "The boundary of the boundary is empty; this is the great axiom of topology, of differential geometry in mathematics, but it expresses the spatial integrity of the edge of the organism." (Thom 1991c, p. 111)

Thom spent some time trying to teach biologists the fundamental notions of topology that he thought they may need. Several sessions of the courses he gave at Solignac were devoted to this subject (Thom 1981). His 1982 lecture *Les réels et le calcul différentiel, ou la mathématique essentielle* (Thom 1990d) at the École Normale contains a summary of what is needed for the biologist as a minimal mathematical background.

In the next section, we talk more about the status of singularities in Thom's philosophical system.

8 Philosophy of Singularities

The goal of Thom's article *Philosophie de la singularité* (Thom 1988a) is to show how the mathematical notion of singularity is used in our perception and description of the world. Not only this, but from Thom's point of view, our perception is essentially a perception of form, a transfer of topological structure from matter to our mind. Thom starts by making a connection between the mathematical notion of analytic continuation in the holomorphic setting and prediction in science. In complex analytic geometry, analytic continuation allows, from the knowledge of a function on some open set, to extend it to a larger set, the domain of holomorphy of the function. Thom views this as a sort of prediction: we predict the value of the function on a larger subset. He notes that what prevents a function from being extended further, depends on the singularities, on the boundary of the domain of holomorphy.

Another well-known notable fact is that a holomorphic function is determined by its singularities. This was already known to Bernhard Riemann. One form of the latter's famous *Existence theorem* says that one can reconstruct a meromorphic function from the knowledge of its singularities and its topological behavior at these singularities. Paul Montel, who has been Henri Cartan's doctoral advisor, emphasized the importance of characterizing a function of a complex variable in terms of its singularities. In the introduction to his *Leçons sur les fonctions entières ou méromorphes* (Montel 1932), he writes: "What characterizes a function, what distinguishes it from others, is the set of its singularities. What holds for functions also holds for concrete objects and living beings" (Montel 1932, pp. vi–vii).[25]

[25] Ce qui caractérise une fonction, ce qui la distingue des autres, c'est l'ensemble de ses singularités. Il en est des fonctions comme des objets concrets et des êtres vivants.

Another idea related to the fact that the singularities of a function give global information on it comes from Morse theory, one of whose fundamental principles is that a Morse function allows the reconstruction of a space out of some finite number of cells of a fixed type and the singularities of the function. Such ideas were part of the background that accompanied Thom throughout the rest of his life.

It is during his time in Strasbourg that Thom began to reflect on Morse theory, which, in fact, is the subject of his first article, a note in the Comptes Rendus (Thom 1949) published in 1949.[26] What fascinated Thom in this theory was, once again, the fact that one can use the singular points of a function to decompose a manifold into cells that have a standard type, and conversely, one can reconstruct the space from the information on the cells and their order in the values of the function. Talking about this kind of situation, Thom recalls an old formula of Descartes: "To reduce everything to situations that are simple enough to be described" (Thom 1992d). This is another idea that never ceased to fascinate him: how to make constructions out of simple things. Regarding simplicity, he writes again in *Paraboles et catastrophes* (Thom 1980, p. 66): "The whole 'philosophy' of catastrophe theory, its general scheme, is precisely this: *It is a hermeneutic theory which tries, in front of any experimental data, to build the simplest mathematical object that can generate it*."

In the article *Philosophie de la singularité* that we already mentioned, Thom motivates the use of analytic functions in physics, and the fact that they constitute the basis for prediction (via analytic continuation) by the fact that the symmetry groups of the laws of physics, such as those that govern particles and their interactions, are analytic. He also recalls that the notion of stratification, in mathematics, arises precisely in an analytic setting. Singularities govern the propagation of analyticity and at the same time they play the role of obstacle for this propagation.

Regarding singularities, we also mention that Thom also studied the notion of apparent contour of objects, contours which present singularities that he can classify. Observing apparent contours reminds him of the inhabitants of Plato's cave, chained and unable to move, only capable of watching shadows of world events that take place outside the cave, projected on the wall in front of them by the light of a fire burning at the entrance of the cave. He says that this is how we should interpret Hironaka's desingularization theorems.[27] He notes incidentally that the singular sets of apparent contours are naturally stratified spaces in which a hierarchy of singularities is manifested, and he mentions a relation between such a situation and the notion of unfolding, a central notion in his theory of elementary catastrophes. Talking about the biological and human world, he mentions the "archetypal

[26] Haefliger, in his article (Haefliger 1988), writes, before presenting the ideas contained in this note and its later developments: "This was not too early, because the administrators at CNRS wondered if it was necessary to continue to support this young mathematician who was so little productive. [Il était temps, car les responsables du CNRS se demandaient s'il fallait continuer à soutenir ce jeune mathématicien si peu productif.]

[27] In the 1970s, Heisuke Hironaka built a theory on the desingularization of of sub-analytic sets, showing that locally, any such set is a finite union of images of spheres glued together by real-analytic maps. The theory had been foreseen by Thom. See Hironaka's exposition in (Hironaka 1973).

singularities" that appear in embryology. He concludes the article by pointing out an opposition between singularities created as defects of an ambient propagative structure and singularities as a source of the propagative effect, a problem that, he says, is at the source of all scientific disciplines.

9 Topology and Ethology

In this short section, I briefly review the point of view of a researcher in bioethics, Yanick Farmer on the importance of Thom's morphological theories in epistemology and in the theory of ethics. This view is expressed in the article *Topologie et modélisation chez René Thom: l'exemple d'un conflit de valeurs en éthique* (Topology and modelization after René Thom: The example of a conflict of values in ethics) (Farmer 2010).

Farmer considers that the abstract (topological) setting of Thom's approach to science, based uniquely on morphology, answers the problem of the need for universality in the particular setting of ethology, by avoiding the problem of relativization that is inherent to the use of specific languages that are proper to particular cultures. The author mentions as an example the notion of *autonomy*, which is central in bioethics but which has different meanings according to the language and the culture where it is used. Thus, from the ethologist point of view, to study a particular phenomenon, one must first differentiate it from its environment, and here, the author, like Thom, speaks of the need for this phenomenon or event to have a *boundary*, and the *substrate* of the discussion becomes a mathematical space which may be assimilated to a 3-dimensional Euclidean space.

I refer the interested reader to Farmer's article, in which he develops his ideas on the subject, which constitute a step toward the use of Thom's theory in ethology, but I would like to quote the conclusion of this article, where the author refers again to Plato's cave passage (I am translating again from the French) (Farmer 2010, p. 386):

> The philosophical positions that stand in the background of Thom's theory of models seem to me to be faithful to the essence of the spirit of philosophy conceived by Plato in his allegory of the cave. They act like a liberation for those among us who, by the chains of knowledge and perceptual limitation, are bind to the surface of familiar objects described by ordinary language. For these prisoners who do not have at their disposal any other impressive technical mean than the strength of their gaze, Thom's philosophy offers an epistemological way of salvation. By challenging the methodological omnipotence of Reductionism, which forces knowledge to reveal the depthless secret of elementary entities, René Thom reminds us, like Plato before him, that the thinker of everyday life and of the familiar things, as long as he is willing to become a geometer, can also try to understand better, in order to act better.

10 Linguistics and Morphology

For René Thom, every science is determined by a morphological scheme, and linguistics is no exception. In his hands, linguistics becomes a topological field. This is the subject of the present section.

Thom was particularly interested in linguistics. He saw in this field, as in biology, a privileged ground for the application of his ideas on morphology, with simple archetypes, where words are considered as living organisms, starting with the initial cry of the newborn, passing through the first stammering and stuttering, and culminating in the articulated and reasoned language of the adult. Aristotle, who studied language in his treatises on animals, considered that it is this articulated language, and the fact that it makes it possible for man to hold a reasoned discourse, that separates man from the other animals.[28] In this respect, Thom also spoke of language as a "boundary," which, as we already pointed out, is a notion that accompanied him during all his life: Following Aristotle, he considered language as the boundary that separates man from the rest of the animals.

In 1971, Thom published his article *Topologie et linguistique* (Thom 1971), which also appeared 5 years later in Russia, translated by Yuri Manin. The article concerns the application of catastrophe theory to linguistics. Thom starts by exposing the idea developed by Ferdinand de Saussure, that linguistics is a system of signs that associates two elements: the "signifier," which, here, is the sound, or the written word, and the "signified," which is of psychological nature, and which is the mental representation of the concept associated with the sign. The signifier and the signified are considered by Thom as two morphological processes, and the whole approach of Thom is based on a transformation between morphologies, which allows us to give an explanation of the syntactic structure of a sentence thanks to its signified. He writes: "For lack of sufficient mathematical knowledge, semanticians did not get the idea that the morphology of the signified could be supported by a multi-dimensional space. Therefore they were reduced to very poor descriptions, inspired by the logical analysis of language." His first objective was then to describe the morphological structure of the signified. For this purpose, he used a known model, the neurophysiological model of Zeeman, which leads to a geometrical model of the succession of ideas in the "stream of consciousness" of the philosophy of introspection, with the existence of a structurally stable attractor. To summarize, Thom's major contribution in linguistics was certainly the fact of bringing into this field the morphological models that he used in a field like biology. We refer the reader to Thom's article *La linguistique, discipline morphologique exemplaire* (Linguistics, an ideal morphological discipline) (Thom 1974a).

Regarding the significant and the insignificant, Thom liked to repeat this fragment of Heraclitus, the philosopher of absolute motion: "The master whose oracle is at Delphi does not say, does not hide, but he signifies," which Thom paraphrases by: "Nature sends us signs that we have to interpret."[29] In this regard, Thom often recalls Aristotle, for whom Art imitates Nature and not the other way around.

[28] In the *Politics* (Aristotle 1908) 1253a9-10, Aristotle considers that what separates man from other animals is the fact that man is a reasonable animal, λόγον ἔχον, "endowed with logos," the word *logos* meaning here speech, or reasoned speech.

[29] Fragment B93, cf. (Héraclite 1988).

Let me note also, talking about language, form and Aristotle, that the question of form in language is also present in the *Problems* (Aristote 1993), a compendium attributed to Aristotle, made up of questions, some of them simple and others that are real open research problems, encompassing all topics of science (mathematics, biology, physics, optics, psychology, astronomy, etc.).[30] In problem 23 of section XI of the collection, it is said that "the voice is air receiving a certain form," – a form which often changes, and which then vanishes. Similarly, problem 57 states that the voice of a human being – who, among the animate beings, is distinguished by the fact that he possesses the gift of speech – can take a multitude of nuances and forms.

Between Aristotle and Thom, I wish to mention another mathematician-philosopher, Marin Mersenne, whose *Harmonie universelle* (1636) contains a book titled *Treatise of voice and of singing* (Mersenne 1965, pp. 1–88). This book consists of an exposition of the physiology of the voice, both in humans and in animals, in the pure Aristotelian tradition. The statement of Proposition IX[31] of this book is the following: "Voice is the matter of speech, and only man can speak." In the demonstration of this proposition, Mersenne develops the Aristotelian idea of matter and form in speech, writing in particular: "We make use of the voice to form speech, as sculptors make use of wood and stones to make images; for images or statues are made by the different figures which are given to the matter of which they are made: and speech is a harmonic perspective, to which the voice serves as a picture to receive all kinds of images, since words are the images of the notions of the mind." Speech, he adds, "is the form, ornament, and perfection of the voice, which can only be formed and figured into speech by man, as speech can only be formed into discourse by the mind."

All this is close to what Thom said four centuries later.

11 Psychoanalysis

In his *Sketch of a semiophysics*, Thom writes that one of his main objectives is "to clarify the analysis of the original psychical mechanisms of our species" (Thom 1990b, p. 16).

In 1990, Michèle Porte defended a doctoral dissertation in Paris whose title is *Psychanalyse et sémiophysique: Études épistémologiques en métapsychologie et en dynamique qualitative* (Psychoanalysis and semiophysics: Epistemological studies in metapsychology and qualitative dynamics) (Porte 1990). In this dissertation, the author uses Thom's catastrophe theory and his ideas on qualitative dynamics, and in particular, structural stability, to try to reformulate metapsychology. Her main thesis

[30] Although, as in well known, there is an uncertainly concerning the authorship of the whole treatise, there is not doubt on the fact that it belongs to the Aristotelian school and that it was inspired, if not completely written, by Aristotle himself.

[31] Mersenne's *Harmonie universelle* is organized into propositions and proofs, like in mathematical treatises.

is that Sigmund Freud's theories cannot be understood in a formalist and structural setting, but, on the contrary, in a dynamical setting of birth, persistence and vanishing of forms and the dynamics behind them. Thus, she proposes an interpretation of Freud's ideas using a morphodynamical model of psychical events, in the tradition of Thom. At the same time, she makes a confrontation between Freud's psychoanalysis and Thom's semiophysics.

Later, the same author published a systematic essay titled *La dynamique qualitative en psychanalyse* (Qualitative dynamics in psychoanalysis) (1994), of which Thom wrote the preface. She writes (Porte 1994, p. 20):

> It is not forbidden to think that in the no man's land that our culture has constituted between the two slopes of the logos, that of number and that of word in general, besides catastrophe theory and semiophysics, a transposed psychoanalysis has its share and its role to play . . . Thus, we give a reason for the "exterritoriality" of psychoanalysis, a theme as hackneyed as it is obscure. "When we have a thought, the meaning of that thought is the form of the underlying neurophysiological process", according to one of the formulae by which Thom has often paraphrased Bernhard Riemann. Thom has taken this statement seriously; Freud's work does not object to it; we propose that both analysts and other researchers take up the program of work that this statement imposes.

Talking about Riemann, let me mention that Thom was also a reader of the latter's philosophical fragments. In his article (Thom 1994), he writes:

> As for myself, it is in reading in Riemann's *Philosophisches Nachlass* the sentence "When we think about thought, the meaning of this thought is nothing but the 'Form' of the underlying physiological process" that I see how the concept of Form can help bridge the gap between the two types of science [Geisteswissenschaften and Naturwissenschaften]. In this light we may reasonably evoke a revival of *Naturphilosophie*.

Porte's reading of Freud using Thom's background brings a fresh point of view on the work of the Viennese Master.

Let me also mention that Thom was aware of the applications of catastrophe theory to psychoanalysis since the beginning of catastrophe theory. We already quoted his comments on Zeeman's article *Topology of the brain* (Zeeman 1965) which gave him an impetus for his ideas on biological modeling; see also Thom's comments in (Thom 2004) (1989). Zeeman's work included the application of catastrophe theory to psychology, ethology and sociology. In the same article, Thom writes that he noticed that in Riemann's *Complete works* edition, there are a few pages where the latter expresses his view on the relation between spirit and matter, and where he claims that one can conceive that if an activity of the brain corresponds to an idea that we think, then the meaning is necessarily defined by the *form* of the corresponding physico-chemical process.

Let me mention, to close this section, that Thom gave a talk at a conference on psychoanalysis titled "L'inconscient et la Science," in which he outlined a possible use in psychoanalysis of his ideas on morphology in the way he applied it in biology and linguistics, see (Thom 1991b).

12 A Metaphysics of Analogy

In the 1970s, Thom started a systematic use of analogies and metaphors in his writings. In fact, his approach to phenomena of science and morphology as taking place in an abstract substrate favored an extensive use of analogy. He declares in the interview *To predict is not to explain* (Thom 1991c, p. 122): "There could be science only if we immerse the real in a controlled virtual. And it is through the extension of the real into a larger virtual that one can study the constraints that define the propagation of the real within that virtual."[32] On a related subject, talking about the notion of "boundary," he declares in another interview: "To reach the limits of what is possible, you have to dream the impossible, and it is really the interface between the possible and the impossible that is important because if we know it, we know exactly the limits of our power" (Nimier).[33]

In the field of analogy, Thom considered himself once again the heir of Aristotle. In a 1976 conference, titled *Le statut épistémologique de la théorie des catastrophes* (Thom 1978d), he writes that "the simple fact of being able to classify analogous situations is a considerable achievement: there has been no theory of analogy since Aristotle." Later, in the interview titled *La théorie des catastrophes* (1992) (Thom 1992d), he declares that catastrophe theory is "a theory of analogy," the first, of such breadth, since Aristotelian logic.

From a strictly mathematical point of view, the word "analogia" is used by Euclid in the *Elements* to denote an equality between two ratios. Thom quotes at several occasions Aristotle's example of analogy saying that old age is to life the same thing as evening is to day, in the form of an equality between two ratios:

$$\frac{\text{old age}}{\text{life}} = \frac{\text{evening}}{\text{day}}.$$

From this follows the metaphor: "old age is the evening of life."[34] In the same passage, Aristotle writes: "The cup is to Bacchus what the shield is to Mars. We shall therefore say: "the shield, cup of Mars," and "the cup, shield of Bacchus."[35]

In his *Esquisse d'une sémiophysique* and in other works, Thom highlighted Aristotle's analogies concerning animals. Let us especially mention an analogy that Aristotle makes in his *Parts of the Animals* (668a10-13), comparing the vascular system of an animal to an irrigation channel in a garden. Thom emphasized this

[32] Il n'y a de science que dans la mesure où l'on plonge le réel dans un virtuel contrôlé. Et c'est par l'extension du réel dans un virtuel plus grand que l'on étudie ensuite les contraintes qui définissent la propagation du réel au sein de ce virtuel.

[33] Pour atteindre les limites du possible, il faut rêver l'impossible, et c'est réellement l'interface entre le possible et l'impossible qui est important parce que si nous le connaissons, nous connaissons exactement les limites de notre pouvoir.

[34] The example is often attributed to Aristotle, but the latter, in a passage of the *Poetics*, quoting this sentence, refers to Empedocles, who said that old age is the sunset of life.

[35] Cf. *Poetics* (Aristotle 1909) 1457b20-21, and *Rhetorics* (Aristotle 1924) 1412b35 and 1413a1-2.

analogy in his 1988 lectures titled *Structure and Function in Aristotelian Biology* (Thom 1990e).[36] Among the other examples of analogies made by Aristotle are the one between the organization of the parts of an animal and the political organization of the city (*Politics*, 1290b25-35), the one saying that sight is to the body what intelligence is to the soul (*Nicomachean Ethics*, 1096b28), the one saying that the soul is to the body what the pilot is to the ship (*De Anima*, 413 a 8-9), the one saying that the scale is to the fish what the feather is to the bird (*History of Animals*, 486b15), and there are many others. Thom, on several occasions, spoke of an analogy between the development of the embryo and that of a Taylor series function. I have commented on other instances of analogy as a method of thinking in the works of Aristotle and Thom in my article (Papadopoulos 2018c).

Thom also talked about Aristotle's theory of analogy in terms of boundary. He writes in (Thom 1991c, p. 75):

> Analogy is not something arbitrary. Analogy, or metaphor, contrary to the common vision that makes it something approximate, fuzzy, appears to me as a strict relation that, in many cases, we can express mathematically, even if this mathematical expression in itself is not interesting in the mental process that makes you consider the analogy.

In the same passage, speaking again of Aristotle's analogy between evening with respect to day, and old age with respect to life, he writes: "The formal structure of this analogy is simply the notion of boundary. You have here a time interval; this interval has an end; one calls 'evening' or 'old age' the tubular neighborhood, if I dare say of the word 'end', and the corresponding catastrophe is, for me, a fold."

Concerning the general role of analogy in science, Thom writes, in his article *The twofold way of catastrophe theory* (Thom 1990a) that this role is very much contested and that it is the object of a vast misunderstanding among professional scientists, most of whom – imbued with a positivist spirit – see it as a source of illusion, claiming that science must be quantitative whereas analogy is a qualitative thought. He reminds the reader that Konrad Lorenz, the famous biologist, in his reception speech of the Nobel Prize, stated that in a certain sense, all analogies are true, by definition. Thom, more precisely, distinguishes two cases of analogies: a "mathematically formalized analogy, associated with an organizing catastrophe," which is always "true, but can be sterile (it can only generate more or less poetic metaphors)," and a "partial" analogy, whose algebraization is incomplete, and he says about the latter: "It is precisely by striving to specify an analogy that one can come across interesting data; it is the incompleteness of the analogy that offers the best possibilities of synthesis. It is only because we accept the risk of error that we can reap new discoveries."

Concerning analogy, I would like also to mention Kurt Gödel, whom Thom often quoted for having shown that a formal system of arithmetic is either inconsistent or incomplete. Gödel made an analogy between this formal side of mathematics and a rigid human society. In a draft letter to David Scurlock, quoted by Pierre Cassou-

[36] A similar agricultural metaphor was made by Plato in the *Timaeus* (Plato 1937) (77c7-8), about blood irrigation in humans.

Noguès in his book *Les Démons de Gödel: Logique et folie* (Gödel's Demons: Logic and Madness) (Cassou-Noguès 2007, p. 164), Gödel writes that one may expect from a society without any freedom at all (i.e., proceeding in everything according to strict rules of "conformity") to be either inconsistent or incomplete in its behavior, i.e., unable to solve certain problems, perhaps of vital importance.[37] He also says that a similar remark can be applied to human beings individually.

Talking about metaphor, I cannot but point out the collection of articles *Mathematics as metaphor* by Manin (2007).

13 Another Return to the Classics

Thom had the particular wish to make topology and the theory of singularities a universal framework which would provide models for all phenomena in the natural and social sciences. Was he convincing? That is another question, but it is a well-known fact that history of knowledge is full of examples of sketches of fundamental theories that took centuries to convince, because they were ahead of their time.

Among the mathematicians who preceded Thom and whose work had a strong philosophical dimension, one can think of Descartes, Leibniz, Euler, Lambert, Riemann, Poincaré, Grassmann, Brouwer, and Weyl, to mention only some of the most famous. All of them considered that the language of mathematics, and in particular that of geometry, goes beyond the mathematical sciences. To mention, while remaining close to Thom's ideas, a famous case of a preeminent mathematician who proposed an idea that eventually was not understood by his contemporaries, one may think of Leibniz's *universal characteristic*, a theory which he intended to develop in reaction to the methods of algebra, which did not satisfy him.[38] The universal characteristic is closely connected with topology in many ways. Leibniz described this field as being more general than that of rigid Euclidean geometry and of Descartes's analytic geometry. He declared that this domain would be purely qualitative and that it was concerned with the study of figures independently of their metrical properties.

Thom refers to Leibniz's universal characteristic in the book *Morphogénèse et imaginaire* (Thom et al. 1978, p. 23):

[37] P. Cassou-Noguès quotes this letter as being part of the Gödel collection held in the library of the Institute of Advanced Study in Princeton. The letter is also quoted by Hao Wang in (Wang 1996, p. 4).

[38] Leibniz used several names for the new field he had in mind: *analysis situs, geometria situs, characteristica situs, characteristica geometrica, analysis geometrica, speciosa situs*, etc.

> In fact, the ultimate ambition of catastrophe theory is to abolish the distinction between mathematical language and natural language which has been rampant in science since the Galilean break. [. . .] A geometrical modeling of ordinary verbal thought will only be of interest if we can, thanks to it, arrive at assertions that the usual logic of natural language cannot provide. This supposes that we are able to:
> 1. Modelize geometrically all the (rigorous) deductions of ordinary thought. In other words: to realize the Leibnizian dream of "universal characteristic".
> 2. Go beyond this.

Leibniz had brought together a number of ideas on this subject and submitted them to Christiaan Huygens, whom he considered as his mentor, but the latter did not find these ideas worth exploring.[39] Leibniz, who had an infinite admiration for his former master, abandoned the project. It is interesting to read here an excerpt from a letter written by Leibniz to Huygens, dated September 8, 1679:

> I am still not happy with Algebra, since it provides neither the shortest ways nor the most beautiful constructions of Geometry. This is why when it comes to that, I think that we need another analysis which is properly geometric or linear, which expresses to us directly *situm*, in the same way as algebra expresses *magnitudinem*. And I think that I have the tools for that, and that we might represent figures and even engines and motion in words, in the same way as algebra represents numbers in magnitude. I am sending you an essay, which seems to me worth considering; there is no other person than you, Sir, who can better judge it, and your opinion will count for me more than many others.

Leibniz sent to Huygens his manuscript titled *Characteristica Geometrica*, dated August 10, 1679. The piece is published in Volume II of Leibniz' *Mathematische Schriften*, edited by Gerhardt (Leibniz 1849), Vol. II, pp. 141–168). and it is also reproduced in Vol. XIII of Huygens' *Collected Works* (Huygens 1899), pp. 219–224). It starts with the words:

> I found some elements of a new characteristic, completely different from Algebra and which will have great advantages for the mental representation in an exact and natural manner, although without figures, of everything that depends on the imagination. Algebra is nothing but the characteristic of undetermined numbers or magnitudes. But it does not directly express place, angles and motions, from which it follows that it is often difficult to reduce, in a computation, what is in a figure, and that it is even more difficult to find geometrical proofs and constructions which are enough practical even when the Algebraic calculations are done. But this new characteristic, based on figures we can see, cannot fail to give at the same time the solution and the construction of geometric proof, in a natural manner, and by analysis, that is, by determined ways. Algebra is obliged to suppose the Elements of Geometry; instead, this characteristic pushes the analysis to the end: if it were completed in the way I conceive it, one could make in characters which will be only letters of the alphabet the description of a machine whatever compound it could be, which would give the means to the mind to know it distinctly and easily with all the parts and even with their use and movement, without using figures or models and without hampering the imagination.

[39] The two men had met for the first time in 1672, in Paris, where Huygens was settled since 1666. Huygens was 17 years older than Leibniz, and for some time he was his private teacher in mathematics.

Leibniz then explains in more detail his vision of this new domain of mathematics, and where it stands with respect to algebra and geometry, giving several examples of a way to denote loci, showing how this allows the expression of statements such as the fact that the intersection of two spherical surfaces is a circle, and the intersection of two planes is a line. The piece ends with the words (Huygens 1899, p. 224):

> I have only one remark to add, namely, that I see that it is possible to extend the characteristic to things which are not subject to imagination. But this is too important and it would lead us too far for me to be able to explain that in a few words.

Huygens was not convinced by Leibniz's project. He responded to him in a letter dated November 22, 1679 ([176] p. 577):

> I have examined carefully what you are asking me regarding your new characteristic, and to be frank with you, I cannot not conceive the fact that you have too much expectations from what you spread on me. Your example of places concerns only realities that were already perfectly known, and the proposition saying that the intersection of a plane and a spherical surface makes the circumference of a circle does not follow clearly. Finally, I cannot see in what way you can apply your characteristic to which you seem you want to reduce all these different matters, like the quadratures, the invention of curves by the properties of tangents, the irrational roots of equations, Diophantus' problems, the shortest and the most beautiful constructions of the geometric problems. And what still appears to me stranger than anything else is the invention and the explanation of machines. I say it to you unsuspiciously: in my opinion this is only wishful thinking, and I need other proofs in order to believe that there could be some reality in what you present. I would nevertheless restrain myself from saying that you are mistaken, knowing the subtlety and the deepness of your mind. I only beg you that the magnificence of the things you are searching won't let you postpone from giving us those which you already found, like this Arithmetic Quadrature you discovered, concerning the roots of the equations beyond the cubical, if you are still satisfied with it.

On continuity in nature, we already mentioned that Thom constantly referred to Aristotle. Talking about Leibniz, we may recall that he declared (Leibniz 1840): "Nothing is done all at once, and it is one of my great and most verified maxims that nature never makes jumps. I called this the law of continuity, when I spoke of it in the news of the republic of letters; and the use of this law is very considerable in Physics," a sentence that reminds us of what Tom declared three centuries later.

The essay that Leibniz sent to Huygens remained practically unknown. But it drew the attention of a few nineteenth-century mathematicians, including Hermann Grassmann (1809–1877), who is considered as the founder of abstract multi-dimensional linear algebra, and, most of all, of a geometric interpretation of this algebra. He was among the first to stress on the importance of Leibniz' *Analysis situs*. We refer the interested reader to the articles (Heath 1917) by Heath and (Echeverrìa 1979) by Echeverrìa.

There exist two recent commented editions of Leibniz's work on the geometric characteristic, both included in doctoral dissertations, the one by J. Echeverrìa (Leibniz 1995), in France, and the other one by de Risi, (de Risi 2007), in Germany.

Like Thom did after him, Grassmann also stressed on the fact that mathematics is the adequate language for understanding the physical world. He formulated a "universal geometric calculus" which was successfully used in the foundations of mechanics, relativity, electrodynamics and quantum mechanics. We refer the reader to the survey article (Hestenes 1996) by D. Hestenes, in which the author gives a glimpse of Grassmann's vision and concludes by the words "An adequate history of geometric calculus remains to be written."

We note in passing that Thom also talked about a universal grammar, see his paper (Thom 1978a).

14 Conclusion

Thom was a singular figure in the history of mathematics, and one of his important contributions was to challenge all dogmas, especially those of science. The fact that he had shifted his focus to philosophy and the consequences of this change were not understood by most of his colleagues: criticisms were voiced here and there, by mathematicians and others, even though the Fields medal assured him an immunity from any dispute of his mathematical talents.

It is not surprising that some colleagues of Thom considered his philosophical reflections as delirium. Some of the criticisms that came from the mathematical community were deeply unfair. In the introduction to his book on catastrophe theory (Arnold 1984), Vladimir Arnold writes (p. viii):

> Neither in 1965 nor later was I ever able to understand a word of Thom's own talks on catastrophes. He once described them to me (in French?) as "bla-bla-bla," when I asked him, in the early seventies, whether he had proved his announcements. Even today, I don't know whether Thom's statements on the topological classification of bifurcations in gradient dynamical systems depending on four parameters are true. [...] Nor am I able to discuss other, more philosophical or poetical declarations by Thom, formulated so as to make it impossible to decide whether they are true or false (in a manner typical rather of medieval science before Descartes or (the) Bacon(s)).[40]

Thom did not waver, but deep down, the reactions to his ideas hurt him. He wrote in *La science malgré tout* (Science despite everything) (Thom 1975a):

[40]It is fair to add that several years later, Arnold changed his mind concerning Thom. In an interview published in avril 1997 (Liu 1997), he writes: "I am deeply indebted to Thom, whose singularity seminar at the Institut des Hautes Études Scientifiques, which I frequented throughout the year 1965, profoundly changed my mathematical universe. I was always delighted by the way in which Thom discussed mathematics, using sentences obviously having no strict logical meaning at all. While I was never able to completely free myself from the straitjacket of logic, I was forever poisoned by the dream of the irresponsible mathematical speculation with no exact meaning."

> It used to be fashionable—and probably still is—in scientific circles, to rave about philosophy. And yet, who could deny that the only really important problems for man are philosophical problems? But philosophical problems, being the most important ones, are also the most difficult ones; in this field, to show originality is very difficult, a fortiori to discover a new truth. This is why the society, very wisely, has given up subsidizing research on philosophical subjects, where the profit is too uncertain, to devote its efforts to scientific research, where, thank God, it is not necessary to be a genius to do a "useful work".

Did the community of philosophers understand Thom? In some sense, the answer is simple: unlike the situation in mathematics, there is no philosophical community. In his article *Leaving mathematics for philosophy*, (Thom 1992c), Thom explains that in philosophy, things are not as simple as in mathematics: there are schools, cliques, etc.

Regardless of his involvement in philosophy, Thom had a personal view on mathematics and on mathematicians and why we do mathematics, which he expressed sometimes in a very metaphorical mode. In his paper *De l'icône au symbole. Esquisse d'une théorie du symbolisme* (From icon to symbol. A sketch of a theory of symbolism) (Thom 1973), he writes:

> I would like to see the mathematician as a perpetual newborn babbling in front of nature; only those who know how to listen to Mother Nature's answer will later manage to open a dialogue with her, and to master a new language. The others only babble, buzzing in the void — *bombinans in vacuo*. And where, you may ask, could the mathematician hear the answer of nature? The voice of reality is in the sense of the symbol.

Manin is among the mathematicians who understood Thom's approach since the beginning. In his contribution to Kyoto's ICM (1990), he writes (Manin 1990):

> [. . .] What is relevant is the imbalance between various basic values which is produced by the emphasis on proof. Proof itself is a derivative of the notion of "truth." There are a lot of values besides truth, among them "activities," "beauty" and "understanding," which are essential in the high school teaching and later. Neglecting precisely these values, a teacher (or a university professor) tragically fails. Unfortunately, this also is not universally recognized. A sociological analysis of the controversies around Catastrophe Theory of René Thom shows that exactly the shift of orientation from formal proof to understanding provoked such a sharp criticism. But of course, Catastrophe Theory is one of the developed mathematical metaphors and should only be judged as such.

In any case, Thom felt out of step with the mathematical milieu of his time, which was rather very monolithic. He writes (Thom 1980, p. 29):

> I never really thought of myself as a mathematician. In fact, a mathematician must have, in my opinion, a taste for difficulty, for beautiful, rich and deep structures. I do not have that taste at all. The ultra-refined structures that fascinate my colleagues—Lie groups, simple finite groups, etc., in short, all these kinds of mathematical mythologies—have never really interested me. On the other hand, I like things that move, flexible things that I can transform to my liking. I prefer the field of mathematics where one does not really know what he is doing! This is the reason why I consider mathematics today with a certain detachment, and I

> cannot say whether there is currently a strictly mathematical problem for which I have a deep interest.

Thom declared several times that even though he became a mathematician by accident, mathematics brought him great satisfaction. In the interview (Thom 1991c), he says: "Really, when one has found a theorem in his lifetime, he can tell that he participated in a certain form of immortality, whatever he did. An illusion perhaps, but among all the fictitious immortalities by which we are deluded, this is still one of the most solid" (Thom 1991c, p. 73)]. In the article *La science malgré tout* (Thom 1975a), he writes: "I know only one really difficult science: mathematics."

Appendix: Jean Cavaillès and Albert Lautman

Let me say a few words on the two philosophers of science, Jean Cavaillès (1903–1944) and Albert Lautman (1908–1944) to whom Thom was attracted as a young student.

Cavaillès is the author of an important philosophical corpus on the foundations of mathematics in which the stress is on a dynamical evolution of these foundations. Among the multitude of philosophical schools of the first half of the twentieth century, Cavaillès was situated between Hilbert's formalism and Brouwer's intuitivism, a trend that he used to call "modified formalism." He was close to the mathematician Émile Borel, one of the French representatives of the "semi-intuitionist" current. As a matter of fact, in 1940, Cavaillès wrote an authoritative synthesis of Borel's theory of the quantification of chance (Cavailles 1940), which constitutes the latter's point of view on the philosophy of probability theory expressed in Vol. IV of his *Traité du calcul des probabilités et de ses applications* (Borel 1939). The volume is titled *Valeur pratique et philosophie des probabilités*.

Cavaillès taught at the École Normale Supérieure before Thom entered there. His philosophical work, which was in part the result of a close collaboration with Emmy Neother, had a non-negligible impact on the mathematical research done at the École.

Lautman had been Cavaillès's student at the École Normale. In 1937, he defended a doctoral dissertation, in two parts, the first one titled *Essai sur les notions de structure et d'existence en mathématiques* (Essay on the notions of structure and existence in mathematics), and the second one, *Essai sur l'unité des sciences mathématiques dans leur développement actuel* (Essay on the unity of the mathematical sciences in their present development).[41] Following the path of Cavaillès

[41] The French doctoral system (doctoral d'état) required the defense of two dissertations. The first one was the main thesis, and the subject of the second one was proposed by the jury of the doctorate, a few months before the defense. Its preparation was supposed to take only a few months (usually 3 to 4).

(and before him, that of Poincaré), Lautman considered that both the formalist and the intuitionist movements were a failure. He was an advocate of structuralism in the tradition of Bourbaki. He was a promoter of the concept of unity of mathematics, see the collection of articles (Lautman 1977).

Cavailles and Lautman found the sources of their theories in the recent developments in mathematics and physics (notably quantum physics). They both wondered about the role and the consequences of the various movements of thought that had appeared at the end of the nineteenth century on the philosophy of mathematics (conventionalism, logicism, constructivism, formalism, etc.). Even if they diverged on some points, and in particular on the organization of mathematics as a system of thought, they both considered that the philosophy of mathematics must necessarily be at the center of any metaphysical theory, as it was already for Plato, Heidegger and other philosophers. Their work is an embodiment of this approach. Both Cavailles and Lautman had close relations with mathematicians of the Bourbaki group such as Cartan, Chevalley, Dieudonné, Ehresmann, and Weil, and they were very much interested in Boubaki's project of writing complete treatises on the foundations of several topics in mathematics.

During the Second World War, Lautman et Cavaillès joined the resistance to the German occupation of France. They were both shot by the Nazis in 1944. The first was 36 years old and the second 40.

The French Society of Philosophy devoted its session of February 4, 1939 to the discussion of Cavaillès and Lautman's works. Several mathematicians, including Henri Cartan, Paul Dubreuil, Paul Lévy, Maurice Fréchet, Charles Ehresmann, and Claude Chabauty, were present at that session. A report on this session was published after the liberation of France, in the Bulletin de la Société française de philosophie (La pensée mathématique 1946) (1946). Thom may have been present at this debate, but being still too young, his name is not mentioned in the Annals.

References

Aristote (1993) Problèmes, texte établi et traduit par Pierre Louis, 3 vol. Les Belles Lettres, Paris

Aristotle (1908) The politics (trans. Jowett B). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press

Aristotle (1909) Poetics (trans. Bywater I). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press, Oxford

Aristotle (1910) History of animals (trans. D'Arcy Wentworth Thompson). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press

Aristotle (1924) Rhetorics (trans. Rhys Roberts W). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press, Oxford

Aristotle (1927) On the soul (trans. JA. Smith). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press, Oxford

Aristotle (1928), Metaphysics (trans. Ross, WD). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press, Oxford, 2d

Aristotle (1930) The Physics (trans. Hardie P, Gaye RK). In: Ross WD, Smith JA (eds) The works of Aristotle translated into English. Clarendon Press, Oxford

Arnold VI (1984) Catastrophe theory, translated from the Russian (Original version, 1981). Springer, Berlin/Heidelberg/New York/Tokyo
Atiyah M (2002) special article. In Mathematics in the 20th Century. Bull Lond Math Soc 34:1–15
Atiyah M et Iagolnitzer D (ed.) (1977) Fields medalists' lectures. World Scientific Series in 20th Century Mathematics. World Scientific et Singapore University Press, Singapore
Atiyah M et al (1994) Responses to: A. Jaffe and F. Quinn, Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics (Bull. Amer. Math. Soc. (N.S.) 29, 1993, No. 1, p. 1–13), Bull. Amer. Math. Soc. (N.S.), 30(2):178–207
Borel É (1939) Traité du calcul des probabilités et de ses applications, t. 4, fasc. 3: Valeur pratique et philosophie des probabilités. Gauthier-Villars, Paris
Cassou-Noguès P (2007) Les démons de Gödel: Logique et folie. Seuil, Paris
Cavailles J (1940) Du collectif au pari. Rev Métaphys Morale 47:139–163
de Risi V (2007) Geometry and Monadology: Leibniz's Analysis Situs and philosophy of space. Birkhäuser
Echeverrìa J (1979) L'Analyse géométrique de Grassmann et ses rapports avec la caractéristique géométrique de Leibniz. Studia Leibnoziana 11:223–273
Farmaki V, Negrepontis, S (2021) The paradoxical nature of mathematics. In: Papadopoulos (ed) Topology and geometry. A collection of essays dedicated to Vladimir G. Turaev. European Mathematical Society, Berlin, pp 599–642
Farmer Y (2010) Topologie et modélisation chez René Thom: l'exemple d'un conflit de valeurs en éthique. Philosophiques 37(2):369–386
Haefliger A (1988) Un aperçu de l'œuvre de Thom en topologie différentielle (jusqu'en 1957). Publ Math IHES 68:13–18
Hamburger J (ed.), *La philosophie des sciences aujourd'hui*, Paris, Gauthier Villars, 1986.
Heath AE (1917) The geometrical analysis of Grassmann and its connection with Leibliz's characteristic. The Monist 27:35–62
Héraclite (1988) Fragments (trans. Weil S), In Dieu dans Héraclite, Œuvres complètes de Simone Weil, éd. A. A. Devaux et F. de Lussy, IV. 2. Gallimard, Paris, pp 144–146
Hestenes D (1996) Grasmann's vision. In: Schubring G (ed) Hermann Gunther Grasmann (1809–1877): Visionary Mathematician, Scientist and Neohumanist Scholar. Kluwer Academic Publishers, Dordrecht/Boston, pp 191–201
Hironaka H (1973) Introduction aux ensembles sous-analytiques (rédigé par André Hirschowitz et Patrick Le Barz). In Singularités à Cargèse, 1972, Astérisque, Nos. 7-8, Soc. Math. France, Paris, pp 13–20
Huygens Ch. (1899) Œuvres complètes, t. VIII. Correspondance 1676–1684 (ed. Johannes Bosscha jr.). Martinus Nijhoff, Den Haag
Jaffe A, Quinn F (1993) Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics. Bull Am Math Soc 29:1–13
La pensée mathématique, Séance du 4 février 1939, exposés de Jean Cavaillès et Albert Lautman, discussion avec H. Cartan, C. Chabauty, P. Dubreil, C. Ehresmann, M. Fréchet, J. Hyppolite, P. Lévy et P. Schrecker, *Bulletin de la Société française de philosophie*, 40e année, 1946, pp 1–39
Lautman A (1977) Essai sur l'unité des mathématiques et divers écrits. Union générale d'éditions, Paris
Leibniz GW (1840) Nouveaux essais. Avant-Propos. Opera philosophica 1:198
Leibniz GW (1849) Leibnizens Mathematische Schriften. Edited by C. I. Gerhardt. Verlag von A. Asher & Comp, Berlin
Leibniz GW (1995) La caractéristique géométrique, text edited and annotated by J. Echeverrìa, translation, notes and postface by M. Parmentier, Coll. Mathesis. Vrin, Paris
Lejeune C (1978) Du point de vue du tiers . . ., in Circé. Cahiers de recherche sur l'imaginaire, 8-9, éditions Lettres Modernes, Paris, pp 91–118
Liu SH (1997) An interview with Vladimir Arnold. Notices of the AMS 44:432–438

Manin YI (1990) Mathematics as metaphor. In: Proceedings of ICM, Kyoto, Mathematical Society of Japan, vol II, pp 1165–1171
Manin YI (2007) Mathematics as metaphor: Selected essays of Yuri I. Manin American Mathematical Society, Manin
Mersenne M (1965) Harmonie universelle, contenant la théorie et la pratique de la musique, 3 vol., Paris, Sébastien Cramoisy, 1636; reprint. CNRS Éditions, Paris
Montel P (1932) Leçons sur les fonctions entières ou méromorphes. Gauthier-Villars, Paris
Negrepontis S, Plato on geometry and the geometers, In Geometry in history, ed. S. G. Dani and A. Papadopoulos, Springer, Cham 2019, p. 1-88.
Nimier J (ed) Entretiens avec des mathématiciens (l'Heuristique mathématique). Publication de l'IREM de Lyon, Villeurbanne
Papadopoulos A (2018a) René Thom, des mathématiques à la biologie et à la philosophie. In: Thom R, Portait mathématique et philosophique. CNRS éditions, pp 19–149
Papadopoulos A (ed) (2018b) René Thom, Portait mathématique et philosophique. CNRS éditions, 460 pages
Papadopoulos A (2018c) Logos et analogie: notes sur deux thèmes récurrents dans la pensée de Thom, In René Thom, Portait mathématique et philosophique (A. Papadopoulos, ed.). CNRS éditions, pp 261–321
Plato (1937) Timaeus translated with commentary by F. M. Cornford. Kegan Paul, London
Porte M (1990) Psychanalyse et sémiophysique: Études épistémologiques en métapsychologie et en dynamique qualitative, thèse de doctorat, Université de Paris 7
Porte M (1994) La dynamique qualitative en psychanalyse, with a Preface by René Thom. Presses Universitaires de France, Paris
Porte M. Notule biographique (notes pour une biographie de Thom), unpublished, available on the internet
Thom R (1949) Sur une partition en cellules associée à une fonction sur une variété. C. R. Acad. Sci. Paris 228:973–975
Thom R (1952) Espaces fibrés en sphères et carrés de Steenrod. Ann. Sci. Éc. Norm. Supér 69: 109–182
Thom R (1962) Sur la théorie des enveloppes. J Math Pures et Appliquées, t. XLI, fasc. 2, pp 177–192
Thom R (1970) Les mathématiques modernes: une erreur pédagogique et philosophique? L'Âge de la Science, III, 3, Paris, Dunod, pp 225–242, Modern mathematics: an educational or philosophical error? Amer. Sci., 59, 1971, pp. 695–699. New directions in the philosophy of mathematics, pp. 67–78, Birkhäuser Boston, Boston, MA (1986); Translated from the French in Fiz.-Mat.-Spis.-Blgar.-Akad.-Nauk., 15 (48) (1972), pp. 151–161 (Bulgarian); Translated from the French in JN: Obzornik-Mat.-Fiz., 19, 4, pp. 161–168 (Slovenian). (There are other translations.)
Thom R (1975) Topologie et linguistique. In: Essays in topology and related topics, mémoirs dedicated to Georges de Rham, Springer, 1971, pp 226–248. Russian version, transl. by Yu. Manin, in Uspehi Mat. Nauk 30, n° 1, 181, pp 199–221
Thom R (1972) Stabilité structurelle et morphogénèse: Essai d'une théorie générale des Modèleo, Mathematical physics monograph series. W. A. Benjamin, Inc., Reading. Advanced Book Program, Paris, Ediscience
Thom R (1973) De l'icône au symbole. Esquisse d'une théorie du symbolisme. Colloque de Mons, Le rôle du symbole en mathématiques, 28-29 avril 1972, Belgique, in Cahier intern. du symbolisme, 22/23, pp 85–106
Thom R (1974a) La linguistique, discipline morphologique exemplaire. Critique 322 René, pp 235–245
Thom R (1974b) Modèles mathématiques de la morphogenèse, 1st edn. Christian Bourgois, Paris
Thom R (1975a) La science malgré tout. Encyclopaedia Universalis, Organum
Thom R (1975b) Les mathématiques et l'intelligible. *Dialectica* 29:71–80

Thom R (1978a) La double dimension de la grammaire universelle, AKUP, Cologne University, 1977, Published also in Circé. Cahiers de recherche sur l'imaginaire, 8-9, éditions Lettres Modernes, Paris, pp 78–90

Thom R (1978b) Les archétypes en biologie et ses avatars modernes, Bulletin de la société ligérienne de philosophie, 1976. Published also in Circé. Cahiers de recherche sur l'imaginaire, 8-9. éditions Lettres Modernes, Paris, pp 52–64

Thom R (1978c) La notion d'archétype en biologie et ses avatars modernes, Bulletin de la société ligérienne de philosophie, 1976. Published also in Circé. Cahiers de recherche sur l'imaginaire, 8-9. éditions Lettres Modernes, Paris, pp 25–39

Thom R (1978d) Le statut épistémologique de la théorie des catastrophes, exposé au séminaire de philosophie et mathématique, ENS, Paris, 21 mars 1976. Circé, Cahiers de recherche sur l'imaginaire, 8-9, Paris, éditions Lettres Modernes, pp 7–24

Thom R (1978e) Les racines biologiques du symbolisme, Conférence à l'Institut de recherche sur l'imaginaire, Chambéry 21-22 mai 1976, Circé, *Cahiers de recherche sur l'imaginaire*, n° 8-9, Paris, Éditions Lettres Modernes, pp 40–51

Thom R (1980) Paraboles et catastrophes, Entretiens sur les mathématiques, la science et la philosophie réalisés par G. Giorello et S. Morini, Paris, Flammarion, collection Dialogues, 1983. First Italian edition, italienne, Milan, Saggiatore

Thom R (1981) Mathématiques essentielles, École d'automne de biologie théorique, Abbaye de Solignac, 14 septembre - 2 octobre

Thom R (1985) La méthode expérimentale: un mythe des épistémologues (et des savants)?, Comptes Rendus de l'Académie des Sciences, série générale, La vie des sciences, II, 1, Paris, Gauthier-Villars, janv.-fév. pp 59–68

Thom R (1987a) Considérations sur la finalité: d'Aristote à Geoffroy Saint-Hilaire. Hommage au Pr. Pierre-Paul Grassé. Evolution, Histoire, Philosophie, Paris, Masson, pp 155–160. D'après S. Khaznadar, conférence à Osaka, colloque Structuralist Biology, décembre 1986

Thom R (1987b) Mathématique et réalité: faut-il croire Aristote? Colloque de la Société Wittgenstein, Kirchberg am Wechsel, 4-13 août 1986. Publié in Logik, Wissenschaftstheorie und Erkenntnistheorie. Akten des 11. internationalen Wittgenstein Symposium, Wien, Holder-Pichler-Tempsky, 1987, pp 115–121

Thom R (1988a) Philosophie de la singularité. Singularities, Banach Cent. Publ. 20, 7–14

Thom R (1988b) Les intuitions topologiques primordiales de l'aristotélisme, Revue Thomiste, 88, 1988b, n° 3, p. 393-409.

Thom R (1988c) Un rapport problématique: la rencontre théorie-expérience. In The role of experience in science, Actes d'une conférence de l'Académie internationale de philosophie des sciences tenue à Heidelberg en 1986. de Gruyter, Berlin, pp 181–189

Thom R (1989a) Causality and finality in theoretical biology: a possible picture. Proceedings of the conference at Abisko, Sweden, 4-8 May 1987. In Newton to Aristotle. Toward a theory of models in living systems, J. Casti et A. Karlqvist ed. Springer, New York, pp 39–45

Thom R (1989b) Problèmes rencontrés dans mon parcours mathématique: un bilan. Publications mathématiques de l'IHÉS 70:199–214

Thom R (1976) The twofold way of catastrophe theory, In Structural stability, the theory of catastrophes and applications in the sciences, Springer Verlag, Lecture notes in mathematics, no. 525, pp 235–252. French transl. Les deux voies de la théorie des catastrophes, In Apologie du Logos, Paris, Hachette, 1990a, pp 373–394

Thom R (1990b) Esquisse d'une sémiophysique: Physique aristotélicienne et théorie des catastrophes, Paris, InterEditions, 1988. English translation by Meyer V, Semio Physics: A Sketch. Aristotelian Physics and Catastrophe Theory, Addison-Wesley

Thom R (1990c) Homéomères et anhoméomères en théorie biologique d'Aristote à aujourd'hui. Conférence à l'île d'Oléron, Biologie aristotélicienne, 26–30 juin 1987. Biologie, Logique et Métaphysique chez Aristote, éd. P. Pellegrin, Éditions du CNRS, Paris, pp 491–511

Thom R (1990d) Les réels et le calcul différentiel, ou la mathématique essentielle, Talk given at the Séminaire de philosophie et mathématiques, École Normale Supérieure, 17 mars 1982, in Apologie du Logos, Paris, Hachette, 1990d, pp 314–331
Thom R (1988) Structure et fonction en biologie aristotélicienne, in Biologie théorique, Solignac, Paris, éditions du CNRS, 1993, also published in Apologie du logos, Paris, Hachette, 1990e, pp 247–266
Thom R (1991a) Matière, forme et catastrophes. In: Sinaceur MA (ed.) Penser avec Aristote. érès, Paris, pp 367–398
R. Thom (1991b) Saillance et Prégnance, In: L'inconscient et la Science, ed. R. Dorey, Dunod Paris, 64–82
Thom R (1991c) Prédire n'est pas expliquer. Entretiens avec Émile Noël, Paris, Eshel, coll. La Question
Thom R (1992a) *La théorie des catastrophes*, interview with Francis Bouchet. INA, Eshel, Paris
Thom R (1992b) Actualité de la physique aristotelicienne. In: Geometry and analysis in nonlinear dynamics Proc. Workshop Chaotic Dyn. Bifurcations, Groningen/Neth. 1989, Pitman Res. Notes Math. Ser. 222, 85–96
Thom R (1992c) Leaving mathematics for philosophy. In: Mathematical research today and tomorrow (Barcelona, 1991), 1-12, Lecture notes in Math., n° 1525. Springer, Berlin
Thom R (1992d) La théorie des catastrophes, entretien avec Francis Bouchet. INA, Eshel, Paris
Thom R (1994) The task of Mathematics in describing the world, Proceedings of the symposium 25 Años de Matematicas en la Universidad de la Lagune, Secretariado de Publicaciones de la Universidade de La Laguna, Tenerife, 21–25 novembre, pp 83–86
Thom R (1995) Biologie Aristotélicienne et topologie. Boletim da Sociedade Portuguesa de Matemática 32:3–14
Thom R (1996a) Histoire de la transversalité, d'Aristote à la théorie des catastrophes. In: Universalité de la transversalité, a talk given at the Journée de réflexion du 6 février: La transversalité dans la Géométrie Moderne, Maison des Sciences de l'Homme, Paris
Thom R (1996b) Qualitative and quantitative in Evolutionary Theory with some thoughts on Aristotelian Biology. In Systematic Biology as an Historical Science. Memorie della Società Italiana di Scienze Naturali e del Museo Civico di Storia Naturale di Milano XXVII(I):115–117
Thom R (1997) The hylemorphic schema in mathematics, In Philosophy of mathematics today, Episteme 22. Kluwer Academic Publishers, Dordrecht, pp 101–113
Thom R (1998) Comment la biologie moderne redécouvre la kinèsis d'Aristote. Bull Hist Epistém Sc Vie 4(3):197–208
Thom R (1999) Aristote topologue. Rev synth 4^{e} s., n° 1, pp 39–47
Thom R (1989) Exposé introductif. In: Logos et théorie des catastrophes, Petitot J (ed.), Actes du colloque international de Cerisy de 1982, Centre culturel international de Cerisy-la-Salle, Genève, Patiño. Published also in Gazette des mathématiciens, special edition, 2004, pp 7–22
Thom R (2017) Œuvres mathématiques, vol I. Société mathématique de France
Thom R (2021) Métaphysique extrême; entretiens. In: Onicar? Foucault, Duby, Dumézil, Changeux, Thom: Cinq grands entretiens au champ freudien, Navarin éd., pp 51–101
Thom R, Lejeune C, Duport J-P (1978) Morphogenèse et imaginaire. Lettres modernes, Circé, Coll. Méthodologie de l'imaginaire, Paris
Wang H (1996) A logical journey: from Gödel to philosophy. The MIT Press, Cambridge, MA
Zeeman EC (1965) Topology of the brain. In: Mathematics and computer science in biology and medicine. H.M. Stationary Office, London

Heuristics and Mathematical Practice

Otávio Bueno

Contents

1 Introduction 432
2 Norms in Mathematics 433
3 Kinds of Mathematical Proofs and Their Roles 434
3.1 Kinds of Mathematical Proofs 434
3.2 Roles Played by Mathematical Proofs 435
4 Justificatory Proofs 436
5 Explanatory Proofs 437
6 Unification-Inducing Proofs 439
7 Concept-Refining Proofs 440
8 Conclusion: Pluralism about Proofs 441
References 441

Abstract

Proofs are central to mathematical practice in large part due to the heuristic role that some of them play. Not only do they help establish a result, but often provide new avenues of mathematical research. Jody Azzouni has argued that underlying the practice of creating mathematical proofs there is a very specific norm: to each proof there should be a corresponding algorithmic derivation, a derivation in an algorithmic system. Here a framework is provided to classify and assess mathematical proofs. It is argued that there is a plurality of kinds of proofs in mathematics and a plurality of roles these proofs play. In the end, mathematical practice is far less unified in this respect than it may seem to be.

For helpful discussions, I would like to thank Joe Auslander, Jody Azzouni, Alan Baker, Peter Caws, Michael Clarkson, Mark Colyvan, Kenny Easwaran, Ed Erwin, Michele Friend, Elaine Landry, Marc Lange, Peter Lewis, Catarina Dutilh Novaes, Andrea Pedeferri, Eric Saidel, Harvey Siegel, Susan Vineberg, Charlotte Werndl, and Tadeusz Zawidzki.

O. Bueno (✉)
Department of Philosophy, University of Miami, Coral Gables, FL, USA
e-mail: obueno@miami.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_42

Keywords

Proofs · Explanation · Unification · Set theory

1 Introduction

Mathematical proofs have multiple functions in mathematics. They are, of course, primarily used to establish the relevant theorems. But it would be inadequate to conclude that this exhausts the role they play. Some proofs also provide explanation – an understanding of why a particular theorem holds, rather than simply establishing that it holds. These proofs often involve some form of construction or, in some cases, some visualizable aspects as an integral part of it. Proofs also develop ways of refining concepts, typically in the context of alleged counterexamples that need to be taken into account and avoided. Concept refinement in mathematical theorem proving is something that Imre Lakatos insightfully highlighted in his work (see Lakatos 1976, 1978). Other proofs connect different parts of mathematics that were thought of as being unconnected. Andrew Wiles' proof of Fermat's last theorem is often presented as providing that sort of unification (Wiles 1995). Some accounts of mathematical explanation connect the explanatory role of mathematical theorizing with this unificatory role; but not all do. In the end, different types of proof play different roles in mathematics.

In what follows, these issues are explored and a framework to classify proofs and assess proofs is offered. It is argued that there is a plurality of kinds of proofs in mathematics and that mathematical practice is far less unified in this respect than it has been suggested. Jody Azzouni argues that, underlying the practice of creating mathematical proofs, we find the norm to the effect that there should always be a corresponding algorithmic derivation, a derivation in an algorithmic system (Azzouni 1994, 2004, 2006). In many instances this is the case, but it need not always be so. Some informal proofs do not seem to be guided or constrained in any way by an underlying algorithmic derivation. As will become clear, there is far more plurality in mathematical proofs than meets the eye. In particular, there is not a single answer to the question of what a mathematical proof really proves. In different contexts, different kinds of answers are offered, with different goals and distinctive norms. It is pluralism all the way down.

I will start by considering some difficulties inherent in Azzouni's formulation of the relevant norm. I will then offer a taxonomy of different kinds of proofs, their roles and the norms they are expected to satisfy.

2 Norms in Mathematics

Azzouni (1994, 2004, 2006) argues that from the beginning of the twentieth century on, the norm of mathematical proofs requires that they correspond to algorithmic mathematical derivations. (I will call this the *algorithmic norm*.) Before the last century, he notes, the norm was simply to avoid gaps in the proofs. The new norm, which was made possible by the development of mathematical logic, is used to explain a significant aspect of mathematical practice: the stability of mathematical theorems. These theorems are preserved over time and turn out not to need to be revised when mechanical (algorithmic, automated) proofs of the relevant theorems are developed and ultimately implemented.

But why is the algorithmic norm generally not acknowledged as a norm in mathematical practice? This is because of its tacit character: mathematicians do not explicitly articulate the norm as such. The exception, of course, is found among computer scientists and those who work in automated theorem proving, since for them the search for algorithmic derivations is indeed the norm. It is telling that the norm is then explicitly acknowledged, articulated and implemented by the community. There is no tacit character in this case.

There are, however, several difficulties with this way of conceptualizing this norm of mathematics. First, a methodological problem emerges. It is unclear how a norm that is not explicitly acknowledged as such can actually play the role of a norm. It could be simply an accident that the various mathematicians have conformed to the norm rather than have actively shaped their practice by following it. The same practice can be structured and understood as emerging from a very different norm. If, in turn, that norm is explicitly acknowledged, it is more likely that it has played (or could have played) the role of a norm in configuring and structuring the practice. Moreover, a tacit norm is one for which it is difficult to provide evidence, since in order for one to establish that a certain norm is operative in the practice, one needs to determine to what extent the relevant practitioners acknowledge it and have acted in accordance to it. In fact, as I will argue below, this is precisely what has happened in the case of mathematics: when norms are active, they are acknowledged as such.

Second, can the stability of mathematical theorems be explained without the algorithmic norm? As will become clear, there are additional norms (each corresponding to the kind of proofs in question) that can be invoked as having played a role in shaping the relevant community. For these norms are not tacit and have been explicitly acknowledged by mathematicians. It is far more reasonable to countenance them as being norms than the tacit algorithmic norm.

Finally, even if the algorithmic norm had been explicitly formulated, it is unclear that it really explains the stability of mathematical theorems. For, as will become clear, the diversity of styles of mathematical proofs (that is, the various kinds of mathematical proofs we will discuss) challenges that a single norm that fits properly only a very narrow range of proofs could account for all the other kinds of proofs and their roles within mathematical practice. And since the algorithmic norm is tacit, Azzouni's account is unable to rule out that other norms are actually in place and structuring the practice.

I assume with Azzouni that some norm is active in mathematical practice. The alternative would be to consider that practice as entirely unregulated and unstructured. If this were the case, it would be very difficult to explain the stability of mathematics. So, I agree with Azzouni that some norm is needed, but I think he gave a disproportionate role to one particular norm, whereas there are many others in place.

For these reasons it is better to conceptualize mathematical practice in terms of other norms: those that are explicitly voiced by the relevant mathematical community. As will become clear, there is a plurality of such norms, each associated with a different kind of mathematical proof.

3 Kinds of Mathematical Proofs and Their Roles

3.1 Kinds of Mathematical Proofs

There are two main kinds of mathematical proofs: informal and formal proofs. Formal proofs are those familiar from formalized languages. Informal proofs comprise the bulk of proofs in mathematics. Among the informal proofs there are: descriptive (or language-based) proofs and diagrammatic proofs. Diagrammatic proofs are those that are implemented via a diagram rather than a linguistic description.[1] Descriptive proofs are the usual proofs mathematicians devise to establish theorems. They are the most common resource employed in mathematical practice and the central and distinctive feature of mathematical activity. As will become clear, proofs are valued for the information they provide about certain mathematical objects,[2] for allowing for conceptual refinement, and for the development of new techniques and the possibility of connecting fields that were not seen as being related before.

Among descriptive proofs, we find: truly informal proofs, pre-formal proofs, and post-formal proofs. Truly informal proofs need not correspond to any formalized counterpart. A significant portion of informal proofs is of this kind. Pre-formal proofs are regular informal proofs that could, with some work, be transformed into a formal one. Finally, post-formal proofs are proofs that emerge after a formal proof is provided. Some belong to the truly informal category, whereas others are pre-formal.[3]

[1] A discussion of diagrams as full-blown proofs can be found, e.g., in Brown (1999) and Azzouni (2013) (see also references therein). Nelson (1993) provides a number of examples of diagrammatic proofs.

[2] Whether these objects exist (or not) independently of their description by suitable comprehension principles is an additional issue, largely of philosophical interest, but not of much mathematical relevance.

[3] I am invoking here a distinction between pre-formal and post-formal proofs that Lakatos once introduced (1978), p. 61). He identified three kinds of mathematical proofs: pre-formal proofs, formal proofs, and post-formal proofs, where pre-formal and post-formal proofs belong to the

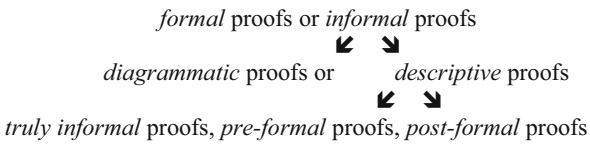


Fig. 1 Kinds of Mathematical Proofs

Schematically, we have the figure above (Fig. 1):

3.2 Roles Played by Mathematical Proofs

This characterization of kinds of proofs does not indicate the roles played by proofs. This motivates a different categorization in which proofs are grouped by their *function* rather than by their *kind*. The functional characterization goes as follows:

(a) *Justificatory proofs*: These proofs have only one purpose: to establish that a certain result holds given the assumptions in the theorem. And that is exactly what they do – nothing more, nothing less. Simple computational proofs are of this kind.
(b) *Explanatory proofs*: These proofs not only establish the result, but also help one understand why the result holds. When following the argument presented in the proof, the reader can see why the result is the case. These proofs are particularly valuable given that they provide additional information that illuminates the topic at hand. Constructive proofs are often of this kind.
(c) *Unification-inducing proofs*: These proofs bring together different areas or sub-fields of mathematics, or rely on methods that are commonly used in one area and apply them to a different area. They are also valuable for the increased information they provide.
(d) *Concept-refining proofs*: These proofs involve the refinement of concepts in order to adjust the target theorem, which only holds for the proper specification of the concepts involved. This is another feature that makes proofs particularly valuable – they are a significant source of conceptual innovation and refinement.

Typically, formal proofs only satisfy the first role. Formal proofs are not explanatory, do not induce unification nor do they lead to concept refinement. They are stripped down to the bare bones. The only trait they exhibit is the capacity to establish inferential links from one line in the proof to the next, with nothing by way of illumination, suggestive links to new sub-fields, introduction of new techniques, or daring ideas. They lack precisely those features that make mathematical

category of informal proofs. His classification did not explicitly accommodate diagrammatic proofs, although these proofs would clearly be informal. Some of the examples that Lakatos discussed involve diagrams, but only as aids rather than full-blown proofs. So a broader characterization than the one he offered is needed.

proofs so valuable to mathematicians. In contrast, rigorous informal proofs, those that are commonly found in mathematical practice, can satisfy all of the roles above. Not every proof will exemplify all the four roles, but the fact that they are capable of doing more than just justifying the result contributes significantly to their relevance.

4 Justificatory Proofs

Mathematical proofs, are traditionally conceptualized in terms of deductive, formal arguments. The standard used in these conceptualizations is the one offered by mathematical logic, in which the steps in a proof follow explicitly given syntactic patterns, and the meaning of the expressions, terms or sentences in the proof are ignored. Of course, a semantic counterpart to these expressions can be given, and if the logic is complete, there is a perfect match between the sentences that can be (syntactically) derived from the principles in question, and the sentences that follow logically (semantically) from such principles.

It turns out, however, that preciously little of this formal approach to mathematical proofs is of much help when we try to make sense of mathematical practice (Azzouni 2006). First, the meaning of mathematical expressions is central in actual mathematical proofs – i.e., those proofs that are actually used in mathematical practice. Often, the validity of the proof will turn on the interpretation of the concepts that are used in the proof. A purely formal argument is hardly ever used.

The mismatch between formal and semantic arguments should not be downplayed. Several significant results in mathematics only hold for logics that are incomplete, such as second-order logic: they allow for true results not to be derivable in the relevant system. A mathematical result that illustrates this situation is the categoricity of real analysis – the fact that every model of real analysis is isomorphic. This result only holds if it is assumed that the underlying logic is second-order,[4] and hence incomplete. Thus, we have here an illustration of the fact that the mismatch between semantic and syntactic components of mathematical proofs not only should be expected, but it is also central to mathematical practice.

More generally, the mismatch between a logical conception of proof and the standards for proof in mathematical practice has a long tradition. Consider, as a simple illustration, that the proofs in Euclid's *Elements* could not be faithfully formulated in syllogistic form. Interestingly enough, the fact that the most important form of logic in Greek antiquity could not be used to formulate and assess the actual proofs offered in the most important companion of geometry of its time was not considered to be a problem for either discipline. The mismatch has been with us for a long time.

There are, of course, areas in mathematics where formal proofs are the main goal: computer-generated proofs and automated theorem proving are significant

[4] For first-order theories, given that the Löwenheim-Skolem theorem holds, there will be non-standard models of real analysis, and the categoricity no longer holds.

developments in computer science. These branches of mathematics could not be developed and implemented without formal proofs. But, as far as mathematical practice is concerned, they are the exception rather than the norm.

5 Explanatory Proofs

Let us consider three very different examples of mathematical proofs:

(i) The Löwenheim-Skolem theorem: This is a model-theoretic argument that involves the construction of various kinds of models for first-order theories.
(ii) Proofs by mathematical induction: This is a structural argument, in the sense that the structure of the relevant objects – in this case, the natural numbers – is taken into account in order to implement the proof. It is due to the way in which natural numbers are built that such proofs can be run.
(iii) The intermediary-value theorem: This is a straightforward feature of continuous functions in the real line, and the proof explores this fact.

Do the proofs of these theorems *explain* the results besides establishing them? The answer depends on the conditions under which mathematical proofs explain the relevant results. There are different conceptions of explanation and different ways of characterizing what makes a mathematical proof explanatory (Mancosu (2011) provides a good survey, which I follow here). Some views insist that an explanatory proof should identify some intrinsic, characterizing feature of the objects under study. As Mark Steiner notes: "an explanatory proof depends on a characterizing property of something mentioned in the theorem" (1978, p. 147; italics omitted). This property characterizes a certain entity or structure, and the proof makes it clear that the result relies on that property.

But here we have a difficulty. Proofs by mathematical induction are typically not thought of as being explanatory (although there is some controversy about this issue; see Lange (2009) and Baker (2010). However, pace Steiner, proofs by mathematical induction seem to be obvious examples of explanatory proofs via a characterizing property, since they rely on the characterizing property of natural numbers. But are they really explanatory? Moreover, neither the Löwenheim-Skolem nor the intermediary-value theorems are established via a characterizing property.

Other views of mathematical explanation insist on the unifying role that some proofs provide (Kitcher 1981, 1989; Friedman 1974). On this model, the goal is to identify "the best tradeoff between minimizing the number of patterns of derivation employed and maximizing the number of conclusions generated" (Kitcher 1989, p. 432). This seems to suggest that finite axiomatizations of certain theories would be preferred as being more explanatory, since such axiomatizations minimize the number of patterns of derivation (the axioms) while maximizing the number of conclusions that are obtained. But is it clear why Zermelo-Fraenkel set theory is less explanatory than von Neumann-Bernays-Gödel, just because the latter is finitely axiomatizable and the former is not? It is also unclear whether any of the proofs of

the theorems mentioned above, (i)–(iii), are explanatory at all. None of them involve such unifying role.

Another possibility is to use the pragmatics of explanation as a framework to account for mathematical explanations. On this model, an explanation is an answer to a why question, and it involves three components (van Fraassen 1980, Chapter 5): a *topic* (what the explanation is about), a *contrast class* (the alternatives with respect to which the question is specified), and a *relevance relation* (that involves what would count as an acceptable answer to the question). In the case of the intermediary-value theorem, why does a continuous function that has negative and positive values also need to have a value equal to zero (rather than no value at all)? One could say that if we were to deny the statement of the theorem, a contradiction would be derived. This answers the question as to why the theorem holds. But clearly this would not be an explanation specific enough for the theorem in question: it holds for every mathematical theorem! How can the theorems mentioned above be illuminated? More generally, there are concerns as to the adequacy of using the theory of why questions in mathematics (Sandborg 1998).

Some diagrammatic proofs can also be explanatory in that they allow one to experience visually how the theorem in question can be true. Consider the following theorem:

Theorem: $1 + 2 + 3 + \ldots + n = n^2/2 + n/2$.

Its usual proof is given by mathematical induction. But it can also be given via a diagram (Brown 2005, p. 64) (Fig. 2):

Although the diagram only illustrates the case $n = 5$, it is clear that it can be extended without difficulty for any natural number. The problem, however, is that the explanatory nature of this proof only holds for a limited class of objects (diagrammatic proofs). And even these proofs may not be explanatory, since one cannot generalize from the diagrams.

As a result, we find here a significant pluralism about mathematical explanations (Mancosu 2011). It is unlikely that a single notion of explanation is capable of accommodating the variety of instances of explanation in mathematical practice. But no proposal invokes anything remotely near the algorithmic norm as part of what makes a proof explanatory.

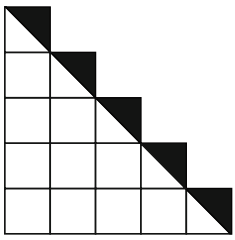


Fig. 2 Diagrammatic Proof

For those who think that mathematical theorems are necessarily true, the idea of an explanatory proof may sound bizarre. What else could be required from a proof but the fact that it establishes the target result, which holds no matter what? But the fact that mathematicians often acknowledge that some proofs are better than others on the grounds that they allow one to understand why the result holds (rather than only that it does hold) is an indication that the distinction between explanatory and non-explanatory proofs is acknowledged.

I do not actually think that mathematical theorems are, strictly speaking, necessary. The results hold given certain comprehension principles that characterize a certain domain and given the underlying logic. If either of them changes, the result in question may no longer hold. This view, which admittedly is not in the majority, accommodates very naturally the distinction between explanatory and non-explanatory proofs. Even though a certain theorem holds – given the comprehension principles and the underlying logic – there may be different strategies of establishing that this is the case. Some of these strategies may be more illuminating or informative than others, thus leading to explanatory proofs rather than non-explanatory ones. The algorithmic norm is not part of this process.

6 Unification-Inducing Proofs

Consider the case of Skolem's theorem regarding the relativity of set-theoretic notions: a countable model can satisfy the first-order sentence that states that there are uncountably many objects (such as uncountably many real numbers). When Thoralf Skolem first introduced his result, it was offered as an attempt to refute the adequacy of set theory as a suitable foundational framework for mathematics (Skolem 1922/1976). With the relativity of set-theoretic notions, Skolem thought, it should be clear that set theory fails to yield the uniqueness that would be expected in a suitable foundation for mathematics. In the end, instead of generating a *reductio* of the use of set-theoretic notions in mathematics, Skolem's result was simply incorporated as a significant feature of set theory. The relativity of set-theoretic notions to the model in which they are expressed is a central aspect of the theory – and it is a feature that has a significant heuristic role. This feature allows set theory to be used so fruitfully in mathematics, offering a rich framework to represent various results.[5]

There are, of course, cases in which proofs can play multiple roles, e.g., being both unification inducing and explanatory. In his examination of the role of a new set-theoretic axiom, Kurt Gödel notes:

[5] Despite set theory's powerful expressive resources, alternative frameworks have been developed as well. In some of them, sets do not feature at all. The use of category theory or second-order mereology to formulate mathematical theories provides obvious examples. No sets need to be invoked there.

> even in case it (a new set-theoretic axiom) had no intrinsic necessity at all, a (probable) decision about its truth is possible also in another way, namely, inductively by studying its 'success', that is, its fruitfulness in consequences demonstrable without the new axiom, whose proofs by means of the new axiom, however, are considerably simpler and easier to discover, and make it possible to condense into one proof many different proofs. The axioms for the system of real numbers, rejected by the intuitionists, have in this sense been verified to some extent owing to the fact that analytical number theory frequently allows us to prove number theoretical theorems which can subsequently be verified by elementary methods. (Gödel 1964, p. 265)

We find here the idea that a new axiom may be heuristically so fruitful that it can "condense into one proof many different proofs". In this case, we obtain unification at the level of proofs rather than at the level of theories or methods. This form of unification also provides additional justification to the relevant results if the proofs obtained using the new axiom can also be obtained with weaker – and less contentious – axioms.

Gödel also is reported as having drawn an interesting distinction between foundational and explanatory principles in mathematics (see Lakatos 1978):

> The role of alleged 'foundations' is rather comparable to the function discharged, in physical theory, by explanatory hypotheses (. . .). The so-called logical or set-theoretical 'foundation' for number-theory or of any other well established mathematical theory, is explanatory, rather than really foundational, exactly as in physics where the actual function of axioms is to *explain* the phenomena described by the theorems of this system rather than to provide a genuine 'foundation' for such theorems. (Mehlberg 1962, p. 86)

In this passage, we find the contrast between a foundational and an explanatory role of axioms. The foundational role is implicitly being questioned, since it seems to require a commitment to the certainty of the principles in the foundation, something that is ultimately incompatible with the empirically-oriented, and ultimately contingent, explanatory role. The sense of explanation in this context seems to be something akin to making sense of how the relevant results emerge (or could be derived) from the principles in question. The revisability of physical principles is thus transferred and assimilated to mathematics. The unification provided by the relevant axioms also brings some contingency, and there is very little role for the algorithmic norm in this context.

7 Concept-Refining Proofs

Lakatos has highlighted the importance of conceptual shifts in mathematics (Lakatos 1976, 1978). He also emphasized the significance of proof analysis as a tool in the formulation and development of mathematical proofs. This approach tends to downplay the importance of a purely formal approach to the understanding of mathematical reasoning, emphasizing the analysis of concepts, conceptual shifts, and conceptual refinement.

On this approach, mathematical results are relativized: they hold in virtue of certain principles that are assumed to hold in a given context, and fail otherwise. Consider, for example, the issue as to whether every set can be well ordered. The answer, ultimately, is: it depends. If the axiom of choice is assumed, then the result does hold, since the latter follows from the axiom. If the axiom does not hold, however, the result fails. Part of the process of proof analysis involves the identification of hidden premises in mathematical arguments, and in particular in attempts to prove certain results. This is particularly clear in the case of Ernst Zermelo's original axiomatization of set theory, and the resulting uncovering of the axiom of choice (Zermelo 1904/1976, 1908/1976).

Formal proofs cannot be the ultimate goal (or the norm) of mathematics. For even after a formal proof is established, informal proofs about the resulting proof need to be addressed. This is illustrated, for instance, by the existence of nonstandard models of arithmetic or analysis. It is realized that results that were proved formally depend on the model that was used to establish them: new results emerge in nonstandard models, but proofs about such models are typically done informally (see Lakatos 1978, p. 69). The resulting interplay between formal and informal proofs does not fit very well with a single algorithmic norm.

8 Conclusion: Pluralism about Proofs

The considerations above suggest that proofs play multiple roles in mathematical practice: they are sources of evidence (justificatory role); they explain why a result holds (explanatory role); they unify diverse domains (unification-inducing role); and they lead to the development and refinement of new concepts (concept-refining role).

Formal proofs are unable to deliver three out of four of these roles. It is thus unclear how the search for algorithmic systems should be the norm for mathematical proofs, as Azzouni has suggested. In the end, a norm that flies on the face of some much mathematical activity does not seem adequate. Mathematical proofs are far more varied and diverse than such a norm would recommend. As a result, a more pluralist view seems to provide a better alternative.

References

Azzouni J (1994) Metaphysical myths, mathematical practice: the ontology and epistemology of the exact sciences. Cambridge University Press, Cambridge

Azzouni J (2004) The derivation-indicator view of mathematical practice. Philos Math 12:81–105

Azzouni J (2006) Tracking reason. Oxford University Press, New York

Azzouni I (2013) The relationship of derivations in artificial languages to ordinary rigorous mathematical proof. Philos Math 21:247–254

Baker A (2010) Mathematical induction and explanation. Analysis 70:681–689

Benacerraf P, Putnam H (eds) (1964) Readings in the philosophy of mathematics. Basil Blackwell, Oxford

Brown JR (1999) Philosophy of mathematics. Routledge, London

Brown JR (2005) Naturalism, pictures, and platonic intuitions. In Mancosu, Jørgensen, and Pedersen (eds.) pp 57–73
Friedman M (1974) Explanation and scientific understanding. J Philos 71:5–19
Gödel, K. (1964) What is Cantor's continuum problem. In Benacerraf and Putnam (eds.) pp 258–273
Kazemier B, Vuysje D (eds) (1962) Logic and language: studies dedicated to professor Rudolf Carnap on the occasion of his seventieth birthday. Reidel, Dordrecht
Kitcher P (1981) Explanatory unification. Philos Sci 48:507–531
Kitcher P (1989) Explanatory unification and the causal structure of the world. In Kitcher and Salmon (eds.) pp 410–505
Kitcher P, Salmon W (eds) (1989) Scientific explanation. University of Minnesota Press, Minneapolis
Lakatos I (1976) In: Worrall J, Zahar E (eds) Proofs and refutations. Cambridge, Cambridge University Press
Lakatos I (1978) In: Worrall J, Currie G (eds) Mathematics, science and epistemology. Cambridge University Press, Cambridge
Lange M (2009) Why proofs by mathematical induction are generally not explanatory. Analysis 69: 203–211
Mancosu P (2011) Explanation in mathematics. The Stanford encyclopedia of philosophy (Summer 2011 edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2011/entries/mathematics-explanation/
Mancosu P, Jørgensen K, Pedersen S (eds) (2005) Visualization, explanation, and reasoning styles in mathematics. Springer, Dordrecht
Mehlberg, M. (1962). The present situation in the philosophy of mathematics. In Kazemier and Vuysje (eds) pp 69–103
Nelson R (1993) Proofs without words. Mathematical Association of America, Washington
Sandborg D (1998) Mathematical explanation and the theory of why-questions. Br J Philos Sci 49: 603–624
Skolem T (1922/1976) Some remarks in Axiomatized set theory. In van Heijenoort (ed) pp 290–301
Steiner M (1978) Mathematical explanation. Philos Stud 34:135–151
van Fraassen B (1980) The scientific image. Clarendon Press, Oxford
van Heijenoort J (ed) (1976) From Frege to Gödel: a source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA
Wiles A (1995) Modular elliptic curves and Fermat's last theorem. Ann Math 142:443–551
Zermelo E (1904/1976) Proof that every set can be well-ordered. In van Heijenoort (ed) pp 139–141
Zermelo E (1908/1976) Investigations in the foundations of set theory I. In van Heijenoort (ed) pp 199–215

Finding a Fit Among Philosophical Finitisms

Eamon Darnell and Aaron Thomas-Bolduc

Contents

1 Introduction 444
2 Historical Background: Hilbert and Gentzen 445
2.1 Hilbert's Program 445
2.2 Gentzen's Proof Theory 446
3 Gentzen's Proof 447
3.1 Ordinal Notations 447
3.2 The Consistency of Pure Number Theory 448
4 Takeuti's Proof 449
4.1 Toy Example: Proving the Well-Ordering of Positive Numerals 449
4.2 The 1-Eliminator 450
5 Terms and the 1-Eliminator 450
6 Proving a Well-Order with the 1-Eliminator 451
7 Defining Finitism After Hilbert 452
7.1 More than PRA? 453
7.2 The Hilbert-Gentzen Standpoint 453
8 The Takeuti-Veblen Finitist Standpoint 454
8.1 The Limits of the Extended Program 454
8.2 The Importance of Takeuti-Veblen Standpoint 458
References 460

Abstract

David Hilbert sought to secure the epistemic foundations of mathematics by providing consistency proofs of axiomatized mathematical theories from within the *finite standpoint*. This standpoint requires concrete constructions without reference to completed infinities. In 1938, Gerhardt Gentzen proved the

E. Darnell (✉)
University of Toronto, Toronto, ON, Canada
e-mail: eamon.darnell@mail.utoronto.ca

A. Thomas-Bolduc
Independent Researcher, Port Alberni, BC, Canada
e-mail: athomasb@ucalgary.ca

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_95

consistency of first-order Peano Arithmetic relying on the well-ordering of certain ordinal notations. This was thought by Gentzen and Paul Bernays to be finitistically acceptable. However, a finitistically acceptable proof of the relevant well-ordering was not available until Gaisi Takeuti's proof in 1978. Beginning with the proposition that Takeuti's proof conforms to the finite standpoint, this chapter explores the limits of a finitism based on the methods of Gentzen and Takeuti. It is argued that the upper limit is the Feferman-Schütte ordinal, and thus that finitism may be extended to strong subsystems of second-order arithmetic-systems strong enough to provide a theory of the ordinals for which Gentzen employed notations.

Keywords

Consistency · Constructive mathematics · Finitism · Gentzen · Hilbert's program · Ordinal notations · Predicativism · Proof theory · Takeuti · Transfinite induction

1 Introduction

The goal of Hilbert's Program, roughly speaking, was to develop epistemically secure foundations for classical mathematics by means of constructive, finitist consistency proofs for as much of mathematics as possible. A plausible candidate for such a proof is Gentzen's (second) proof of the consistency of pure number theory (first-order arithmetic with full induction). However, Gentzen's proof involves induction on ordinal notations less than ε_0: a move that is potentially controversial depending upon what the bounds of finitist acceptability are taken to be. Yet, there is no consensus with respect to the limits of the finitist standpoint. At the lower end is the standpoint advanced by W. W. Tait (1981, 2002) and others, who argue that Primitive Recursive Arithmetic (PRA) exhausts finitism. Remarks from Hilbert, Bernays, and others suggest that finitism extends beyond PRA: that induction up to certain transfinite ordinals is finitistically acceptable. Unless already considered finitistically acceptable, a constructive finitist proof of the well-ordering of ordinal notations up to a transfinite ordinal is necessary and sufficient for determining that induction up to that ordinal is finitistically acceptable. Relative to this standpoint, it remains unclear how far the bounds of finitism will extend. This thread is pursued below, and drawing upon Gaisi Takeuti's (1987) work, a new potential upper limit for the bounds of finitist acceptability is identified.

This chapter begins with a brief introduction to the historical background of the finitist standpoint: an outline of its origins in Hilbert's Program and its intended epistemic role with respect to consistency proofs. The discussion then turns to Gentzen's proof theoretic work and his proof of the consistency of pure number theory. Establishing that Genzen's consistency proof is finitistically acceptable is the motivation behind Takeuti's (1987) proof of the well-ordering of ordinal notations (in Cantor normal form) less than ε_0. That proof is outlined, and the key methods and

operations that the proof involves are introduced. Takeuti claims that all of those methods and operations are acceptable according to what he calls the "Hilbert-Gentzen finitist standpoint." Yet, it is straightforward to illustrate how the methods and operations used in Takeuti's well-ordering proof can be amended to demonstrate the well-ordering of ordinal notations (given using Veblen functions and symbols for zero and ordinal addition) for ordinals less than Γ_0. Accordingly, if Takeuti's claims about his proof's finitistic acceptability are correct, the Hilbert-Gentzen finitist standpoint can be extended to the *Takeuti-Veblen* standpoint: induction up to Γ_0 is finitistically acceptable. This chapter concludes with a brief examination of the philosophical and mathematical significance of the Takeuti-Veblen standpoint.

2 Historical Background: Hilbert and Gentzen

The late nineteenth and early twentieth centuries saw monumental developments in mathematical logic, especially set theory. But the foundations of those discoveries were shaken by the discovery of "antinomies" like Russell's paradox and Cantor's paradox. These paradoxes motivated the development of programs in the philosophy of mathematics aimed at avoiding inconsistencies and securing the epistemic foundations of mathematics. Russell's logicism and use of the theory of types, the intuitionism of Brouwer and Heyting, the predicativism of Kroeneker and Poincaré, and the finitism of Hilbert and Bernays have all been developed and defended in some form throughout the twentieth century and into the twenty-first. (See the first section of Benacerraf and Putnam (1983) for a historical introduction, and Shapiro (2005) and Lindström et al. (2009) for more recent developments.)

2.1 Hilbert's Program

At its most general, Hilbert's Program involves three components:

1. The axiomatization of mathematical theories.
2. The concept of finitistic mathematics or the finite standpoint as epistemically privileged and secure.
3. A search for finitisically acceptable consistency proofs for axiomatized mathematical theories.

Remark 1 Hilbert's formalism was also an important part of his philosophical and mathematical work at the time, and while that aspect of his work is closely tied to finitism and constructive mathematics, it will be less important in what follows.

Hilbert's axiomatization of Euclidean geometry (Hilbert 1899) was an early important success for the first component, while an axiomatization of arithmetic was most famously developed by Guiseppe Peano (1889) and Richard Dedekind

Dedekind (1888). The axiomatization of set theory was under active development in the early twentieth century, as well (though that project is arguably ongoing).

The second component – finitism – is the core of the program and will be discussed in detail in this section and below. However, given an acceptance of component 2, it is easy to see that component 3 – finitisically acceptable consistency proofs for mathematical theories – will provide an epistemically secure foundation for those theories. If we can prove that a theory is consistent using tools that we already know to be acceptable and secure, we can then be sure that that theory will not lead us astray.

Remark 2 It should be noted that a defense of component 2 will not be undertaken herein; the starting point for this discussion is at least the weak position that some version of finitism can plausibly do the epistemic work needed to motivate the search for finitisically acceptable consistency proofs. That said, the proofs of Gentzen and Takeuti, and their implications, are of significant theoretical interest outside of a commitment to finitism or other constructivist philosophies of mathematics.

As the reader may already have noticed, however, there is a potential problem for Hilbert's Program that needs addressing: Gödel's Theorems. In 1931, Kurt Gödel proved that for any sufficiently strong mathematical theory Γ, any theory that can prove the consistency of Γ must be *strictly* stronger than Γ. This means that the finitist cannot proceed simply by proving the consistency of stronger and stronger theories based on already established consistency results. To prove the consistency of Peano Arithmetic, the finitist cannot rely on, say, PRA, or even on an equiconsistency result; they must use more powerful resources that nevertheless conform to the finite standpoint. Such resources were arguably developed by Gerhard Gentzen in the 1930s. It is the acceptability of those resources and the consequences of their finististic *bona fides* that are the concern of much of the rest of this chapter. But first, a brief discussion of Gentzen's proof theory is in order.

2.2 Gentzen's Proof Theory

Gerhard Gentzen can rightly be seen as the father of modern proof theory, developing the techniques that formed the foundation of later developments. He is perhaps most famous for his proof of the cut-elimination theorem which says that for any proof in the sequent calculus using the cut rule, there is an analogous proof without cut. This is key to the consistency proof discussed below. Gentzen also made significant developments in both intuitionistic and classical sequent and natural deduction calculi, including the normalization theorem for the latter, which is closely related to the cut elimination theorem (see Mancosu et al. (2021) for an introduction to Gentzen's proof theory, and Gentzen (1969) for translations of his important works). While the details of Gentzen's contributions to proof theory (besides the proof discussed in the following section) are far beyond the scope of this chapter, it is important to note his influence on later proof theorists, including Takeuti, as well as

the recent resurgence of interest in his work (see Kanckos 2011, and related publications).

3 Gentzen's Proof

Gentzen published three consistency proofs for what was called, at the time, pure arithmetic-first order Peano Arithmetic (PA). The strategy of the first two proofs is similar, and indeed Akiyoshi and Takahashi (2013) show that the second proof is a special case of the first. Nevertheless, it was, and still is, the second proof, published in 1938, that has been the most influential. It is the second proof that underlies the importance of Takeuti's well-ordering proof.

3.1 Ordinal Notations

A theorem due to Georg Cantor (1897) says that every ordinal can be written as a sum of 0 and exponents of ω, i.e.,

$$\alpha = \omega^{\beta_1} + \omega^{\beta_2} + \omega^{\beta_3} + \omega^{\beta_4} + \ldots$$

α, β ordinals, the βs either in this form, or 0s, and

$$\beta_1 \geq \beta_2 \geq \beta_3 \geq \ldots$$

Since Gentzen's proof requires only ordinals less the ε_0 and ε_0 is the first ordinal such that $\omega^{\alpha} = \alpha\,(\text{i.e.}\omega^{\varepsilon_0} = \varepsilon_0)$, it is guaranteed that $\alpha > \omega^{\beta_i}$ for all $i > 1$, and $\alpha = \omega^{\beta_1}$ only in the case where the $\omega^{\beta_{i>i}}$ are empty.

An ordinal written in this notation is in *Cantor normal form*. For simplicity, ω^0 can be written as 1, $\omega^0 + \omega^0$ as 2, and so on for the natural numbers. The further condition that any "+0" terms are deleted guarantees that each notation is unique.

The *natural sum* of two ordinals, α and μ, written $\alpha\#\mu$, is a (possibly) new ordinal in Cantor normal form constructed by interleaving the monomials – terms of the form ω^{β_i} – so that the β's are decreasing. For example, given the following ordinal notations, ξ and γ

$$\xi = \omega^{\omega^{\omega^3}} + \omega^{\omega+\omega} + \omega^2 + 2$$

and

$$\gamma = \omega^{\omega^{\omega^{\omega}}} + \omega^{\omega+\omega} + \omega$$

their natural sum, $\xi\#\gamma$, is the following ordinal notation:

$$\xi\#\gamma = \omega^{\omega^{\omega^{\omega}}} + \omega^{\omega^{\omega^{3}}} + \omega^{\omega+\omega} + \omega^{\omega+\omega} + \omega^2 + \omega + 2$$

It is the well-ordering of these notations that Gentzen used for the transfinite induction steps in his proof. Thus, for Gentzen's consistency proof to be truly finitistic, the well-ordering of these ordinal notations must be established from within the finite standpoint. This is Takeuti's essential contribution.

3.2 The Consistency of Pure Number Theory

The general strategy of Gentzen's consistency proof is as follows. Take an arbitrary proof of the empty sequent in the sequent calculus, with arithmetical initial sequents, and the inference rule for (full) arithmetical induction (i.e., first-order Peano Arithmetic formulated in the sequent calculus) and show that such a proof cannot exist.

Remark 3 The empty sequent is a contradiction, so its derivation implies the inconsistency of the theory in which it has been derived. Importantly, though, the empty sequent, unlike any other contradictions, contains no logical connectives.

To show that no such proof exists, the end part of the proof, defined as the largest segment of the proof looking "up" from the end-sequent, that contains only structural rules (including inductions), is pushed all the way to the top of the proof. This results in a proof containing no introduction or elimination rules. This is possible, because the end-sequent contains no connectives, so any complex formulae will have to have been removed by a cut at some point.

Inductions are reduced to sequences of cuts, and cuts on complex formulae are reduced to cuts on atomic formulae. This is all done to a "regular" proof, in a principled manner, but the details need not concern us here. (See Mancosu et al. 2021, for a detailed introduction to Gentzen's proof.)

Remark 4 A regular proof is one in which all of the noneigen variables have been replaced with 0 s and the eigenvariables have been replaced with appropriate arithmetic terms.

The key to the proof, and the part relevant here, is showing that the procedure terminates. Gentzen does this by assigning ordinal notations to each sequent and inference, and showing that each reduction step decreases the ordinal notation assigned to the end-sequent of the proof. Although the procedures for assigning ordinal notations to inductions and cuts are quite complex, the operations involved are just the stacking of finitely many $\boldsymbol{\omega}$ exponents, the taking of natural sums, and ordinary, arithmetical subtractions.

What is left is to show that the reduction procedure terminates, or, in other words, to show that the ordinal notations in Cantor normal form less than ε_0 are well ordered.

Of course, from a nonfinitist perspective, it is immediately clear that the ordinal notations are well-ordered: Each notation represents a unique ordinal less than ε_0, which is itself an ordinal. This entails that the ordinal notations are well-ordered by the set membership relation, "$\in$." But this requires reasoning about completed infinities, which the finitist eschews.

4 Takeuti's Proof

Takeuti proposes a concrete method for demonstrating that the ordinals less than ε_0 are well-ordered. His demonstration consists in applications of a series of concrete methods, called "eliminators." Eliminators are methods for taking any given strictly decreasing sequence of ordinals and constructing new strictly decreasing sequence of ordinals such that if the latter contains no infinitely descending chains, neither does the former. Using eliminators, Takeuti gives a demonstration that the ordinals $<\varepsilon_0$ are well-ordered.

Remark 5 Takeuti uses the term "concrete" in much the same way the term "constructive" might be used. While constructive mathematics or concrete proof procedures apply outside the finitist domain, in the context of Takeuti's proof, it can be safely assumed that "concrete" implies "conforms, in Takeuti's view, to the finite standpoint."

With the uncontroversial starting assumption that the natural numbers (and thus $\boldsymbol{\omega}$) are well-ordered, the basic idea is to separate parts of ordinal notations that have already been proven to be well-ordered, to show that the larger ordinal notations are well-ordered. This allows the well-ordering of larger and larger ordinal notations to be established.

An example from Takeuti's proof is provided below, but a toy example may be helpful first.

4.1 Toy Example: Proving the Well-Ordering of Positive Numerals

Take as your goal, rather than the well-ordering of notations for transfinite ordinals, the well-ordering of the notations for natural numbers – the positive numerals. Assume, as a starting point, that you know, perhaps from counting your fingers, that the numerals 1–10 are well ordered. The next step is to establish the well-ordering of the numerals 11–20. Write these as $10 + 10{,}10 + 9{,}10 + 8 \ldots 10 + 1$.

Remark 6 Note that the notations are written in *descending* order, so that we can more easily see that there are no infinitely descending chains of notations.

Obviously $10 = 10$, so, given the usual definition of addition, we can immediately see that the notations for natural numbers less than 20 are well-ordered. Now that we

know that, we can use essentially the same procedure to get us to 100 in 8 more steps.

The next set of steps gets us to 1000. Create a descending sequence as follows $100 + 100{,}100 + 90 + 9100 + 90 + 8...100 + 80 + 9...100(+0) +1$.

We can see that the second two terms ($90 + 8$, etc.) are the same as our notations for natural numbers less than 100, so we only need to deal with the sequence $100 + 100{,}100 + 90{,}100 + 80...100 + 10$, and then use our earlier procedure for the notations in between. But we can rewrite our sequence as $100 + 10{\cdot}10{,}100 + 10{\cdot}9...100 + 10{\cdot}1$. Since $100 = 100$ and $10 = 10$, by our initial assumption, the notations for natural numbers less than 1000 are well-ordered.

We can continue on like this, adding another layer for each power of 10, and using our previous results to show that our base-10 notations for natural numbers less than $\boldsymbol{\omega}$ are well-ordered.

Because Takeuti is working with transfinite ordinals with a lexigraphic ordering, his proof is significantly more complex, but the idea is the same: Separate the notations into parts that are equal, and parts we already know are well-ordered.

The following subsection is a reconstruction of the first step of Takeuti's proof, adapted from Darnell and Thomas-Bolduc (2018, §4) and Darnell and ThomasBolduc (2022, §3).

4.2 The 1-Eliminator

Takeuti begins his proof by supposing that the natural numbers are well-ordered (Takeuti 1987, pp. 92–93). Takeuti's *eliminators* capitalize on the well-ordering of the natural numbers and enable him to demonstrate that, strictly decreasing sequences of ordinals beginning with any ordinal, $<\varepsilon_0$ must be finite.

5 Terms and the 1-Eliminator

(Adapted from Takeuti 1987, p. 93.) The first eliminator Takeuti introduces the *1-eliminator*. The 1-eliminator is a concrete method for constructing a *1-sequence*, S_0' from a given decreasing sequence of ordinals, S_0, such that the first ordinal in S_0' is the *1-major part* of the first ordinal in S_0 and if S_0' is finite, then S_0 is finite. Takeuti explains these terms by way of an illustration:

Consider a strictly decreasing sequence of ordinals,

$$(S_0) \quad a_0 > a_1 > \ldots$$

such that each a_i in S_0 is written in Cantor normal form and a_0 is not a natural number. A 1-eliminator is a method for taking S_0 and producing a decreasing 1-sequence:

$$(S'_0) \quad b_0 > b_1 > \ldots$$

that satisfies the following conditions:

(CI) $b_0 = a_0^*$ (i.e., b_0 is the 1-major part of a_0; see below).
(CII) If S'_0 is finite, then S_0 is finite.

Takeuti calls the part of a_i that does not contain k_i the *1-major part* of a_i:

$$\underbrace{\omega^{\mu^i_1} + \omega^{\mu^i_2} + \ldots + \omega^{\mu^i_n}}_{1-\text{major part of } a_i} + k_i$$

where each $\mu^i_m > 0, \mu^i_{m-1} > \mu^i_m$ and k_i is a natural number (the i's are just indices). Let the 1-major part of a_i be written as: a'_i. Takeuti calls a sequence that consists of just the 1-major parts of each a_i in S_0 a *1-sequence*.

The 1-eliminator is shown to satisfy (CI) as follows. By definition, each ordinal in S_0 is identical with its 1-major part plus a given natural number. That is, $a_i = a'_i + k_i$ for all a_i in.

S_0. Hence, S_0 can be written as:

$$a'_0 + k_0 > a'_1 + k_1 > \ldots$$

applying a 1-eliminator to this sequence still produces the 1-sequence, S'_0. Now, take some finite part of S'_0, say the sequence:

$$b_0 > b_1 > \ldots > b_m$$

where, $b_0 = a'_0$ and $b_m = a'_i$ (for some i). So, either $a'_i = a'_{i+1} = \ldots = a'_{i+p}$ (for some p) and a_{i+p} is the last term in S_0 (the last term with a 1-major part), or $a'_{i+p} > a'_{i+p+1}$. If the former is the case, then stop. If the latter is the case, then make $b_{m+1} = a'_{i+p+1}$ and repeat. If one arrives at a sequence of a'_i's such that $b_m = a'_i$ and $a'_i = a'_{i+1} = \ldots = a'_{i+p} = \ldots$, it follows that S_0 must be finite. Since S_0 is a strictly decreasing sequence, $a'_i = a'_{i+1} = \ldots = a'_{i+p} = \ldots$ entails that $k_i > k_{i+1} > \ldots > k_{i+p} > \ldots$ Given the well-ordering of the natural numbers, the latter sequence must be finite. Hence, S_0 must be finite. Given the definition of S'_0 (i.e., a strictly decreasing 1-sequence constructed from S_0), if S'_0 is finite, then there is a last term, b_m, in S'_0 such that $b_m = a'_i$ and $a'_i = a'_{i+1} = \ldots = a'_{i+p} = \ldots$ Therefore, if S'_0 is finite, so is S_0.

6 Proving a Well-Order with the 1-Eliminator

(Adapted from Takeuti 1987, p. 93.) Takeuti uses the 1-eliminator to show that all decreasing sequences of ordinals that begin with an ordinal $<\omega^2$ must be finite. Let

$$(S_1) \quad a_0 > a_1 > \ldots$$

be a strictly decreasing sequence of ordinals such that $a_0 < \omega^2$. Applying a 1-eliminator to S_1 enables the construction of the 1-sequence:

$$\left(S_1^*\right) \quad b_0 > b_1 > \ldots$$

such that S_1^* satisfies conditions (CI) and (CII) and where $a_0 \geq b_0$. Since $a_0 < \omega^2$, $b_0 < \omega^2$. It follows that each ordinal b_i in S_1^* has the form, $\omega \cdot k_i$ (where k_i is a natural number). Accordingly, since a_0 is not a natural number, each b_i in S_1^* will be of the form, $\omega \cdot k_i$ (where k_i is a natural number). Hence, S_1^* can be written as:

$$\omega \cdot k_0 > \omega \cdot k_1 > \ldots$$

It must then be the case that $k_0 > k_1 > \ldots$, which must be finite. Therefore, $b_0 > b_1 > \ldots$ must be finite. Since S_1^* satisfies (CI) and (CII), S_1 must be finite. S_1 was arbitrary, so this result generalizes: Any decreasing sequence of ordinals which begins with an ordinal $<\omega^2$ must be finite.

The above procedure nicely illustrates Takeuti's strategy throughout his proof. He continues with this strategy such that with each new eliminator, combined with the well-ordering of the sequences of ordinals already established, he can show that sequences of ordinals with higher and higher limits must be well-ordered, ending with a demonstration of the well-ordering of the ordinals $<\varepsilon_0$.

7 Defining Finitism After Hilbert

As might be expected, there has been a lack of consensus as to the limits of the finite standpoint. One way to think about this is in terms of an upper bound on the strength of mathematical theories that can be properly thought of as finitist. Although there have been attempts to delineate the upper bound of the finite standpoint, including a new proposal put forth below, there is no consensus as to what exactly should count as finitistic mathematics. Indeed, after the publication of Gödel's second incompleteness theorem, it is unclear what Hilbert and Bernays considered to be the limits of finitism. There are indications that they considered finitism to extend beyond Primitive Recursive Arithmetic, and it seems Bernays considered Gentzen's second consistency proof to be finitisically acceptable (see Hilbert and Bernays (1939, p. 340–348), Zach (1998), and Darnell and Thomas-Bolduc (2018)).

Below, one thread in the development of finitism flowing through the work of Gentzen and Takeuti is taken up and developed. This thread focuses on constructive proof-theoretical tools, and the limits those tools put on the development of finitisically acceptable consistency proofs. Before delving into that, however, a brief look at the lower bound of finitism will be instructive.

7.1 More than PRA?

Primitive recursive arithmetic (PRA) – a relatively weak theory containing all of the primitive recursive functions, including those for 0, +, ×, and the successor function, as well as the quantifier-free induction scheme – is solidly within the realm of finitism. As the value of a primitive recursive function can always be found in finitely many steps, there is no threat of impermissible infinite reasoning. And because the induction schema is quantifier free, there is no risk of quantifying over completed infinite totalities.

Tait (1981, 2002) argues that finitism is completely exhausted by PRA, though he explicitly distances himself from the epistemological underpinning of Hilbert's finitism. As Stenlund (2009) argues, the purely technical aspects and epistemological goals of finitism cannot be so easily separated.

Given that the development of finistism is motivated, in large part, by a desire to establish epistemic security for a useful portion of mathematics, it is best not to be guided too strictly by intuitive notions of what counts as "finitistic." To do so would beg the question against the finitist. Nevertheless, PRA is a good lower bound for which mathematical theories conform to the finite standpoint (Darnell and Thomas-Bolduc 2018).

7.2 The Hilbert-Gentzen Standpoint

Bernays himself acknowledges that Gentzen's proofs were finitiscally acceptable:

> During the printing of this report the proof for the consistency of the full number theoretic formalism has been presented by G. Gentzen, using a method that conforms to the fundamental demands of the finite standpoint. (Bernays 1935, final paragraph)

It is likely, then, that Bernays had some familiarity with Gentzen (1943), wherein Gentzen discusses the finitistic acceptability of the induction on notations for transfinite ordinals. Gentzen tentatively concludes that his methods are finitistically acceptable, as did Bernays.

Thus, even before Takeuti published his well-ordering proof, there was a conception of finitism that included Gentzen's methods. This is the basis of what Takeuti calls the "Hilbert-Gentzen" finitist standpoint (Takeuti (1987, p. 101), Feferman (2005, p. 601)).

More specifically, there are two core components to the Hilbert-Gentzen standpoint, as discussed by Takeuti. The first is the limited use of transfinite induction, such as that employed in Gentzen's proof (Takeuti 1987, p. 88). The second is that it involves "'Gedankenxperimente' only on clearly defined operations applied to some concretely given figures and on some clearly defined inferences concerning these operations" (Takeuti 1987, p. 97).

The Hilbert-Gentzen standpoint, then, is roughly the standpoint that the construction of finitisically acceptable consistency proofs may involve limited use of

transfinite induction on concretely given figures, and *Gendankenexperimente* on finitistically acceptable operations and inferences. While this may seem vague, it circumscribes the methodology available under this standpoint. It also has the unsurprising consequence of encompassing Gentzen's proof via Takeuti's proof.

Takeuti's proof solidifies Gentzen's use of ordinal notations, and by extension PA, as finsitically acceptable (see Darnell and Thomas-Bolduc 2018, for an argument for this claim). However, Takeuti mentions that his proof could be extended beyond ε_0 (Takeuti 1987, p 97). He does not provide a proof and concentrated his later efforts on ordinal diagrams (see Takeuti (1957, 1958)). Ordinal diagrams are an interesting and complex ordinal notation system, but it will suffice to note that, despite significant efforts, it has not been shown that ordinal diagrams are "concrete" (and therefore finitisically acceptable).

8 The Takeuti-Veblen Finitist Standpoint

Taking Takeuti's assertion that his method of eliminators can be extended beyond ε_0, it is natural to ask just how far those methods can be taken. Since ε_0 is the first ordinal such that $\varepsilon_0 = \omega^{\varepsilon_0}$, new tools are needed. However, given those tools, as described below, Takeuti's method can be extended to *concretely* show that notations for ordinals less that Γ_0, the Feferman-Schütte ordinal, can be well-ordered. This opens up the possibility of finitism encompassing vastly more powerful mathematics.

The finite standpoint so extended might then be called the Takeuti-Veblen standpoint. The significance of this extended standpoint will be elaborated on below; however, its feasibility must first be established.

8.1 The Limits of the Extended Program

Takeuti mentions that the method of eliminators can be extended beyond ε_0, but does not conjecture as to an upper limit. In order to prove the well-ordering of ordinal notations up to some limit $\lambda > \omega$ (we write α_λ for the limit ordinal with notation λ) viz. the method of eliminators, the following must be satisfied.

(N1) For every ordinal α such that $0 \neq \alpha < \alpha_\lambda$, there is a unique constructive notation for α.
(N2) For every ordinal notation $\boldsymbol{\theta} < \lambda$, there is a concrete (or constructive) method for obtaining $\boldsymbol{\theta}$.
(N3) For every ordinal α such that $1 < \alpha < \alpha_\lambda$, the notation for α can be given using finitely many notations for ordinals $<\alpha$, "0," and appropriate symbols for ordinal addition, multiplication, or exponentiation.

Remark 7 (N1)–(N3) are mutually reinforcing. Each is needed for the purposes of: (i) maintaining the arithmetic relations (symbolized) between ordinal notations that

are required for using Takeuti's eliminators (some of which – the use of addition, multiplication, etc. – are highlighted in previous sections), and (ii) ensuring that it is possible to construct new eliminators from eliminators that have been concretely given (for details, see (Takeuti 1987, p. 96) and (Darnell and Thomas-Bolduc 2022, §3.4)).

Feferman (Feferman 1964) and Schütte (Schütte 1965) (independently) established that a unique constructive notation for every ordinal $<\Gamma_0$ can be given using Veblen functions (introduced by Veblen in (Veblen 1908)), and symbols for zero and ordinal addition (a detailed exposition of this result is given in Schütte (1977)). Their work satisfies (N1)–(N3) for ordinals $<\Gamma_0$; thus, the method of eliminators can be used to show the well-ordering of the relevant notations for ordinals $<\Gamma_0$.

Remark 8 For brevity, a complete proof of this claim is not provided herein. Instead, a sketch of a representative part of its demonstration is provided. The complete proof is now being prepared and will appear elsewhere.

Following Feferman (1964, p. 15), we can construct unique notations for ordinals $0 < \alpha < \Gamma_0$.

Remark 9 The notation employed below is a modification Feferman's Feferman (1964) notation in order to make it more obvious how Takeuti's eliminator method can be amended and extended beyond ε_0.

For ordinal notations ξ, θ, and γ, and limit notation λ, we write $\phi(\gamma, \theta)$ for the normal functions (i.e., continuous and strictly increasing functions) from ordinals to ordinals in the following transfinite sequence of normal functions,

$$\phi(\gamma, 0) = \omega^{\gamma};$$

$$\phi(\gamma, \theta + 1) = \text{the } \gamma^{\text{th}} \text{ solution } \xi \text{ of } \phi(\xi, \theta) = \xi;$$

$$\phi(\gamma, \lambda) = \text{the } \gamma^{\text{th}} \text{ solution } \xi \text{ of } \phi(\xi, \theta) = \xi \text{ for all } \theta < \lambda.$$

The functions generated in this way make up the *Veblen hierarchy*. Using the above (and the previously introduced symbols 0 and +), we can construct ordinal notations in a variation of Cantor normal form (hereafter, VCNF). To illustrate, ω^0 is written as $\phi(0,0)$ in VCNF, $\omega^0 + \omega^0$ is written as $\phi\ (0,0) + \phi\ (0,0)$ in VCNF, and so on (continuing to use numerals as notations for ordinals $< \boldsymbol{\omega}$). We write the natural sum of ordinal notations in VCNF analogously to before. For example, given the ordinal notations, ξ and γ in VCNF,

$$\xi = \phi(\phi(\phi(3,\ \ 0),\ 0), 0) + \phi(\phi(1,\ 0) + \phi(1,\ 0), 0) + \phi(2, 0) + 2$$

(in Cantor normal form, ξ is written, $\omega^{\omega^{\omega^3}} + \omega^{\omega+\omega} + \omega^2 + 2$) and

$$\gamma = \phi(\phi(\phi(\phi(1, \quad 0), \quad 0), \quad 0), 0) + \phi(\phi(1, \quad 0) + \phi(1, \quad 0), 0) + \phi(1, 0)$$

(in Cantor normal form, γ is written, $\omega^{\omega^{\omega^\omega}} + \omega^{\omega+\omega} + \omega$) their natural sum, $\xi \# \gamma$, is the following ordinal notation in VCNF:

$$\begin{aligned} \xi\#\gamma = \quad & \phi(\phi(\phi(\phi(1, \quad 0), \quad 0), \quad 0), 0) + \phi(\phi(\phi(3, \quad 0), \quad 0), 0) \\ & +\phi(\phi(1, \quad 0) + \phi(1, \quad 0), 0) + \phi(\phi(1, \quad 0) + \phi(1, \quad 0), 0) \\ & +\phi(2, 0) + \phi(1, 0) + 2 \end{aligned}$$

(in Cantor normal form, $\xi \# \gamma$ is written, $\omega^{\omega^{\omega^\omega}} + \omega^{\omega^{\omega^3}} + \omega^{\omega+\omega} + \omega^{\omega+\omega} + \omega^2 + \omega + 2$).

Remark 10 In Miller (1976, §6), Miller uses notations similar to VCNF to give a characterization of Takeuti's ordinal diagrams. The relationship between ordinal diagrams and Schütte's notations is examined in Levitz (1970).

Takeuti's method of eliminators can be straightforwardly amended to show the well-ordering of ordinal notations in VCNF for ordinals $<\varepsilon_0$ (i.e., for notations $\theta < \phi(0, 1)$). It is also not difficult to extend the amended method of eliminators to show the well-ordering of ordinal notations in VCNF for ordinals $<\Gamma_0$. For example, in order to show that all decreasing sequences of ordinal notations that begin with notations $\phi(0, 1) < \theta < \phi(\phi(0, 1) + 1, 0)$ (i.e., VCNF notations for ordinals α where, $\varepsilon_0 < \alpha < \omega^{\varepsilon_0+1}$) are finite, Takeuti might use a 1^1_0-eliminator.

Associated with 1^1_0-eliminators would be analogues of the concepts of a 1-major part and a 1-sequence. In particular,

Definition 1 For any ordinal $\alpha_i < \Gamma_0$ with VCNF notation a_i, the 1^1_0-*major part* of a_i is the part of its VCNF that contains no monomial $<\phi(\phi(0, 1), 0)$ (i.e., no monomial $< \omega^{\varepsilon_0}$).

For every ordinal $\alpha_i < \Gamma_0$ with VCNF notation a_i, we write a_i^{1*} for the 1^1_0-major part of a_i.

Definition 2 For any strictly decreasing sequence of VCNF notations $a_0 > a_1 > \ldots$, the sequence $a_0^{1*} > a_1^{1*} \ldots$ is the 1^1_0-*sequence* corresponding to $a_0 > a_1 > \ldots$, just in case for any $i \geq 0$, every a_i^{1*} in $a_0^{1*} > a_1^{1*} > \ldots$ is the 1^1_0-major part of a_i in $a_0 > a_1 > \ldots$.

For any strictly decreasing sequence of ordinals S_j, we write S_j^{1*} for the 1^1_0-sequence corresponding to S_j. For any strictly decreasing sequence of ordinals S, we say that S is a 1^1_0-sequence just in case there is a decreasing sequence of ordinals S_j such that $S = S_j^{1*}$ (i.e., S is the 1^1_0-sequence corresponding to S_j).

Definition 3 A 1_0^1-*eliminator* is a (concrete) method for constructing a 1_0^1-sequence S_j^{1*} from a (concretely) given strictly decreasing sequence of ordinal notations S_j, such that the first ordinal notation in S_j^{1*} is the 1_0^1-major part of the first ordinal notation in S_j, and if S_j^{1*} is finite, then S_j is also finite.

To illustrate, let

$$\left(S_0^1\right) \qquad a_0 > a_1 > \ldots.$$

be a strictly decreasing sequence such that, for any $i \geq 0$, every a_i in S_0^1 is of the form $a_i^{1*} + c_i$, and

$$\underbrace{\phi\left(\mu_1^i, 0\right) \cdot k_i^1 + \ldots + \phi\left(\mu_{m_i}^i, 0\right) \cdot k_i^m}_{a_i^{1*} = 1_0^1 - \text{major part of } a_i} + \underbrace{\phi\left(\mu_{m_i+p_i}^i, 0\right) \cdot k_i^{m+p} + \ldots + \phi\left(\mu_{m_i+p_i+l_i}^i, 0\right) \cdot k_i^{m+p+l}}_{c_i}$$

only if

$$\phi\left(\mu_1^i, 0\right) \cdot k_i^1 \geq \phi(\phi(0,\ 1), 0), \ldots, \phi\left(\mu_{m_i}^i, 0\right) \cdot k_i^m \geq \phi(\phi(0,\ 1), 0)$$

And

$$\begin{aligned} \phi\left(\mu_{m_i+p_i}^i, 0\right) \cdot k_i^{m+p} &< \phi(\phi(0,\ 1), 0), \ldots, \phi\left(\mu_{m_i+p_i+l_i}^i, 0\right) \cdot k_i^{m+p+l} \\ &< \phi(\phi(0,\ 1), 0). \end{aligned}$$

The 1_0^1-eliminator is a method that takes S_0^1 and concretely produces a 1_0^1-sequence,

$$\left(S_0^{1*}\right) \qquad b_0 > b_1 > \ldots$$

such that S_0^{1*} satisfies each of the following conditions:

$\left(\text{CI}_0^1\right)$ $b_0 = a_0^{1*}$ (i.e., b_0 is the 1_0^1- major part of a_0).
$\left(\text{CII}_0^1\right)$ If S_0^{1*} is finite, then (it can be concretely shown that) S_0^1 is finite.

An analogue of the procedure used earlier to show that the 1-sequence S_0' generated by the 1-eliminator (applied to S_0) is finite and satisfies (CI) can be used to show that S_0^{1*} is finite and satisfies $\left(\text{CI}_0^1\right)$. Once adequately justified by previous results (and buttressed by *Gedankenexperimente*), an appropriate analogue of

Takeuti's (1987, p. 94) induction hypothesis (e.g., *Any descending sequence of ordinal notations $d_0 > d_1 > \ldots$, with $d_0 < \phi(\phi(0, 1) + 1, 0)$ is finite.*) can then be used to demonstrate that S_0^{1*} satisfies $\left(\mathrm{CII}_0^1\right)$.

Remark 11 For Takeuti, adherence to the Hilbert-Gentzen finitist standpoint permits omitting the actual execution of a step in his proof when that step of the proof or its relevant consequence(s) can be known viz. *Gedankenexperimente* restricted to "only on clearly defined operations applied to some concretely given figures and on some clearly defined inferences concerning [those] operations" (Takeuti 1975, p. 97).

The 1_0^1-eliminator can be used to demonstrate the well-ordering of ordinal notations (in VCNF) for ordinals $\alpha < \omega^{\varepsilon_0+1}$. Assume we have demonstrated the well-ordering of VCNF notations for ordinals $\alpha < \varepsilon_0$. Let

$$\left(S_e^1\right) \quad a_0 > a_1 > \ldots.$$

be a strictly decreasing sequence such that, $\phi(0, 1) < a_0 < \phi(\phi(0, 1) + 1, 0)$ (i.e., a_0 is the VCNF notation for some ordinal α such that $\omega^{\varepsilon_0} < \alpha < \omega^{\varepsilon_0+1}$). Applying a 1_0^1-eliminator to S_e^1 generates the 1_0^1-sequence:

$$\left(S_e^{1*}\right) \quad b_0 > b_1 > \ldots$$

such that S_e^{1*} satisfies each of $\left(\mathrm{CI}_0^1\right)$ and $\left(\mathrm{CII}_0^1\right)$ and where $a_0 \geq b_0$. By supposition, $a_0 < \phi(\phi(0, 1) + 1, 0)$, and so $b_0 < \phi(\phi(0, 1) + 1, 0)$. By Definitions 1 and 2, for any $i > 0$, every b_i in S_e^{1*} is of the form $\phi(0, 1) \cdot k_i$ (where k_i is a natural number). Hence, S_e^{1*} can be written as:

$$\phi(0, 1) \cdot k_0 > \phi(0, 1) \cdot k_1 > \ldots$$

Thus, $k_0 > k_1 > \ldots$ The natural numbers are well-ordered, and so $k_0 > k_1 > \ldots$ is finite. Hence, S_e^{1*} is finite. Hence, by Definition 3, S_e^1 is finite. Therefore, ordinal notations $< \omega^{\varepsilon_0+1}$ are well-ordered.

Appropriate analogues of the steps in Takeuti's Takeuti (1987) well-ordering proof can be used to show that an eliminator can be concretely given for every ordinal notation in VCNF for every ordinal $<\Gamma_0$. Once this is shown, a procedure like the one just sketched above can be used to show the well-ordering of ordinal notations in VCNF for ordinals $<\Gamma_0$.

8.2 The Importance of Takeuti-Veblen Standpoint

Γ_0 is the proof-theoretic ordinal of the subsystem of second-order arithmetic called ATR_0 – Arithmetic Transfinite Recursion. While that certainly does not sound finitistic, it means that given a suitable axiomatization and proof system, ATR_0, or

a system of equal consistency strength, ought to be provably consistent within the Takeuti-Veblen finite standpoint.

This would be a significant result for a number of philosophical and mathematical reasons. Most generally and most basically, this result could be seen as a vindication of Hilbert's Program. If consistency proofs for strong subsystems of second-order arithmetic are available with the finite standpoint, Hilbert's Program can be seen as a bridled success a century later.

More specifically, the system ATR_0 can prove, or is equivalent to, a number of important theorems, including from descriptive set theory (see Simpson 2009, C V for details and proofs). One result of particular interest is that "the set existence axioms of ATR_0 are indispensable for a decent theory of countable ordinals" (Simpson 2009, p. 195). In other words, the Takeuti-Veblen standpoint allows for the development of a theory of countable transfinite ordinals. This is a particularly pleasing result, as it entails that ordinals themselves that Gentzen used notations for in his consistency proof are actually available within the finite standpoint.

On the other hand, from a certain philosophical perspective, these results might lead to an opposite conclusion. For some finitists, such strong results might be proof that something has gone wrong and the Hilbert-Gentzen standpoint should be rejected. The apparent availability of such intuitively nonfinitistic resources may be taken as evidence that PRA is the *upper* as well as lower bound of finitism. However, such a position would require finding fault with Takeuti's proofs, likely focusing on the use of *Gedankenexperimente*.

The last consequence to be discussed here is less certain and has to do with the conjectured ordinal upper bound of the Takeuti-Veblen standpoint, $\Gamma_0 \cdot \Gamma_0$, which is arguably a natural stopping point because it is thought to also be the limit of predicative arithmetical theories.

Remark 12 Predicativism, very roughly, is the philosophico-mathematical idea that to secure the epistemic foundations of mathematics, we must avoid definitions where the definiens range over the entities in the definiendum. The idea is to ban circularity in definitions and axioms and thereby avoid paradoxes – like Russell's paradox – arising from impredicative axioms like naive set comprehension.

That finitism and predicativism might coincide in such a way would be a remarkable convergence of the two most significant strands of classical constructive mathematics (see Kreisel 1958; Feferman 2005; Simpson 2009, for discussion of the relationships between finitism and predicativism). This coincidence is little more than conjecture at this point; however, both its proof and dis-proof would be interesting results. If the two programs were shown to coincide in their upper limits, that would be powerful evidence of a rich and useful theory of constructive classical mathematics, and potentially greatly extend the tools available for the development of constructive mathematics.

On the other hand, a proof or argument that the upper limits of predicativism and finitism do not coincide would reveal important information about both programs and their relationship. The obvious way to go about this would be to determine upper

bounds from within each respective system – finitism using finititist means, a predicativism using predicativist means. These would be important results in their own rights. However, if a comparison were carried out that conformed to the strictures of *both* programs, the formal nature of the divergence would provide insight into classical constructive mathematics that would shed light on the nature of the epistemic issues undergirding its development. In either case, because of the constructive nature of the proofs, specific information would be forthcoming beyond the existence or not of the coincidence.

References

Akiyoshi R, Takahashi Y (2013) Reading Gentzen's three consistency proofs uniformly (in Japanese). Kagaku kisoron kenkyu ā (J Japan Assoc Philos Sci) 41(1):1–22

Benacerraf P, Putnam H (eds) (1983) Philosophy of mathematics, selected readings, 2nd edn. Cambridge University Press

Bernays P (1935) Hilbert's Investigations into the Foundations of Arithmetic, "Hilberts Untersuchungen über die Grundlagen der Arithmetik" in Hilbert's Gesammelte Abhandlungen, vol 3. Springer, Berlin, pp 196–216. Dirk Schlimm (Trans)

Cantor G (1897) Beiträge zur Begründung der transfiniten Mengenlehre. Mathematische Annalen 49:207–246, part II

Darnell E, Thomas-Bolduc A (2018) Takeuti's well-ordering proof: finitistically fine? In: Zach M (ed) Research in history and philosophy of mathematics: Proceedings of the CSHPM annual meeting in Toronto, Ontario. Springer, Toronto

Darnell E, Thomas-Bolduc A (2022) Takeuti's well-ordering proof: an accessible reconstruction. Australas J Log 19(1)

Dedekind R (1888) Was sind und was sollen die Zahlen. Friedrich Vieweg & Sohn, Braunschweig

Feferman S (1964) Systems of Predicative Analysis. J Symb Log 29:1–30

Feferman S (2005) Predicativism. In: Shapiro (2005), pp 590–624

Gentzen G (1938) New version of the consistency proof for elementary number theory. In: Gentzen (1969), m.E. Szabo (Ed., Trans)

Gentzen G (1943) Provability and nonprovability of restricted transfinite induction in elementary number theory. In: Gentzen (1969), m.E. Szabo (Ed., Trans)

Gentzen G (1969) The collected papers of Gerhard Gentzen. North-Holland, Amsterdam, m.E. Szabo (Ed., Trans)

Gödel K (1931) Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. Monatshefte für Mathematik und Physik 38:173–198

Hilbert D (1899) Grundlagen der geometrie. In: Festschrift zur Feierder Enthllung des GaussWeber-Denkmals in Göttingen. Teuber, Leipzig, pp 1–92

Hilbert D, Bernays P (1939) Grundlagen der Mathematik, vol II. Springer, Berlin, Heidelberg

Kanckos A (2011) A possible and necessary consistency proof. PhD thesis. University of Helsinki, Helsinki

Kreisel G (1958) Hilbert's programme. Dialectica 12(3–4):346–372

Levitz H (1970) On the relationship between Takeuti's ordinal diagrams $O(n)$ and Schütte's system of ordinal notations $\Sigma(n)$. In: Kino A, Myhill J, Vesley RE (eds) Intuitionism and proof theory. North-Holland, Amsterdam, pp 377–405

Lindström S, Palmgren E, Segerberg K, Stoltenberg-Hansen V (eds) (2009) Logicism, intuitionism, and formalism: what has become of them?, Synthese library, vol 341. Springer

Mancosu P, Galvan S, Zach R (2021) An introduction to proof theory. Oxford University Press

Miller LW (1976) Normal functions and constructive ordinal notations. J Symb Log 41:439–459

Peano G (1889) Arthimetices principia. Bocca, Turin

Schütte K (1965) Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik. Archiv für mathematische Logik und Grundlagenforschung 7:45–60
Schütte K (1977) Proof theory. Springer, New York, translation from the German by J.N. Crossley
Shapiro S (ed) (2005) Oxford handbook of philosophy of mathematics and logic. Oxford University Press, New York
Simpson SG (2009) Subsystems of second order arithmetic. Cambridge University Press
Stenlund S (2009) Hilbert and the problem of clarifying the infinite. In: Lindström et al (2009), pp 485–503
Tait WW (1981) Finitism. J Philos 78(9):524–546
Tait WW (2002) Remarks on Finitism. In: Sieg W, Sommer R, Talcott C (eds) Reflections of the foundations of mathematics: essays in honor of Solomon Feferman, Lecture notes in logic, vol 15. Assoc Symb Logic, Ubana, pp 410–419
Takeuti G (1957) Ordinal Diagrams. J Math Soc Japan 9(4)
Takeuti G (1958) On the formal theory of the ordinal diagrams. Ann Japan Assoc Philos Sci 1(3): 151–170
Takeuti G (1975) Consistency proofs and ordinals. In: ISILC Proof Theory Symp. Springer, pp 365–369
Takeuti G (1987) Proof theory and ordinal analysis, studies in logic and the foundations of mathematics, vol 81, 2nd edn. Elsevier, New York. first published in 1975
Veblen O (1908) Continuous increasing functions of finite and transfinite ordinals. Trans Am Math Soc 9:280–292
Zach R (1998) Numbers and functions in Hilbert's Finitism. Taiwanese J Philos Hist Sci 10:33–60

Section II

Mathematician's Reflections on Practice and Working Philosophy of Mathematics

Athanase Papadopoulos

Introduction to *Mathematicians' Reflections on Practice and Working Philosophy of Mathematics*

Athanase Papadopoulos

Abstract

This is the introduction to the section "*Mathematicians' Reflections on Practice and Working Philosophy of Mathematics*" of the *Handbook of the History and Philosophy of Mathematical Practice*. It essentially consists of descriptions of the various chapters that constitute this section.

Keywords

Plato's philosophy · Indivisible lines · L. E. J. Brouwer · Mathematical practice as philosophy · Anthyphairesis · Pavel Florensky · Mathematics in education · Intuitionism · Russian Golden Age · Russian Silver Age

AMS Classification

00A30 · 01A60 · 01A20 · 00A30 · 03A10 · 03A05 · 97E20

The second section of this Handbook, titled "Mathematicians' Reflections on Practice and Working Philosophy of Mathematics," consists of 10 chapters. All are written by mathematicians involved in philosophical questions, and they are concerned with the relation between mathematical and philosophical practices. The interested reader, whether he is knowledgeable or not in philosophy, will find a large spectrum of ideas and points of view on how mathematics and philosophy together can help understanding the world of ideas.

A. Papadopoulos (✉)
Institut de Recherche Mathématique Avancée (IRMA) and Centre de Recherche et d'Expérimentation sur l'Acte Artistique (ITI CREAA), Université de Strasbourg and CNRS, Strasbourg Cedex, France
e-mail: athanase.papadopoulos@math.unistra.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_128

The subject of the first three chapters in this section is Plato's philosophy. The first two are written by Stelios Negrepontis and the third one by Stelios Negrepontis and Demetra Kalisperi. I would like to emphasize that Plato was a mathematician; probably not in the modern sense of the word, since no theorem can be attributed to him, but because all his thinking was oriented towards mathematics. He was fully engaged in questions such as what is mathematics, why and how can we acquire mathematical knowledge, what is the nature of mathematical ideas and of the various mathematical objects, what kind of mathematics is suitable for education and for the well-being of the soul, etc. It is also worth remembering that Plato was the founder of a school, *the Academy*, probably the most important school that ever existed, and that in that school, Plato was essentially teaching mathematics. I like to recall a sentence by Aristoxenus of Tarentum, the famous fourth century BC music theorist, about Plato with whom he was quite familiar. In his *Harmonics*, reporting on the teachings of Plato, Aristoxenus writes that the latter used to entice his audience by announcing lectures on the Good or similar philosophical topics, but that eventually these lectures turned out to be on geometry, number theory or astronomy, with the conclusion that "the Good is One." Let me also quote the following passage from Plato's *Philebus* (51b9–c7): "For when I say beauty of form, I am trying to express, not what most people would understand by the words, such as the beauty of animals or of paintings, but I mean, says the argument, the straight line and the circle and the plane and solid figures formed from these by turning-lathes and rulers and patterns of angles; perhaps you understand. For I assert that the beauty of these is not relative, like that of other things, but they are always absolutely beautiful by nature."

Understanding Plato's philosophy has always been, since Plato's times, the most challenging problem in Philosophy. Obviously, the problem has to be addressed by mathematicians and not only by philologists or philosophers as it is usually done. Negrepontis, who is an author of the chapters on Plato's philosophy, is among the very few mathematicians of the modern period who studied Plato thoroughly. He has a fresh point of view on Plato's writings, which I will present briefly in these chapters.

The first of these chapters by Negrepontis is titled ▶ "The Anthyphairetic Interpretation of Knowledge as Recollection in Plato's *Meno* 80d-86e and 97a-98b." The title refers to the *Meno*, a dialogue of Plato whose subject is *knowledge*. Socrates, as in many of Plato's dialogues, is one of the participants to the dialogue. The dialogue starts with a question about the nature of virtue, more precisely, whether it can be taught or whether it is acquired by nature, and if it is not innate, how can one acquire it. The dialogue turns then to the more general question of knowledge, whether it is innate or acquired. Socrates' aim is to prove that knowledge is recollection. Naturally, it appears that "knowledge" for Socrates (and for Plato) is knowledge about mathematics and about abstract and intelligible entities. For this purpose, Socrates calls one of Meno's slave boys, and by way of a succession of questions, he shows that this boy gradually remembers mathematical notions that he has never been taught since birth. The notions concern the irrationality of the diagonal to the side of the square.

Negrepontis, in this chapter, analyzes the two passages of the *Meno* mentioned in the title of the chapter, of which he gives an anthyphairetic interpretation, in line with

other anthyphairetic interpretations he has developed for several other dialogues of Plato. Thus, he shows that Plato's philosophical notion of knowledge in the *Meno*, which amounts to that of the knowledge of a concrete geometrical question, namely, the diameter a of a square given that we know its side b, is in fact the knowledge of the anthyphairesis of a line a satisfying $a^2 = 2b^2$ for a given line b. In philosophical terms, this is the knowledge of True Opinion plus Logos/Logismos. This anthyphairetic interpretation of Plato's philosophy is a recurrent theme in Negrepontis' commentaries on Plato.

The second chapter by Negrepontis, titled ▶ "The Indivisible Line Coincides with the Platonic Intelligible Being," is concerned with the concept of indivisible line, famously referred to by Aristotle who attributes it to Plato and Xenocrates. This interpretation is based on the identification of the indivisible line with Plato's intelligible Being, and a geometric/anthyphairetic interpretation of the nature of the intelligible Being, namely, an intelligible Being is the philosophic analogue of a dyad of opposite kinds in periodic anthyphairesis. The arguments use passages from the *Parmenides*, *Phaedrus*, and *Republic*. In his identification between the indivisible line and the intelligible Being, Negrepontis takes into account ancient sources, first of Plato himself, but also Zeno (especially the second version of the Dichotomy Paradox), Aristotle (the famous and mysterious *Metaphysics* 992a19–22 passage, the *Physics* in his use of the potential infinite as the main argument against Plato and Zeno, and the pseudo-Aristotelean work *On Indivisible Lines*), Proclus, and Simplicius.

The next chapter in this section, by Stelios Negrepontis and Demetra Kalisperi, is titled ▶ "The Mystery of Plato's Receptacle in the *Timaeus* Resolved." In the *Timaeus*, Plato declares, surprisingly, that it is necessary to introduce, beside the intelligibles and the sensibles, a Third Kind, the Receptacle. The nature of the Receptacle and the reason why Plato introduces it has always been a mystery for all Plato's scholars. Negrepontis and Kalisperi give an original response to these two questions, based on well-hidden deep mathematics of incommensurability and of infinite anthyphairesis that underlie Plato's account. The authors argue that these mathematical notions are fundamental not only for Plato's philosophical system of the unchanging intelligible Beings but also for his description of the permanently changing sensible, physical world. The starting point of the article is the interpretation of the intelligible Being as the philosophical analogue of a dyad in periodic anthyphairesis. In his earlier work (e.g., in the *Republic* 475e-480a), Plato identifies the knowledge of a sensible entity with a True Opinion of the intelligible entity N in which this sensible entity participates, namely with a finite initial segment of the infinite periodic anthyphairesis of N. But in the *Timaeus*, Plato's concern is to avoid a Third Man Argument leading to infinite Worlds, and for this Plato needs to transform every finite initial segment of the infinite anthyphairesis of the intelligible into its equivalent, a double tight inequality involving only the two elements of the initial intelligible dyad. Plato realizes that while such a transformation, involving mathematical manipulations, is possible for a geometric anthyphairesis, it makes no sense for an intelligible anthyphairesis. To overcome this difficulty, Plato is obliged to replace the intelligible dyad that generates all the sensibles by the geometrical

(surfaces of the) four regular solids, forming the canonical dyad cube/earth vs. icosahedron/water, augmented by pyramid/fire and octahedron/air. The resulting anthyphairetic inequalities are interpreted by Plato as turning the (geometrized) earth into the Receptacle of the (geometrized) water, air, and fire.

The next chapter, titled "Mathematics in the Being Mode of Existence Mathematics in Abel, Galois, Noether, and Grothendieck," by Victor Pambuccian, is concerned with the work of L. E. J. Brouwer, the founder of the modern intuitionistic school of mathematics. Brouwer was a brilliant mathematician, in particular a great topologist, who tried throughout his academic life to build firm foundations for intuitionistic mathematics. Pambuccian begins his essay by declaring that Brouwer's intuitionism may be unique in the history and philosophy of mathematics for the quality of the hostility it has always encountered from mathematicians and philosophers alike. He recalls that to this day Brouwer's work on intuitionism has not been fully recognized, and that very few mathematicians and philosophers are interested in his intuitionistic mathematics and philosophy of intuitionism. He wonders why this is so. He reviews the few attempts to link Brouwer's work to the works of Aristotle, Descartes and Kant, and to draw parallels with works of Bergson, Husserl, and other Western philosophers. He also reviews Brouwer's conflicts with Hilbert and others, without finding an answer to the above question. Pambuccian's contribution consists in showing the relationship between Brouwer's philosophy and Eastern philosophy, that of Buddha and Krishnamurti and of the Western philosophers close to these Eastern philosophies: Meister Eckhart, Albert Schweitzer, David Bohm, Spinoza, and others. This chapter sheds new light on Brouwer's philosophy, providing a comprehensive overview of the Dutch philosopher and mathematician's life and work, of his position towards the philosophical trends of his time and of his theory of language.

In the next two chapters, by Arkady Plotnitsky, questions such as *What is mathematical practice as Philosophy?* and *What is the working philosophy of mathematics?* are addressed.

The first of these two chapters is titled ▶ "Mathematical Practice as Philosophy, with Galois, Riemann, and Grothendieck." In this chapter, Plotnitsky examines the notion of "working philosophy of mathematics" or, as he likes to say, of "mathematical practice as philosophy." Based on the examples of three paired prominent mathematicians: Abel–Galois, Lobachevsky–Riemann, and Weil–Grothendieck, he starts by explaining how mathematicians' working philosophy of mathematics is closely connected with their actual mathematical thinking and practice. Granting the differences between the thinking of the above mathematicians, his aim is to explore the differences between their mathematical practices as philosophy. Plotnitsky considers that creative mathematical thinking is usually defined by the invention of new concepts and he describes a mathematical theory as being essentially an organized assemblage of concepts. Central to this article are Galois' concept of group, Riemann's concept of manifold and Grothendieck's concept of topos. Plotnitsky's analysis involves a comparison between general creative mathematical practice and that of the two French philosophers Gilles Deleuze and Félix Guattari, who considered that philosophical practice, or even true philosophy itself, is defined

as the practice of the invention of new concepts, with philosophical concepts defined in parallel to mathematical concepts.

In the next chapter, titled "'That Which a Minority Construct': Abelian Mathematics in Abel, Galois, Noether, and Grothendieck – From Galois to Grothendieck," Plotnitsky analyzes the nature of radical transformation in mathematics, enabled by what he calls a "minority mathematics." His discussion mostly concerns what he calls "modern mathematics," that is, the mathematics which emerged as abstract mathematics, roughly around the year 1800, "abstracted," he says, "from mathematics' relations to the natural world and physics, relations that previously dominated mathematics." Referring again to Deleuze and Guattari, the author considers that a minority mathematics is not defined as something that exists entirely outside a "major mathematics" (to be distinguished here from a majority mathematics), but a mathematics that constructs itself inside and even at the very core of this major mathematics. A hint is made to T. Kuhn's revolutionary vs. normal mathematical practice. The cases of N. H. Abel, É. Galois, E. Noether, and A. Grothendieck, as the main cases of revolutionary practice, are central to the discussion.

The next two chapters in this section, one of them written by the author of this introduction and the other one by Alexey Sossinsky, analyze the life and working condition of two mathematicians, one of them a philosopher as well, in the Soviet Union.

The first of these chapters is titled ▶ "Pavel Florensky and His World." It is concerned with the life and works of Pavel Florensky, an important and singular figure of the so-called *Silver Age of Russian mathematics*, a period which is also part of the *Silver Age of Russian literature, poetry, art and philosophy*. Florensky is certainly among the great philosophers-theologians of the twentieth century, with a very atypical trajectory of life. His work in philosophy, the breadth of which is now being discovered (his works are being edited and published at an exponential rate in several languages), is imbued with mathematical ideas, and I think that it is good that this work has started to be known to and analyzed by mathematicians. Florensky developed his *œuvre* in Russia and the early years of the Soviet Union, during the first third of the twentieth century. Talking about Florensky's life and works is also an opportunity to reflect upon the Russian mathematical school of that period, its philosophical foundations, and the conflicts it experienced. It is also an occasion for discussing poetry and literature of the Russian Silver Age.

In the next chapter, titled ▶ "The Practice of Mathematics in a Totalitarian Society," Alexei Sossinsky, based on his own experience, describes the working atmosphere of a mathematician in the Soviet Union at a later period, from 1957 to 1987. This is a period which belongs to the so-called *Golden Age of Russian mathematics*. This chapter provides us with a first-hand picture of this politically turbulent period, which was also a time when mathematics in Russia experienced a development unrivalled in its scope and importance, probably anywhere else in the world.

The next chapter, by Ken'ichi Ohshika, is titled ▶ "Philosophy Behind Mathematics." The author considers the very fruitful interaction between mathematics and

philosophy. One of the themes he develops is that two important philosophical points of view, namely, conventionalism and logical positivism, which are held by mathematicians, but not only by them, were born in works of mathematicians. This is also the occasion to consider the philosophical and mathematical works of some mathematicians-philosophers including Henri Poincaré, who was one of the main figures of conventionalism, and Hans Hahn and Karl Menger, referring to the so-called logical positivism way of thinking. Ohshika explains how the works of Hahn and Menger helped freeing mathematics and philosophy from old Kantian apriorism, both in geometry and logic. Finally, the author addresses the state of a problem which remains unsolved by both conventionalism and logical positivism, namely, the relation between mathematics and natural science.

In the final chapter of this section, titled ▶ "Some Disillusioned Thoughts About School and Teaching," Valentin Ponaru gives an overview of his interests and his experience, since his youth, first in physics, then in philosophy and mathematics. He describes the path that led him to some important philosophical problems, starting from the questions he formulated as a child about the possibility or impossibility of motion, then his discovery of the Greek philosophers like Zeno of Elea, and his reading of Kant. This leads him to some critical thoughts concerning high-school teaching, particularly in France, which he experienced through his children's education.

The Anthyphairetic Interpretation of Knowledge as Recollection in Plato's *Meno* 80d-86e and 97a-98b

Stelios Negrepontis

Contents

1 Outline 474
2 *Theaetetus' anthyphairetic definition of the proportion of magnitudes, the Logos Criterion for anthyphairetic periodicity, and the reconstruction of Theaetetus' proof of the Pythagorean Proposition: if* $a^2 = 2b^2$*, then Anth (a, b) = [1, period (2)]* 475
3 *Eristic logos (knowledge is impossible) vs. dialectic Logos (knowledge is possible because knowledge is recollection)*. Meno *80e1-81e2* 478
4 *Experimental verification that knowledge is recollection: an uneducated boy will recall the knowledge of line a, whose square* a^2 *with side is the double of the square* b^2 *with given side b*. Meno 81e3-82e2 479
5 *The initial stage of Recollection of the Knowledge of line a in relation to line b, with* $a^2 = 2b^2$*, is identified with the geometrical initial segment of the anthyphairesis of line a to the b, with* $a^2 = 2b^2$. Meno *82e2-83e10* 481
6 *The philosophic term* ***True Opinion*** *about the line a to the given line b, with* $a^2 = 2b^2$*, is identified with* ***the knowledge of an initial segment of the geometric Anthyphairesis*** *of a to b, with* $a^2 = 2b^2$. (Meno *83e11-84d2, 97a9-c5,* and Sumposion *210a4-e6*) 483
7 *Heuristics*: *From the statement that* ***Knowledge is True Opinion plus Something*** *(Meno 97c6-d3) and the anthyphairetic interpretation of True Opinion (Sections 4 and 5), we expect that Knowledge will be the full infinite anthyphairesis of a to b and the "something" needed will be the Logos Criterion* 486
8 *Line a such that its square* a^2 *is double the square* b^2 *of the given line b, is shown, to be the* ***line with the name diameter*** *of the square with side b* 487

2020 Mathematics Subject Classification. Primary 00A30; Secondary 01A20

S. Negrepontis (✉)
Department of Mathematics, Athens University, Athens, Greece
e-mail: snegrep@gmail.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_138

9 *The first rather indirect hint of how True Opinion is to be transformed into Knowledge and Recollection in* Meno *85b8-86b5 is given in terms of the crucial word* ***pollache*** *85c11* 490
10 *For the first time, in the* Meno *86b6-c3, Plato explicitly brings forth* ***"logos"*** *as the fundamental ingredient of* ***Knowledge****, in line with a corresponding role for "logos" in the* Republic 534b3-6, Theaetetus 202c2-5, Sophist 260a5-b2, Statesman 286a4-7, *and elsewhere in Plato. Platonic Scholars render "logos" by terms such as "account," "argument," "discourse," "rational definition," "rational explanation," "reason," and "word"(cf. Sections 10.2 and 10.3). These renderings in our opinion are not in the correct direction and obscure the true meaning of "logos." Nehamas (1984) suggested that the meaning of Logos must come from a study of the Divisions in the* Sophist *and the* Statesman*. This is exactly our approach in Sections 11 and 12* 495
11 *The revealing anthyphairetic interpretation of Logos in the Divisions of the Angler, the Noble Sophistry, and the Sophist in the* Sophist *as the philosophic analogue of Theaetetus' Logos Criterion establishing anthyphairetic periodicity* 498
12 *The Logos/Logismos/analogio as Bond in* Meno *97d4-98a4,* Sophist *251d8-9, 253a4-c1, 268c5-d5, and* Timaeus *31c2-32a7* 507
13 *Knowledge as Recollection in the* Meno *and the whole of Plato's work* 514
14 Summary 518
15 Translations 518
References 518

Abstract

This chapter aims to obtain a novel anthyphairetic interpretation of Knowledge as Recollection in Plato's *Meno* 80d5-86c3 and 97a9-98b6, in a self-contained manner, in line with the anthyphairetic interpretation I have developed for the whole of Plato's work.

Plato sets out to explain his philosophical notion of Knowledge in the *Meno*, by explaining what he means by Knowledge in the concrete geometrical case of line a such that $a^2 = 2b^2$ for a given line b, in fact of the diameter a of a square with side b. For that purpose, Plato introduces the concepts of True Opinion and Logos/Logismos and formulates his answer as follows:

the *Knowledge* of the diameter a to the side b of a square, $a^2 = 2b^2$ is True Opinion plus Logos/Logismos.

Our interpretation of *True Opinion* is that it consists of the first two anthyphairetic relations of a to b

$a = b + c_1$, $b = 2c_1 + c_2$, and

our interpretation of *Logos* is the Theaetetus' Logos Criterion for periodic anthyphairesis, assuming for the case in question the form of a continuous proportion

$b/c_1 = c_1/c_2$.

Knowledge is then the knowledge of the infinite anthyphairesis of a to b

$$a = b + c_1$$
$$b = 2c_1 + c_2$$
$$c_1 = 2c_2 + c_3$$
$$\ldots$$
$$c_{n-1} = 2c_n + c_{n+1}$$
$$\ldots$$

resulting formally from the two relations above, namely the knowledge of the sequence of the anthyphairetic quotients of a to b [1,2,2,2,...], denoted for short [1, period (2)].

Finally, *Recollection* is a (poetic) reformulation of the repetition c_1/c_2 of the ratio b/c_1, instituting the complete knowledge of the anthyphairesis of a to b.

The *Meno* might appear as an isolated work of geometrical Knowledge, of the dyad <diameter, side of a square>, certainly related to the Pythagorean discovery of the incommensurability of the diameter to the side of a square, but, for Plato, the *Meno* is something more, is, in fact, a paradigmatical case for his philosophical theory of Ideas/True Intelligible Beings, and thus the quest for Knowledge in the *Meno* opens the gate and provides a key for the whole of Plato's philosophy.

The philosophical terms that Plato uses to describe the geometrical situation, namely True Opinion and Logos, fit in his philosophical theory and strongly suggest that for Plato, a true Intelligible Being is a dyad, to wit either the <One, Being> in the second hypothesis of the *Parmenides* or the <Being, not-Being> in the *Sophist*, in the philosophical analogue of anthyphairesis, satisfying the philosophical analogue of Theaetetus' Logos Criterion for anthyphairetic periodicity, as it becomes amply clear in the revealing divisions (of the Angler, Noble Sophistry, and the Sophist) in the *Sophist*, and in the description of the paradigmatical true Being in the second hypothesis in the *Parmenides* in terms of the anthyphairetic division of the dyad <One, Being> plus the circularity of the "hapseis"/logoi.

Knowledge of True Being, namely the dyad <Being, not-Being> is *True Opinion*, namely an initial finite segment of the anthyphairesis of the dyad plus *Logos*, namely the philosophic analogue of Theaetetus' Logos Criterion. Theaetetus' proof of the Pythagorean Proposition: if $a^2 = 2b^2$, then Anth (a, b) = [1, period (2)].

The description of an Intelligible Being and its Knowledge takes various forms in Plato's dialogues but remains largely equivalent to the model introduced in the Meno.

[Part I] The first one and a half steps of the anthyphairesis of a to b ($a = b + c_1$, $b = 2c_1 + c_2$) plus

Keywords

Plato · Meno · Diameter and side of a square · True Opinion · Logos · Logismos-Bond-Knowledge · Knowledge · Recollection · Theaetetus · Periodic anthyphairesis

1 Outline

[Section 2] Computation of the Anthyphairesis of a to b, with $a^2 = 2b^2$ *probably by Theaetetus, the model for Plato's theory of Recollection in the* Meno.

We start with Theaetetus' anthyphairetic definition of the proportion of magnitudes, the Logos Criterion for anthyphairetic periodicity, and the reconstruction of

The reconstruction of the proof, employing tools developed by Theaetetus, consists of:

[Part II] The Theaetetus' Logos Criterion ($b/c_1 = c_1/c_2$) (Proposition 2.5)

The proof in Proposition 2.5 will serve as the model for the whole *Meno* passage on Knowledge as Recollection.

[Sections 3 and 4] *Experimental verification that the knowledge of the duplication of a square is recollection.*

The eristic thesis that knowledge is impossible is juxtaposed to the dialectic thesis that knowledge is possible because knowledge is recollection. *Meno* 80e1-81e3 (Section 3).

An uneducated boy will recollect the knowledge of line a, whose square a^2 with side a is double the square b^2 with the given side b. *Meno* 81e3-82e2 (Section 4).

[Sections 5 and 6 on True Opinion] *Anthyphairetic interpretation of True Opinion in the* Meno*, corresponding to Part I of the proof of Proposition 2.5.*

The initial stage of Recollection of the Knowledge of line a in relation to line b, with $a^2 = 2b^2$, is identified with the geometrical initial segment of the anthyphairesis of line a to line b, with $a^2 = 2b^2$. *Meno* 82e2-83e10 (Section 5).

The philosophic term True Opinion about the line a to b, with $a^2 = 2b^2$ is identified with the geometric term Initial Segment of the Anthyphairesis of a to b, with $a^2 = 2b^2$.

Meno 83e11-d2, 97a9-c5, and *Sumposion* 210a4-e6 (Section 6).

[Section 7] *Heuristics.*

From the anthyphairetic interpretation of True Opinion (Sections 5 and 6) and the statement that *Knowledge is True Opinion plus Something*, we expect (a) that the *Something* needed will be (a philosophic analogue of) the *Logos Criterion* of Proposition 2.3, and (b) that the Knowledge of line a with respect to the line b will be the *full infinite anthyphairesis* of a to b. But this remains to be verified. *Meno* 97c6-d3.

[Section 8] *Name only does not constitute Knowledge.*

The line a, such that its square a^2 is the double of the square b^2 of the given line b, is shown, to be the line with *the name diameter* of the square with side b. The unexpected statement in 85b4 that the name of the line is given by the *sophists*, and not, as would be expected, by the geometers, suggests that Plato has in mind the statement, given in the *Republic* 454a-b, that knowledge by *name (onoma) only is eristic/sophistic* and not dialectical. *Meno* 84d3-85b7.

[Sections 9–12 on Logos] *Anthyphairetic Interpretation of Logos. Knowledge and Recollection in the* Meno is *corresponding to Part II of the proof of Proposition 2.5.*

The first indirect hint of how True Opinion is to be transformed into Knowledge and Recollection is given in terms of the crucial word *pollache. Meno* 85b8-86b5 (Section 9).

For the first time, in *Meno* 86b6-c3, Plato explicitly brings forth the fundamental role of *Logos*. This is certainly no isolated occurrence but by what Plato says, elsewhere, including *Republic* 534b3-6, *Theaetetus* 202c2-5, *Sophist* 260a5-b2, and *Statesman* 286a4-7. Scholars render "logos" by such unhelpful terms as "account," "argument," "discourse," "rational definition," "rational explanation," "reason," and "word" (Section 10).

The revealing interpretation of Logos in the Divisions of the Angler, the Noble Sophistry, and the Sophist in the *Sophist* as the philosophic analogue of Theaetetus' Logos Criterion establishes the nature of Plato's intelligible Being as the philosophic analogue of a dyad in anthyphairetic periodicity (Section 11).

The Logos/Logismos/analogio as *bond* in the *Sophist* 268c5-d5 in connection with our interpretation of the Division of the Sophist in Section 11, in the *Sophist* 251d8-9, 253a4-c1, and in the *Timaeus* leads to the conclusion that the Logismos/Bond in the *Meno* 97d4-98a4 passage is the Logos Criterion $b/c_1 = c_1/c_2$ (Section 12).

[Section 13 on Knowledge, Recollection in the *Meno* as the gate for understanding the whole of Plato's philosophy] Knowledge of the diameter a to the side b of a square is True Opinion (the first two anthyphairetic relations of the diameter to the side) plus Logos (the Logos Criterion $b/c_1 = c_1/c_2$), namely the full knowledge of the infinite anthyphairesis of a to b. This Knowledge is described poetically as Recollection, simply in the sense that the second ratio is a repetition/recollection of the first. Understanding Plato's philosophical terminology (Knowledge, True Opinion, Logos, and Recollection) in a concrete geometrical problem provides valuable insight and confirmation for the whole of Plato's philosophy (Section 13).

2 *Theaetetus' anthyphairetic definition of the proportion of magnitudes, the Logos Criterion for anthyphairetic periodicity, and the reconstruction of Theaetetus' proof of the Pythagorean Proposition: if $a^2 = 2b^2$, then Anth (a, b) = [1, period (2)]*

Our reconstruction of the original Pythagorean proof of the incommensurability of diameter to the side of a square, described in Negrepontis and Farmaki (2019) and Negrepontis et al. (2020), involves heavily the Pythagorean Application of Areas, to computing initial two or three anthyphairetic quotients, and in the form of Preservation of Gnomons, to ensure that the anthyphairesis continues *ad infinitum* as in steps 2 and 3.

We know from the famous Aristotle's *Topics* 158b passage that the first definition of the proportion of magnitudes was in terms of equal anthyphairesis:

2.1. Definition (Theaetetus, cf. Aristotle, *Topics* 158b29-36). If a, b are magnitudes and c, d are magnitudes, then the proportion $a/b = c/d$ holds if the sequence of quotients Anth(a,b) of the anthyphairesis of a to b is equal to the corresponding sequence Anth(c,d) of the anthyphairesis of c to d.

2.2. Proposition (Theaetetus), the **Logos Criterion** for the periodicity of anthyphairesis.

If a, and b are magnitudes, a > b, with anthyphairesis

$$a = k_0 b + c_1,$$

$$b = k_1 c_1 + c_2,$$

$$c_1 = k_2 c_2 + c_3,$$

$$\cdots$$

$$c_{m-1} = k_m c_m + c_{m+1}$$

$$\cdots$$

and there are indices m < n, such that

$$c_m / c_{m+1} = c_n / c_{n+1},$$

then the anthyphairesis of a to b is eventually periodic and in fact
Anth(a, b) = $[k_0,k_1,\ldots,k_{m-1}, \text{period}(k_m,\ldots k_{n-1})]$.

Proof By definition of proportion, $\text{Anth}(c_m, c_{m+1}) = \text{Anth}(c_n,c_{n+1})$. Then
Anth (a, b) =
$[k_0,k_1,\ldots,k_{m-1}, k_m,\ldots k_{n-1}, \text{Anth}(c_n, c_{n+1})]=$
$[k_0,k_1,\ldots,k_{m-1}, k_m,\ldots k_{n-1}, \text{Anth}(c_m, c_{m+1})]=$
$[k_0,k_1,\ldots,k_{m-1}, k_m,\ldots k_{n-1},k_{m-1}, k_m,\ldots k_{n-1}, \text{Anth}(c_n, c_{n+1})]=$
$[k_0,k_1,\ldots,k_{m-1}, k_m,\ldots k_{n-1},k_{m-1}, k_m,\ldots k_{n-1}, \text{Anth}(c_m, c_{m+1})]=$
$[k_0,k_1,\ldots,k_{m-1}, k_m,\ldots k_{n-1},k_{m-1}, k_m,\ldots k_{n-1}, k_{m-1}, k_m,\ldots k_{n-1}, \text{Anth}(c_n, c_{n+1})]=$
$\ldots$
$[k_0,k_1,\ldots,k_{m-1}, \text{period}(k_m,\ldots k_{n-1})]$.

We also need the following elementary:

2.3. Proposition If a,b,c,d are lines, such that ad = bc, then Anth(a,b) = Anth(c,d).

Proof Let Anth(a,b) = $[I_0, I_1,\ldots]$. We proceed by induction. Thus $a = I_0 b + e_1$, with $e_1 < b$. Then $cb = ad = I_0 bd + e_1 d$. By Proposition I.44 of the *Elements*, there is a line segment f_1, such that $e_1 d = bf_1$.

Since $e_1 < b$, it is clear that $f_1 < d$. Thus, $cb = I_0 bd + e_1 d = I_0 bd + bf_1$; then $c = I_0 d + f_1$ with $f_1 < d$. Thus Anth(a, b) = $[I_0,I_1,\ldots]$ = $[I_0, \text{Anth}(b, e_1)]$, while Anth(c, d) = $[I_0, \text{Anth}(d, f_1)]$. Thus, we have that $e_1 d = bf_1$ and we must prove that $\text{Anth}(b, e_1) = \text{Anth}(d, f_1)$.

We continue as in the first step and finish the proof by induction.

We note that the earliest recorded use of (incomplete) mathematical induction to our knowledge occurs for the Pythagorean proof of Pell's property for N = 2 of the side and diameter numbers (examined in Negrepontis et al. (2023), Section 9).

2.4. Proposition (*Pythagorean Theorem*). If a, b, and c are the sides of a right triangle, with hypotenuse c, then $a^2 + b^2 = c^2$.

Proof The Euclidean proof in Proposition I.47 of the *Elements* is, as Proclus states in his *Commentary to Euclid* 426,9-14, due to Euclid. The original Pythagorean proof is most likely based on Proposition II.4 of the *Elements* and is a natural generalization of the special isosceles case, proved in the *Meno*. Negrepontis and Farmaki (2019), Section 5.7 gives a reconstruction along these lines.

2.5. Proposition (*Pythagorean*). If b is a line and a is a line such that $a^2 = 2b^2$, then a is the diameter of the square with side b, and Anth(a,b) = [1, period(2)].

Proof *(Theaetetus)*
[Part I]
By the (isosceles) Pythagorean theorem (Proposition 2.4), line a is constructed as the diameter of the square with given side b. Clearly $a > b$,
Since $a^2 = 2b^2$, it follows that $a < 2b$
(if we had $a \geq 2b$, then $a^2 \geq 4b^2 > 2b^2 = a^2$, contradiction).
Thus, $b < a < 2b$. Hence,
$a = b + c_1$, $c_1 < b$.
It follows that $b > 2c_1$. $3b > 2a$
(if we had $b \leq 2c_1 = 2a-2b$, then $3b \leq 2a$, hence $9b^2 \leq 4a^2 = 8b^2$, contradiction). Hence,
$b = 2c_1 + c_2$.
$b > 2(a-b)$, $3b > 2a$.
[Part II]
Then, $c_1 = a-b$, hence $c_2 = b-2c_1 = b-2(a-b) = 3b-2a$, hence
$b.c_2 = b(3b-2a) = 3b^2-2ab$, and
$c_1^2 = (a-b)(a-b) = a^2 + b^2-2ab = 3b^2-2ab$.

Hence, $b.c_2 = c_1^2$. Since $b > c_1$, and
$b.c_2 = c_1^2$,
clearly $c_2 < c_1$,
thus the second step is indeed anthyphairetic
$b = 2c_1 + c_2$, $c_2 < c_1$.
By $b.c_2 = c_1^2$ and Proposition 1.3, we have
Anth (b, c_1) = Anth (c_1, c_2), namely $b/c_1 = c_1/c_2$, according to the Theaetetean definition.
Then Anth (c_1, c_2) = Anth (b, c_1) = [2, Anth (c_1, c_2)]. Hence,
Anth (a, b) = [1, 2, Anth (c_1, c_2)] = [1,2, Anth (b, c_1) = [1,2, 2, Anth (c_1, c_2)] = ... = [1, period (2)].
Notes. There is no such proof preserved in ancient sources. The geometer who would conceive of such a proof would normally be Theaetetus, since Theaetetus was the first geometer to discover a theory of proportion of magnitudes based on equal anthyphairesis, the pre-Eudoxean theory mentioned by Aristotle in *Topics* 158b.

In the *Topics* passage, there is no explicit attribution, but there is something of a consensus among historians of Mathematics that the theory is due to Theaetetus.

By Proposition 2.5, we obtain *the complete knowledge* of the diameter a with respect to the side a of the square. Since the anthyphairesis of a to b is infinite, by Proposition X.2, a and b are incommensurable.

The computation depends on the representation of the first quotient c_1 as a-b, and of the second quotient c_2 as equal to 3b-2a. These representations use the *side and diameter numbers* q_n, p_n; thus, $c_1 = p_1a\text{-}q_1b = a\text{-}b$, $c_2 = q_2b\text{-}p_2a = 3b\text{-}2a$.

Theaetetus' proof provides *the complete knowledge* of the diagonal with respect to the side of a square with one and a half anthyphairetic division steps plus the Logos Criterion (consisting in the *repetition* of Logos). This *proof* will be our guide for deciphering and interpreting Plato's theory of *Knowledge as Recollection* in the dialogue *Meno*, passages 80e1-86c3 and 97a9-98b6. In turn, this interpretation of *Meno* makes clear that indeed Theaetetus had obtained the present proof.

3 *Eristic logos (knowledge is impossible) vs. dialectic Logos (knowledge is possible because knowledge is recollection).* Meno *80e1-81e2*

The eristic argument holds that knowledge is impossible 80e1-81a4.

> [*Socrates*] Do you see what an *eristic argument (eristikon logon 80e2)* you are introducing—that, forsooth, a man cannot inquire either about what he knows or about what he does not know? For he cannot inquire about what he knows, because he knows it, and in that case is in no need of inquiry; nor again can lie inquire about what he does not know, since he does not know about what he is to inquire.
> *Meno* Now does this argument (*logos*) seem to you to be a good argument (kalos legesthai), Socrates?
> *Socrates* It does not.
> *Meno* Can you explain (*legein*) how not?

[Translation of the *Meno* passages in this present chapter, unless otherwise indicated, are from *Plato in Twelve Volumes*, by W. R.M. Lamb, with modifications, including italicization of some words for emphasis, by the author.]

The dialectic argument holds that Knowledge is Recollection 81a5-d5.

> *Socrates* I can; for I have heard from wise men and women who told of things divine that
> *Meno* What was it they said *(tina logon legonton)*?
> *Socrates* Something true, as I thought, and admirable.
> *Meno* What was it? And who were the speakers *(hoi legontes)?*
> *Socrates* The speakers *(hoi men legontes)* were certain priests and priestesses who have studied so as to be able to provide logos *(logon* didonai*)* of their ministry; and Pindar also *speaks (legei)* and many another poet of heavenly gifts.
>
> As to their words (ha de legousin), they are these: mark now, if you judge these sayings to be true (*alethe legein*). They say that *the soul of man is immortal*, and at one time comes to an end, which is called dying, and at another is born again, but never perishes. Consequently, one ought to live all one's life in the utmost holiness.

"For from whomsoever *Persephone (Phersephona)* shall accept requital for ancient wrong, the souls of these she restores in the ninth year to the upper sun again; from them arise" "glorious kings and men of splendid might and surpassing wisdom, and for all remaining time are they called holy heroes amongst mankind."

Seeing then that the soul is immortal and has been born many times, and has beheld all things both in this world and in the nether realms, she has acquired knowledge (*mematheken 81c7*) of all and everything; so that it is no wonder that she should be able *to recollect* (*anamnesthenai 81c8*), all that she knew (*epistato*) before about virtue (*aretes*) and other things. For as all nature is akin, and the soul has learned (memathekuias 81c9) all things, there is no reason why we should not, by *recollecting* (*anamnesthenta 81d2*) but one single thing—an act which men call *learning (mathesin 81d2)*—discover everything else, if we have courage and faint not in the search; since, it would seem, *research (zetein 81d4) and learning (manthanein 81d4)* are wholly *recollection (anamnesis 81d5).* 81a5-d5

The eristic argument vs. the dialectical argument 81d5-e2.

So we must not hearken to that *eristic* argument (*toi eristikoi logoi* 81d6): it would make us *idle* (argous), and is pleasing only to the indolent ear, whereas the other makes us *energetic (ergatikous) and inquiring (zetetikous 81e1).* Putting my trust in its truth, I am ready to inquire with you into the nature of virtue. 81d5-e2

We can summarize the two opposing arguments:

Eristic logos: Humans cannot search for or attain knowledge, because if you [think you] possess knowledge of something, there is no reason to search for it, and if you do not possess knowledge of something, then you do not know what to search for.

Dialectic logos: Humans have the power to search and attain Knowledge because the human soul is immortal and attains knowledge by recollection (anamnesis) of what she has forgotten.

4 *Experimental verification that knowledge is recollection: an uneducated boy will recall the knowledge of line a, whose square a^2 with side is the double of the square b^2 with given side b*. Meno 81e3-82e2

4.1. *Experimental verification that knowledge is recollection with an uneducated boy.* Meno *81e3-82b8*.

Socrates will show experimentally that "knowledge is recollection" by guiding an uneducated boy to recollect knowledge by himself

Meno Yes, Socrates, but what do you mean by saying that we do not learn (hou manthanom) and that what we call *learning* (mathesin) is *recollection* (anamnesis)?
Can you instruct me that this is so?
Socrates I remarked just now, Meno, that you are a rogue and so here you are asking.
if I can instruct you when I say *there is no teaching but only recollection (anamnesin):* you hope that I may be caught contradicting myself forthwith.
Meno I assure you, Socrates; that was not my intention I only spoke from habit. But if you can somehow prove (endeixasthai) to me that it is as you say, pray do so (enseixai).

Socrates It is no easy matter, but still, I am willing to try my best for your sake. Just call one of your own troop of attendants there, whichever one you please, that he may serve for my demonstration (epideixomai).
Meno Certainly. You, I say, come here.
Socrates He is a Greek, I suppose, and speaks Greek (hellenizei);?
Meno Oh yes, to be sure—born in the house.
Socrates Now observe closely whether he strikes you as *recollecting* (*anamimneskomenos*) or as learning μανθάνων from me.
Meno I will.

4.2. *Definition of square; the area of a square with side b is* b^2. Meno 82b9-d8**.**

Socrates Tell me, boy, do you know that a square figure (*tetragonon chorion*) is like this?
Boy I do.
Socrates Now, a square figure has these lines, four in number, all equal?
Boy Certainly.
Socrates And these, drawn through the middle, are equal too, are they not?
Boy Yes.
Socrates And a figure of this sort may be larger or smaller?
Boy To be sure.
Socrates Now if this side (pleura) were two feet and that also two, how many feet would the whole be? Or let me put it thus: if one way it were two feet, and only one foot the other, of course, the space would be two feet taken once?
Boy Yes.
Socrates But as it is two feet also on that side, it must be twice two feet?
Boy It is.
Socrates Then the space is twice two feet? *Boy* Yes.
Socrates Well, how many are twice two feet? Count (logisamenos) and tell me.
Boy Four, Socrates.

Socrates takes b = 2 feet, but it is clearer if we deal throughout with the general side b. The construction of the square with the given side is in Proposition I.46 of the *Elements,* making use of the Fifth Postulate.

4.3. *Statement of the problem: to obtain the knowledge of line a whose area* a^2 *is double* $2b^2$ *the area* b^2 *of the square with side b*. Meno 82d8-e2

Socrates And might there not be another figure twice (diplasion), the size of this, but of the same sort, with all its sides equal like this one?
Boy Yes.
Socrates Then how many feet will it be?
Boy Eight.
Socrates Come now, try and tell me how long (pelike) will each line (gramme) side of that figure be. This one is two feet long: what will be the side of the other, which is double in size?

At this stage, it is not yet clear what would be considered satisfactory *knowledge of this a.*

5 *The initial stage of Recollection of the Knowledge of line a in relation to line b, with $a^2 = 2b^2$, is identified with the geometrical initial segment of the anthyphairesis of line a to the b, with $a^2 = 2b^2$.* Meno *82e2-83e10*

5.1. *The first anthyphairetic relation of a to b:*
$b < a < 2b$, equivalently $a = b + c_1$, $c_1 < b$,
is the result of successive (ephexes) recollection by the boy under questioning.
Meno *82e2-83e10*

Boy Clearly, Socrates, double.

Let $a^2 = 2b^2$, and it is attempted to obtain the knowledge of line a in terms of the given line b.

The initial reply of the Boy that $a = 2b$ shows that the Boy falsely thinks that he knows the lina a.

Socrates Do you observe, Meno, that I am not teaching the boy anything, but merely asking him each time? And now he supposes that he knows about the line required to make a figure of eight square feet; or do you not think he does?

Meno I do.
Socrates Well, does he know?
Meno Certainly not.
Socrates He just supposes it, from the double size required?
Meno Yes.
Socrates Now watch his *recollecting (anamimneskomenon) in succession (ephexes 82e12), as one must recollect (hos dei anamimneskesthai).*
Tell me, boy, do you say we get the double area from the double line? The space I speak of is not long one way and short the other but must be equal each way like this one while being double its size—eight square feet. Now see if you still think we get this from a double length of the line.
Boy I do.
Socrates Well, this line is doubled, if we add here another of the same length?
Boy Certainly.
Socrates And you say we shall get our eight-foot area from four lines of this length?
Boy Yes.
Socrates Then let us describe the square, drawing four equal lines of that length. This will be what you say is the eight-foot figure, will it not?
Boy Certainly.
Socrates And here, contained in it, have we, not four squares, each of which is equal to this area of four feet?
Boy Yes.
Socrates Then how large is the whole? Four times that area, is it not?
Boy It must be.
Socrates And is four times equal to double?
Boy No, to be sure.
Socrates But how much is it?

Boy Fourfold.
Socrates Thus, from the double-sized line, boy, we get an area, not of double, but of fourfold size.
Boy That is true.
Socrates And if it is four times four it is sixteen, is it not?
Boy Yes.
Socrates What line will give us an area of eight feet? This one gives us a fourfold space, does it not?
Boy It does.
Socrates And an area of four feet is made from this line of half the length?
Boy Yes.
Socrates Very well; and is not a space of eight feet double the size of this one, and half the size of this other?
Boy Yes.
Socrates Will it not be made from a line longer than the one of these, and shorter. than the other?
Boy I think so.
Socrates Excellent: always answer just what you think. Now tell me, did we not draw this line two feet, and that four?
Boy Yes.
Socrates Then the line on the side of the eight-foot figure should be more than this of two feet, and less than the other of four?
Boy It should.

Socrates by questioning him guides him to the "recollection in succession," as he calls the process, in fact to the double inequality

$$b < a < 2b,$$

a condition clearly equivalent to *the first anthyphairetic relation of line a to line* b:

$$a = b + c_1, c_1 < b.$$

5.2. *The (first and the) half of the second step of the anthyphairesis of a to b with* $a^2 = 2b^2$ $b = 2c_1 + c_2$, *equivalently* $2a < 3b$, *are the result of walking (badizon, 84a3-4) on [the road of] recollection under questioning*. Meno *83e1-10.*

Socrates Try (peiro) and tell me how much (peliken) you would say it is.
Boy Three feet (tripous).
Socrates Then if it is to be three feet, we shall add on a half to this one, and so make it three feet? For here we have two, and here one more, and so again on that side there are two, and another one; and that makes the figure of which you speak.
Boy Yes.
Socrates Now if it be three this way and three that way, the whole space will be thrice three feet, will it not?
Boy So it seems.
Socrates And thrice three feet are how many?
Boy Nine.
Socrates And how many feet was that double one to be?
Boy Eight.
Socrates So we fail to get our eight-foot figure from this three-foot line.
Boy Yes, indeed. 83e1-10

It is easily seen that the next step consists in proving the inequality

$2a < 3b$.

[Assume that $2a \geq 3b$. But this cannot happen, since then we would have $8b^2 = 4a^2 \geq 9b^2$, impossible.]

Thus, there is a line c_2, such that $3b = 2a + c_2$, equivalently $b = 2(a\text{-}b) + c_2$, and since $c_1 = a\text{-}b$,

$b = 2c_1 + c_2$,

We don't yet know if $c_2 < c_1$, so this is *half of the second anthyphairetic relation of line a to line b*.

It is clear that these two division steps of the anthyphairesis of a to b form the initial steps of the process of recollection; this is stated explicitly at the beginning (82e12) of the passage 82d5-83d5, where the two anthyphairetic division steps are established, and it is also stated in 84a3-4.

6 *The philosophic term True Opinion about the line a to the given line b, with $a^2 = 2b^2$, is identified with the knowledge of an initial segment of the geometric Anthyphairesis of a to b, with $a^2 = 2b^2$.* (Meno *83e11-84d2, 97a9-c5,* and Sumposion *210a4-e6*)

6.1. *The first two steps of the anthyphairesis of a to b are described as the process of* ***walking towards recollection***. Meno 83e11-84b2.

> *Socrates* But, from what line shall we get it? Try and tell us exactly (akribos 83e11); and if you would rather not reckon it out, just show (deixon) what line it is (84a1).
> *Boy* Well, on my word, Socrates, I for one do not know (ouk oida).
> *Socrates* There now, Meno, do you observe who is already *walking in recollection (badizon hode tou anamimneskesthai*, 84a3-4)? At first, he did not know what is the line that forms the figure of eight feet, and he does not know even now (oude nun oiden):
> but at any rate, he thought he knew (eidenai) then, and confidently answered as though he knew (eidos), and was aware of no difficulty; whereas now he feels the difficulty he is in, and besides not knowing (oiden) does not think he knows (eidenai).
> *Meno* That is true. 83e11-84b2.

6.2. ***True Opinions*** *of the boy about the lines a and b, such that $a^2 = 2b^2$ are* ***formed under questioning***. Meno 84b3-d2.

Compared to the initial false belief of the boy that he possessed knowledge already, the aporia/wonder and realization that he had no knowledge but only some partial initial knowledge, even though this comes as a shock, is a positive and necessary step for the eventual acquisition of knowledge.

> *Socrates* And is he not better off in respect of the matter which he did not know?
> *Meno* I think that too is so.
> *Socrates* Now, by causing him to doubt and giving him *the torpedo's (electric ray's) shock* (*narkan*), have we done him any harm?

Meno I think not.
Socrates And we have certainly given him some assistance, it would seem, towards finding out the truth of the matter: for now, he will push on in the search gladly, as lacking knowledge; whereas then he would have been only too ready to suppose he was right in saying, before any number of people any number of times, that the double space must have a line of double the length for its side.
Meno It seems so.
Socrates Now do you imagine he would have attempted to inquire or learn what he thought he knew when he did not know it until he had been reduced to the perplexity of realizing that he did not know, and had felt a craving to know?
Meno I think not, Socrates.
Socrates Then *the torpedo's shock* (*narkesas*) was of advantage to him?
Meno I think so.
Socrates Now you should note how, as a result of this perplexity, he will go on and discover something by joint inquiry with me, while I merely ask questions and do not teach him; and be on the watch to see if at any point you find me teaching him or expounding to him, instead of questioning him *on his opinions. (doxas)* ething by joint inquiry with me, while I merely ask questions. 84b3-d2

6.3. *True Opinions are described as* ***"ephexes"*** *in* Symposium *210a4-e6.*

Plato, in the *Symposium* 210a-e, describes the ascent to ideal Beauty as at first *"ephexes,"* a process gradual, sequential, and painful, gathering many higher and higher true opinions ("in the first place" true opinion about a beautiful body, 210a6; "but next" true opinion about many beautiful bodies, 210a8; "but after these" true opinions about customs and laws, 210b6; "after" true opinions about all the "mathemata," 210c8); and finally, suddenly and unexpectedly as *"exaiphnes,"* culminating with the revelation of the knowledge of the idea Beauty.

6.4. *Knowledge of the road (hodon, 97a9, b1) to Larisa is likened to Knowledge of the diameter* to the side, and ***walking*** *("badizoi," 97a10) on the road to Larisa is likened to the True Opinions about the diameter to the side.* Meno *97a9-c5*

Socrates I will tell you. If a man knew the road to Larisa (eidos ten hodon ten eis Larisan, 97a9) or any other place you please, and walked (badizoi, 97a10) there and led others, would he not give right and good guidance?
Meno Certainly.
Socrates Well, and a person who had a right opinion (orthos men doxazon) as to which was the road (hodos, 97b1, but had never been there and did not really know (med' epistamenos), might give right guidance, might he not?
Meno Certainly.
Socrates And so long, I presume, as he has right opinion (orthen doxan) about the road, which the other man really knows (epistemen), he will be just as good a guide—if he possesses true opinion (oiomenos men alethe) but not the knowledge of it (phronon de me), —as the man who has the knowledge.
Meno Just as good.
Socrates Hence true opinion (doxa alethes) is as good a guide to the rightness of action as knowledge; and this is a point we omitted just now in our consideration of the nature of virtue when we stated that knowledge is the only guide of right action; whereas we find there is also true opinion (doxa alethes).

Meno So it seems.
Socrates Then right opinion (orthe doxa) is just as useful as Knowledge (epistemes). *97a9-c5.*

6.5. *The identification of the philosophic term* ***True Opinions*** *about the line a to b, with* $a^2 = 2b^2$ *with the geometric term* ***Initial Segment of the Anthyphairesis*** *of a to b, with* $a^2 = 2b^2$.

Under Socrates' questioning, the boy was able to find the first steps of the anthyphairesis of line a to the line b if $a^2 = 2b^2$, in 82b9-84c9; and Plato tells repeatedly in 84c10-d2, 85b8-d2 that the boy is forming true opinions about line a in relation to the line b under Socrates' questioning.

The anthyphairetic relations are described as products of the process towards *recollection*, a process that initially is described as *successive* (*ephexes*, 82e12) recollection and then as *walking* (*badizon*, 84a3-4) towards recollection. In the *Symposium* 210a-e, the True Opinions of the Idea of Beauty are described as ephexes, while in the 97a-d passage, True Opinion is likened to walking the road to Knowledge.

Based on these similarities, we have no hesitation to identify the geometric concept of the initial segment of the anthyphairesis of a to b, with $a^2 = 2b^2$, with the philosophical concept of True Opinion of line a in relation to the line b, for $a^2 = 2b^2$.

We expect that in places where Plato is talking about True Opinion about an intelligible Being, as in the *Symposium*, the *Republic*, the *Theaetetus*, and the second hypothesis in the *Parmenides* that there is some philosophic analogue of anthyphairesis related to the intelligible Being, and that the True Opinion will be a finite initial segment of the anthyphairesis of the intelligible Being.

The identification of the philosophical term True Opinion with a finite initial segment of the anthyphairesis

	Initial anthyphairetic relations of a to b with $a^2 = 2b^2$	True opinions about a to b with $a^2 = 2b^2$
Described as in succession (ephexes)	*Meno 82e12*	*Symposium 210a-e*
Likened to walking (badizein)	*Meno 84a3*	*Meno 97a10*
In the road (hodos)		*Meno 97a9, b1*
Formed as answers of the boy to Socrates' questioning	*Meno 82b9-84c9*	*Meno 84c10-d2, 85b8-d2*

6.6. Bluck (1961) *finds no satisfactory explanation why the process of recollection toward Knowledge should be in order "ephexes."*

Bluck (1961) writes:

> The only rule we are given is that the correct way to recollect is to recollect things '*in order*' (ἐφεξῆς, 82e12), which probably means that one thing must be dealt with at a time, and is perhaps also to be associated with the 'kinship of all nature' and the consequent possibility, through remembering one thing, of being reminded of everything else (81d); *but we are not told what a correct order of remembering might be*. p. 16

ἀναμιμνησκόμενον ἐφεξῆς. We have seen above (81d) that it is possible through remembering one thing to be reminded of another, and it is probably with reference to this fact that Socrates now says that one must recollect things *'in order'*. But we are not told what the relationship is between the things here in question, or *in what order* they should be recollected. *One thinks of the various 'steps' on the 'upward path' of the dialectican's progress in the Republic (511b), where again the precise relationship between the steps is not made clear.*

Probably in our present passage, at any rate, Plato has no very definite relationship or *definite order in mind* but wishes to insist that one thing must be dealt with at a time, since only if a thing is properly understood can we be reminded by it of other things. In this passage, the positive procedure of eliciting from Meno a correct response is bound up with the elenctic or refutative procedelencticure designed to show him that his false answers are false, which continues down to 84a. The refutation of the false answers involves showing Meno what the area is of a square with sides of three feet or o four, and the very recollection of these facts may be regarded as steps towards the recollection of the correct answer to the question about a square of eight square feet. *This is largely a method of trial and error*; but o, discovery can lead to another. Having realized that neither a two- nor three- nor four-feet will give a square of eight square feet, and having before us the diagram of a square of sixteen square feet with transversals . . ., one might well think of halving this latter square (which is what Socrates does), and thereby discover the desired line. Recognition of the nature of the four-foot-sided square and of the three-foot-sided square thus forms part of a thought process that could, and does, lead to the discovery of the correct solution. If this is what Plato had in mind, 'in order' will not refer to any particular series or special procedure, but simply to the desirability of understanding one thing at a time, if ἀνάμνησις is to function through the association of ideas. *No particular order is required*, and no particular relationship between the ἀναμνηστά is implied; but only confusion results from trying to take a second step before the first has been completed (cf. *Phaedo* 101 d-e). p. 297–298 [Italics are added by the author.]

Bluck is attempting to explain in what sense Plato describes as "ephexes," in order, or "badizein," walking in steps, the process of recollection in the *Meno*, and also notes similar descriptions in the *Republic* 511b and the *Phaedo* 101d-e, but also, as we have noted above, in the *Symposium* 210a4-e6, but finds no satisfactory explanation.

Our realization that what Plato is describing in the *Meno*, and in fact in the other cases mentioned by Bluck, are the first steps of the geometric anthyphairesis of line a tu the line b, explains satisfyingly the reason why Plato emphasizes the successive nature of the process.

7 Heuristics: From the statement that Knowledge is True Opinion plus Something (Meno 97c6-d3) and the anthyphairetic interpretation of True Opinion (Sections 4 and 5), we expect that Knowledge will be the full infinite anthyphairesis of a to b and the "something" needed will be the Logos Criterion

7.1 *Knowledge is True Opinion plus Something*

Meno With this difference, Socrates, that he who has *knowledge (epistemen)* will forever (aei) hit on the right way, whereas he who has the right opinion (*orthen doxan*) will sometimes do so, but sometimes not.

Socrates How do you mean? Will not he who always has the right opinion (orthen doxan) be right, so long as he opines rightly (ortha doxazoi)?
Meno It appears to this being the case, that *knowledge (episteme) should ever be more prized (polu timiotera) than right opinion (orthes doxas)*, and why they should be two distinct and separate things. *97c6-d3*

The passage 97c-d in the *Meno* clearly states that the Knowledge of line a that doubles the square of line b is True Opinion plus Something. Since we have identified True Opinion with an initial segment of the anthyphairesis of a to b, it is natural to expect that for Plato, the full Knowledge of a to b will be *the knowledge of the full infinite anthyphairesis of a to b*. From the second part of the proof of Proposition 2.5, we can see that the full infinite anthyphairesis of a to b is obtained with the addition of Theaetetus' Logos Criterion to the first two anthyphairetic relations. Thus, we naturally expect that *Knowledge will be defined as True Opinion plus the Logos Criterion*.

8 *Line a such that its square a^2 is double the square b^2 of the given line b, is shown, to be the line with the name diameter of the square with side b*

However, Knowledge by name (onoma) only is eristic/and not dialectical, as stated in the Republic *454a-b; the sophists promote such an eristic knowledge for the case of the diameter according to* Meno *85b4-5.* Meno *84d3-85b7 and* Republic *454a-b.*

8.1. *The line a such that $a^2 = 2b^2$ is identified in* Meno 84d3-85b7 with **the diameter** of the square with side b.

[*Socrates*]. Tell me, boy: here we have a square of four feet, have we not? Do you understand?
Boy Yes.
Socrates And here we add another square2 equal to it?
Boy Yes.
Socrates And here a third, 3 equal to either of them?
Boy Yes.
Socrates Now shall we fill up this vacant in the corner?
Boy By all means.
Socrates So here we must have four equal areas (choria)?
Boy. Yes.
Socrates Well now, how many times larger is this whole area than this other?
Boy Four times.
Socrates But it was to have been only twice, you remember?
Boy To be sure.
Socrates And does this line,1 drawn from corner to corner, *cut in two equal parts (temnousa dicha)* each of these areas?
Boy Yes.
Socrates And have we here four equal lines containing this area (chorion)?

Boy We have.
Socrates Now consider how large (pelikon) this area (chorion) is.
Boy I do not understand (ou manthano).
Socrates Has not each of the inside lines cut off half of each of these four areas?
Boy Yes.
Socrates And how many areas of that size are there in this part?
Boy Four.
Socrates And how many in this?
Boy Two.
Socrates And four is how many times two?
Boy Twice.
Socrates And how many feet is this area?
Boy Eight feet.
Socrates From what line do we get this figure?
Boy From this.
Socrates From the line drawn corner-wise across the four-foot figure?
Boy Yes.
Socrates The sophists (85b4) call (kalousin) it the diameter (diametron, 85b4):
so if *the diameter (diametros)* is its *name* (*onoma*, *85b5*), then according to you, Meno's boy, the double area (*diplasion chorion*) is the square of *the diameter*.
Boy Yes, certainly it is, Socrates. 84d3-85b7

In passage 84d3-85b7, line a such that $a^2 = 2b^2$ is proved to be the diameter/diagonal of the square with side b. The construction in the present passage of the a as the diameter is the answer to Socrates' second leg of his earlier question in 83e11-84a1, where Socrates had asked first for the *exact (akribos, 83e11)* nature of line a, namely for the complete knowledge of the a, and, failing this, for the lesser knowledge consisting in *showing (deixon apo poias, 84a1)* the line geometrically.

8.2. *Note*. The construction is closely related to a proof of the isosceles Pythagorean theorem and provides a clue for the proof of the general Pythagorean theorem with the use of Proposition II.4 of the *Elements*, a proof that most likely is the original Pythagorean proof of the Pythagorean theorem.

8.3. *However, the rigorous identification of line a with the diameter of the square with side b definitely does NOT provide the Knowledge of line a.*

It is noted and realized by most scholars that, even though the experiment with the boy is ending at this point, full Knowledge of the line a has not been achieved yet.

As Nehamas (1985), p. 22, correctly realizes knowledge and recollection have not been achieved yet:

> Socrates now claims that if 'someone asks him the same questions many times and in many ways' he will finally have as much knowledge about these topics as anyone else (85c9-d11). From this, it is clear that *the slave still does not have episteme* of the subject and Socrates drives the point home by locating the slave in the future in his very next question (*epistesetai*, 85d3-4). He then goes on to say that it is just *this recovery of knowledge state which is still all in the future for the slave, which is recollection*
> (85d6-7)

A clear indication that knowledge has not been achieved yet is noted by Fine (2014), p. 153:

> Socrates speaks of 'the knowledge which he [the slave] now has' (85d9). Socrates has been emphasizing the fact that the slave lacks knowledge. Why, then, does he say that the slave now has knowledge?
>
> One possibility is that he means that, though the slave doesn't have explicit knowledge, he has latent knowledge. However, Socrates doesn't explicitly mention latent or explicit knowledge. He simply says that the slave lacks knowledge and then, at 85d9, that he has it.
>
> We need not infer either that Socrates contradicts himself or that, in saying earlier that the slave lacks knowledge, he was speaking loosely. Nor do we need to invoke both explicit and latent knowledge to avoid contradiction. For the context makes it clear that 'now' (nun) is forward-referring, to the time when the slave will acquire the knowledge he has just been said to lack. The slave doesn't have knowledge in the present; but he will acquire it in the future if all goes well.

This agrees with our heuristic discussion in Sect. 7, according to which the full Knowledge sought in the *Meno* is the full infinite anthyphairesis of the diameter a to the side b of a square.

8.4. *According to* Republic *454a1-9,*
Knowledge by ***Name (kata to onoma) only*** *is eristic, anti-logical, and non-dialectic.*

> "What a grand thing, Glaucon," said I, "is the power of the *anti-logical* (antilogikes) art"
> "Why so?"
> "Because," said I, "many appear to me to fall into it even against their wills, and to suppose that they are acting *not eristically (ouk erizein, 454a5) but dialectically (dialegesthai),* owing to their inability to apply the proper *divisions according to kinds (kat'eide diairoumenoi)* to the subject under consideration, but pursuing the opposition on one another *according to name (kat' auto to onoma, 454a7)*, *eristically (eridi, 454a8), not dialectically (ou dialektoi)*." 454a1-9

We may reach an understanding of Plato's thought if we note that Plato employs the term *"name"* to describe the diameter. In the *Republic* 454a-b and the *Sophist* 218b-219a (cf. Sect. 11.4 below), among other places, Plato regards that the knowledge provided *by name only (onoma monon)* is not true dialectic Knowledge but only an *eristic* one, a term that for Plato has very little difference of meaning from the *sophistic*. So, our present knowledge of line a is at the "name only" level, at the level of knowledge practiced by the sophists.

The comparison of the *Meno* passage 84b-85d, especially Plato's statement that "*the sophists* (85b4) use *the name* (onoma, 85b5) diameter" for the line, which at first may strike us as weird, as we would expect Plato to state that "*the geometers* use the name diameter," is well explained, if we set it next to the *Republic* 454a-b passage, bearing in mind that the practice of the sophists is exactly eristic and non-dialectical. We expect that the term anti-logical for the eristic point of view, implying that the dialectical point of view is logical, will mean that for the dialectical method the Knowledge of the diameter is true and complete when, with the aid of the Logos Criterion, the full infinite anthyphairesis is obtained.

9 *The first rather indirect hint of how True Opinion is to be transformed into Knowledge and Recollection in* Meno *85b8-86b5 is given in terms of the crucial word pollache 85c11*

9.1. *It is repeated that the boy is at the stage of possessing True Opinions but no Knowledge 85b8-c8*

> *Socrates* What do you think, Meno? Was there any *opinion* (doxan) that he did not give as an answer of his own thought?
> *Meno* No, they were all his own.
> *Socrates* But you see, he did not know, as we were saying a while since.
> *Meno* That is true.
> *Socrates* Yet he had in him *these opinions (doxai)*, had he not?
> *Meno* Yes.
> *Socrates* So that he *who does not know* (ouk eidei) about any matters, whatever they be, may have *true opinions* (aletheis doxai) on such matters, about which *he possesses no knowledge (me eidei)*?
> *Meno* Apparently. 85b8-c8.

9.2 *The crucial part of the passage 85b8-86b5 is contained in 85c9-d2, stating that acquisition of Knowledge is achieved by being asked the same questions many times* ***in many places*** *(pollachei, 85c11)*

9.2.a. *Passage 85c9-d2 in Greek*

We first quote the Greek passage 85c9-d2 untranslated.

> ΣΩ. Καὶ νῦν μέν γε αὐτῷ ὥσπερ ὄναρ ἄρτι ἀνακεκίνηνται *αἱ δόξαι αὗται*
> εἰ δὲ αὐτόν τις ἀνερήσεται *πολλάκις τὰ αὐτὰ ταῦτα καὶ* ***πολλαχῇ*** *(85c11),*
> οἶσθ' ὅτι *τελευτῶν* οὐδενὸς ἧττον *ἀκριβῶς ἐπιστήσεται* περὶ τούτων.
> ΜΕΝ. Ἔοικεν. 85c9-d2

9.2.b. *Renderings of 85c9d2 by various scholars translate the crucial word* ***pollachei****, 85c11, in an unhelpful way as* ***"in many ways"*** *or* ***"in a number of different ways."***

Bluck (1961):

> At 85c Socrates said that if someone were to ask the slave *the same questions many times over and in many ways,* then whereas he at present possesses only true opinions, he would come to have *sure knowledge*. The idea seems to be that although the slave has *to some extent* grasped each step of the argument, he would need to go over the whole working many times and in many different ways before he grasped each step *fully* and before he grasped the relationship between the steps and appreciated the argument as a whole, so as to see for himself why the results reached must be correct. This is now to be understood as 'working out the cause', a procedure to be accomplished further 'recollection', and culminating in the attainment of inner conviction. p. 31

Nehamas (1985) writes:

> The slave, Socrates says, will come to have knowledge
>
> 'if one asks him *the very same questions [or: questions about the very same things] many times and in many ways*' (85c10-11)
>
> . . .
>
> Certainly, simply having *many doxai* about geometry cannot be itself sufficient for *episteme*. One must also acquire the ability to organize them systematically, to become able to move from one of them to another properly and on one's own, to know how they are supported by one another. . .. Having the answers to *as many questions as one pleases* does not constitute *episteme* unless one is capable of answering ever new questions as well as of formulating questions of one's own. p. 22–23

Scott (2005) renders pollachei "in many ways":

> At present, these beliefs have just been stirred up in him like in a dream. But if someone questions him *about the same matters on many occasions and in many ways*, you can be sure that he will end up with as exact a knowledge of them as anyone. men. So it seems. 85c9-d1 p. 105

and writes:

> Socrates makes the point by saying, 'if someone questions him about the same matters on many occasions and in many ways . . .'. Some translations render the point as 'if the same questions are put to him on many occasions and in different ways. .' [emphasis added]. This is possible given the Greek but makes little sense. Socrates can hardly be saying that, in order to convert your true opinions into knowledge, you need to be asked exactly the same questions, for you would most likely just continue to repeat the same answers.
>
> Rather, the point is that we should revisit *the same topic or problem but from different angles, and with increasing depth and complexity*. The reference to *different contexts* is very important: the more varied the contexts in which one places the original problem, the more interconnections one makes within one's set of beliefs. This ties in, of course, with the emphasis on synoptic reasoning that I have stressed at various points in this book. It will also be important when we come to examine Socrates' distinction between knowledge and true belief at 98a. p. 106–107

Sedley and Long (2011):

> Socrates: Yes, and now these opinions have only just been stirred up for him as if in a dream. But if someone goes on to question him *about these things on many occasions and in many ways*, you know that eventually, he will have knowledge about them as precise as anyone's. Meno: So it seems. p. 22

Landry (2012):

> At the moment, these beliefs have only just been stirred up in him and it all feels like a dream, but if he were to be *repeatedly asked the same questions in a number of different ways*, he'd certainly end up with knowledge of these matters that is as good and as accurate as anyone's. [85d], p. 163

Fine (2014) also renders "pollachei" as "in many ways":

> And indeed, upon being *questioned further*, the slave arrives at the right answer. Socrates says that the slave still doesn't know the right answer; he just has a true belief about it (85b–d). However, Socrates also says that, if the slave were *questioned further*, he could acquire the knowledge he lacks:
> and now these beliefs have just been stirred up in him as if in a dream. But if someone asks him *these same things many times and in many ways*, you know that eventually, he will know about these things no less accurately than anyone. (85c9–d1) p. 119
> At present, these beliefs have just been stirred up in him, like in a dream.
> But if someone questions him about *these same things often and in many ways*,
> you know that in the end, he will know about these things no less accurately than anyone. (85b8–d1). p.148

The standard rendering of 85c9-d1 appears to be well-described by the phrase:

> εἰ δὲ αὐτόν τις ἀνερήσεται *πολλάκις τὰ αὐτὰ ταῦτα καὶ* ***πολλαχῇ***, οἶσθ' ὅτι *τελευτῶν* οὐδενὸς ἧττον *ἀκριβῶς ἐπιστήσεται* περὶ τούτων.
> if questioned these same things many times ***in many ways***, the boy in the end will have obtained true Knowledge.

In the next section, I will propose that *"pollachei"* should be rendered as *"in many places."*

9.2.c. *Our rendering of the word* ***"pollache"*** *and our tentative anthyphairetic interpretation of 85c9-d2.*

The passage, thus rendered, does not appear convincing. As noted by several scholars, it is hard to see how questioning the boy the same things many times will result in knowledge. If the questions are the same, soon the answers will also be the same, and most likely short of knowledge. We believe that the trouble lies with the rendering of the word "*pollachei*" by "many ways" or by "different ways"; a better rendering of the crucial word "pollachei" is, as indeed suggested by LSJ, "*in many places*" or "on many grounds," and thus our rendering of 85c9-d2 is the following:

> *Socrates* And now those opinions have just been stirred up in him, like a dream;
> but if he were asked *these same questions (ta auta tauta) many times (pollakis) and in many places (pollachei)* you know that he will have in the end as exact an *knowledge* of them as anyone.
> *Meno* So it seems. 85c9-d2

Next, we examine the use of "*pollachei*" by Plato, starting with the following *Parmenides* passage:

> "Then unity is an attribute of *every part of the Being part* and is not wanting to *a smaller or larger or any other part*." "True."
> "Can the one be *in many places* (*pollachou,* 144d1) at once and still be a whole? Consider that question." "I am considering and I see that it is impossible."

> "Then it is divided into parts, if it is not a whole; for it cannot be attached to *all the parts of the Being part* at once unless it is divided." "I agree." 144c6-d4

In this passage, "pollachou" clearly refers to the infinite multitude of the parts of the Being part, generated by the anthyphairetic division of the dyad One, Being, as analyzed in Negrepontis (2024), and each part of the Being part is described as one of the many places. The occurrence of "pollachei" in the *Phaedrus*:

> *Socrates* Conceiving and bringing *together (sunoronta) in one idea scattered (diesparmena) in many places* (pollachei) *Phaedrus* 265d3-4

is exactly in the same sense; the passage describes the fundamental property, the collection of the many into one, of an intelligible Being, namely the equalization of the infinitely many parts dispersed in many places.

Returning to the *Meno* 85c9-d2, we recall that the Socrates-boy questions and answers have started with the anthyphairesis of the diameter a to the side b and have produced the anthyphairetic remainders parts c_1 and c_2, exactly as in the *Parmenides* and the *Phaedrus* passages. So the word "pollachei" refers to these parts, and it is for those that Plato is saying that there should be "the same questions many times about the many parts/places" in which the intelligible Being is being divided by anthyphairesis. Now we are told that if we ask the same questions about the places/parts generated by the anthyphairesisof a to b, then we will achieve complete knowledge, and in fact by Recollection. We can guess, considering the geometrical proof of Proposition 2.5, that *the same question asked many times* that Plato has in mind would be:

What is the ratio of part/place b to the next place c_1?
What is the ratio of part/place c_1 to the next place c_2?

Indeed, complete knowledge is obtained, as seen at the moment we realize that these ratios are the same, the ratio of b to c_1 is equal to the ratio c_1 to c_2, namely when we verify the Logos Criterion.

The same exact conclusion results from the study of the Divisions in the *Sophist*, and their interpretation as the philosophical analogues of geometric periodic anthyphairesis (in Sect. 10 below). In passage 85c9-d2 Plato provides the first hint, for the purpose of predisposing the reader that knowledge will involve a repetition and a recollection, in fact of a logos.

9.3. *The anthyphairetic interpretation is in fact suggestive of Recollection; in fact, asking the same questions many times* ***in many places*** *is described not only as Knowledge but more specifically as Knowledge using Recollection 85d3-86b4*

> *Socrates* Without anyone having taught him, and only through questions put to him, he will acquire *knowledge (epistemen)*, *recovering the knowledge (epistesetai)* out of himself?
> *Meno* Yes.

Socrates And is not this recovery of *knowledge (epistemen)*, in himself and by himself, *recollection (anamimneskestai)*?
Meno Certainly.
Socrates And must he not have either once acquired or always had the *knowledge (epistemen)* he now has?
Meno Yes.
Socrates Now if he always had it, he was always in a state of *knowing (epistemon)*; and if he acquired it all sometimes, he could not have acquired it in this life. Or has someone taught him *geometry* (*geometrein*)? You see, *he can do the same as this with all geometry and every branch of knowledge* (*houtos gar poiesei peri pases geometrias tauta tauta kai ton allon mathematon apanton,* 85e1-3). Now, can anyone have taught him all this? You ought surely to know, especially as he was born and bred in your house.
Meno Well, I know that no one has ever taught him.
Socrates And has he *these opinions (tautas tas doxas)*, or has he not?
Meno He must have them, Socrates, evidently.
Socrates And if he did not acquire them in this present life, is it not obvious at once that he had them and learned them during some other time?
Meno Apparently.
Socrates And this must have been the time when he was not a human being?
Meno Yes.
Socrates So if in both of these periods—when he was and was not a human being—he has had *true opinions (aletheis doxai)* in him which have only to be *awakened (epegertheisai)* by questioning to become *knowledge (epistemai)*, his soul must have had this cognizance throughout all time? For clearly he has always either been or not been a human being.
Meno Evidently.
Socrates And if *the truth of all true Beings* is always in our soul, then *the soul must be immortal*; so that you should take heart and, whatever you do not happen *to know (epistamenos)* at present—that is, what you do not *remember (memnemenos)*—you must endeavor to search out and *recollect (anamimneskestai)*? 85d3-86b4

In case our intuition is correct and Knowledge is precisely the knowledge of the full infinite anthyphairesis of the diameter to the side of a square using the Logos Criterion, as we have indicated in the heuristic Sect. 6, then we must have no difficulty describing this process as Recollection. Indeed, Knowledge is obtained at the precise moment that we realize that a previous logos (of successive parts in the anthyphairesis) is equal to a logos (of two successive parts of the anthyphairesis at some later stage). This repetition can be validly and interestingly be somewhat poetically described as Recollection. As we shall see later, Recollection, according to both Plato and Proclus, is specifically the Recollection of Logoi.

9.4. *Despite Plato's broad hint with "pollachei," the method of turning True Opinions into Knowledge and Recollection still remains a mystery to Meno and to most readers of Plato, whose resolution needs more explicit help 86b5*

Meno What you say is for me a good Opinion (eu moi dokeis), Socrates, but I know not how**.** 86b5

So Meno realizes that he is still in the state of Opinion, and a clearer description of the process of acquiring Knowledge by Recollection is necessary.

10 *For the first time, in the* Meno *86b6-c3, Plato explicitly brings forth "logos" as the fundamental ingredient of Knowledge, in line with a corresponding role for "logos" in the* Republic 534b3-6, Theaetetus 202c2-5, Sophist 260a5-b2, Statesman 286a4-7, *and elsewhere in Plato. Platonic Scholars render "logos" by terms such as "account," "argument," "discourse," "rational definition," "rational explanation," "reason," and "word"(cf. Sections 10.2 and 10.3). These renderings in our opinion are not in the correct direction and obscure the true meaning of "logos." Nehamas (1984) suggested that the meaning of Logos must come from a study of the Divisions in the* Sophist *and the* Statesman. *This is exactly our approach in Sections 11 and 12*

10.1. Meno *86b6-c3 on Logos*

We have noted that knowledge has not been acquired at 85b7. It is further clear that even at 86b5, it has not been explained how true opinions turn into recollection and knowledge:

> [*Socrates*] Most of the points I have made in support of ***Logos*** (*huper tou logou*, 86b7) are not such as I can confidently assert; but that the belief in the duty of inquiring after what we do not know will make us better and braver and less helpless than the notion that there is not even a possibility of discovering what we do not know, nor any duty of inquiring after it—this is a point for which I am determined to do battle, so far as I am able, both using ***Logos*** (*kai logoi*, 86c2) and work.
>
> *Meno* There also I consider that *you provide Logos* (legein) well, Socrates. 86b5-c3

We note that in the W.R.M. Lamb classical translation of the passage, the first occurrence 86b7 of *Logos* is rendered as *"argument,"* while the second 86c2 as *"word."*

10.2. *The renderings of the crucial term "logos" in 86b7 and 86c2 by Platonic scholars "argument" and "word," respectively.*

R. E. Allen (1959):

> There are other things about the *argument* that I would not confidently affirm, but that we shall be better men, more courageous and less idle, if we think we ought to inquire into what we do not know, instead of thinking that because we cannot find what we do not know we ought not to seek it – that I would do battle for, so far as possible, in *word* and deed. (86b6-c2, in Halper 2007, p. 239)

J. Klein (1965):

> For he goes on *asserting*, forcefully and without any ambiguity, that in all that was said there was only one point (81 d 5- e 1; 86 b 2 - 4) for which he would fight strenuously, *in word* and deed (kai logdi kai ergoi), whenever he was able to do so. p.182

D. Scott (2005):

As for the other points, I wouldn't absolutely insist on *the argument.*
But I would fight, both in *word* and deed, for the following point:
that we would be better, more manly, and less lazy if we believed that we ought to inquire into what we do not know, than if we believed that we cannot discover what we do not know and so have no duty to inquire. (86b6–c2) p.121.
Logos is rendered as argument and word.

Waterfield (2005):

Socrates: Yes, I think so too, Meno. I wouldn't support every aspect of *the argument* with particular vigour, but there's one proposition that I'd defend to the death, if I could, by *argument and by action*: that as long as we think we should search for what we don't know we'll be better people– less fainthearted and less lazy– than if we were to think that we had no c chance of discovering what we don't know and that there's no point in even searching for it. p. 124
Logos is rendered as argument.

Ebert (2007):

So I seem to myself, Meno. All the other points I have made in support of *the argument* are not such as I can confidently assert; but that, if we are convinced we should inquire after what we do not know, we should get better and braver and less lazy than if we believe that we are neither able nor obliged to inquire after things we do not know – this is something for which I am determined to fight, so far as I am able, both in *word and deed*. (86b6-c2) p. 197

Sedley and Long (2011):

Socrates: I think so too, Meno. In defense of *the argument* I would not affirm the other points very strongly, but that we would be better, more manly, and less lazy by believing that one should search for what one doesn't know than if we believed that we cannot discover what we 86c do not know and should not even search for it – that is something over which I would fiercely contend, if I were able, in both *word and deed*. p. 23

Fine (2014):

As far as the other points (ta alla) are concerned, I wouldn't altogether take a stand on *the argument*. But that we will be better and more manly and less idle if we think one should inquire into that which one doesn't know than if we think it isn't possible to discover what we don't know and that we don't need to inquire into it: that is something I would certainly fight for to the end, if I could, in both *word and deed*. (85d12–86c2) **p. 150**

Benson (2015):

I would not confidently assert the other things said in defense of this *account,* but that we would be better and braver and less idle if we believe that one ought to inquire concerning those things he fails to know than if one believes it is not possible to discover nor necessary to inquire concerning those things one fails to know (εἰ οἰοίμεθα ἃ μὴ ἐπιστάμεθα μηδὲ δυνατὸν εἶναι εὑρεῖν μηδὲ δεῖν ζητεῖν), I would fight for in both ***word*** *and deed* as far as I am able. (Meno 86B6–C2; emphasis added) p. 62

10.3. *At this point, it is necessary to realize that in several other passages in Plato, where "logos" is described as an absolutely fundamental ingredient of Knowledge, the renderings of the term,* as *"account," "discourse," "rational definition," "rational explanation," reasonby various scholars of Plato, seriously obscure its real meaning.*

Republic *534b3-6:*
And do you not also give the name *dialectician (dialektikon)* to the man who is able to exact *an account (logon)* of the true Being of each thing? And will you not say that the one who is unable to do this, in so far as he is incapable of rendering *Logos (logon)* to himself and others, does not possess intelligence[2] about the matter?

Note. *Logos* is rendered in the classical translation by P. Shorey as *account.*

Theaetetus *202c2-5:*
for he who cannot give and receive *Logos* (logon) of a thing is without knowledge (anepistemona) of it; but when he has acquired Logos (logon) he may possibly have become all that I have said and may now be perfect in *knowledge (epistemen).*

Note. *Logos* is rendered in the classical translation by H.N. Fowler as *rational explanation*.

Sophist *260a5-b2. Logos is necessary for the Knowledge of intelligible Beings*
Stranger Our object was to establish that *the Logos (logon)* is one of the kinds of the Beings (ton onton). For if we were deprived of *this [Logos]*, we would be deprived of philosophy, which would be the greatest calamity; moreover, we must at the present come to an agreement (diomologesasthai) about the nature (ti pot' estin) of *Logos (λόγου)*, and if we were robbed of *it [Logos]*, there would be no true Being (med' einai) and we could no longer be able to provide *Logos (legein), and* we should be robbed of *it [Logos]*, if we agreed that there is *no mixture (meixin)* of anything with anything.

Note. The classical translation by H.N. Fowler renders all six references to "Logos" as "discourse."

Statesman *286a4-7. Logos necessary for the Knowledge of Intelligible Beings*
We must therefore study so as to have the power (dunaton einai) *to give and receive (dounai kai dexasthai) the Logos* (logon) of each of them; for the immaterial *[intelligible] Beings*, which are the noblest and greatest, can be made known (saphos deiknutai) *by Logos (logoi) only*, and it is for their *[the intelligibles']* sake (heneka) that all *Logoi* are provided *(legomena).*

Note. The classical translation by H.N. Fowler renders "logon" as "rational definition," logoi as "reason," and "legomena" as "all we are saying."

10.4. *Interestingly*, Nehamas (1984) *suggested that the meaning of Logos in Plato is to be found in the divisions of the* Sophist *and* the Statesman. *This is exactly our approach in Sections 11 and 12.*

Nehamas (1984) expresses the opinion that the meaning that Plato has in mind with Logos must come from a study of the Divisions in the *Sophist* and the *Statesman.*

The answer [to the meaning of Logos in Plato] can only be given using the long and complicated divisions through which the *Sophist* and the *Politicus* attempt to define their subject matter. Both dialogues try to define a species of art (techne) or science (episteme), that of the sophist and that of the statesman (cf. Pol, 258b2-7). In both dialogues, the art in question is defined by being located within a determinate network of other arts and sciences. And the process by means of which the definition is reached provides, at the same time, both an explanation of the aporetic ending of the *Theaetetus* and an Illustration of the obscure

opening of the *Philebus*. Plato, therefore, does try to answer the question of *Theaetetus*. But his answer is not cryptically contained within the dialogue itself, either negatively (äs Cornford argued) or positively (äs Fine suggested). His answer is given in the two dialogues that follow the *Theaetetus*. p. 36

11 *The revealing anthyphairetic interpretation of Logos in the Divisions of the Angler, the Noble Sophistry, and the Sophist in the* Sophist *as the philosophic analogue of Theaetetus' Logos Criterion establishing anthyphairetic periodicity*

11.1. *Plato considers that the clearest demonstration of the nature of Logos is provided in the last division of the Sophist in the* Sophist *264b-268d*

Towards the end of the *Sophist*, after having given a considerable number of hints and suggestions, Plato is not yet sure that he has made his point clear, and Theaetetus is still not convinced of the crucial importance of "logos" for achieving knowledge of intelligible Beings:

Theaetetus That is true enough, but I do not understand why we must come to an agreement about [the crucial importance of] *logos* [H. N. Fowler renders *logon* by *discourse*] just now. *Stranger* Perhaps the easiest way for you to understand is by what follows. 260b3-5

The final step is the demonstration of the power of *logo*s with the definition of the Sophist using a dyad Being not-Being, as noted explicitly in 258eb-7, possessing *logos*, in the *Sophist* 234e-236d & 264b-268d at the very end of the dialogue.

11.2. *The anthyphairetic nature of Logos is revealed in the* Sophist and Statesman *Divisions*

Indeed, the best way to come to an essential understanding of what Plato means with Logos is to study the Divisions in the *Sophist* and in the *Statesman*. The meaning is revealed to be anthyphairetic in S. Negrepontis (1999, 2000a, 2000b, 2012, 2018), p.350–367; (2019), p. 28; and Negrepontis et al. (2023). For completeness, we will include an analysis of the *Sophist* divisions, in the present Section: on the Angler in the *Sophist* 218b-221c (Sect. 11.5), on the Noble Sophistry (Sect. 11.6), and on the Sophist (Sect. 11.7). The anthyphairetic interpretation of the Logos in the *Sophist* Divisions will play a crucial role in our interpretation of the *Meno* passages in the next Sect. 12.

11.3. *According to* Sophist 218b5-c5, 218e2-219a3 **intelligible** Knowledge (episteme) is not only **Name** but also **Logos**

Stranger Very well (Εὖ λέγεις); you will follow your own devices about that as *Logos (tou logou) proceeds*; but now you and I must investigate in common, beginning first, as it seems to me, with *the sophist,* and must search out and make plain *by means of logos (logoi)* what he is *(ti pot' esti)).* For as yet you and I have in common about him [the sophist] *only the name (tounoma monon)*, but as to the Being itself to which we give *the name*, we may perhaps each have a conception of it in our own minds; however, we ought always in every

instance to come to an agreement about it *by means of logoi* rather than about *the name only (tounoma monon) without (choris) Logos (λόγου)*. 218b5-c5

Stranger Well, then, what example can we set before us which is well-known and small, but no less capable of *Logos (logon) than any* of the greater things? Say *an angler (aspalieutes)*; is he not known to all and unworthy of any great interest?

Theaetetus Yes.

Stranger But I hope he offers us *a method (methodon)* and is capable of a *Logos (logon)* not unsuitable to our purpose.

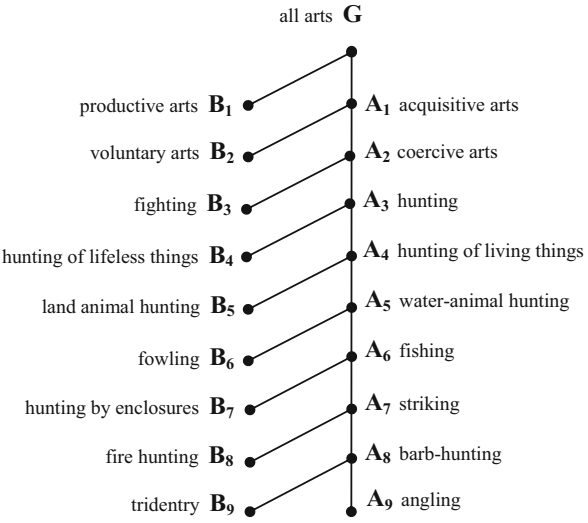
Theaetetus That would be good. 218e2-219a3

Note. "Logos" is rendered as "rational explanation," "argument," and "definition" by the same translators.

11.4. *Name and Logos of the intelligible Being Angler in the* Sophist 218b-221c.

11.4.a. *The Division of the Angler*

The method of "Name and Logos" (cf. *Theaetetus* 201e2-202b5, *Sophist* 218c1-5, 221a7-b2, 268c5-d5) is exemplified at the start of the *Sophist* 218b-221c by the definition of the Being Angler. The Angler is certainly a lowly paradigm of an intelligible Being, but even for lowly entities, there is an intelligible Being, cf. *Parmenides* 130b7-e3. In the scheme below, we reproduce the binary division process which leads to the Angler.



The Division of the Angler (*Sophist* 218b-221c)

11.4.b. *Logos Criterion of the Angler*

The description of the Logos Criterion of the Angler is contained in the *Sophist* 220e2-221c3 passage (which we have divided into two parts [A] and [B] for the sake of convenience):

[A] '*Stranger* Then of striking which belongs to barb-hunting [A8],
that part which proceeds *downward from above* ('anothen eis to kato'),

is called, because tridents are chiefly used in it, tridentry [B9],
I suppose. . ..
Stranger The kind that is characterized by the opposite sort of blow,
which is practiced with a hook and strikes, ...
and proceeds *from below upwards* ('katothen eis tounantion ano'),
being pulled up by twigs and rods.
By what name, Theaetetus, shall we say this ought to be called?
Theaetetus I think our search is now ended and we have found the very thing we set before us a while ago as necessary to find.
Stranger: Now, then, you and I are only agreed about the name of angling, [A9]'
220e2-221b1
[B] 'but *we have acquired also a satisfactory 'Logos'* of the thing itself.
For (gar)
of art as a whole, half was acquisitive,
and of the acquisitive, half was coercive,
and of the coercive, half was hunting,
and of hunting, half was animal hunting,
and of animal hunting, half was water hunting,
and of water hunting [A5]
the whole part from below (to katothen tmema holon) was fishing, [A6]
and of fishing, half was striking,
and of striking half was barb-hunting, [A8]
and of this [A8]
the part in which the blow is pulled
from below upwards *(to peri ten katothen ano)* was angling. [A9] (221b1-c3).
[Translation by H. N. Fowler (1921), with modifications by the author.]

In [A], the opposing relation of Tridentry to Angling is carefully explained: all Fishing with a hook is divided into

Tridentry (=Fishing with a trident), which is described as
Fishing with a hook with an art that proceeds **from above downwards**, and
Angling (=Fishing with a rod), which is described as
Fishing with a hook with an art that proceeds **from below upwards**
We have pushed Division all the way to the Angling; thus, we have certainly found 'the name' of Angling.

But now, in [B], it is claimed that "the Logos" of the Angling has been found, too.

The justification – the proof – that we have indeed found the Logos, too, is contained in the remainder [B], since this remaining part [B] starts with a "for" ("gar"), and this justification can be seen to consist of:

(i) An accurate recounting of all the division steps, abbreviated in the sense that of the two species into which each genus is divided, only the one that contains the Angler is mentioned, while the species opposite to it is omitted.
(ii) A reminder that the last species, the angling, is characterized as the part of its genus that proceeds "from below upwards."
(iii) The ONLY new information (since (i) and (ii) are repetitions of things already contained in the Division and in [A]), which concerns the species of fishing, three steps before angling, and which informs us for the first time that this species is "the whole part from below" of its genus.

Since this is an abbreviated account, there is no explicit information on the species opposite to "fishing" – namely "fowling," but since "fishing" was described not simply as "the part from below" of its genus, but emphatically as "the whole part from below," it follows that the opposite species – "fowling" – must be characterized as "the (whole) part from above" of the same genus. In fact, there can be no other justification for the presence of the term "whole" in the description of "fishing" with a view to arguing that we have obtained "Logos," except to indicate and imply this description for its opposite species, "fowling."

We recall that the part [B] from the word "for" ("gar") is explicitly a justification of the claim that we have succeeded in finding the "Logos" of the Angling. We may then ask: what is the "Logos" of the Angler that reasonably results from such a justification? There can really be only one answer: the "Logos" we are looking for is the equality of the "philosophic ratio" of Tridentry to Angling, namely the equality of the ratio "of from above downwards to from below upwards," to the ratio "of Fowling to Fishing."

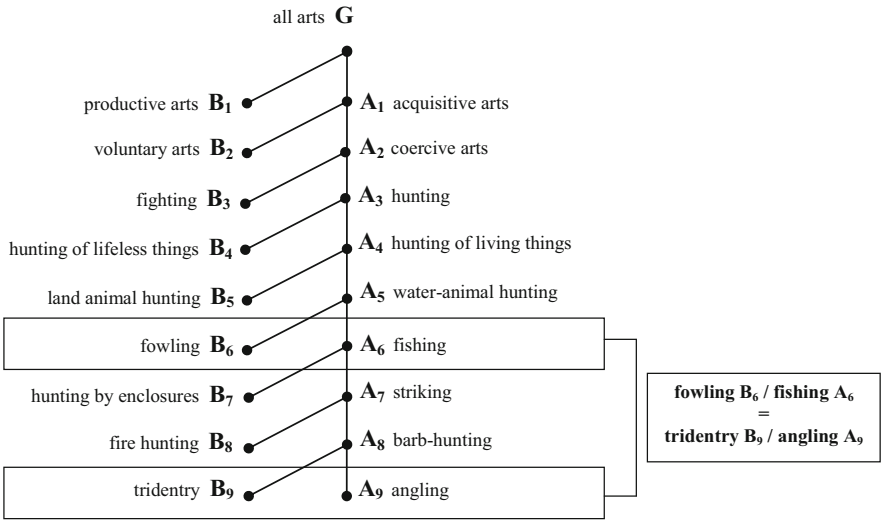
Thus,

fowling B6/fishing A6 = tridentry B9/angling A9 =
from above downwards/from below upwards.

Since the species Tridentry and Angling form a pair of opposite species, and the species Fowling and Fishing form another pair of opposite species in the Division Scheme for the Angler, the resulting "Logos" bears a most uncanny similarity to the Logos Criterion for the periodicity of the anthyphairesis of geometric magnitudes and of geometric powers in particular.

11.4.c. *The Name and Logos of the Angler.*

The Name and Logos of the Angler thus takes the following form:



Thus, the Name and Logos of the Angler consists of the anthyphairetic division together with the Logos, analogous to the Logos Criterion for periodicity for geometric anthyphairesis.

11.5.*The Interpretation of Logos in the Definition of the* ***Noble Sophistry*** *by* ***Name and Logos*** *in the* Sophist 226c-231b as the philosophic analogue of the Logos Criterion establishing periodic anthyphairesis.

We outline the Name and Logos of the intelligible Being *Noble Sophistry* in the *Sophist*.

G The art of discrimination/separation (diakritike), 226c	
a1 Separation of like from like, 226d	b1 Separation of worse from better, 226d=purification (kathartikon), 226d

b1 Purification (kathartikon)	
a2 Purification of the body, 227c	b2 Purification of the soul, 227c

b2 Purification of the soul	
a3 Purification from wickedness, 228d=justice (kolastike) 229a Rougher purification of the soul (trachuteron)	b3 Purification from ignorance, 228d=instruction (didaskalike) 229a Smoother purification of the soul (leioteron)

b3 Instruction (didaskalike)	
a4 Instruction in handicraft, 229d	b4 Instruction in logoi, 229d =education (paideia) 229d

[Question whether education is “already indivisible” or still admits of division 229d]

b4 Education (paideia)	
a5 Admonition (nouthetike) 230a Rougher education (trachuteron),229e	b5 Noble sophistry, dialectic 231b Smoother education (leioteron) 229e

Logos Criterion for anthyphairetic periodicity:

a3/b3 = a5/b5 = rougher/smoother

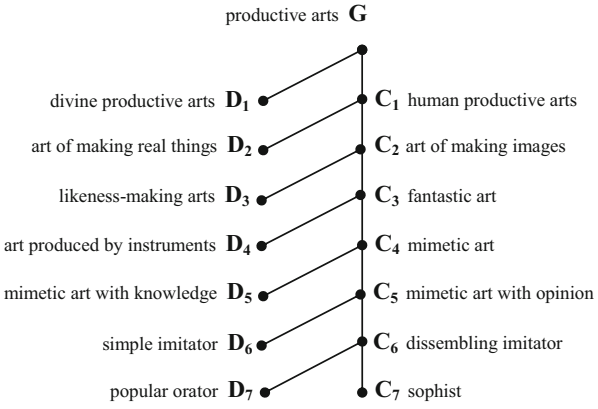
The definition of the Noble Sophistry by Name and Logos exhibits the same structure of anthyphairetic periodicity as the Angler.

11.6. *The Name and Logos of the Intelligible Being* ***Sophist*** *in the* Sophist *234e-236d and 264b-268d*

It appears that the Name and Logos of an intelligible Being is very much like the anthyphairetic Division and the Logos Criterion of a geometric "power." When Socrates expressed, in the *Theaetetus* 45c7-148e5, his exhortation to imitate the geometric situation, it seems that he meant a much closer imitation than anybody has thus far suspected! But before proceeding to wide-ranging conclusions, it would be prudent to examine whether the Name and Logos of the Sophist, in the *Sophist* 264b-268d, has the same kind of structure and, in particular, whether there is a similar type of "Logos." So we shall take a look at the Name and Logos of the Sophist.

11.6.a. *The Division of the Sophist*

The Division for the Sophist follows the same pattern as the Division of the Angler, starting with a Genus, in this case "all the productive arts," proceeding by binary division of each Genus to two species, where the next Genus is that species of the previous step in the Division that contains the entity to be defined, in this case, the Sophist, and ending with the division step that produces the Sophist as a species. The whole division scheme is as follows:



Division of the Sophist (*Sophist* 264b-268d)

11.6.b. *The fundamental analogy in the Divided Line of the Republic 509d-510b and the Logos Criterion of the* Sophist

We will make essential use of the fundamental analogy in the Divided Line of the *Republic* 509d-510b:

the opinable/the knowable = the likeness/that of which it is a likeness.

A strong indication that Plato's intention was to use the *Republic*'s ratio of the Divided line is the extra care he takes to justify the equality of the *Republic* ratio, namely the ratio of all likenesses-images to all real things, all meaning both divine and human, to the ratio in the Sophist, namely the ratio of human likenesses-images (C2) to human real things (D2). He does this meticulously in *Sophist* 265e8-266d7,

in which there is a division not only of C1 into D2 and C2 but also of D1 into the arts producing real things (say D2') and into the arts producing images (C2'), a division of a "rejected" species, something unexpected and unique in Plato's divisions. The proof of equality, involving "breadth" and "width," employs an analogue of Proposition VI.1 of Euclid's *Elements*.

The Logos Criterion for the Sophist is now immediately recognized:

Proposition.

The ratio of the second step in the Division of the Sophist, namely the ratio of the art of making *real things* **D2** to the art of making *images* **C2**	is equal to	the ratio of the fifth step in the Division of the Sophist, namely the ratio of the mimetic art *with knowledge* **D5** to the mimetic art *with opinion* **C5**

11.6.c. *The next two Logoi of the Name and Logos of the Sophist are analogous to the geometrical ratios of periodic anthyphairesis.*

The definition of the Sophist presents an additional feature that ties it even closer to the mathematical model: there are two Logoi after the Logos Criterion, and if the mathematical anthyphairetic model is indeed followed, we would expect to have two further equalities of Logoi: namely the logos of the third step should be equal to the ratio of the sixth division step, and the logos of the fourth division step should be equal to the ratio of the seventh and final division step.

(c1) *The equality of the third and sixth ratios of the Division of the Sophist*. In the third division step, the Genus "humanly produced images" is divided into two species,

> *Stranger* I see the likeness-making art (eikastiken) as one part of [the image-making art]. This is met with, as a rule, whenever anyone produces the imitation by following the proportions of the original in length, breadth, and depth, and giving, besides, the appropriate colors to each part. 235d6-e2
> *Stranger* Now then, what shall we call that which appears, because it is seen from an unfavorable position, to be like the beautiful, but which would not even be likely to resemble that which it claims to be like if a person were able to see such large works adequately? Shall we not call it, since it appears, but is not like, an appearance?
> *Theaetetus* Certainly. 236b
> Stranger And to the art which produces appearance, but not likeness (phantasma all' ouk eikona), the most correct name we could give would be "fantastic art," (phantastiken) would it not?
> *Theaetetus* By all means.
> *Stranger* These, then, are the two forms of the image-making art (eidolopoiikes) that I meant, the likeness-making (eikastikes) and the fantastic (phantastikes). 236c3-7
> *Stranger* We must remember that there were to be two parts of the image-making class (eidolourgikes), the likeness-making (eikastikon) and the fantastic (phantastikon). 266d8-9

Thus, the likeness-making art is an image that does not distort the original object but rather expresses a *faithful* likeness of that object, while in the sixth division step, described in 267e7-268a8, the Genus opinionated-imitator is divided into the two

species, "the simple" (haploun) and "the dissembling" (eironikon) opinionated-imitator, the simple imitators being those who

> think they know that about which they have only opinion,

while the "dissembling" are those who

> because of their experience in the rough and tumble of arguments, strongly suspect and fear that they are ignorant of the things which they pretend before the public to know.

Thus, the simple imitators *do not distort* their opinion but rather express a *faithful* likeness of their opinion, while the dissemblers *distort and disguise* their opinion behind a false appearance.

Therefore, we have

Proposition.

The ratio of the third step in the Division of the Sophist, namely the ratio of *the likeness-making arts* D3 to *the fantastic arts* C3	is equal to	the ratio of the sixth step in the Division of the Sophist, namely the ratio of *the simple imitator* D6 to *the dissembling imitator* C6

(c2) *The equality of the fourth and seventh ratios*. In the fourth division step, described in 267a1-b3, the Genus, fantastic arts, is divided into two species, as follows:

> *Stranger* Let us, then, again bisect the fantastic art.
> *Theaetetus* How?
> *Stranger* One kind is that produced by instruments, the other that in which the producer of the appearance offers himself as the instrument.
> *Theaetetus* What do you mean?
> *Stranger* When anyone, by employing his own person as his instrument, makes his own figure or voice seem similar to yours, that kind of fantastic art is called mimetic.
> 267a1-b3

Thus, in mimetic art, the instrument of imitation is the imitator himself, while in the nameless opposite art, the instrument of imitation is other than the imitator.

In the seventh division step, described in 268a9-c4, the dissembler is divided into the demagogue and the sophist, the demagogue being

> one who can dissemble in long speeches in public before a multitude,

while the sophist is someone

> who does it in private in short speeches and forces the person who converses with him to contradict himself.

Thus, he who listens to a dissembler is deceived, if that dissembler is a demagogue, and contradicted not by himself but by another instrument of deceit (namely the demagogue himself), while if the dissembler is a sophist, the listener is forced by the sophist to himself become the instrument of deceit.

Thus, we have:

Proposition.

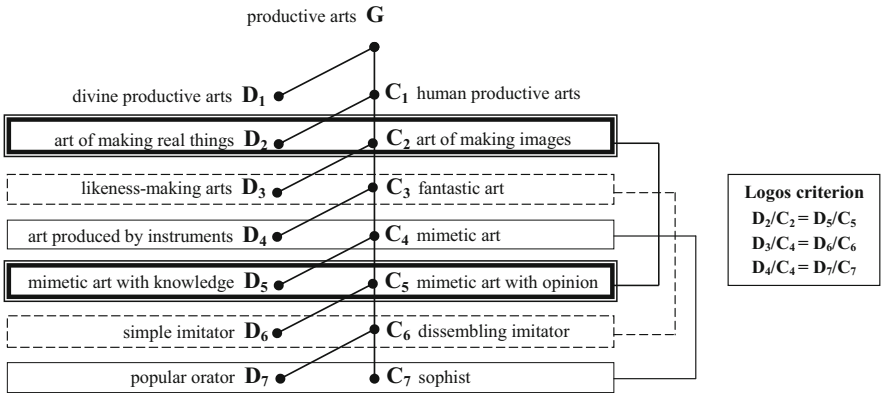
The ratio of the fourth step in the Division of the Sophist, namely the ratio of the imitator who uses *not himself, but other* instruments of imitation D4 to the imitator who is *himself* the instrument of imitation C4	is equal to	the ratio of the seventh step in the Division of the Sophist, namely the ratio of the demagogue, whose listener is contradicted and deceived *not by himself but by other* D7 to the sophist, whose listener is contradicted and deceived *by himself* C7

Propositions (C1) and (C2) provide powerful additional evidence in favor of the anthyphairetic interpretation of Division and Collection.

11.6.d. *The Name and logos of the Sophist*

The complete Name and Logos of the Sophist (*Sophist* 264b-268d) can thus be summarized in the following scheme:

The definition of the Sophist by Name & Logos
as the philosophic analogue of periodic anthyphairesis 264b-268d, 258b5-7



Thus, we have full confirmation for the anthyphairetic meaning of Logos that we proposed in the Name and Logos of the Angler. This time, the Logos is not of the simple-minded, obvious type that appeared in the definition of the Angler (from above downwards/from below upwards), but of a more sophisticated and philosophical kind, playing a central role in the dialectics of the *Republic.*

We note that the correlation of the *Republic* ratio in the divided line with the Name and Logos of the Sophist in the *Sophist* is hinted at by Proclus in his *Commentary to the Republic* 1290,7-10.

12 *The Logos/Logismos/analogio as Bond in* Meno *97d4-98a4,* Sophist *251d8-9, 253a4-c1, 268c5-d5, and* Timaeus *31c2-32a7*

Bond is the analogy, the proportion with three terms Meno.

Knowledge of the diameter a to the side b of a square is the knowledge of the infinite anthyphairesis Anth(a, b) = [1, period(2)].

12.1. *The True Opinion is transformed into Knowledge, with the addition of* ***"causal logismos" (logisms aitias)***, acting as ***bond (desmos)*** *on True Opinion* Meno 97d4-98a4.

Socrates Well, do you know why it is that you wonder, or shall I tell you?
Meno Please tell me.
Socrates It is because you have not observed with attention *the sculptures of Daedalus.* But perhaps there are none in your country.
Meno What is the point of your remark?
Socrates That if they are *not fastened* up *(me dedemena, 97d9)* they play truant and run away; but, if *fastened (dedemena, 97d10),* they stay (paramenei) where they are.
Meno Well, what of that?
Socrates To possess one of his works which is let loose (lelumenon) does not count for much in value; it will not stay with you any more than a runaway slave: but when *fastened up (dedemenon, 97e4)* it is worth a great deal, for his productions are very fine things.
And to what am I referring in all this? To *the true opinions (tas doxas tas alretheis).* For these, so long as *the true opinions (hai doxai hai aletheis)* stay with us, are a fine possession, and affect all that is good; but they do not care to stay for long, *and run away out* of the human soul, and thus are of no great value until one *binds (desei, 98a3)* them *using causal "logismos"* (*desei logismoi aitias).* 97d4-98a4

The True Opinion becomes knowledge if provided with a bond, a tie. This bond is described as *logismos aitias, causal logismos*, a crucial but as yet unclear description, related, however, to the *Logos* in the earlier *Meno* passage 86b5-c3. It needs to be carefully interpreted.

12.2. *Knowledge, just obtained as True Opinion plus the* ***bond Logismos/Logos*** *(section 12.1), is Recollection.* Meno 98a4-b6

And *this process (touto)*, friend Meno, is *recollection (anamnesis)*,
as in our previous talk, we have agreed.
But when once they are *fastened (dethosin, 98a5),*
in the first place they turn into *knowledge*, and in the second, are *constant (monimoi).*
And this is why *knowledge* (episteme) is more prized than *right opinion (orthes doxes):*
Knowledge transcends *right opinion* by the *bond (desmoi,98a8).*
Meno Upon my word, Socrates, it seems to be very much as you say.
Socrates And, indeed I set *logos (lego)* as one who does not know but only conjectures:
yet that there is a difference between right opinion (orthe doxa) and knowledge (episteme) is not at all a conjecture with me but something I would particularly assert that I knew: there are not many things of which I would say that,
but this one, at any rate, I will include among those that I know.
Meno Yes, and you are right, Socrates, in setting "logos" (legeis). 98a4-b6

Thus Knowledge and Recollection are acquired from True Opinion by the addition of the bond, an entity described as *causal logismos*, previously as *logos*. In the two passages in the *Meno* 97d-98b, the role of the *logismos/logos* is repeatedly described as that of a *bond.* We will look carefully into the precise meaning of bond.

12.3. *Interpretation of "logismos aitias" Logismos aitias 98a4 by earlier scholars*
J. Klein (1965):

All that is applicable to "true opinions." For they, too, are a possession to be highly treasured as long as they stay put: they make for all that is good and beneficent (panta tagatha ergazontai). Unfortunately, they by themselves do not want to stay for any length of time but run out of the human soul. Thus they are not worth very much until one "binds" them. This binding consists in finding *reasons for them in one's own thinking (aitias logismoi).*

Nehamas (1985):

True beliefs, he claims, like Daedalus' statues, are always escaping from the soul. But when they are bound down by *an 'account of the explanation'*, which, 'as we earlier agreed is recollection (98a3-5), they are transformed into episteme and become permanent. p.7

The former, he says, 'is worth little until it is tied down by reasoning about the explanation'(98a3-4). And it is this (touto, 98a4), he continues, that, 'as we agreed earlier, is recollection' (98a4-5). But the *aitias logismos*, as far as I can see, corresponds to nothing at the first stages of the slave's interrogation. The only process to which it can be connected is the repeated questioning that will eventually lead to the recovery of *episteme* (85c-d) p. 22

What brings about the *aitias logismos* and transforms *doxa* into *episteme* is not a new operation, additional to the eliciting of true *doxa* but rather the eliciting of enough true doxa about the subject to make having them *constitute* the *aitias logismos*. The very same true beliefs the slave now has, Socrates claims at 86a7-8, having been aroused by questioning, become knowledge (*epistemai*). . . . Plato does not explain how this transformation is to occur, and it is very difficult to know what is involved in the transition. p. 22–23.

The *aitias logismos* and recollection, strictly speaking, consists in this ability, which transcends merely having answers to different questions. . .

Waterfield (2005):

What am I getting at? I mean this to be an analogy for true beliefs. As long as they stay put, true beliefs too constitute a thing of beauty and do nothing but good. The problem is that they tend not to stay for long; they escape from the human soul and this reduces their value unless they're anchored by *working out the reason*. p. 139

Ebert (2007):

So true belief has to be bound by the "reckoning of the reason" (aitia logismou) (98a3-4). And Socrates continues: "This is, Meno my friend, recollection as was agreed between us earlier on." (98a4-5). Now this certainly does not support the idea that what we learn whenever we learn something comes to us by recollection. The passage Socrates seems to refer to here is 85c9-d1; there it has been agreed that a true belief like the one now engendered in the slave boy can be turned into knowledge by further questioning, what one has to add is the fact that by this process of questioning the reason for the geometrical theorem will be found. p. 198

Sedley and Long (2011)

> For true opinions are also a thing of beauty, as long as they stay with one, and all their consequences are good. But they're not prepared to stay with one for long. Instead, they run away from the person's soul. As a result, they are not worth very much until someone ties them down *by reasoning out the cause*. p. 37

Fine (2021):

> Another reason to think that Plato is defining knowledge is that, in the *Theaetetus*,
> 'knowledge is true belief plus a logos'
> is suggested as an answer to the question 'What is knowledge?'
> One might, to be sure, argue that
> 'knowledge is true belief plus a logos'
> is not equivalent to
> 'knowledge is true belief tied down with an aitias logismos.'
> Be that as it may, the two claims are similar enough in shape or form for it to be reasonable to think that, since the former is explicitly offered as an answer to the question 'What is knowledge?' the latter is also intended to answer it. p. 74

Roughly speaking, Fine accepts that *logos* and *logismos* in the *Meno* have essentially the same meaning. Thus, in this passage, Plato comes very close to the statement that Knowledge of the diameter to the side of a square may be achieved by true opinion plus Logos.

> Although our discussion so far leaves many questions unanswered, it at least suggests that a good translation of *aitias logismos* is: '*reasoning about the explanation*.' Plato's claim is then that to know that p is (a) to have the true belief that p and (b) for that true belief to be bound by reasoning about the aitia of p. That is, for one's true belief that p to constitute knowledge that p, one must bind it properly by being able to explain why p is so. p. 79

Fine renders *logismos aitias* as *reasoning (logismos) about the explanation (aitia).*

The phrase has been translated in many ways. Indeed, some authors (including myself in 'Inquiry in the Meno') use more than one translation, without comment. Among the translations I have seen are the following:

> 'by a chain of *causal reasoning*' (Thompson, *The Meno of Plato*);
> 'by *calculation of cause*' (Bluck, *Plato's Meno*; Sedley, '*Three Platonist Interpretations of the Theaetetus*,' has '*by calculation of the cause*');
> 'the *calculation of the reason*' (Vlastos, '*Anamnesis*');
> '*working out the explanation*' (Burnyeat, '*Socrates and the Jury*'; R. W. Sharples, *Plato: Meno* (Warminster: Aris and Phillips, 1985));
> '*reasoning out the explanation*' (Nehamas, '*Meno's Paradox*');
> '*reasoning about the explanation*' (Nehamas, '*Meno's Paradox*');
> '*account of the explanation*' (Nehamas, '*Meno's Paradox*'). p. 79

Thus, *logismos* is variously rendered as *reasoning, calculation, or account*, while *aitia* as *cause, explanation, or even reason.*

12.4. *The* ***circular/revolutionary*** *meaning of* ***bond*** **(desmos)** in the Republic 616 b7-c5

We first consider the use of the term *bond (desmos)* in the *Republic* 616b-c:

> and they saw there at the middle of the light the extremities of its *bonds (desmon)* stretched (tetamena) from heaven; or this light was *the bond (sundesmon)* of the heavens like the undergirders (hupozomata) of triremes (trieron), holding together in like manner the entire revolving vault (periphoran). And from the extremities was stretched the spindle (atrakton) of Necessity, through which all the orbits revolve (periphoras).
> *Republic* 616 b7-c5

It follows from the *Republic* 616b-c passage that for Plato a *bond* is something that holds the parts of an entity together, in order that a circular revolution in that entity may take place. The elaborate circular motions of the stars in heaven, comprising a variation of the Pythagorean Harmony of the Spheres, are held together using *light (phos)*, acting as a bond; more down to earth, Plato employs the striking metaphor of the cyclical revolutions of the oars of a moving trireme, held together in unison by the *undergirders* of the trireme acting as the firm bond.

12.5. *The* ***bond*** *268d5, formed by the intertwining (sumploken) of the names (onomaton), described in the summary 268c5-d5 of the definition of the Sophist by the method of Name and Logos, examined in Section 11.7, is the* ***Logos Criterion****, holding together the periodic anthyphairesis by which the* Sophist *is defined.*

Next, we examine the *bond (268c5)* described in the summary 268c-d and note that its function is, as the function of the bond described in the *Republic* 616b-c, to hold together a revolutionary process, namely a periodic anthyphairesis.

Sophist 268c5-d5

ΞΕ. Οὐκοῦν ***συνδήσομεν*** αὐτοῦ, καθάπερ ἔμπροσθεν, ***τοὔνομα συμπλέξαντες*** ἀπὸ τελευτῆς C7 ἐπ᾽ ἀρχήν G; ΘΕΑΙ. Πάνυ μὲν οὖν.		
ΞΕ. Τὸ δὴ τῆς ἐναντιοποιολογικῆς C7		
	εἰρωνικοῦ μέρους C6	
		τῆς δοξαστικῆς C5
μιμητικόν C4,		
	τοῦ φανταστικοῦ γένους C3	
		ἀπὸ τῆς εἰδωλοποιικῆς C2

(continued)

οὐ θεῖον D1ἀλλ' ἀνθρωπικὸν C1 τῆς ποιήσεως G
ἀφωρισμένον ***ἐν λόγοις***
τὸ θαυματοποιικὸν μόριον C7,
"ταύτης τῆς γενεᾶς τε καὶ αἵματος" ὃς ἂν φῇ ***τὸν ὄντως σοφιστὴν*** C7 εἶναι τἀληθέστατα, ὡς ἔοικεν, ἐρεῖ. ΘΕΑΙ. Παντάπασι μὲν οὖν.

Nominative	Genitive	Genitive of Genitive
Stranger Shall we then **bind up (sundesomen, 268c5)** his **name (tounoma, 268c6)** C7, as we did before, **intertwining (sumplexantes, 268c6)** from **the end** C7 to **the beginning** G **Theaetetus** By all means.		
The juggling part, that of the art of contradiction C7,		
	of the dissembling part C6	
		of the art of opinion C5 which is part of
the imitative kind C4		
	of the fantastic class C3	
		of the image-making art C2,
and is not divine D1 but human C1, productive art G		
defined in logoi —he who says that the true sophist is of this descent and blood will, in my opinion, speak the exact truth.		

In the *Sophist* 268c5-d5, the entity in question is the definition of the Sophist by means of the method of Name and Logos, given in Sect. 11.

[1] Successive *names* of the Division, in the relation of heteron to heteron, have been interpreted as the remainders of an anthyphairetic Division, starting with the Genus of all productive arts G, and ending with the Sophist C7 (in the notation of Section 11.7):

$$G > D1 > C1 > D2 > C2 > D3 > C3 > D4 > C4 > D5 > C5 > D6 > C6 > D7 > C7.$$

need some bonds, and these are the logoi, intertwining/bonding successive names with each other; thus, the end name of the Sophist is fastened with the beginning name of all productive arts

[2] We next form the ratios of successive names
(*hapseis* in the second hypothesis in the *Parmenides*,
logoi in the *Theaetetus*:

> but the things composed (sugkeimena) of these are themselves *intertwined (peplektai)*,
> and so their *names are intertwined (sumplakenta)*
> and form a *logos (logon)*;
> for *the intertwining (sumploken) of names is the true Being (ousian) of logos (logou)*.
> *Theaetetus* 202b2-5

[*Note*. The classical H.N. Fowler translation renders *logon* as *rational explanation* and *reasoning*],

as in the chain below, where a * between two successive names x and y indicates the intertwining/*sumploke* x*y leading to the generation of the ratio/*logos* x/y, formed of one name x with respect to the other name y:

$$G^*D1^*C1^*D2^*C2^*D3^*C3^*D4^*C4^*D5^*C5^*D6^*C6^*D7^*C7.$$

We note that
C7 and C4 are in the Nominative case,
C6 and C3 in the Genitive of the Nominative case, and
C5 and C2 in the Genitive of the Genitive of the Nominative case,
indicating the coupling affected by the *Logos Criterion*; thus,

$$D7/C7 = D4/C4, D6/C6 = D3/C3, D5/C5 = D3/C3,$$

still further confirmation of the anthyphairetic interpretation of the Division.

We can interpret this summary as follows:

[3] In fact, Plato more likely means the end C7 not of the Division but of the period, and thus, the beginning is not the initial genus with the name G, but the genus with the name D2, from where the period starts:

$$C7^*\ D7^*\ C6^*\ D6^*\ C5^*\ D5^*\ C4^*\ D4^*\ C3^*\ D3^*\ C2^*\ D2.$$

[4] The *bond (desmos)* of the logoi formed is exactly Theaetetus' Logos Criterion, as explained in Section 11.7:

$$[[[C7^*\ D7]]]^*\ [[C6^*\ D6]]^*\ [\mathbf{C5}^*\ \mathbf{D5}]^*\ [[[C4^*\ D4]]]^*\ [[C3^*\ D3]]^*\ [\mathbf{C2}^*\ \mathbf{D2}],$$

namely

D2/C2 = D5/C5, hence D3/C3 = D6/C6 and D4/C4 = D7/C7.

[5] The *bond* functions exactly as in the *Republic* 616b-c passage, considered in Section 12.4, namely, it holds together a circular, periodic revolution, namely, as we saw in Section 11, the Logos Criterion is the bond holding together the periodic anthyphairesis that defines the Sophist. Thus, the interpretation of the term *'bond'* in 268c5 is anthyphairetic.

12.6. The ***Logos acting as Bond collects*** *the parts generated by the dyad Being, not-Being* ***into the same/tauton*** (Sophist 253a4-c1, 251d8-9) and thus **equalizing** them (257a-e)

It follows from the *Sophist* 253a4-12, b8-c1 that, as *the vowels* (253a4) are the *bonds* (desmos, 253a5) possessing the power (dunata, 253a8) to make letters (grammata, 253a1) communicate (koinonein, 253a8) other to other (heteron heteroi, 253a6), in the same way, *the logoi* (253b10) are the *bonds*, possessing the power to make genera (gene, 253b8) communicate other to other, explained earlier (*251d8*) to mean *to collect the genera in the same/tauton*, (*eis tauton sunagagomen*), more explicitly to *equalize the Being part and the not-Being* part (*257a-e*).

This collection into the same/one and the equalization in the *Sophist* have anthyphairetic content, as follows from our analysis and interpretation of the corresponding equalization of the parts One and Being in the second hypothesis of the *Parmenides*, using the anthyphairetic circularity/periodicity of the "hapseis"/ contacts, exactly the Logos Criterion, as analyzed by Negrepontis (2024).

12.7. *Analogio is the most beautiful bond in the* Timaeus 31c2-32a7

The anthyphairetic interpretations of the collection into the same and equalization using Logos acting as a Bond in the Sophist *forces an anthyphairetic interpretation of* ***the most beautiful bond (kallistos desmos)*** *in the* Timaeus*, a proportion (analogio) with three terms x/y = y/z, collected into the same/into one.*

> And *the fairest (kallistos) of* ***bonds (desmon)*** is *that which most perfectly collects into one itself and the things which it binds together (sundoumena)*; and to effect this in the fairest manner is *the natural property of* ***proportion (analogio)***. For whenever the middle term of any three numbers, *or volumes (ogkon) or powers (dunameon)*, is such that as the first term is to it, so is it to the last term, and again, conversely, as the last term is to the middle, so is the middle to the first, then the middle term becomes, in turn, the first and the last, while the first and last become in turn middle terms, and the necessary consequence will be that all the terms are *the same (ta auta)*, and being *the same (ta auta) they all are one (hen panta estai). Timaeus* 31c2-32a7 [Translated by W.R.M. Lamb, with modifications by the author.]

The most beautiful *bond* (desmos) in the *Timaeus* 31c2-32a7

(a) Is a *bond*
(b) Is *a proportion (analogio)* in three terms of the form x/y = y/z
(c) Possesses the power to *collect* the terms x, y, and z into the same/one

Since it is a bond (a), we expect (by the *Republic* passage, Section 12.4) that it holds together a revolutionary/circular process, since it possesses the power to collect (c) and has the form of a proportion (c), we expect (by the *Sophist*'s passages, Sections 12.5,6) that it be the Logos Criterion for some periodic anthyphairesis. In fact, the Logos Criterion $b/c_1 = c_1/c_2$ in the calculation of the anthyphairesis of the diameter to the side of a square (Proposition 2.5) ideally fits with all three conditions.

12.8. *The Knowledge of the diameter* a to the side b of a square in the Meno 80d-86c and 97a-98b is True Opinion plus Logos, namely the infinite anthyphairesis of a to b

Anth(a,b) = [1, period(2)], while
the Logos $b/c_1 = c_1/c_2$

(a) is a bond (according to Section 11.7), and
(b) provides full Knowledge of the diameter to the side of a square when added to the True Opinion of the diameter to the side of a square (according to Proposition 1.5)

We have shown that the Knowledge of the line a that duplicates the square of a given line b in the *Meno* is the knowledge of the complete infinite anthyphairesis of the diameter a to the side b of a square; and that this Knowledge is achieved by first obtaining the two initial steps of the anthyphairesis of the duplicating line a with respect to b, a *True Opinion*, and then showing Theaetetus' *Logos* Criterion. Thus, it is fully confirmed that Plato's philosophical discussion on Knowledge in the *Meno* 80d5-86c3 and 97a9-98b6 passages follows closely and has as a model our reconstruction of Theaetetus' geometrical computation of the anthyphairesis of the diameter to the side of a square (in Section 2).

13 *Knowledge as Recollection in the* Meno *and the whole of Plato's work*

13.1. *Our anthyphairetic interpretation of philosophical Knowledge* as *Recollection* in the Meno is a reference to the Repetition of Logos using which the Theaetetus' Logos criterion establishes anthyphairetic periodicity and Knowledge of every dyad of lines commensurable in square only and of every intelligible Being.

The Knowledge obtained in the *Meno* 80d-86c and 97a-98b is described, not only as True Opinion and Logos but also as Recollection. But why such a proportion is called recollection/anamnesis? To understand in what sense the knowledge of the infinite anthyphairesis of the diameter to the side of a square may reasonably be described as Recollection, it is helpful to note that, as it is stressed both in Plato's *Epinomis* and in Proclus *Comments to Euclid*, *Recollection* is specifically *Recollection of Logoi*.

13.2.1. *According to the* Epinomis *977b-c, Recollection is specifically Recollection of Logoi*

> Moreover, turning back some little way,
> let us *recollect (anamnesthomen) the logoi* (*tois logois*, 977b9-c1)

13.2.2. *According to Proclus,* Comments to Euclid *44,25-47,8, Knowledge (episteme) is Recollection (Anamnesis) of Logoi*

> According to the tradition, the Pythagoreans recognized that everything we call *learning (mathesis)* is *recollection (anamnesis), not* something placed in the mind from without, like

> the images of sense pictured in the imagination, *nor* transitory (episodeiodes), like *the knowledge of true opinion (doxastike gnosis), but*, though *awakened (anegeiromene)* by sense-perception, [learning] has its source *within us*, by *the dianoia returning (epistrephomenes) to itself.*
>
> They realized too that, although evidence of such *recollections (anamneseis)* can be seen in many areas, it is especially evident in *mathematics*, as *Plato* also remarks. 'If you guide a person to the *geometric diagrams (diagrammata)*, he says, "then you can show most clearly that *leaming (mathesis) is recollection (anamnesis).*" This is why *Socrates in the Meno* uses this kind of argument to prove that *learning is* nothing but the soul's *recollecting (anamimneskesthai) its own logoi (ton heautes logon) 45,5-21*
>
> This, then, is what *learning (mathesis)* is, *recollection (anamnesis)* of *the* eternal *logoi (ton logon) in the soul;* and this is why the study that especially brings us *the recollections (tas anamneseis) of these logoi (tas ekeinon ton logon) is called the science concerned with learning/mathematics.*·46,15-18
>
> [Translations by Morrow (1970), with modifications by the author.]

13.3. *Our anthyphairetic interpretation of Recollection in the* Meno.

Proclus, in the passage cited above, provides a further hint: the mind acquires knowledge by returning to itself, thus strongly suggesting the periodicity of the ratios/logoi, using which, namely using the Logos Criterion, Knowledge is indeed acquired:

> *[μάθησις]* προβαλλομένη δὲ ἔνδοθεν ἀπ' αὐτῆς τῆς διανοίας εἰς ἑαυτὴν ἐπιστρεφομένης, [learning] has its source *within us*, by *the dianoia returning to itself.*

It is then clear that Recollection is simply the repetition of the second ratio/Logos, which indeed provides, according to Theaetetrus' Logos Criterion, full knowledge of the anthyphairesis of the diameter to the side of a square. Thus, Recollection is a *poetic reformulation* of Theaetetus' Logos Criterion.

13.4. *In all previous interpretations of Plato's theory of Knowledge as Recollection, there is no correlation of Recollection with the Theaetetus' Logos Criterion for anthyphairetic periodicity.*

Landry (2012):

> A recurring question about Plato's Meno is: what is the status and role of recollection? For philosophers of mathematics, in particular, this is an important question. Typically, and *more often dismissively*, recollection is offered up as *the platonist's 'mysterious' method, akin to intuition*, whereby we come to grasp both mathematical objects and philosophical Forms (see [Shapiro 2000]).
>
> In the more general philosophical literature, we find three detailed responses to this question.
>
> In line with the typical interpretation noted above, Guthrie [1956], Benson [1990], and Scott [2009] argue that we ought to take recollection as a method for knowledge and interpret the slave-boy situation as, respectively, an experiment in, or demonstration of, this method.
>
> On the other hand, Sternfeld and Zyskind [1986] suggest that *recollection be taken as a mere myth* that Socrates uses to motivate inquiry.

> Against both of these interpretations, Moravcsik [1971] claims that, whereas Plato intended recollection to be an a priori truth, we ought to view it as lying *somewhere between a very general empirical theory about the structure of the mind and a metaphor for how we learn*.

Shapiro (2000):

> The Meno suggests another epistemology. There, Plato has Socrates lead a slave to the theorem that the square on the diagonal of a given square is double the area of the original square. Socrates emphasizes that neither he nor anyone else taught the theorem to the slave. By asking carefully chosen questions, and pointing to aspects of a drawn diagram, Socrates gets the slave to discover the theorem for himself.
>
> Plato uses the experiment to support a doctrine that when it comes to geometry—or the world of Being generally—what is called 'learning' is actually remembering from a past life, presumably a time when the soul had direct access to the world of Being. Scholars disagree on the nature and role of this *'recollection'* in Plato's epistemology, *and most subsequent Platonists demur from it*. In any case, Plato did hold that the soul is in a third ontological category, with the ability to apprehend both the world of Being and the world of Becoming. With or without *the 'mystical' elements of the epistemology*, one gets the impression from the dialogues that the physical world is constructed as it is just so that we will be driven beyond our senses to investigate the world of Being. For Plato, mathematics is a key step in this process. It elevates the soul, reaching beyond the material world to the eternal world of Being. p. 52

Moravscik (1971):

> Important philosophic theses often have the disturbing character of falling somewhere between the areas of apriori propositions, empirical theories, and mere metaphors. One such thesis is Plato's proposal, expounded most fuHy in Meno 80D-86C, that learning of a certain kind is recollection. p. 53
>
> As a hypothesis about the structure of the mind, the recollection thesis seems to fall between being a very general empirical theory and being a metaphor. The reason for this is the inherent vagueness of the thesis. p. 69

13.5. *Meno's Knowledge as Recollection is most valuable for understanding the anthyphairetic structure of Plato's intelligible Being.*

We observe that Plato takes great interest to describe in philosophical terms the geometric knowledge of the anthyphairesis of the diameter to the side of a square. Furthermore, we can understand Plato's philosophical interest in this special case of geometric knowledge, since we have seen (in Sections 11 and 12) that in the *Sophist* Plato's intelligible Being is a philosophical imitation of a dyad of geometric lines having an anthyphairesis that satisfies Theaetetus' Logos Criterion and is thus in periodic anthyphairesis. Thus in effect, *Meno*'s theory of Knowledge as True Opinion plus Logos and as Recollection can be considered as a revealing introduction to Plato's theory of Knowledge of an intelligible Being.

The anthyphairetic nature of the philosophic term *True Opinion* in the *Meno* for a geometrical problem has implications for the use of the term True Opinion in purely philosophic dialogues, such as *Symposium*, *Republic, Timaeus Parmenides's* second hypothesis, *Theaetetus*, *Sophist*.

Thus in the *Symposium,* the Knowledge of a true Intelligible Being is given in terms of True Opinion plus Logos, exactly as in the *Meno*; in the *Republic* True Opinion, an initial segment of the philosophic anthyphairesis of an intelligible Being equals the full knowledge of a sensible participating in the intelligible being, something that is modified in the *Timaeus*, with the introduction of the Receptacle (as studied in the work by Negrepontis and Kalisperi 2023).

In particular, the Knowledge of an intelligible Being can have such an anthyphairetic description only if it is a dyad in the philosophic analogue of periodic anthyphairesis, something that is fully confirmed independently in the cases of the One, completed to the periodic anthyphairetic dyad One, Being of the second hypothesis in the *Parmenides* (Negrepontis 2024) and of the Being, completed to the periodic anthyphairetic dyad Being, not-Being in the *Sophist* (Negrepontis in preparation).

The following diagram shows in simplified form the various equivalent descriptions of Knowledge (episteme) of intelligible Beings in Plato's work.

Knowledge (episteme) in Plato				
Recollection	True Opinion +Logos	Name +Logos	Division in Kinds	Division +Collection
Meno 80d5-86c3, 97a9-98b6	*Meno* 80d5-86c3, 97a9-98b6			
	Symposium 201e10-202a10 209e5-212a7			
	Republic 475e-480a	*Republic* 454a1-b10	*Republic* 454a1-b10 509d-511e, 531e-534d	
	Timaeus 27d5-28a4 51d3-52a7			
	Parmenides 2nd *hypothesis* 155d3-e2	*Parmenides* *2nd hypothesis* 155d3-e2		
	Theaetetus 201d9-202c6	*Theaetetus* 201d9 -202c6		*Theaetetus* 147c-d
	Sophist 260b-264b	*Sophist* 218b-231b 264b-268d	*Sophist* 264d10-265a2	
		Statesman The whole dialogue	*Statesman* 285a4-b8 286d7-287b3	*Statesman* 285a4-b8
Phaedrus 249b6-c4			*Phaedrus* 273d2-274a7 265d3-266c8	*Phaedrus* 273d2-274a7 265d3-266c8 249b6-c4

14 Summary

Plato sets out to explain his philosophical notion of Knowledge in the *Meno*, by explaining what he means by Knowledge in the concrete geometrical case of the diameter to the side of a square. For that purpose, Plato introduces the concepts of True Opinion and Logos/Logismos and formulates his answer as follows: the Knowledge of the diameter a to the side b of a square is True Opinion plus Logos/Logismos.

Our interpretation

of *True Opinion* is that it consists of the first two anthyphairetic relations of the diameter to the side of a square,

of *Logos* is the Theaetetus' Logos Criterion for periodic anthyphairesis,

of *Knowledge* is the knowledge of the infinite anthyphairesis of the diameter to the side of a square, obtained formally as True opinion plus Logos, and

of *Recollection* is a (poetic) reformulation of the repetition of the Logos/ratio, resulting in the complete knowledge of the diameter to the side of a square.

15 Translations

G.R. Morrow, 1970. *Proclus: A Commentary on the First Book of Euclid's Elements, Translated, Introduction*, *Notes*. Princeton University Press, Princeton

Plato Twelve Volumes:

Vol. 1 translated by Harold North Fowler; Introduction by W.R.M. Lamb. Cambridge, MA, Harvard University Press; London, William Heinemann Ltd. 1966 (including *Phaedo*).

Vol. 3 translated by W.R.M. Lamb. Cambridge, MA, Harvard University Press; London, W. Heinemann Ltd. 1967 (including *Meno*).

Vols. 5 & 6 translated by Paul Shorey. Cambridge, MA, Harvard University Press; London, W. Heinemann Ltd. 1969 (including *Republic*).

Vol. 9 translated by W.R.M. Lamb. Cambridge, MA, Harvard University Press; London, W. Heinemann Ltd. 1925 (including *Symposium, Phaedrus, Timaeus*).

Vol. 12 translated by Harold N. Fowler. Cambridge, MA, Harvard University Press; London, W. Heinemann Ltd. 1921 (including *Parmenides, Theaetetus, Sophist, Politicus, Philebus, Epinomis*).

[Translations of passages from Plato in the present chapter, unless otherwise indicated, are from *Plato in Twelve Volumes*, with modifications, including italicization of some words for emphasis, by the author.]

References

Allen RE (1959) Anamnesis in Plato's Meno and Phaedo. Rev Metaphys 13:165–174

Benson HH (2015) Clitophon's Challenge. Dialectic in Plato's Meno, Phaedo, and Republic. Oxford University Press

Bluck RS (1961) Plato's Meno. Edited with Introduction and commentary. Cambridge University Press

Ebert T (2007) The theory of recollection in Plato's Meno: against a myth of platonic scholarship. In: Erler M, Brisson L (eds) Gorgias-Menon. Selected papers from the seventh symposium Platonicum. Academia Verlag Sankt Augustin, pp 184–199
Fine G (2014) The possibility of inquiry Meno's Paradox from Socrates to Sextus. Oxford University Press
Fine G (2021) Knowledge and true belief in the Meno. In: Essays in ancient Epistemology. Oxford University Press, pp 63–93
Halper EC (2007) A lesson from the Meno. In: Erler M, Brisson L (eds) Gorgias-Menon. Selected papers from the seventh symposium Platonicum. Academia Verlag Sankt Augustin, pp 234–242
Klein J (1965) A commentary on Plato's Meno. The University of North Carolina Press, Chapel Hill
Landry E (2012) Recollection and the mathematician's method in Plato's Meno. Philos Math 20: 143–169
Moravcsik J (1971) Learning as recollection. In: Vlastos G (ed) Plato. A collection 01 critical essays I: metaphysics and epistemology. Palgrave Macmillan
Negrepontis S (1999) Η Διαλεκτική του Πλάτωνος υπό την Ανθυφαιρετικήν Βάσανον, στον τόμο (επιμέλεια Β. Σύρος, Α. Κουρής, Η. Καλοκαιρινού) Διαλέξεις, ακαδημαϊκό έτος 1996–97. Όμιλος Φιλοσοφίας Πανεπιστημίου Κύπρου, Λευκωσία [in Greek]
Negrepontis S (2000a) Η Ανθυφαιρετική Φύση της Διαλεκτικής του Πλάτωνος, στον τόμο (επιμέλεια Φ. Καλαβάση & Μ. Μεϊμάρη) Διεπιστημονική Προσέγγιση των Μαθηματικών και της Διδασκαλίας τους, Εκδόσεις Gutenberg, pp.15–77 [in Greek]
Negrepontis S (2000b) Ένας Μαθηματικός Ρίχνει Φως στον Πλάτωνα. Ο καθηγητής Στέλιος Νεγρεπόντης ξαναδιαβάζει τον φιλόσοφο. Συνέντευξη Σ.Ν. στον Κώστα Κάρη στην Εφημερίδα Αυγή, Κυριακή 20 Ιουλίου 2000, pp. 10–11 [in Greek]
Negrepontis S (2012) Plato's theory of knowledge of forms by division and collection in the Sophistes is a philosophic analogue of periodic anthyphairesis (and modern continued fractions), arXiv:1207.2950, 12 Jul 2012
Negrepontis S (2018) The Anthyphairetic revolutions of the platonic ideas. In: Sialaros M (ed) Revolutions and continuity in Greek. Mathematics. Walter de Gruyter, pp 335–381
Negrepontis S (2019) Plato on geometry and the geometers. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Cham, pp 1–88
Negrepontis S (2024) The periodic Anthyphairetic nature of the one of the second hypothesis in Plato's Parmenides. In: Hascher X, Papadopoulos A (eds) Proceedings of the conference Mathématiques et musique des Grecs à Euler, Strasbourg, 10–11 September, 2015. Hermann, Paris. (112 p)
Negrepontis S (in preparation) The Anthyphairetic interpretation of not-being in the Sophist
Negrepontis S, Farmaki V (2019) Ιστορία των Αρχαίων Ελληνικών Μαθηματικών, τόμος Ι, Εκδόσεις Εκκρεμές.[In Greek]
Negrepontis S, Kalisperi D (2023) The Mystery of Timaeus' receptacle resolved. In: Sriraman B (ed) Handbook in history and philosophy of mathematical practice. Springer Nature, Cham
Negrepontis S, Farmaki V, Brokou M (2020) The central role of incommensurability in Pre-Euclidean Greek mathematics and philosophy. Ganita Bharati, Bull Indian Soc Hist Math 42:1–34
Negrepontis S, Farmaki V, Brokou M (2023) The Pell equation in the Pythagoreans, Theaetetus, and Hindu mathematics. In: Sriraman B (ed) Handbook of history and philosophy of mathematical practice. Springer Nature, Cham
Nehamas A (1984) Episteme and Logos in Plato's later thought. Archiv für Geschichte der Philosophie 66:11–36
Nehamas A (1985) Meno's Paradox and Socrates as a teacher. Oxf Stud Anc Philos 3:1–30
Scott D (2005) Plato's Meno. Cambridge University Press
Sedley D, Long A (2011) Plato, Meno, and Phaedo. Cambridge University Press
Shapiro S. (2000) Thinking about mathematics. Oxford University Press
Waterfield R (2005) Plato. Meno and other dialogues Charmides, Laches, Lysis, Meno translated with an introduction and notes. Oxford University Press

The Mystery of Plato's Receptacle in the *Timaeus* Resolved

Stelios Negrepontis and Demetra Kalisperi

Dedicated with admiration

ὁ δὲ [Θεαίτητος] οὕτω λείως τε καὶ ἀπταίστως καὶ ἀνυσίμως ἔρχεται ἐπὶ τὰς μαθήσεις τε καὶ ζητήσεις μετὰ πολλῆς πρᾳότητος,
οἷον ἐλαίου ῥεῦμα ἀψοφητὶ ῥέοντος,
ὥστε θαυμάσαι τὸ τηλικοῦτον ὄντα οὕτως ταῦτα διαπράττεσθαι.
But [Theaetetus] advances toward mathematics (τὰς μαθήσεις) and research (ζητήσεις) smoothly and surely and successfully (λείως τε καὶ ἀπταίστως καὶ ἀνυσίμως), with perfect gentleness, like a stream of oil that flows silently (οἷον ἐλαίου ῥεῦμα ἀψοφητὶ ῥέοντος), so that one marvels (θαυμάσαι) how he accomplishes all this at his age (τὸ τηλικοῦτον ὄντα). Theaetetus 144b3-7.

Abstract

Plato, most unexpectedly, in the middle of the *Timaeus* (48e2-49a7) declares that the sensible bodies cannot be explained solely by their participation in the intelligible, as we were led to believe by reading the long succession of all his previous dialogues, but that it is now *necessary* to introduce, beside the intelligibles and the sensibles, a Third Kind, the *Receptacle*.

We must, however, in beginning our fresh account of the Universe make more distinctions than we did before; for whereas then we distinguished two Kinds (τότε μὲν γὰρ δύο εἴδη διειλόμεθα), we must now declare another *Third Kind* (τρίτον ἄλλο γένος). For our *former exposition*, those two were sufficient, one of them being assumed as a *Paradigm* (παραδείγματος εἶδος), intelligible, and always self-similar Beings (νοητὸν καὶ ἀεὶ κατὰ ταὐτὰ ὄν), and the second as *the Paradigm's Image* (μίμημα δὲ παραδείγματος), subject to generation and visible (γένεσιν ἔχον καὶ ὁρατόν). A *Third Kind* we did not at that time

S. Negrepontis (✉) · D. Kalisperi
Department of Mathematics, Athens University, Athens, Greece

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_73

distinguish, considering that those two were sufficient, but now the argument seems *to compel us* (ἔοικεν *εἰσαναγκάζειν*) to try to reveal by words a Kind that is baffling and obscure (χαλεπὸν καὶ ἀμυδρὸν εἶδος). What essential property, then, are we to conceive it to possess? This in particular – that it should be *the receptacle* (ὑποδοχὴν), and as it were *the nurse of all Generation* (πάσης εἶναι γενέσεως αὐτὴν οἷον τιθήνην).Yet true though this statement is, we need to describe it more plainly 48e2-49a7.
(Emphasis by italics or bold here and in future occurences is ours)

It is fair to state that both the reason that Plato felt compelled to introduce the Third Kind and its nature have so far eluded understanding and remained a mystery, and this is certainly not because of any lack in efforts.

In the present work, we aim to obtain a new and, we believe, definitive understanding of the nature of the Receptacle and of the reason that Plato felt obliged to introduce it. This resolution is possible only because we have uncovered, under Plato's fascinating account, the well-hidden deep Mathematics of incommensurability and of infinite anthyphairesis, some going back to the Pythagoreans, some discovered recently by Theaetetus, the great resident mathematician of the Academy, and which are fundamental, not only for Plato's philosophical system of the unchanging intelligible Beings, but also for his ambitious description of the incessantly changing sensible, physical World, as well. Without reaching an understanding of the underlying Mathematics, it is impossible to gain any true understanding of the philosophy based on it; it is very much like trying to understand Newtonian physics without calculus or quantum physics without Hilbert space operators.

The starting point of our work is the interpretation of the intelligible Being as the philosophical analogue of a dyad in periodic anthyphairesis, and the pre-*Timaeus* interpretation of a sensible body participating in the intelligible, in terms of an initial finite segment of the infinite intelligible anthyphairesis. The serious problem that results from the built-in intelligible anthyphairetic *infinity* is that this infinity leads to the formation of *infinite kosmoi*, a problem akin to the intelligible Third Man Argument; it has not been realized that the *sole* reason for the necessity of introducing the Third Kind in the *Timaeus* is the necessity to avoid infinite kosmoi. At this point, it must be emphasized that the intelligibility is neutral on this problem, its only necessary condition being that the intelligible must retain the control of the sensibles, whether these sensibles live in infinitely many kosmoi or in one kosmos.

Plato attempts to deal with the problem, by eliminating the infinite multitude of anthyphairetic remainders, replacing them by the mathematically equivalent tight double inequalities and generalized side and diameter numbers (the "convergents" of modern continued fractions), which are vividly and aptly described, in order to represent every sensible body, as a dyad consisting of Content & Receptacle, but he realizes that such arithmetical methods, while they work perfectly well for the geometric anthyphairetic remainders, fail to make sense for the intelligible philosophical analogues of the geometric anthyphairetic remainders.

It is precisely because of this differentiation that Plato finds it necessary to introduce the geometric Third Kind; it consists of the *four primary bodies*, each identified with one of the surfaces of the *four* (minus the 12hedron) *canonical solids*, *primary earth* PE/cube, *primary water* PW/20hedron, *primary air*

PA/8hedron, *primary fire* PF/pyramid, studied in the Theaetetean Book XIII of Euclid's *Elements*. The *comprehension of the canonical solids in the sphere*, a central theme in Book XIII, in fact the world sphere in Plato's mind, provides several crucial advantages for this choice, to wit:

(a) It makes possible the *canonical comparison* by which the primary earth PE turns to be *incommensurable*, in fact in periodic anthyphairesis to each of the other three, mutually commensurable, primary bodies, and thus allows for the *substitution* of the intelligible dyad <intelligible earth, intelligible fire> by the Third Kind dyad <PE, PW> in its role of generating and being the *cause* of the sensible bodies and their motions.
(b) ("*intelligence*") It makes possible the *participation*, via the Soul of the World, of the Third Kind into the intelligible, thus turning the Third Kind into a "*bastard intelligible*" and salvaging Plato's absolutely essential philosophical requirement that the intelligible governs over the sensibles (by means of True Opinion).
(c) ("*necessity*") It entraps and imprisons the Third Kind in the World sphere, not allowing the generation of infinitely many remainders and Kosmoi above and beyond the initial one, but turning, equivalently, the participation into sensible dyads, vividly and accurately described as having the form Content & Receptacle, underlying the tight remainder-free double inequalities and anthyphairetic "convergents," thus saving *the One Kosmos* with the traditional four (not elements any more, but) derived kinds.

Keywords

Plato · Timaeus · Infinite kosmoi · Receptacle · Theaetetus · Periodic anthyphairesis · Books X, XIII of Euclid's *Elements*

Introduction

Plato's theory of intelligible Beings, or Ideas, is systematically presented in the *Parmenides*, *Sophistes*, and *Politicus* and is present in all platonic dialogues, including the *Politeia*, *Phaedo*, *Meno*, *Sumposion*, and *Theaetetus*. According to the interpretation of Plato's philosophy given by Negrepontis 2012, 2018, 2019, TA, an intelligible Being is a dyad in a philosophic analogue of periodic, hence, infinite, anthyphairesis (its *Oneness* being secured by the equalization of the elements of the dyad by periodicity, its *periodicity* by *Logos*, and the Logos Criterion of periodicity) (Sect. 1).

Before *Timaeus*, there is no systematic treatment of the sensibles. Every sensible body participates in an intelligible Being. The full knowledge of such a sensible is obtained as a True Opinion of the intelligible Being in which it participates (*Sumposion* 210e-211a, *Politeia* 475e-480a, 509d-510b), and a True Opinion of an intelligible Being is identified with an initial finite segment of the intelligible infinite anthyphairesis (*Theaetetus* 189e-190a, *Sophistes* 264a-b) (Sect. 2).

Thus, an intelligible Being has a built-in, intrinsic *infinity, anthyphairetic infinity*, and *infinite multitudes of anthyphairetic remainders.* We claim that it is precisely this

intelligible infinity that poses *the problem of infinite kosmoi*, stated twice in the *Timaeus* (31a2-b3, 55c7-d2), and that, although usually treated by modern scholars as a peripheral and irrelevant one, this is the *central*, in fact the *sole*, problem that makes *necessary* (ἀνάγκη, 48a1, 49a3, 52b3) the introduction of the Third Kind/Receptacle (Sect. 3).

At this point, it is argued *heuristically* that Plato's "cure" (ἴαμα, in Plutarch's words) of avoiding the infinitely many anthyphairetic remainders consists in passing from an initial finite segment of the anthyphairetic division to an *equivalent* representation, in terms of some remarkable *tight double remainder-free mathematical inequalities* (generalizing the obvious equivalence of an anthyphairetic relation $a = kb + c$, in which the remainder c is present, with the double inequality $kb \leq a \leq (k + 1)b$, in which the remainder c has been eliminated). This representation is closely related to the Pythagorean side and diameter numbers, greatly extended by Theaetetus to all quadratic incommensurabilities, the precursors of the *convergents* of modern continued fractions. As we shall see in Sect. 9, the impressive tightness of these double inequalities is described vividly and validly in the *Timaeus* with the term *Receptacle* and with a variety of other terms (Sect. 4).

However, Plato realizes and explains lucidly in passages 51e6-52a4 and 52c5-d1 the *impossibility* of applying these geometrical techniques and manipulations, directly to the dyad of the pure intelligible Being. The necessary algebraic manipulations make no sense for an intelligible dyad a, b (such as the dyad One, Being in the second hypothesis in the Parmenides), for which it makes no sense to form a multiple of a, and even less sense for a to enter into, to have a included in, another multiple of b (Sect. 5).

Plato introduces *Triangle A*, which he uses to generate the surface of the cube, and *Triangle B*, which he uses to generate the surfaces of the pyramid, 8hedron, and 20hedron, thus introducing four of the five *canonical solids*, constructed in the Theaetetean Book XIII of Euclid's *Elements* (53c4-d7, 54a1-d7, 54d5-55c6) (Sect. 6); the surfaces of these bodies, simply renamed, form *the four primary bodies*, to wit primary earth PE, the surface of a cube, consisting of 24 (right angle isosceles) Triangles A, and primary fire PF, air PA, and water PW, the surface of the pyramid, 8hedron, 20hedron, consisting of 24, 48, 120 (half of the equilateral triangle) Triangles B, respectively. The introduction of these geometric bodies strongly suggests that Plato is acting in the direction of the heuristics we explained in Sect. 4, and that he introduces the *Third Kind/Receptacle* (48e2-49a7) for the *sole* reason of avoiding the infinite kosmoi. At this point, we have no indication if Plato's choice of the geometric Third Kind is a good one. *Four requirements* must be met before Plato's choice can be deemed successful.

The *first requirement* is that the primary earth PE must be *incommensurable* to, in fact in periodic anthyphairesis with, any of the other (clearly *commensurable* to each other) three primary bodies PW, PA, and PF. A dyad among the elements of the Third Kind must *replace* the intelligible dyad in its role of generating the sensibles – and for this certainly infinite anthyphairesis is necessary. Plato clearly hints at their incommensurability in 54b5-d3, 56d1-6, but does not explain how this comes about. For this to happen, Triangle A and Triangle B must be *incommensurable,* in fact in periodic anthyphairesis. We show that there is a *canonical way* of comparing these triangles, the following: We regard Triangle A as half of the face of a cube,

Triangle B as half of the face of a pyramid, and we comprehend this cube and this pyramid in the same sphere. It is shown with the help of careful computations mostly carried out in Propositions 13, 15, and 18 of Book XIII of the *Elements*, and with the help of Theaetetus' theorem on the periodicity of the anthyphairesis of quadratic irrationals, that indeed any multiple of Triangle A is in periodic anthyphairesis with any multiple of Triangle B (Sect. 7).

The *second requirement*, completing the first, is that the Third Kind will generate all the sensibles and their motions, replacing the intelligible. Indeed, according to the fundamental (and ill-understood thus far) passage 57d-58a, one of the incommensurable dyads of primary bodies will be the basis for the *Third Kind*, substituting for the intelligible Being as the *cause* for the generation of all sensible bodies. Plato selects the incommensurable dyad PE, PW, with the aid of the commensurable auxiliary ratios PW/PA = 5/2 and PW/PF = 5/1; thus, a sensible body is determined by an initial finite anthyphairetic segment of the infinite anthyphairesis PE to PW. This implies that all sensible motions and generations of sensibles, including those by the Receptacle, are between dyads of primary bodies (Sect. 8).

We next present the main argument that the description of the Third Kind is equivalently given by the Receptacle and by the "diakena" of the primary bodies (an initial argument was already given just above). The argument consists in comparing the description of the generation of the *Timaeus* sensibles in terms of the "diakena" of the primary bodies with the corresponding description in terms of the Receptacle. The four steps, by means of which the generation of the sensibles and their motion takes place in and by the Receptacle, as described in the passages 50a-51a, 52d4-53a6, and 57a7-c6, are so uncannily similar to the four steps, by means of which the generation of the sensibles and their motion takes place in and by the "diakena" of the primary bodies, as described in the passages 58a4-c2, 61a7-b3, that the conclusion, that the Receptacle and the primary bodies are equivalent descriptions of the Third Kind, is inescapable. The same argument both makes clear the mathematical origin of the term Receptacle (as a reflection of the double tight inequalities) and leads to the satisfaction of the *third requirement*, namely the fact that the description of the sensibles in terms of the Third Kind does indeed succeed in eliminating the infinite remainders and Kosmoi, securing a One Kosmos. But we still have not pinpointed what is special with the primary bodies that makes the One Kosmos choice not simply possible, but mandatory and necessary. We will take up this issue, thus completing the third requirement, in the next section (Sect. 9).

The *fourth requirement* is that the Third Kind *participates in the intelligible*, something necessary in order to salvage the sine qua non of Plato's philosophic system, namely the domination of intelligible Beings over sensibles, namely the *participation of these modified sensibles in the intelligible.* This is achieved with (a) the participation of the Soul of the World into the intelligible (with the sphericity of the World 33a6-b5, and the revolutions of the Soul 36d8-e5, imitating intelligible periodicity), and (b) the *comprehension* of the primary bodies in the sphere, as shown in the second part of Propositions 13–16 of the Theaetetian Book XIII of the *Elements*.

With the fourth requirement satisfied, the Third Kind does indeed assume (paraphrasing somewhat 52b1-2) the status of a *"bastard intelligible"* entity, of an

entity which, though far removed from the intelligible nevertheless, *participates* in the intelligible Being in a most perplexing ("δυσαλωτότατον," 51a7-b1) manner, a *special* geometric dyad in periodic anthyphairesis, that substitutes the intelligible in the role of being the cause and generator of the sensible bodies and of their motions.

We still need to *satisfy the third requirement*. A geometric dyad in periodic anthyphairesis is equally capable of generating either remainders and infinite kosmoi, like the intelligible, or remainder-free dyads of the form Receptacle & Content. We have seen in Sects. 8 and, especially, 9 that the primary bodies, and specifically the dyad PE, PW, do indeed generate sensibles in the desired Receptacle & Content form, leading to One kosmos. But we have not yet pinpointed the reason that this desired form is achieved. This is indeed achieved again by the *comprehension* of the primary bodies *in the sphere*. According to passage 58a4-7, the resulting confinement, imprisonment, and entrapment of the primary bodies into the spherical Soul of the World results in the inability of the Third Kind to employ the transmitted, participated intelligible anthyphairetic information for the generation of many, potentially infinite anthyphairetic remainders and kosmoi, and the *necessity* to generate instead sensible bodies of the form Content & Receptacle, by the filling and tightening of the small primary body into the great primary body (Sect. 10).

We are now, at long last, able to describe fully the nature and the structure of the *Timaeus* sensibles: The sensibles participate in the intelligible via the Third Kind, every sensible is a syllable of the form Receptacle & Content, and there are four kinds of sensibles in One Kosmos.

Sensible water or earth is defined in terms of a "convergent" M_n/N_n of the dyad PE, PW; for odd index, the "convergent" satisfies the Content & Receptacle type of inequality $M_{2n-1}PW < N_{2n-1}PE < (M_{2n-1} + 1)PW$, and thus the defeated PE becomes dissimilar to itself, *dispersed* by and similar to the victorious PW, and the whole dyad is a *waterlike sensible body,* while for even index the "convergent" satisfies the Content & Receptacle $(M_{2n} - 1)PE < N_{2n}PW < M_{2n}PE$, and thus the defeated PW becomes dissimilar to itself, *consolidated* by and similar to the victorious PE, and the whole dyad is an *earthlike sensible body.* For sensible air or fire, we must consider the triad PA or PF, PW, and PE, instead, in which the masses of primary air PA or primary fire PF enter and fill the "diakena" of the masses of primary water PW, which have already filled the "diakena" of the masses of primary earth PE. With this stratagem,[1] the One World with four kinds of sensible bodies is now finally achieved (Sect. 11).

The essential correctness of our interpretation is further confirmed by the fully satisfying answers it can provide for a host of *questions* that up to now had no satisfactory answer or no answer at all:

[1]. The reason for stating in the first place in 31a-b, and for reopening the *"infinite kosmoi"* question at 55c-d

[2]. The reason why Plato defines the primary bodies as the *surface*, rather than the volume, as it would have been more natural, of the four regular solids

[1] From Greek στρατήγημα; the official spelling "stratagem" does not appear correct.

[3]. In which way primary earth and primary fire are *incommensurable*, as hinted but not explained by Plato
[4]. The necessity of incommensurability of the primary earth to the other three primary bodies and the crucial importance of passage 57d-58a, particularly the meaning of *Inequality*, in understanding the construction of the Third Kind
[5]. *"Diakena," first* occurring in 58b5, replace and are identical with *Receptacle last* occurring in 57c3
[6]. The filling of the Receptacle and the "diakena" "fully" "so that not even one part is left out" is a vivid description of the Mathematics of infinite anthyphairesis (modern "convergents" of continued fractions)
[7]. The difference between sensible *"such"* and Receptacle *"this,"* 49e7-50a2
[8]. Every sensible participating in the intelligible Being with dyad F, notF, is *sometimes F, sometimes notF*, a puzzling earlier description of the sensibles (*Sumposion* 210e-211a, *Politeia* 475e-480a), explained in terms of *Dispersion and Consolidation*
[9]. The reason that the Universe does not lapse into eternal Rest, but will *keep moving incessantly*, is intimately related with the incommensurability and the participation of the *"bastard intelligible"* entity into the intelligible, by means of being *comprehended* in the Sphere, 58a2-7, c2-4.

In fact, the *Timaeus* is an intellectual triumph, because it represents the first attempt to explain mathematically the physical world in accordance with the experimental data available, specifically to explain the incessant motion observed in it, employing the deep mathematical achievements of his era, namely the quadratic incommensurability and periodic anthyphairesis, achieved initially by the Pythagoreans and then greatly extended by Theaetetus. It is the first instance of a long series of the "Unreasonable Effectiveness of Mathematics in Natural Sciences," an effectiveness that was enunciated as a mysterious problem by Wigner 1960, and to which an answer is being attempted, based on the Paradoxical Nature of Mathematics (Negrepontis and Farmaki 2021).

We should note that the whole platonic conception of the Receptacle in the *Timaeus* relies heavily on the deep mathematical achievements of the great mathematician Theaetetus. The main reason for the difficulties of the scholars on the *Timaeus* to reach, despite intensive study, essential understanding is that there appears to have been little appreciation of the underlying deep mathematics, especially the Mathematics of incommensurability, on which Plato's philosophic structure is founded.

Section 1 *The Platonic Theory of the Intelligible True Beings: The Philosophical Analogue of a Dyad in Periodic Anthyphairesis*

According to our interpretation, an *intelligible (Platonic true) Being* consists of a dyad of parts satisfying the *philosophical analogue of geometric periodic anthyphairesis*. The paradigmatical true Being is the One of the second hypothesis in the *Parmenides*, consisting of the dyad <One, Being> (studied in Negrepontis TA), or of the dyad <Being, Not-Being>, such as <Just, Unjust>, in the *Sophistes* and *Politicus* (studied in Negrepontis 2012, 2018), in a philosophical analogue of geometric anthyphairesis, for which the exact analogue of the Logos Criterion for periodicity holds.

1.1. *Geometric anthyphairesis*

Let us recall the *geometrical concept of anthyphairesis* of a dyad of opposite magnitudes, the greater a and the smaller b, $a > b$:

$$\begin{aligned}
&a = k_0 b + a_1, \quad \text{with } a_1 < b,\\
&b = l_0 a_1 + b_1, \quad \text{with } b_1 < a_1,\\
&\ldots\\
&a_{n-1} = k_{n-1} b_{n-1} + a_n, \text{with } a_n < b_{n-1},\\
&b_{n-1} = l_{n-1} a_n + b_n, \text{with } b_n < a_n,\\
&a_n = k_n b_n + a_{n+1}, \text{with } a_{n+1} < b_n,\\
&\ldots
\end{aligned}$$

Here

$$k_0, l_0, k_1, l_1, \ldots, k_n, l_n, \ldots,$$

denotes the sequence of successive quotients (natural numbers), and

$$a > b > a_1 > b_1 > \ldots > a_n > b_n > \ldots$$

the sequence of successive remainders (magnitudes) of the anthyphairesis of a to b.

In case a remainder never divides the immediately previous remainder, the anthyphairesis is said to be *infinite*. If the anthyphairesis of a to b is infinite, then a and b are *incommensurable* (Proposition X.2 of Euclid's *Elements*).

The anthyphairesis of a to b is *periodic* if there is an index n, such that the ratio a/b is equal to the ratio a_n/b_n (or to the ratio b_n/a_{n+1}) (*Logos Criterion* for periodicity).

1.2. *A platonic intelligible true Being is the philosophical analogue of a dyad in periodic anthyphairesis*

According to our interpretation, an intelligible Platonic true Being consists of a dyad of parts satisfying the philosophical analogue of geometric periodic anthyphairesis.

The paradigmatical true Being is the One of the second hypothesis in the *Parmenides*, consisting of the dyad < One, Being >, or the dyad < Being, Not-Being > in the *Sophistes*, in the philosophical analogue of periodic anthyphairesis. The anthyphairetic interpretation of a platonic intelligible true Being has been obtained by our analysis of the following:

- The dyad < One, Being > in the second hypothesis of the *Parmenides* 142b1-155e3, and of *Philebus* 15a-17a (Negrepontis TA)
- The Division and Collection of the Angler and the Sophist in the *Sophistes*, and of the Statesman in the *Politicus* (Negrepontis 2012, 2018)
- The dyad < beautiful, ugly > in the *Politeia* 475e9 (Negrepontis 2019)

According to this interpretation, a platonic idea is the philosophic analogue of a dyad of opposite parts; periodicity, as described in the *Parmenides* 144c2-e1, *equalizes* the two opposite and unequal elements of the dyad; thus, a platonic idea qualifies to be called a *One*, as indeed it is explained in the *Plilebus* 15a-17a. This sophisticated sense of Oneness is completely lost among scholars on Plato. Thus, e.g., Annas 1981, Chap. 9, claims that there are two classes of platonic ideas, one coming from opposites, and one not.

Thus, an intelligible Being is *a dyad a, b* of parts, in the philosophical analogue of periodic anthyphairesis. The model of an *intelligible Being* is *the dyad* a = One, b = Being in the second hypothesis of the *Parmenides*.

It is shown there that the *anthyphairesis*

$$
\begin{aligned}
&a = b + a_1, \text{with } a_1 < b, a_1 \text{ a smaller One,}\\
&b = a_1 + b_1, \text{with } b_1 < a_1, b_1, \text{a smaller Being,}\\
&\ldots\\
&a_n = b_n + a_{n+1}, \text{with } a_{n+1} < b_n,\\
&b_n = a_{n+1} + b_{n+1}, \text{with } b_{n+1} < a_{n+1},\\
&\ldots
\end{aligned}
$$

is *infinite*, and in fact *periodic*; namely, there is an index n, so that

$$a_n/b_n = a/b \ (\textit{Logos Criterion} \text{ for periodicity}).$$

1.3. *The difference between a geometrical anthyphairesis and its philosophical analogue*

For an understanding of the *Timaeus*, it is of importance to have in mind *the difference* between a geometrical anthyphairesis and its philosophical analogue.

The philosophical analogue has *no quotients*; since it makes no sense to talk about two times Beautiful or two times Ugly, it certainly makes sense to consider two times a magnitude. This difference will turn out to be crucial in Sects. 4 and 5, when we will realize that geometrical anthyphairesis can be manipulated mathematically, in ways that do not make sense and hence are not allowed for its philosophical intelligible analogue. Plato undoubtedly would regard the philosophical anthyphairesis as the pure, simple, and genuine superior entity, and the geometrical as the rather corrupted inferior one.

Section 2 *[Pre-Timaeus Sensible Bodies] A True Opinion of an intelligible Being is identified with an initial finite segment of the infinite anthyphairesis of the intelligible Being, and the Sensibles participating in an intelligible Being are fully known by a True Opinion of the intelligible Being.*

We outline our interpretation of Plato's pre-*Timaeus* theory of *sensible bodies* (Negrepontis 2012; Negrepontis and Pappa Birba TA). According to the *Timaeus*

30c-31a, there is a unique intelligible Being, identified as the dyad <IE, IF>, consisting of intelligible earth IE and intelligible fire IF, in which every sensible body *participates*. A sensible body, identified as a *"doxaston,"* is an entity whose full knowledge is a true opinion *(doxa)* of the intelligible Being (*Politeia* 475e6-480a3, 509d1-510b1), namely as a finite initial system of the *infinite anthyphairesis* of the intelligible Being (*Sumposion* 202a, 209e-211d, *Meno* 97e-98c, *Sophistes* 264a-b).

2.1. *The anthyphairetic interpretation of True Opinion ("alethes doxa"): The True Opinion of an intelligible Being is identified with an initial finite segment of the infinite anthyphairesis of the intelligible Being.*

The anthyphairetic interpretation of the intelligible Beings makes possible a deeper and more precise understanding of True Opinion. Another equivalent description of "Division and Collection" is given by "True Opinion ("alethes doxa") and Logos" in *Sumposion* 202a2-10*,* 209e5-211d1, *Meno* 97e2-98c4 (where instead of "logos" the equivalent expression "logismos aitias" is employed) and in *Theaetetus* 201c8-d3*,* 202b8-d7. Furthermore, given that "true opinion" is described as "*alogos*" in *Sumposion* 202a6, *Theaetetus* 201d1, and *Timaeus* 28a3, 51e4, "opinion" is related only to anthyphairetic division (and not at all to the Logos Criterion for anthyphairetic periodicity).

What is *that which always is (τὸ ὂν ἀεί)* and has no becoming (γένεσιν δὲ οὐκ ἔχον)?
And what is that which is *always becoming* (τὸ *γιγνόμενον* μὲν *ἀεί*)
and never is (ὂν δὲ οὐδέποτε)?
Now the one of these (τὸ μὲν) is apprehended (περιληπτόν) by means of noesis (νοήσει)
with logos (μετὰ λόγου), since it exists always (ἀεὶ κατὰ ταὐτὰ ὄν);
whereas the other (τὸ δ') is an *object of opinion (δόξῃ) alogos (ἀλόγου)*
opinable (δοξαστόν) with the aid of sensation (μετ' αἰσθήσεως),
since it becomes (γιγνόμενον) and perishes (ἀπολλύμενον)
and never really **is** (ὄντως δὲ οὐδέποτε ὄν). *Timaeus 27d6-28a4*
[the translation of all passages from Plato in this work is, unless otherwise indicated, by Plato, Translation of Plato in twelve volumes, with modifications by the authors]

On account of "opinion" being "alogos," the most natural conclusion would be to equate "true opinion" with "division." However, division is infinite, while the two descriptions of "opinion" in Plato, *Theaetetus* and *Sophistes* (*Theaetetus* 189e4-190a6, *Sophistes* 264b1, *Scholia in Eucliden* X.2, 1–6) clearly show that "opinion" is a process finite in nature.

But the soul, as *the image* presents itself to me when it thinks,
is merely conversing with itself,
asking itself questions and answering, affirming and denying.
When it has arrived at a decision (ἐπάξασα), whether slowly or with a sudden bound,
and is at last agreed, and is *not* in need of a *dichotomic decision* (μὴ διστάζῃ),
we call that its *Opinion* (δόξαν) *Theaetetus* 189e4-190a6

There is considerable backing for interpreting "διστάζῃ" as a *dichotomic decision:*

(i) Hesuchios *Lexicon*, in Delta 1976 and 2021, [distazei] is rendered as "dichonoei" (being of a mind divided in half), while in Delta 2086, "distazei" is equated with "amphidoxei" (having both opinions).
(ii) *Magnum* 278, 35–36 explains "distazein" as thinking in a double manner ("dichos bouleuesthai") or as "to have two stations in thinking" ("duo staseis echein kata dianoian").
(iii) Philoponus notes in *Comments to Aristotle's On Soul* 15,78,11-26 that judgment can be compared to "distazei," which is likened to traveling a divided road ("schisten odon"), considering to go in one or in the other direction.

Thus, "μὴ διστάζῃ" signifies a stop at an initial finite stage of the infinite intelligible anthyphairetic division.

Finally, in *Sophistes 264a9-b1* "thought is the conversation of the soul with itself, and opinion is the *final* result ('apoteleutesis') of thought."

Moreover, the *Anonymous Scholion X.2 to Euclid's Elements* describes "opinion" as the philosophical analogue of *finite anthyphairesis*. It is stated there that it is in the nature of numbers to have finite anthyphairesis,

> those [the numbers] when divided end in a definite end, unity

[cf. Propositions VII.1, 2 of the *Elements*]; this is so, the *Scholion* continues, because

> numbers act in an opinion-like manner ('tous de arithmous doxastikos'), and
> opinion is of a finite nature ('peperastai gar mallon he doxa').

It now follows, from the interpretation of division and collection, that a true opinion of an intelligible Being is the knowledge provided by a finite initial segment of the infinite anthyphairetic division of the intelligible Being.

There is of course not just one, but *an infinite multitude of true opinions*, each one corresponding to a finite initial segment of the infinite anthyphairesis of the intelligible dyad. Details can be found in Negrepontis (2012, Section 7).

2.2. *A sensible body participating in an intelligible Being is fully known by a True Opinion of that intelligible Being (Politeia 475e6-480a13)*

A sensible entity participating in an intelligible idea is generally described by Plato as an image, a copy of the intelligible Being:

> the one kind (ἐν μὲν εἶδος) being assumed (ὑποτεθέν) as a *model* (ὡς παραδείγματος), *intelligible (νοητὸν)* and always (ἀεὶ) unchangingly (κατὰ ταὐτὰ) existent (ὄν), and the second (δεύτερον) as *the model's copy*, subject to being generated and visible.
> *Timaeus* 48e3-49a1

But this is too indefinite. Since the intelligible idea has the structure of a dyad in periodic anthyphairesis, we might ask what would constitute a copy, an image, and an approximation of this idea?

From the *Politeia* 475e6-480a13 ("doxaston" 478a11, b2,b3, e3, especially 479d7), 509d1-510b1 ("doxaston" 510a9, used in the analogy of the Divided Line), 534a2-9, and the *Timaeus* 28a2-4 ("doxaston" 28a3), 28c1-2 and 51d3-7, and 51e6-52a6, we present below a condensed version of the argument in the *Politeia* 475e-480a, as reconstructed in Negrepontis and Pappa Birba TA, that a sensible body coincides with a "doxaston."

Basic Proposition in the *Politeia* 475e6-480a3. A sensible body participating in an intelligible Being is fully known by a true opinion of that intelligible Being.

Proof Let S be a sensible body participating in an intelligible Being, with dyad

$$< \mathrm{F}, \mathrm{notF} >.$$

By the crucial property of sensible bodies, S is an entity which *sometimes is F, and sometimes not F* (479a5-b2, d3-6).[2]

Hence, S is a *Being between* the intelligible Being (τοῦ εἰλικρινῶς ὄντος) with dyad <F, notF> and the absolute not-Being (τοῦ μηδαμῇ ὄντος) (477a6-8).

Hence, the complete knowledge (dunamis, 477d2) of S is a *knowledge between* ignorance and the knowledge of the intelligible Being (478e4-5).[3]

Hence, by 2.1, the complete knowledge of S is a true opinion about the intelligible Being.[4]

Hence, by definition, S is *opinable* (doxaston) with respect to the intelligible Being.[5]

[2] Exactly the same description of sensibles is also stated in the *Sumposion* 210e-211a, but in neither *Sumposion* nor *Politeia* is the inference shown or adequately explained. It is explained in the *Timaeus* 58a-c, in terms of the representation of a sensible body as a dyad Receptacle & Content either in Dispersion (and then not F) or in Consolidation (and then F), clearly an anthyphairetic property (cf. Sects. 9, 10, and 11).

[3] A *between Being corresponds* to a *between Knowledge by which it is fully known,* since

the extreme Beings correspond (ἀποδιδόντες) to the extreme Knowledges (τοῖς μὲν ἄκροις [extremes Being] τὰ ἄκρα [extremes Knowledges], and

the Between Beings to the Between Knowledges (τοῖς δὲ μεταξὺ [Between Being] τὰ μεταξὺ [Between Knowledge]) 478e4-5.

[4] A true opinion (δόξα) about an intelligible Being is a *knowledge* (about the intelligible Being) *between (μεταξὺ)* ignorance (ἀγνοίας) and intelligible knowledge (episteme, γνώσεως). 478c13-d4. This indeed follows from our account of true opinion in 2.1.

[5] A *Being Between (μεταξὺ)* the intelligible Being (τοῦ εἰλικρινῶς ὄντος) with dyad <F, notF> and the absolute not-Being (τοῦ μηδαμῇ ὄντος) is an opinable entity with respect to an intelligible Being with dyad <F, notF> 477a6-8,479d7-10.

Thus, a sensible body participating into an intelligible Being is opinable and fully known by a true opinion of that intelligible Being.[6]

2.3. The *anthyphairetic interpretation of the sensibles and of their participation in the intelligible Being*

By Sect. 2.2, the complete knowledge of a sensible body participating in an intelligible platonic Being is a true opinion of the intelligible Being, and therefore, according to the anthyphairetic interpretation of true opinion, by Sect. 2.1, it is identified with the knowledge provided by some first finite steps of the anthyphairetic division of the intelligible Being.

2.4. *The intelligible* ***is the cause of*** *every sensible body which participates in the intelligible (Phaedo 100c4-8).*

> I think that if anything is beautiful (*ἄλλο καλὸν*) besides πλὴν intelligible beauty (*αὐτὸ τὸ καλόν*), it is beautiful *for* no other reason than because it participates (*μετέχει*) in intelligible beauty *(ἐκείνου τοῦ καλοῦ)*; and this applies to everything.
> Do you assent to this view of ***cause (τῇ τοιᾷδε αἰτίᾳ)***?" "I do," said he. *Phaedon* 100c4-8

The intelligible Being is the ***cause*** of every sensible which participates in that intelligible. Thus, for a sensible body to have as ***cause*** an intelligible Being means that the knowledge of the sensible is determined by an initial finite segment of the intelligible infinite anthyphairesis.

This sense of "cause" of the sensibles will be crucial in understanding the "cause" appearing in 57d-58a in Sect. 8.

2.5. *A simplification: All sensibles participate in a unique single intelligible Being (Timaeus 30c2-31a1).*

All sensibles participate in and resemble (ἐοικὸς, ὁμοιῶσαι) a single, unique intelligible Being (the intelligible living creature, "panteles zoon")

There is a *living creature*

[1]. of which all other living creatures (τἆλλα ζῷα), severally and generically (καθ' ἐν καὶ κατὰ γένη), are parts (μόρια),
[2]. which embraces and contains within itself (ἐν ἑαυτῷ περιλαβὸν ἔχει) all the intelligible living creatures (τὰ γὰρ δὴ νοητὰ ζῷα),
[3]. which is intelligible,
[4]. which is fairest of all (*καλλίστῳ*) and in all ways most perfect (καὶ κατὰ πάντα τελέῳ) (30d1-31a1).

[6] An entity S is *opinable ("doxaston," δοξαστὸν)* with respect to an intelligible Being if the complete knowledge (γνῶναι ὡς ἔχει) of S is a true opinion of that intelligible Being. 478a6-11.

This "*unique*" intelligible being seems, somehow, at odds to what we would assume that there would be many platonic ideas. Besides, this is what we understand when reading, for example, *Phaedo* 100d where Plato says that "beautiful" things are beautiful by participating in "beauty," (ἀποκρίνασθαι ὅτι τῷ καλῷ τὰ καλὰ [γίγνεται] καλά), "great" things are great by participating in "greatness," and "smaller" things are smaller by participating in "smallness." At this point, we accept that there is a unique intelligible Being, such that all the sensible entities in the Heaven participate in and resemble to it. Thus, according to 2.1 and 2.2, every sensible can be fully described by an initial finite segment of the anthyphairesis of the unique intelligible dyad.

2.6. *The tentative identification of the unique Platonic intelligible Being in which all sensible bodies participate*

Plato, as we saw in 1.6, makes clear in the *Timaeus* 30c-31a that there is a single, unique platonic intelligible idea ("panteles zoon") into which all sensibles, all parts, and all entities in the *"kosmos"* participate. From our analysis of intelligible Beings (Sect. 1.1), we conclude that this idea, the unique intelligible Being that shapes the material world, consists of a dyad of opposites in periodic anthyphairesis.

In the *Parmenides* 130c1-2, there is talk about *intelligible fire or water* (αὐτό τι εἶδος . . . πυρὸς ἢ καὶ ὕδατος).

In the *Philebus* 29a6-d6, Plato distinguishes between sensible fire and earth and intelligible fire and earth and states that the sensible fire and earth participate in the intelligible fire and earth (οἷον πῦρ ἔστι μέν που παρ' ἡμῖν, ἔστι δ' ἐν τῷ παντί. . . τε τῆς ἐν τοῖς ζῴοις γῆς τῆς ἐνθάδε καὶ τῆς ἐν τῷ παντί).

In the *Timaeus* 51b7, he speaks of *intelligible fire* (ἆρα ἔστιν τι *πῦρ αὐτὸ ἐφ' ἑαυτοῦ* καὶ πάντα περὶῶν ἀεὶ λέγομεν οὕτως αὐτὰ καθ' αὑτὰ ὄντα ἕκαστα) in which the sensible fire participates. From the four traditional sensible bodies, Plato clearly, according to *Timaeus* 30c-31a, regards the earth and the fire as the more fundamental ones.

From the above, it is reasonable to conclude that the unique intelligible Being that shapes all "kosmos" is the dyad

$$< \text{Intelligible Earth IE}, \text{Intelligible Fire IF} >,$$

in a philosophic analogue of *periodic anthyphairesis:*

$$\begin{aligned}
&\mathrm{IE} = \mathrm{IF} + \mathrm{IE}_1, \mathrm{IE}_1 < \mathrm{IF},\\
&\mathrm{IF} = \mathrm{IE}_1 + \mathrm{IF}_1, \mathrm{IF}_1 < \mathrm{IE}_1,\\
&\mathrm{IE}_1 = \mathrm{IF}_1 + \mathrm{IE}_2, \mathrm{IE}_2 < \mathrm{IF}_1,\\
&\mathrm{IF}_1 = \mathrm{IE}_2 + \mathrm{IF}_2, \mathrm{IF}_2 < \mathrm{IE}_2,\\
&\ldots\\
&\mathrm{IE}_n = \mathrm{IF}_n + \mathrm{IE}_{n+1}, \mathrm{IE}_{n+1} < \mathrm{IF}_n,\\
&\mathrm{IF}_n = \mathrm{IE}_{n+1} + \mathrm{IF}_{n+1}, \mathrm{IF}_{n+1} < \mathrm{IE}_{n+1},\\
&\ldots
\end{aligned}$$

The sequence of anthyphairetic remainders is strictly decreasing:

$$\mathrm{IE} > \mathbf{IF} > \mathbf{IE_1} > \mathrm{I}_1 > \mathrm{IE}_2 > \mathrm{IF}_2 > \ldots > \mathrm{IE}_n > \mathbf{IF_n} > \mathbf{IE_{n+1}} > \mathrm{IF}_{n+1} > \ldots$$

(x > y means x is greater/heavier than y, y is equal to a part of x). This anthyphairetic division (of IE into IF and a smaller part IE_1 of the type IE, and so on) is formed in strict analogy to the division of the paradigmatical dyad One, Being in the second hypothesis of the *Parmenides*.

2.7. *The image account "εἰκός λόγος" of the sensibles 29c1-3*

The fact that a sensible body participates in the intelligible, as an approximate imperfect copy of the intelligible, is expressed systematically in the *Timaeus* by the expression "according to the image" ("kata to eikos," 29c). The expressions εἰκότας λόγους (29c2), κατά τό εικός (53b3) occur quite often throughout *Timaeus*, and thus it is important to have a correct rendering.

Lamb's translation of the *Timaeus* 29c1-3 is as follows:

> whereas the *accounts [λόγους]* of
> that which is *copied (τοῦ ἀπεικασθέντος) after the likeness* of that Model,
> and is itself a *likeness (εἰκόνος),*
> will be *analogous (ἀνὰ λόγου)* thereto and *possess likelihood (εἰκότας)*;
> for as Being is to Becoming, so is Truth to Belief.

As far as we can tell, this rendering of "eikos logos" has originated with Cornford 1937, and is probably based on the following misunderstanding:

> In Plato's sense of the word, we cannot have knowledge of the physical world. Plato's use of the word 'knowledge' is very strong, however, and there are differing degrees of opinion we can have concerning the physical world; we may even hold true opinions about it. Gregory 2008, p. xxxvi

This is a misunderstanding because *the complete knowledge* of a sensible participating in an intelligible Being is equivalent to the true opinion of the intelligible Being, something certainly accessible to humans.[7] This misunderstanding produces the one about "eikos logos":

[7] In *Meno,* the slave-boy is able to obtain *full knowledge* of the diameter to the side of a square (and not merely of some simple form of the Pythagorean theorem as is usually thought) by "*recollecting*," namely by repeating the logos, namely by applying the Logos Criterion for obtaining full knowledge of its periodical anthyphairesis (cf. Negrepontis, 2024); in the *Sophistes*, the full knowledge of the Angler and of the Sophist (Negrepontis 2012, 2018, 2019), and in the *Politicus* of the Statesman (Negrepontis 2018), is obtained by reaching and obtaining the logos of anthyphairetic periodicity. The *full knowledge* of a sensible equals a true opinion of the intelligible in which this sensible participates, and this is just a small "alogos" (not involving logos) part of the full knowledge of the intelligible.

> "The 'likely' part of this description stems from a word-play in the Greek, between *eikones* 'likenesses' and *eikos* 'likely'. One of the important metaphysical themes of *Timaeus* is that the cosmos is in some sense a copy or a likeness."; and "What the 'likely account' cannot do and could never do is produce knowledge, in the sense that Plato believes we can have knowledge of the forms." Gregory and Waterfield 2008, p. xxv:

M. F. Burnyeat 2005, has argued for a modification of Cornford's rendering, somewhat away from "likely" or "probable," but the move is still not enough.

Our rendering of "eikos logos" is precisely and simply as "image account," and of 29c1-3 as follows:

> the knowledge (accounts, λόγους) of the sensible
> is an *image (εἰκότας)* of the knowledge (accounts) of the intelligible
> in the same *(analogous, ἀνὰ λόγον)* way that
> the sensible bodies
> is the *likeness (τοῦ ἀπεικασθέντος)*, the image *(εἰκόνος)* of the intelligible (πρὸς μὲν ἐκεῖνο),
> **as** intelligible Being **to** sensible generation,
> **so** full knowledge of intelligible Being to true opinion of intelligible Being. **29c1-3**

The passage states that the sensible is an image of the intelligible, both ontologically and epistemologically, an echo of the Divided Line (*Politeia* 509b6-d1).

Section 3 ***[The Problem of Infinite Kosmoi] The participation of the sensibles in the intelligible infinite anthyphairetic division results in infinite "kosmoi" and in a sensible version of TMA, a consequence seemingly incompatible with the traditional One Kosmos with four kinds of sensible bodies.***

We identify *the principal,* in fact *the sole Problem,* that Plato faces with his pre-*Timaeus* theory of Sensibles, outlined in Sect. 2. The problem is that there is a built in, intrinsic infinity in every intelligible Being, and this infinity is manifested in the totality of the sensible bodies as a potential infinity: Given that a sensible entity is a "doxaston," fully knowable by an initial finite segment of the infinite anthyphairesis of the intelligible Being, we run into the difficulty of potentially an *infinite multitude of kosmoi* and the sensible version of the *Third Man Argument* (*Timaeus* 31a-b, 55c-d; Plutarch, *Obsolescence of the Oracles* 424C, 425A), both arising from the infinite multitude of anthyphairetic remainders in the intelligible Being. It seems almost impossible to reconcile and make compatible Plato's anthyphairetic theory with the older *traditional* theory, according to which there is just one Universe, one kosmos, consisting of four *elements*, earth, fire, air, and water, having no participation whatsoever in Plato's intelligible Being.

Cornford 1937 (p. 42) and other scholars do not realize that Plato introduces the Third Kind precisely in order to salvage the One Kosmos model.

> He is not offering a proof that there cannot be more than one world; he merely asserts that only one was made, because it seemed better that the copy should he unique, like the model.

3.1. *The problem of infinite "kosmoi" and of a sensible form of the Third Man Argument in the Timaeus 31a2-b3*

Right after Plato has completed his description of a sensible body as an approximate image of the intelligible Being, according to our interpretation, outlined in Sect. 2, fully knowable by an initial finite, but arbitrarily long, initial segment of the intelligible infinite anthyphairesis of the dyad earth, fire, in the passage *Timaeus* 31a-b, he states a question, essentially raising a problem.

Plato, seemingly mysteriously and out of context, raises the question whether there are infinitely many universes or finitely many. He hastens to answer that there must be only one Heaven ("ouranos") or Universe ("kosmos") . His justification is certainly unexpected, in that it involves some version of the Third Man Argument (TMA) , first presented in the *Parmenides* 131e8-132b2 Introduction. Vlastos (1954) and others believed that the argument is valid and destroy Plato's theory of intelligible Beings. The TMA has been analyzed in the context of the second hypothesis of the *Parmenides* in the Negrepontis TA paper, showing that the Argument poses no problem at all for the theory of ideas. Vlastos' reconstruction was based, among others, on the Non-Identity (NI) principle, which would be valid if an intelligible Being consisted of a dyad in merely infinite anthyphairesis, and not a periodic one. It is shown that the periodicity of the anthyphairesis in an intelligible Being, and the resulting equalization of parts, invalidates the NI principle.

The present passage in the *Timaeus*, clearly written years after the *Parmenides,* is in fact by its very presence a strong internal indication that Plato does not consider the TMA as an argument dangerous for his theory of ideas. However, if we transfer from the intelligible Being to the sensible bodies, then the Logos Criterion for periodicity is no more present, and a sensible body is completely known by a finite initial segment of the intelligible anthyphairesis. In such a situation, the TMA is valid for the sensibles, which are based only on the infinite intelligible anthyphairetic division. In fact, the validity of TMA is essentially just another form of the problem with infinite "kosmoi." Both are valid because of infinite anthyphairetic intelligible division.

> Are we right, then, in describing the Heaven as *one (ἕνα οὐρανὸν)*,
> or would it be more correct to speak of heavens
> as many or *infinite in multitude* (πολλοὺς καὶ *ἀπείρους*)?
> *One* it must be termed,
> if it is to be framed after its [intelligible] Pattern *(κατὰ τὸ παράδειγμα)*.
> For that which embraces περιέχον all intelligible Living Creatures could *never* be
> *second (δεύτερον)*, with *another* (μεθ' ἑτέρου) beside it;
> for if so, there must needs exist *yet another (ἕτερον) [third]* Living Creature,
> which should embrace them both,
> and *of which* they two would each be a part;
> in which case this Universe
> could no longer be rightly described as modelled on *these two (ἐκείνοιν),*
> but rather on *that third (ἀλλ' ἐκείνῳ)* Creature which contains them both.
> Wherefore, in order that this Creature might resemble the all perfect Living Creature
> in respect of its uniqueness (κατὰ τὴν μόνωσιν), for this reason its Maker made
> *neither two (οὔτε δύο) Universes* (κόσμους) *nor an infinite multitude (οὔτ' ἀπείρους),*
> but there is and will continue to be this **one** generated Heaven (οὐρανὸς),
> *unique* of its kind (μονογενὴς). *Timaeus* 31a2-b3

Plato hastens to answer in favor of one "kosmos," not of an infinite multitude of them, and denying the possibility of a sensible form of the TMA, but he does not really explain why. In fact, he returns to the question in the *Timaeus* 55c7-d2, exactly before introducing the four primary bodies, with which the problem is solved.

> Now in reasoning about all these things, a man might question
> whether he ought to affirm the existence of
> an *infinite multitude* of Universes *(ἀπείρους χρὴ κόσμους)*
> or a finite number *(ἢ πέρας)*;
> and if he questioned aright he would conclude that
> the doctrine of an *infinite multitude (*τὸ μὲν *ἀπείρους)* is that of
> a man unversed (τὸ μὲν *ἀπείρους*) in matters wherein he ought to be versed ***(ἔμπειρον)*;**
> *Timaeus* 55c7-d2

We will return later to study Plato's solution of the problem. For the time being, let us make sure we understand why the problem is not a whimsical, out of the blue question, but an essential and necessary question.

3.2. *The problem of infinite "kosmoi" understood, with the help of Plutarch's account, as an essential feature of the built-in infinity in Plato's dialectics*

Plutarch in *The Obsolescence of Oracles* provides a much more detailed and vivid description of the question on infinitely many "kosmoi" and helps make the question understandable.

> For if each of the bodies has *its own particular space (τόπον οἰκεῖον)*, as he asserts,
> *the earth* must of necessity turn toward the centre (*ἐπὶ τὸ μέσον*) from all directions and
> *the water* be above it *(ἐπ' αὐτῆς)*, settling below (ὑφιστάμενον)
> *the lighter elements* (τοῖς κουφοτέροις) because of its weight (διὰ βάρος).
> Plutarch, *The Obsolescence of Oracles* 424B10-C2

> If, therefore, there be *more worlds than one (πλείονες κόσμοι),*
> it will come to pass that
> *in many places the earth will rest above (πολλαχοῦ μὲν ἐπάνω) the fire and the air,* and
> *in many places below (πολλαχοῦ δ' ὑποκάτω)* them;
> and the air and the water likewise,
> in some places existing in positions in keeping with nature (κατὰ φύσιν) and
> in other places in positions contrary to nature (παρὰ φύσιν).
> Plutarch, *The Obsolescence of Oracles* 424C2-6
> [translation of the passages from Plutarch's The Obsolescence of Oracles, by F.C. Babbitt, 1936, with modifications by the authors].

The anthyphairetic interpretation of Plato's question on infinite "kosmoi" may now be obtained. According to our interpretation, described in Sect. 1, a platonic sensible entity is fully described by a finite, but possibly arbitrarily large, initial segment of the infinite intelligible anthyphairesis of the dyad <IE, IF>, generating a strictly decreasing sequence of anthyphairetic remainders:

$$IE > IF > IE_1 > IF_1 > IE_2 > IF_2 > \ldots > IE_n > IF_n > IE_{n+1} > IF_{n+1} > \ldots$$

Plutarch's account of infinite "kosmoi" combined with our interpretation of pre-*Timaeus* sensibles in Sect. 2.2 makes understandable the problem of infinite "kosmoi." The statement

> in many places (πολλαχοῦ μὲν) *the earth will rest above (ἐπάνω) the fire*
> Plutarch, *The Obsolescence of Oracles* 425A5-8

seems paradoxical at first, but if we apply faithfully the anthyphairetic interpretation outlined in Sects. 1 and 2, then it appears to be a natural consequence of the anthyphairetic interpretation of the intelligible and sensible entities. Indeed, in our interpretation of sensibles,

> the earth IE_1 will rest above the fire IF, since IE_1 is lighter than IF,
> the earth IE_2 will rest above the fire IF_1, since IE_2 is lighter than IF_1,
> . . .
> the earth IE_{n+1} will rest above the fire IF_n, since IE_{n+1} is lighter than IF_n,
> . . .,

is in excellent agreement with Plutarch's description of how the potentially "infinite kosmoi" are formed within Plato's dialectics.

Plato's anthyphairetic theory of intelligible and sensibles appears to indeed result in an infinite multitude of "kosmoi" and in particular to be incompatible with One World with four kinds of sensible bodies. The notion that there are infinitely many "kosmoi" is undesirable for Plato, as becomes clear from these two passages in the *Timaeus*. In fact, Plato is committed to One Kosmos, as is made clear in the passages of the *Timaeus* 31a2-b3 and 55c7-d2.

Plato's anthyphairetic model in its natural divisive formulation appears to be incompatible with the One World theory. Plutarch, in his account in *The Obsolescence of Oracles* 425E8-F2, makes abundantly clear this contradictory situation.

Section 4 *Every finite initial segment of an infinite* ***geometric*** *anthyphairesis can be represented by a double tight remainder-free inequality, as a dyad Receptacle-Content either in Dispersion or in Consolidation. Thus,* ***if the intelligible dyad was a geometrical anthyphairetic dyad, then every sensible body would be represented*** *by a double tight remainder-free inequality* ***as a Content*** *&* ***Receptacle dyad****.*

Thus Plato's theory of sensibles is in need of revision and modification. We present now the simple but basic mathematical ideas that Plato had in mind in his attempt to eliminate the problem of anthyphairetic infinity that enters in his description of the Sensibles, outlined in Sect. 3, and his Ingenious and imaginative description of this idea Content & Receptacle. The idea is to replace any initial finite system of an infinite anthyphairesis of two magnitudes a and b with a single double tight inequality not involving remainders. Before stating the general Basic Proposition, in 4.2, the simple manipulations can be best seen with some simple examples.

4.1. *Examples*

Example 1 Let us consider the first four initial steps in the infinite anthyphairesis of a diameter to the side of a square, whose sequence of successive quotients is equal to [1, 2, 2, 2,. . .].

With simple calculations, we find the following equivalent double inequalities:

1. The anthyphairetic relation $a = b + a_1$, $a_1 < b$ is equivalent to the double remainder-free inequality $b < a < 2b$. ($M_1 = 1$, $N_1 = 1$).
   Note that $a_1 = a - b$.
2. The system of the first two anthyphairetic equations
   $a = b + a_1$, $a_1 < b$; $b = 2a_1 + b_1$, $b_1 < a_1$
   is equivalent to the double remainder-free inequality $8a < 12b < 9a$ ($M_2 = 12$, $N_2 = 9$).
   [Indeed, $a = b + a_1$, $a_1 < b$; $b = 2a_1 + b_1$, $b_1 < a_1$ if and only if $2a_1 < b < 3a_1$ if and only if
   [by (1)] $2(a - b) < b < 3(a - b)$ if and only if $2a < 3b$ & $4b < 3a$ if and only if $8a < 12b$ & $12b < 9a$ if and only if $8a < 12b < 9a$, the double tight, remainder free inequality].
   Note that $b_1 = b - 2a_1 = b-2(a - b) = 3b - 2a$.
3. The system of the first three anthyphairetic equations
   $a = b + a_1$, $a_1 < b$; $b = 2a_1 + b_1$, $b_1 < a_1$; $a_1 = 2b_1 + a_2$, $a_2 < b_1$
   is equivalent to the double remainder-free inequality $49b < 35a < 50b$. ($M_3 = 49$, $N_3 = 35$).
   [Indeed, $a = b + a_1$, $a_1 < b$; $b = 2a_1 + b_1$, $b_1 < a_1$; $a_1 = 2b_1 + a_2$, $a_2 < b_1$ if and only if $2b_1 < a_1 < 3b_1$ if and only if [by (1) and (2)] $2(3b - 2a) < a-b < 3(3b - 2a)$ if and only if $7b < 5a$ & $7a < 10b$
   if and only if $49b < 35a$ &$35a < 50b$ if and only if $49b < 35a < 50b$, the double tight, remainder-free inequality].
   Note that $a_2 = a_1-2b_1 = (a - b)-2(3b - 2a) = 5a-7b$.
4. The system of the first four anthyphairetic equations
   $a = b + a_1$, $a_1 < b$; $b = 2a_1 + b_1$, $b_1 < a_1$; $a_1 = 2b_1 + a_2$, $a_2 < b_1$; $b_1 = 2a_2 + b_2$, $b_2 < a_2$
   is equivalent to the double remainder-free inequality $288a < 408b < 289a$. ($M_4 = 408$, $N_4 = 288$).
   [Indeed, $a = b + a_1$, $a_1 < b$; $b = 2a_1 + b_1$, $b_1 < a_1$; $a_1 = 2b_1 + a_2$, $a_2 < b_1$; $b_1 = 2a_2 + b_2$, $b_2 < a_2$ if and only if $2a_2 < b_1 < 3a_2$ if and only if [by (2) and (3)] $2(5a - 7b) < 3b - 2a < 3(5a - 7b)$ if and only if $12a < 17b$ &$24b < 17a$ if and only if $24.12a < 24.17b$ & $17.24b < 17.17a$ if and only if $288a < 408b < 289a$, the double tight, remainder-free inequality].

Example 2 We indicate the simple manipulations needed for the first two steps of any infinite anthyphairesis.

Here is how the first anthyphairetic relation

$$a = k_1 b + a_1, \text{with } a_1 < b,$$

containing the unwanted anthyphairetic remainder a_1, is transformed into the equivalent tight double inequality involving only multiples of a and b,

$$k_1 b < a < (k_1 + 1)b,$$

in which the anthyphairetic remainder a_1 has disappeared. Setting $M_1 = k_1$, $N_1 = 1$, we have

$$M_1 b < N_1 a < (M_1 + 1)b.$$

We next find (with some amazement) that the system of the first two anthyphairetic relations

$$a = k_1 b + a_1, \text{with } a_1 < b,$$

$$b = l_1 a_1 + b_1, \text{with } b_1 < a_1$$

is, setting $N_2 = (kl + 1)\ (l + 1)\ M_2 = (kl + 1)\ (k(l + 1) + 1)$, equivalent to the double tight remainder-free inequality

$$(N_2 - 1)a < M_2 b < N_2 a.$$

4.2. *The basic geometrical proposition expressing an initial finite segment of an infinite anthyphairesis in terms of a double tight inequality*

The Pythagoreans discovered the anthyphairesis of the diameter to the side of a square and subsequently studied the side and diameter numbers. Furthermore, Negrepontis (2018), based on the *Theaetetus* 147-8 passage describing Theodorus and Theaetetus' achievement on incommensurabilities, on an analysis of the fascinating palindromically periodic division of the statesman in the *Politicus,* and on the tools in the Theaetetean Book X of the *Elements*, has shown that Theaetetus had a proof of the following highly nontrivial Theorem (used in Sect. 7 below): "for every pair $a > b$ of lines such that a,b are incommensurable but a^2,b^2 are commensurable, the anthyphairesis of a to be is palindromically periodic." In the reconstruction of the proof, essential use has been made of *the generalized side p_n and diameter q_n numbers*, in order to prove the representation of every anthyphairetic remainder as an *apotome* line, in the sense of Book X of the *Elements.*

Thus, we can be reasonably certain that, thanks to Theaetetus, Plato had at his disposal the powerful mathematical techniques that are the direct progenitors of the "convergents" in modern continued fraction theory:

$$q_1/p_1 < q_3/p_3 < \ldots < q_{2n-1}/p_{2n-1} < \ldots < a/b$$
$$< \ldots < q_{2n}/p_{2n} < \ldots < q_4/p_4 < q_2/p_2.$$

Let us note here in passing that the transformation of a finite initial system of the anthyphairetic equations of a to b to a convergent (in the modern terminology of continuous fractions) provides a link for the transition from Theaetetus' theory of ratios of magnitudes based on equal anthyphairetic to the Eudoxian theory based on Dedekind-cuts.

Plato employs the following two tight double inequalities in terms of the numbers M_n and N_n which are closely related to the generalized side and diameter numbers:

$$\mathbf{M_{2n-1}b} < \mathbf{N_{2n-1}a} < (M_{2n-1} + 1)b,$$
$$(N_{2n} - 1)a < \mathbf{M_{2n}b} < \mathbf{N_{2n}a}.$$

Basic Proposition *(Pythagoreans, Theaetetus)*.

Let a, b, a > b, be a dyad of *incommensurable* magnitudes. Denote by

$$k_0, l_0, k_1, l_1, \ldots, k_n, l_n, \ldots$$

the (infinite) sequence of successive *quotients* of the anthyphairesis, and by

$$a > b > a_1 > b_1 > \ldots > a_n > b_n > \ldots$$

the (infinite) sequence of successive *remainders* of the anthyphairesis of a to b.

[1]. *Generalized Side and Diameter Numbers, Convergents*

Define the double sequence of the *generalized side and diameter numbers* of the anthyphairesis of a to be:

$$p_1 = 1, \qquad q_1 = k_0,$$
$$p_2 = l_0, \qquad q_2 = 1 + k_0 l_0,$$
$$\ldots$$
$$p_{2n-1} = k_{n-1}p_{2n-2} + p_{2n-3}, q_{2n-1} = k_{n-1}q_{2n-2} + q_{2n-3},$$
$$p_{2n} = l_{n-1}p_{2n-1} + p_{2n-2}, \quad q_{2n} = l_{n-1}q_{2n-1} + q_{2n-2},$$
$$\ldots$$

Then for all n = 1,2,...

$$a_n = p_{2n-1}a - q_{2n-1}b, \qquad b_n = -p_{2n}a + q_{2n}b,$$
$$(q_{n+2} - q_n)/q_{n+1} = (p_{n+2} - p_n)/p_{n+1},$$
$$\text{Anth}\,(q_{2n-1}, p_{2n-1}) = [k_0, l_0, \ldots, k_{n-1}],$$
$$\text{Anth}\,(q_{2n,}p_{2n}) = [k_0, l_0, \ldots, k_{n-1}, l_{n-1}],$$
$$p_{n+1}q_n - p_nq_{n+1} = (-1)^n, \text{and}$$
$$q_1/p_1 < q_3/p_3 < \ldots < q_{2n-1}/p_{2n-1} < \ldots < a/b$$
$$< \ldots < q_{2n}/p_{2n} < \ldots < q_4/p_4 < q_2/p_2.$$

[2]. *Representation of every finite initial segment of a geometric infinite anthyphairesis as a tight inequality of the form Receptacle & Content, alternately in Dispersion or in Consolidation*
We set

$$M_1 = q_1, \qquad N_1 = p_1,$$
$$M_{2n-1} = q_{2n-1}(p_{2n-2} + p_{2n-1}), \qquad N_{2n-1} = p_{2n-1}(p_{2n-2} + p_{2n-1}); \text{and},$$
$$M_{2n} = q_{2n}(q_{2n-1} + q_{2n}), \qquad N_{2n} = p_{2n}(q_{2n-1} + q_{2n}).$$

Then, for every natural number n,

$$\text{Anth}\,(M_{2n-1}, N_{2n-1}) = \text{Anth}\,(q_{2n-1}, p_{2n-1}) = [k_0, l_0, \ldots, k_{n-1}],$$
$$\text{Anth}\,(M_{2n}, N_{2n}) = \text{Anth}\,(q_{2n,}p_{2n}) = [k_0, l_0, \ldots, k_{n-1}, l_{n-1}],$$
$$M_1/N_1 < M_3/N_3 < \ldots < M_{2n-1}/N_{2n-1} < \ldots < a/b < \ldots < M_{2n}/N_{2n}$$
$$< \ldots < M_4/N_4 < M_2/N_2; \text{ and}$$

(i) The system of the first 2n + 1 anthyphairetic equations a and b

$$a = k_0b + a_1, b = l_0a_1 + b_1, \ldots, a_{n-1} = k_{n-1}b_{n-1} + a_n,$$

is *equivalent* to the double inequality of *Dispersion (diakrisis)*

$$\mathbf{M_{2n-1}b} < \mathbf{N_{2n-1}a} < (M_{2n-1} + 1)\ b; \text{ and},$$

(ii) The system of the first 2n steps of the anthyphairesis a and b

$$a = k_0b + a_1, b = l_0a_1 + b_1, \ldots, a_{n-1} = k_{n-1}b_{n-1} + a_n, b_{n-1} = l_{n-1}a_n + b_n,$$

is *equivalent* to the double inequality of *consolidation (sugkrisis)* even order-*n*

$$(N_{2n} - 1)a < \mathbf{M_{2n}b} < \mathbf{N_{2n}a}.$$

The (elementary) *proof* of the basic Proposition will not be reconstructed here.

4.3. *Plausible interpretations and heuristics*

The problem of "infinite kosmoi" is the infinity of anthyphairetic remainders, and Plato intends to "cure" it, in Plutarch's words. But Plato is obliged because of a serious problem at 48ff to introduce the mysterious Third Kind/Receptacle. Our informal reasoning suggests that Plato's introduction of the Receptacle is somehow related to his intention to "cure" the infinite kosmoi/infinite anthyphairetic remainders problem.

The infinity of anthyphairetic remainders of a geometric anthyphairesis can be "cured" by reverting to the equivalent tight remainder-free double inequalities, closely related to generalized side and diameter numbers-modern convergents. So our plausible reasoning is guiding us to the conclusion that Plato's "cure" is by means of such tight double inequalities.

Now our heuristic reasoning suggests that the concepts introduced by Plato to deal with the problem, such as

> Receptacle & Content,
> filled quite fully so that not even one more can be added,
> defeated and victorious part,
> becoming dissimilar to itself and similar to others,
> dispersion or consolidation,

actually fit surprisingly well as interpretations of the mathematical inequalities. But before proceeding with our informal, plausible heuristics, we want to consider the sense of Receptacle in Aristotle and in Plato, through Aristotle's critique.

The initial idea, in early form, that the basic geometric Proposition 4.2 is at the basis of the Receptacle in the *Timaeus*, appeared in Negrepontis 2005, (F).

4.4. *The clarification of Plato's concept of receptacle/place-content by Aristotle's criticism*

Aristotle sets six conditions (which we numbered [1] to [6]) for something to be *immediate place* (τόπος) or container of a sensible entity.

But what actually is *place* (*τόπος*)? Here is a way to find an answer. Let us take as premisses all the properties which seem genuinely to belong to *place* in its own right. This is what we expect to be true of *place* (*τόπον*)[8]:

[8] The Third Kind is described as "chora" (τρίτον δὲ αὖ γένος ὃν τὸ τῆς χώρας, 52a8; ὄν τε καὶ χώραν καὶ γένεσιν εἶναι, 52d3; σφίγγει πάντα καὶ κενὴν χώραν οὐδεμίαν ἐᾷ λείπεσθαι. 58a7). Thus "chora" is identical with the Receptacle, the place in which the generation of the sensibles takes place. In 52b4-5, the sentence "τὸ ὂν ἅπαν ἔν τινι τόπῳ καὶ κατέχον χώραν τινά" makes clear that "topos" is still another term for the Receptacle.

Aristotle in *Physics* 209b11-16 clearly identifies platonic "chora" and "topos," with the Receptacle.

[1] that it is the immediate container (πρῶτον . . .περιέχον) of
that of which it is the place *(τόπος)*,
[2] that it is not a part of the object it contains,
[3] that a thing's *immediate place (πρῶτον [τόπον])* is exactly the same size as it,
[4] that it can be left behind (ἀπολείπεσθαι) by the object and is separable (χωριστὸν)
from it,
[5] and also that every *place (τόπου)* admits of the distinction between above and below,
[6] and that every body naturally moves up or down to its own *proper place (οἰκείοις τόποις)*
and stays (μένειν) there. *Physics* 210b32-211a5
[translation of Aristotle's Physics passages in the paper by Bostock, Waterfield, 1996, with modifications by the authors]

By [1] and [3], the place must contain the content and nothing else. They are made explicit also in 209a32-b1.

Place (τόπος) may refer
either to the shared (κοινός) place which contains all bodies
or to the particular (ἴδιος) place which *immediately contains (ἐν ᾧ πρώτῳ)* a body.
For instance, you are now in the world, because you are in the air and the air is in the world;
and you are in the air because you are in the earth;
and by the same token you are in the earth because
[3] you are in this *particular place (ἐν τῷδε τῷ τόπῳ)*,
which contains nothing more than you (ὃς περιέχει οὐδὲν πλέον ἢ σέ). 209a32-b1

By [2] and [4], the place is not a part of the contained, and so it can be left behind; in this, Aristotelian "*topos*" differs from the Platonic. This is also made explicit in 209b28-33:

In fact, people do think of *place (τόπος)* as being like a vessel *(τὸ ἀγγεῖον)*
(since a vessel is a *movable place (τόπος μεταφορητός)* and
[2, 4] a vessel is not a part of the object οὐδὲν τοῦ πράγματός it contains.

The following passage is a critique of Plato's sense of place in the *Timaeus*, where the Receptacle/place is an inseparable part of the sensible entity.

however true it may be that one is in the other,
the jar *(ἀμφορεὺς)* will still *contain (δέξεται)* the wine (οἶνον)
not because it is itself wine (αὐτὸς οἶνος), but because it is *a jar (ἐκεῖνος),* and
the *wine (ο νος)* will be in (ἐνέσται) *the jar (ἐν τῷ ἀμφορεῖ)*
not because it is itself a jar, but because it is wine. *Physics* 210b13-17

Conditions [5] and [6] about "proper places" find Plato in agreement.

Thus, for both Plato and Aristotle the basic property of the Receptacle is that it contains the Content and nothing more, or, more precisely, *not even one part more*. (We will indicate the difference of the two both in 4.5 below, but especially in Sect. 9, [3].)

A main feature and peculiarity of Plato's Receptacle/Place turns out to be that surprisingly it is *not separate* from the Content, the "matter" (hule), but an integral

part of it, while on the contrary Aristotle considers, as we would find reasonable, that the containing Receptacle is separate from the thing it contains, and criticizes the platonic sense of Chora.

4.5. *More heuristics*

We wish next to obtain a plausible interpretation of the tight double inequalities obtained in 4.2 as a dyad Receptacle & Content.

We claim that in trying to make more vivid and colorful, but still faithful to the mathematics involved, the nonmathematical language, Plato introduces the term *Receptacle* (*ὑποδοχὴ*) or *place* (*τόπος, χώρα*).

The Receptacle or Place of an object is the immediate Container of another object, the Content. Recall that according to Aristotle's *Physics* 210b-211a, as we saw in 4.4, the Receptacle contains the object, its Content, it is filled by it, and it can contain nothing more, not even one thing more (for the exact uses of this last statement by Plato, we will have to wait till Sect. 9).

In both of these types of tight double inequalities

$$[1]\ \mathrm{Mb} < \mathrm{Na} < (\mathrm{M}+1)\mathrm{b},\ [2]\ (\mathrm{N}-1)\mathrm{a} < \mathrm{Mb} < \mathrm{Na},$$

the inequality **Mb < Na** holds, which may be interpreted as saying that the mass Mb of the small b, the Content, enters the mass Na of the great a, the Receptacle/Place; both of these inequalities and entrances are tight in the sense that the Content *completely fills* the Receptacle, but in different senses:

In [1] in the sense that attempting *to add even one* b makes the mass (M + 1)b go above Na, namely in the sense that the *Content is maximal*, while
In [2] in the sense that attempting *to subtract even one* a makes the mass (N − 1)a go below Mb, thus in the sense that the *Receptacle is minimal*

It follows that

In [1], b is the active/victorious part and a is the passive/defeated part; the mass Na of great parts becomes similar to, takes the shape of its opposite mass Mb of small parts, and can thus be described as *Dispersion (διάκρισις)*; while
In [2], the mass Mb of small parts becomes similar to, takes the shape of its opposite mass Na of great parts, and can thus be described as *consolidation (σύγκρισις)*.

It is then an excellent terminology to think of the mass Na as the *Receptacle* and the mass Mb as its *Content.*

We will gradually in the sequel support our heuristics with strong evidence, culminating in Sects. 8 and 9.

Section 5 ***But unfortunately, the elimination of anthyphairetic remainders and their replacement with tight double inequalities make no sense for the intelligible dyad, hence the representation of a sensible body as Content & Receptacle of the intelligible dyad is impossible.***

The reader might have noticed that the idea we mentioned in Sect. 4 was strictly for an infinite anthyphairesis of magnitudes, of *geometric* entities, and we were careful not to mention its philosophic analogue, the intelligible Being, and the platonic idea. Were it possible to assume that the transformation, by means of mathematical manipulations, of an initial finite system of relations of an anthyphairesis of a geometrical dyad works equally well and produces corresponding philosophical *intelligible* Receptacle & Content pairs, we would be able to obtain a theory that combines one kosmos and four sensible bodies (in a modification of the traditional theory of four elements) with participation of the sensibles in the intelligible Being <IE, IF>.

But, this is definitely NOT the case! We can see this clearly and directly as follows. The mathematical manipulations, involving the generalized side and diameter numbers, equivalently the modern convergents, leading to the tight double remainder-free inequalities and the Receptacle & Content representation are valid if a and b are geometric magnitudes, but they definitely make no sense at the philosophical, intelligible level. Thus, for example, the first intelligible anthyphairetic relation

$$\mathrm{IE} = \mathrm{IF} + \mathrm{IE}_1, \text{with } \mathrm{IE}_1 < \mathrm{IF},$$

cannot possibly be translated to the double inequality

$$\mathrm{IF} < \mathrm{IE} < 2.\mathrm{IF}.$$

Plato states clearly this problem with his two passages (*Timaeus* 51e6-52a4, 52c5-d1), according to which a mass of the intelligible fire, the small, cannot enter into a mass of the intelligible earth, the great.

Thus Plato's idea to eliminate the infinity of anthyphairetic remainders works for a geometric pair in infinite anthyphairesis but fails for its philosophic analogue, the intelligible dyad <IE, IF>.

5.1. Plato states clearly this problem with his two passages (*Timaeus* 51e6-52a4, 52c5-d1), according to which the intelligible dyad <IE, IF> does not satisfy the Receptacle-Content inequalities

> This being so, we must agree that one is
> the self-identical Kind (τὸ κατὰ ταὐτὰ εἶδος) [=the intelligible true Being],
> ungenerated and indestructible,
> *neither receiving* into itself any other from elsewhere
> (οὔτε *εἰς ἑαυτὸ εἰσδεχόμενον* ἄλλο ἄλλοθεν)
> *nor* itself *entering* anywhither into another
> (οὔτε αὐτὸ εἰς ἄλλο ποι *ἰόν*),
> invisible and in all ways imperceptible by sense,
> it being the object which it is the intelligence *(νόησις)* to contemplate 51e6-52a4

Thus, according to 51e6-52a4, the dyad of the intelligible Being, whether it is the dyad < One, Being > of the second hypothesis of the *Parmenides*, or the dyad <IE, IF>, cannot possibly generate a sensible by having a mass of the smaller part, acting as Content, enter a mass of the greater part, acting as Receptacle.

> whereas to the aid of the true Being (τῷ δὲ ὄντως ὄντι) [=the intelligible true Being]
> there comes the *accurately true Logos* (ὁ δι' ἀκριβείας ἀληθὴς λόγος),
> that so long as (ὡς ἕως ἄν)
> one thing is one thing, and another something different (τὸ μὲν ἄλλοῇ, τὸ δὲ ἄλλο),
> it will *never (ποτὲ) happen* that
> one thing will be *generated in the other thing* (οὐδέτερον *ἐν οὐδετέρῳ γενόμενον*)
> so that they *become* (γενήσεσθον) the same thing *(ταὐτὸν)*
> simultaneously *(ἅμα)* both one and two *(ἓν καὶ δύο).* 52c5-d1
> [[translation by Lamb, with modification by the authors]]

As explained in 5.1 this is so because for the *intelligible* dyad a,b, a mass Na or Mb makes no sense, and hence the entrance of the mass Mb of small b's into the mass Na of great a's makes even less sense.

5.2. The Content & Receptacle transformation of geometric anthyphairesis, described in Sect. 4, is quite attractive, and Plato would be happy to adopt it for his theory of the sensibles, in order to overcome the difficulty with infinite kosmoi and sensible TMA. But unfortunately, a platonic true Being is a *philosophical analogue* of a geometric periodic anthyphairesis between two magnitudes, and not a bona fide geometric one, and this presents a serious problem.

The arithmetical sequence of approximating generalized side and diameter numbers' tight inequalities we derived in Sect. 4,

$$\begin{aligned} \mathbf{M_{2n-1}b} \quad &< \mathbf{N_{2n-1}a} < (\mathrm{M}_{2n-1} + 1)\,\mathrm{b} \;\; (\textit{Dispersion}), \text{and,} \\ (\mathrm{N}_{2n} - 1)\mathrm{a} \quad &< \mathbf{M_{2n}b} < \mathbf{N_{2n}a} \;\; (\textit{Consolidation}), \end{aligned}$$

starting with a finite initial segment of the infinite anthyphairesis of a to b, were the result of perfectly allowable mathematical manipulations. These mathematical manipulations are valid, if a and b are magnitudes, but definitely make NO sense, if a and b are the parts of an indefinite philosophical dyad, such as the dyad <One, Being> in the *Parmenides*, the dyad <Being, Not-Being> in the *Sophistes*, and the intelligible dyad <IE, IF>. As already indicated above, it makes no sense to attempt to transform the intelligible anthyphairetic relation $\mathrm{IE} = \mathrm{IF} + \mathrm{IE}_1$, with $\mathrm{IE}_1 < \mathrm{IF}$, into the double remainder-free inequality IF <IE <2.IF of the form Content & Receptacle; more generally since M.IF, N.IE make no sense, the Content & Receptacle inequalities also make no sense, and thus we cannot get rid of the infinitely many intelligible remainders in the same way that we can get rid of the geometrical remainders.

Section 6 ***The geometric dyad of Triangle A and Triangle B (53c4-d7, 54a1-d7), and the generation of the surfaces of the four canonical solids, of cube from Triangle A, and of pyramid, 8hedron, and 20hedron from Triangle B (54d5-55c6)***

We describe the introduction of

Triangle A, the isosceles right angle triangle which is, half of the square, from which the *surface* of the cube/6hedron is generated, and
Triangle B, the orthogonal triangle having the ratio 2/1 of the hypotenuse to one side, half of the equilateral triangle, from which the **surface** of the pyramid/4hedron, 8hedron, and 20hedron are generated.

6.1. *The dyad Triangle A and Triangle B is introduced as the principle of the four primary bodies (Timaeus 53c4-d7, 54a1-d7).*

The two *principles* of the four primary bodies, triangles A and B, are obtained by a sequence of five reductions:

> –*First of all* (Πρῶτον μὲν), it is obvious to everyone (δῆλόν που καὶ παντί), that
> *fire, earth, water, and air* (πῦρ καὶ γῆ καὶ ὕδωρ καὶ ἀὴρ) are
> –*bodies* (ὅτι σώματά ἐστι); and all body (τὸ δὲ τοῦ σώματος εἶδος πᾶν) has
> – *depth* (βάθος). Depth (βάθος), moreover, must be (αὖ πᾶσα ἀνάγκη) bounded by
> – *surface* (ἐπίπεδον περιειληφέναι),

[here Plato limits consideration to rectilinear surfaces only, excluding the circle, cf. definition I.19 of Euclid's *Elements*]

> and any *surface* bounded by straight lines (ὀρθὴ τῆς ἐπιπέδου βάσεως) is composed of
> – *triangles* (τριγώνων). Now all *triangles* derive their origin from *two triangles*, each having one angle right and the others acute; and the one of these triangles has on each side half
> **a** *right angle* marked off by equal sides, while the other has the *right angle* divided into unequal parts by unequal sides.

[here Plato prepares the way to further limit consideration to the surfaces of the four of the five canonical solids, studied by Theaetetus in Book XIII of Euclid's *Elements*]

> These we lay down as *the principles (ἀρχὴν)* of *fire and all the other bodies*, proceeding according to a method in which the probable is combined with the necessary;
> –but *the principles (ἀρχὰς) which are still higher* than these
> are known only to God and the man who is dear to God. 53c4-d7

Definition
Triangle A is the isosceles right triangle (54a1-2).
Among the infinitely many scalene right triangles, Plato, guided by criteria and considerations of *beauty and fairness* (54a2-b1), determines that

Triangle B is the right triangle that generates the equilateral triangle, namely the one which has its hypotenuse twice as long as its lesser side (54a5-b1, 54b4-5, 54d5-7).

> Now of the two triangles, *the isosceles* possesses one single nature, 54a1-2
> but *the scalene (πρόμηκες)* an infinite number;
> and of these infinite natures we must select *the fairest (τὸ κάλλιστον),*
> if we mean to make a *suitable principle* (εἰ μέλλομεν *ἄρξεσθαι κατὰ τρόπον*).
> If, then, anyone can claim that he has chosen one that is *fairer (κάλλιον)* for the construction
> of these bodies, he, as friend rather than foe, is the victor. 54a2-5
> We, however, shall pass over all the rest
> and postulate as *the fairest (κάλλιστον)* of the triangles
> *that triangle* out of which, when two are conjoined,
> the *equilateral (ἰσόπλευρον) triangle* is constructed as a third. 54a5-b1
> and the other [triangle] that which always has the square on its greater side (τὴν μείζω πλευρὰν) *three times* the square (τριπλῆν κατὰ δύναμιν ἔχον) on the lesser side (τῆς ἐλάττονος) 54b4-5
> that triangle which has its hypotenuse (τὸ τὴν ὑποτείνουσαν)
> *twice as long* (διπλασίαν ἔχον μήκει) as its lesser side (τῆς ἐλάττονος πλευρᾶς) 54d5-7

The dyad of triangles A and B is a dyad of opposites, since Triangle A is characterized by *Equality* while Triangle B by *inequality* of their perpendicular sides. Together they form *the most beautiful* geometric dyad. Both features are basic; we will be able to explain in the sequel their significance (Sect. 8.4 for inequality, Sect. 10 for beauty).

6.2. Proposition *Triangles A and B are the* ***principles and elements*** *of (the surfaces of) four of the five canonical solids:*

The (surface of the) canonical 6hedron/cube is equal to 24A.
The (surface of the) canonical 4hedron/pyramid is equal to 24B.
The (surface of the) canonical 8hedron is equal to 48B.
The (surface of the) canonical 20hedron is equal to 120B (54d5-55c6).

Once the triangles A and B have been introduced and defined, Plato proceeds to generate from them, considered as principles and elements, the surface of the four of the canonical polyhedra. The four, of the five (the 12hedron is mentioned at 55c4-6), canonical polyhedra are constructed and *comprehended in a sphere*, the 4hedron/pyramid (in Proposition XIII.13 of the *Elements*), the 8hedron (in XIII.14), the 20hedron (in XIII.16), and the 6hedron/cube (in XIII.15). The comprehension of the canonical polyhedra in a sphere comes up in passage 58a4-7, analyzed in Sect. 10.

The surface of a pyramid consists of four equilateral triangles. An equilateral triangle is divided into two triangles of the form Triangle B; thus, the surface of a pyramid consists of 8 triangles B. But in fact, Plato divides the equilateral triangle into six rather than two triangles B, by dividing the equilateral triangle by means of the three heights. Thus, Plato divides the surface of the pyramid into 24 triangles B. By dividing

- The surface of the pyramid into 24 triangles B, rather than 8 triangles B, and
- The surface of the cube into 24 triangles A, rather than 12 triangles A,

Plato succeeds in dividing the pyramid and the cube into an *equal* number of their corresponding elements, with the result that the relation of the two fundamental regular solids (cube and pyramid) is precisely the same as the relation of their respective elements (Triangle A and Triangle B).

It certainly comes as a surprise that Plato treats throughout *Timaeus* the four canonical solids, **not** as solids with volume, but considering *ONLY their surfaces.* The reason will become evident in the next section.

Section 7 *The incommensurability, in fact periodicity of the anthyphairesis, of Triangle A to Triangle B (54b5-d3, 56d1-6)*

Plato insists on the exchangeable nature of the surfaces of the three canonical solids (4hedron, 8hedron, 20hedron), according to the arithmetical ratio 4:8:20 = 1:2:5, because they have a common unit, measure Triangle B, and are thus *commensurable* to each other,

but declares impossible the exchangeability presumably according to an arithmetical ratio, with the cube, because of the impossibility to exchange a mass of Triangle A with a mass of Triangle B (54b5-d3, 56d1-6). Such impossibility can only mean that the areas of the triangles A and B do not have a common measure, namely that they are *incommensurable*.

Note that these triangles are defined only up to their similarity equivalence class. Plato says nothing about the absolute size of these two triangles, and we might easily arrange so that some size of Triangle A is even equal to some other size of Triangle B. But how can it be established that Triangle A is *incommensurable* to Triangle B?

We now describe how we are able to establish the *incommensurability* of Triangle A to Triangle B, and *a fortiori* the incommensurability of the surface of the cube to the surface of each of the canonical solids pyramid, 8hedron, 20hedron. We were indeed able to discover the *canonical comparison* of the two triangles:

We consider Triangle A as the half of the square face of the cube.
Triangle B as the half of the equilateral triangle face of the pyramid.
We *comprehend* the pyramid and the cube *in the same sphere*, as Theaetetus does in Propositions (13, 15 &) 18 of Book XIII of the *Elements*; it can then be proved that the hypotenuse of Triangle A and of Triangle B are equal.

The canonical comparison of Triangle A and Triangle B and the resulting equalization of the hypotenuses of the two triangles reveal that the (linearized) ratio of the areas of the two triangles is indeed incommensurable, but nevertheless *commensurable in square*. The study of these *quadratic incommensurabilities* was started by the Pythagoreans, continued later by Theodorus, and culminated again by Theaetetus, as reported by Plato in the dialogues *Theaetetus*, *Sophistes*, and *Politicus*. As shown in Negrepontis 2018, Theaetetus developed the necessary

tools in Book X of the *Elements* and established the highly nontrivial theorem that the anthyphairesis of every quadratic incommensurability (namely, an incommensurability of lines, whose squares are commensurable) is *palindromically periodic*, a theorem imitated by Plato to define, by division and collection, the statesman in his dialogue *Politicus*.

Proposition The geometric dyads <Triangle A, Triangle B> and <surface of cube = 24A, surface of 20hedron = 120B> are incommensurable, and in fact in periodic anthyphairesis.

There are two crucial advantages gained by treating these canonical solids as surfaces:

The first is the incommensurability, and in fact the *anthyphairetic periodicity*, of the dyad <cube, 20hedron>, a fact whose importance will be seen in the generation of the sensible bodies (Sects. 8 and 9); and

The second is the mutual *commensurability* of the three solids 4hedron, 8hedron, and 20hedron, a fact that will allow Plato to subsume the traditional four body theory in modified form in his account of the sensibles (Sects. 8 and 9).

7.1. *The impossibility of exchanging a mass of triangles A with a mass of triangles B, (54b5-d3), or a mass of primary earth with a mass of primary water, or air, or fire, (56d1-6), can only mean the incommensurability of Triangle A to Triangle B.*

> Moreover, a point about which our previous statement was obscure must now be defined more clearly. It appeared as if the four Kinds *(τὰ τέτταρα γένη)*, in being generated (γένεσιν ἔχειν), all passed through one another into one another, but *this appearance was deceptive*.
> For out of the triangles [triangles A and B] which we have selected
> four Kinds are generated,
> *three of them* [4hedron, 8hedron, 20hedron] out of that one triangle which has its sides unequal [triangle B], and
> *the fourth Kind* [6hedron] *alone* composed of the isosceles triangle [triangle A].
> Consequently, *they are not all capable* of being dissolved into one another so as to form a few large bodies composed of many small ones, or the converse;
> *but three of them* do admit of this process (**τὰ δὲ τρία οἷόν τε**). For these three are all (ἅπαντα) naturally compounded of one triangle [[triangle B]], so that when the larger bodies ***(τῶν μειζόνων)*** are dissolved (λυθέντων) many small ones will form themselves from these same bodies (*πολλὰ σμικρὰ ἐκ τῶν αὐτῶν συστήσεται*), receiving (δεχόμενα) the shapes that befit them (τὰ προσήκοντα ἑαυτοῖς σχήματα); and conversely, when many small bodies (σμικρὰ ὅταν αὖ πολλὰ) are resolved (διασπαρῇ) into their triangles *(κατὰ τὰ τρίγωνα),* they will produce (ἀποτελέσειεν ἄν), when unified, one single large mass of another Kind (ἕκαστον αὐτῶν γέγονεν εἶδος). So let thus much be declared concerning their generation into one another. 54 b5-d3

The same exchangeability between the three canonical bodies and the same impossibility of exchange between each of the three and the fourth are described in physical terms for the four primary bodies (to be defined in Sect. 8 in terms of the

surfaces of the four canonical solids). In geometrical language, the rule of intergeneration of the three regular solids lies on the simple fact that their surfaces have as common unit Triangle B. Thus

4hedron = 24 Triangle B, 8hedron = 48 Triangle B, and 20hedron = 120 Triangle B.

Therefore,

2. 4hedron = 1. 8hedron, 5. 4hedron = 1.20hedron, and 1.20hedron = 2. 8hedron + 1. 4hedron.

> Given all we have said so far about *the kinds (τὰ γένη)*, the following account
> [of their generation to each other] is according to the image account (κατὰ τὸ εἰκὸς):
> When [primary] *earth* encountering [primary] fire and being *dissolved (διαλυθεῖσά)*
> by [primary] fire's sharpness is set in motion, being *dissolved (λυθεῖσα)*
> either within [primary] fire itself, or within a mass of [primary] air or [primary] water,
> until (μέχριπερ) its parts, encountering themselves somewhere, being refitted together (συναρμοσθέντα) again, become (γένοιτο) [primary] *earth*.
> The reason is that
> *the parts of [primary] earth will never pass into another [primary] kind (ε δος)* 56d1-6

Plato emphasizes that the "*earth will never pass into another kind.*" Since the generation "according to the triangles" (54c4) that works for the other three solids is rejected for the earth, we understand clearly that the impossibility of exchange of solid earth with the other three solid kinds is due to the impossibility of exchange of their surfaces. In effect, these exchanges are impossible, because

the exchange of Triangle A with Triangle B is impossible.

Plato states that the earth will never turn into any of the other three solids, meaning that such turning into is impossible in the sense of the generation according to the triangles, namely by *equating* the surface areas of the corresponding masses of earth and say water. This can be expressed as follows:

The surface area of a mass of earth/cubes will never be *equal* to the surface area of a mass of fire/pyramids. Since the surface area of a mass of earth/cubes is equal to M.A for some number M, and the surface area of a mass of fire/pyramids is equal to N.B for some number N, it follows that

there are no natural numbers M and N such that
the two areas M.A and N.B are equal, M.A = N.B.

Mathematically, this is expressed in terms of the concept of *commensurability* that appears in the Theaetetean Book X of the *Elements*:

Definition X.1 of Euclid's *Elements*.

Those magnitudes are said to be *commensurable (σύμμετρα μεγέθη)*
which are measured by the same measure (τὰ τῷ αὐτῷ μέτρῳ μετρούμενα), and those *incommensurable (ἀσύμμετρα)*
which cannot have any common measure (μηδὲν ἐνδέχεται κοινὸν μέτρον).

Thus Plato's statement that primary earth cannot be transformed to one of the other three primary bodies according to the triangles is equivalent to the statement that

(the area of) Triangle A is *incommensurable* to (the area of) Triangle B.

But how can it be established that Triangle A is *incommensurable* to Triangle B? Plato provides no information about the absolute size of these two triangles, and we might easily arrange so that some size of Triangle A is even equal to some other size of Triangle B. Thus Plato's claim of incommensurability must rely on a *canonical* way of comparing Triangle A to Triangle B.

7.2. *The **canonical comparison** of the triangles A and B is achieved by inscribing pyramid and cube **in the same sphere**, as in Propositions XIII. Total (13, 15 &) 18 of the Elements, by equating their hypotenuses, and the resulting **periodicity of the anthyphairesis of Triangle A to Triangle B**.*

Plato does not explain how these two elementary triangles are to be compared so that they are incommensurable, but since he clearly implies that they are incommensurable, he must have some specific "*canonical*" comparison in mind. In order to discover this canonical comparison, we must take into consideration that Plato is following the mathematical description of the canonical polyhedra, as presented in the Theaetetean Book XIII of the *Elements*.

Below we present that the canonical comparison of the areas of the triangles A and B is achieved (when the regular solids are comprehended in the same sphere) by equating *their hypotenuses.* We want to compare the hypotenuses of the two triangles, the hypotenuse p of the Triangle A is the *diameter* of the sphere comprehending the square of a cube, satisfies $p^2 = 2\,k^2$ (k the side of the square), and the hypotenuse q of the Triangle B is the side of the equilateral triangle of a pyramid. We recall that Plato calls "diameter" the hypotenuse of the triangle and builds the solids as inscribed in a sphere (55a3).

*Pyramid and cube constructed and comprehended in the **same** sphere:*

In Book XIII, these computations are first performed for each canonical solid separately, in Proposition XIII.13 for the pyramid and in Proposition XIII.15 for the cube, and then in Proposition XIII.18 all canonical polyhedra are comprehended and compared in *the same* sphere. These Propositions suggest the way for the canonical comparison of the areas of the triangles A and B. Triangle A will be compared to

Triangle B, when
Triangle A is considered as a constituent part of the surface of a cube.
Triangle B is considered as a constituent part of the surface of a pyramid.
AND these two solids, the cube and the pyramid, **are** *comprehended in the same sphere* (σφαίρᾳ περιλαβεῖν), as in fact in *Propositions XIII.13, 15, 18.* We obtain the following

Proposition
If a pyramid and a cube are comprehended *in the same sphere,* then the side of the triangle face of the pyramid, namely *the hypotenuse of triangle B is equal to* the diameter of the square face of the cube, namely *the hypotenuse of triangle A*.

Proof
Indeed, in the common sphere of diameter D,
the side q of the pyramid satisfies $q^2 = (2/3)D^2$
(by Proposition XIII.13 of the *Elements*), and
the side k of the cube satisfies $k^2 = (1/3)D^2$
(by Proposition XIII.15 of the *Elements*).

This allows for the *canonical comparison* of Triangle A and Triangle B: The hypotenuse of the Triangle B is the side q, and the hypotenuse p of the Triangle A satisfies (by the Pythagorean theorem I.47 of the *Elements*) $p^2 = 2\ k^2$.
Thus, we have $q^2 = (2/3)D^2 = 2\ (1/3)D^2 = 2\ k^2 = p^2$. Hence q = p.
Thus, we are led to compare the areas of triangles A and B *by identifying their hypotenuses.*

7.3. Proposition *Under the* ***canonical comparison*** *Triangle A and Triangle B are incommensurable. In fact, the anthyphairesis of any mass N.A to any mass M.B (where M and N are natural numbers) is periodic, hence infinite, hence the dyad NA, MB is incommensurable.*

This crucial property of the triangles A and B, and of the surfaces of any mass of primary earth/cube to the surfaces of any mass of, say, primary water, explains fully why Plato resorted to the *surfaces* of the regular solids, *not* considering at all *their volumes:* The reason lies exclusively with the remarkable fact that these surfaces are related to *periodic anthyphairesis*, and thus connect closely with *intelligible Being*.

Proposition

(i) (area of) triangle A > (area of) triangle B,
(ii) the area of triangle A is *incommensurable* to the area of triangle B,
(iii) if r and s are line segments with r:s = area of triangle A: area of triangle B, then the ratio $r^2:s^2$ is commensurable, and
the anthyphairesis of (area of) triangle A to (area of) triangle B is infinite, in fact periodic.

(iv) [Theaetetus][9] the *anthyphairesis* of (area of) any mass NA of triangles A (where N is any natural number) to (area of) any mass MB of triangles B (where M is any natural number), with MB < NA, is infinite, in fact *periodic.*
(v) The anthyphairesis of the surface of the cube to the surface of the 20hedron is periodic.

Proof
Ad (i). Triangle A and Triangle B have a common diameter, say D, as seen in the previous Proposition. Since Triangle A is isosceles and Triangle B scalene, it is clear that area of Triangle A > area of Triangle B.
Ad (ii). Setting x for side of Triangle A,
we have that the area of Triangle A is $x^2/2$; since $2x^2 = D^2$,
the area of the Triangle A is equal to $D^2/4$; and,
setting y,z for the sides of Triangle B,
we have the area of Triangle B is yz/2, and since $z^2 = 3y^2$ and $4y^2 = D^2$, it follows that $2y = D$, $y = D/2$, and $z^2 = 3y^2 = 3/4\ D^2$,
hence (in modern notation) $z = \sqrt{3}/2\ D$; thus,
the area of the Triangle B is equal to $\sqrt{3}/8\ D^2$.
Hence (area of) Triangle A: (area of) Triangle B = $2{:}\sqrt{3}$.
Thus, the ratio area of Triangle A: Area of Triangle b is incommensurable (by Theodorus' theorem); by Proposition X.2 of the *Elements*, the anthyphairesis of Triangle A to Triangle B is infinite.
Ad (iii). If r and s are line segments, such that $r{:}s = 2{:}\sqrt{3}$, then $r^2{:}s^2 = 4{:}3$, commensurable.
By Theaetetus' theorem, this anthyphairesis is (eventually) periodic.
Ad (iv). By Theaetetus' theorem.
Ad (v). Special case of (iv).

The incommensurability of the surface of the cube to the surface of the 20hedron is fundamental and absolutely necessary for the introduction of the Third Kind, defined and examined in Sects. 8 and 9, by means of which the infinite kosmoi are avoided.

[9] Theaetetus' Theorem states: If a and b are line segments, such that a > b, a,b incommensurable, and a^2, b^2 commensurable, then the anthyphairesis of a to b is palindromically periodic. In Negrepontis 2018, it is shown (a) how the study of the *Politicus* dialogue leads to the inescapable conclusion that Plato had perfect knowledge of this Theorem, and (b) how the Theaetetean Book X of the *Elements* contains all the tools necessary for the proof of this Theorem, namely the notion of the *apotome* line and of the line *of two names*, and principally Propositions X.112–114. A complete proof of the palindromic periodicity theorem using only the tools of Book X of the *Elements* is given in Negrepontis (Notes by V. Kleftaki) 2003.

Section 8 ***The four primary bodies defined in terms* of *the surfaces of the four canonical solids (55c7-57d6), toward the resolution of the problem of infinite kosmoi; the first argument, based on the fundamental passage 57d7-58a2, that the incommensurable dyad of primary earth PE and primary water PW constitutes, with the aid of the ratios PW to primary air PA = 5/2 and to primary fire PF = 5/1, the Third Kind, substituting the intelligible dyad, as the cause of all sensibles***

Let us reiterate that Plato, as explained in Sects. 4 and 5, is *obliged* (εἰσαναγκάζειν, 49a3), for the sole reason of ensuring the One Kosmos, to introduce in 48e-49a the Third Kind/Receptacle. In this section, we introduce *the four primary bodies*, defined by Plato by means of the corresponding surfaces of the four canonical solids (8.2), and present our interpretation of passage 57d7-58a2 (in 8.4-6), a fundamental passage that has not been understood properly so far, according to which the incommensurable, infinite anthyphairetic dyad of primary earth PE to primary water PW, aided by the commensurable dyads PW/PA = 5/2 and PW/PF = 5/1 (where PF is primary fire and PA primary air), is the *Third Kind*, coinciding with the Receptacle and *substituting* the intelligible dyad IE, IF, as *the cause and generator* of the sensible bodies.

Having prepared the ground, with the geometry of the (surface of the) four canonical solids and their commensurability or incommensurability, in 53c4-55c6, described in Sects. 6 and 7, Plato now is about to introduce, in 55d7-56b6, his most crucial *definition of the primary bodies* (ἄκρατα καὶ πρῶτα σώματα, 57c7).

8.1 *The seeming puzzle of Plato's return to the problem of infinite Kosmoi in* ***55c7-d6***

But precisely after 55c6 and before the definition of the primary bodies starting at 55d7, Plato interposes a seemingly unrelated paragraph 55c7-d6, in which he returns to the question he had first posed in 31a-b (passage given in Sect. 3) on whether there is one kosmos or infinite kosmoi.

Cornford 1937, finds the passage "extremely puzzling," because "There is nothing in the previous context to suggest a plurality of worlds…" (p. 220). "the ancients,…have found the passage as obscure as it is to us." (p. 221). Gregory 2001, pp. 42–43, calls the passage "rather puzzling."

But our analysis shows that the reopening of the question and problem on infinite kosmoi at this point is quite appropriate and justified, because Plato *exactly at this point*, with the definition of the primary bodies he is about to give, aims to succeed to get rid of the infinite Kosmoi and ensure the desired One Kosmos.

8.2. *The* ***definition*** *of the four primary and uncompounded four bodies (ἄκρατα καὶ πρῶτα σώματα, 57c7)*

> To [primary] *earth* let us assign the cubical kind (κυβικὸν εἶδος); 55d8

> Wherefore, we are preserving the image account (τὸν εἰκότα λόγον διασῴζομεν) when we assign this figure [*cube*] to [primary] *earth (γῇ)* 55e7-56a1

> Thus, in accordance with the right and the image (κατά τὸν εἰκότα) account, that solid which has taken the kind (εἶδος) of a *pyramid* shall be the element and seed of [primary] *fire* (πυρὸς στοιχεῖον καὶ σπέρμα); the second in order of generation [*8hedron*] we shall affirm to be [primary] *air (ἀέρος),*
> and the third [*20hedron*] [primary] *water (ὕδατος).* 56b3-6

Plato calls these four bodies the *uncompounded primary bodies* (ἄκρατα καὶ πρῶτα σώματα, 57c). There are some heuristic arguments in favor of these assignments, but in reality they form a

Definition
primary earth PE = the surface of the cube = 24A,
primary fire PF = the surface of the pyramid = 24B,
primary air PA = the surface of 8hedron = 48B, and
primary water PW = the surface of the 20hedron = 120B.

Thus the primary bodies are defined as *purely geometrical rectilinear areas* (55d7-56b6, *57c7-d6*).

8.3. *Convenient modifications of the primary bodies (57c7-d6)*

By definition PE = 24A. But Triangle A can be divided into two triangles of the same type, and we can continue dividing as long as we wish. Thus $A = 2^{n\cdot}A'$, where A' is a triangle of the same form, similar to A, and n is any natural number. Thus

$$PE = 24.2^{n\cdot\cdot}A'.$$

Also by definition PW = 120B. But Triangle B can be divided into three triangles of the same type, and we can continue dividing as long as we wish. Thus $B = 3^{m}B'$, where B' is a triangle of the same form, similar to B, and m is any natural number. Thus,

$$PW = 120.3^{m\cdot\cdot}B'.$$

Similar remarks hold for PA, PF. The incommensurability (and in fact periodicity of anthyphairesis) of PE to the other three primary bodies and the commensurability of the three, minus PE, primary bodies are not affected by considering triangles A and/or B greater or smaller in the above sense. This is the content of *Timaeus 57c7-d6.*

8.4. *We interpret the fundamental passage 57d7-58a2 according to which the inequality is the cause of all motion:*

The inequality refers to the inequality of the sides of Triangle B, and in fact to the incommesurability of the dyad Equality of Triangle A to inequality of Triangle B; it is the cause of all sensible motion, substituting for the intelligible Being.

There results the interpretation of the Receptacle in terms of the primary bodies, and a first argument identifying the Third Kind in terms of the primary bodies.

8.4.1. *The fundamental passage Timaeus 57d7-58a2*

> Now, unless we can arrive at some agreement concerning Motion and Rest,
> as to how and under what conditions they are generated,
> our subsequent argument (λογισμῷ) will be greatly hampered (ἐμποδὼν).
> The facts about them have already been stated in part;
> but in addition thereto we must state further the following, that
> motion never consents to exist in a state of *non-opposition* (ἐν μὲν ὁμαλότητι).
> For it is difficult, or rather impossible,
> for that which is to be moved (τὸ γὰρ κινησόμενον) to exist without that which is to move (τοῦ κινήσοντος), or that which is to move without that which is to be moved;
> but in the absence of [the dyad of] these there is no motion,
> and that these should ever be unopposed (ὁμαλὰ) is a thing impossible.
> Accordingly, we must always (**ἀεὶ**) place
> rest in the state of non-opposition (ἐν ὁμαλότητι), and
> motion in the state of opposition (εἰς ἀνωμαλότητα); and
> *the cause (αἰτία)* of the nature of the state of opposition (τῆς ἀνωμάλου φύσεως) is *inequality (ἀνισότης)*. We have already explained the generation of *inequality*. 57d7-58a2

Thus for *motion* (κίνησις) to be possible, there must be
a dyad in a state of opposition (εἰς *ἀνωμαλότητα*), namely a dyad consisting of
that part which is to move (τὸ κινῆσον), namely the active part of the motion, and
that part which is to be moved (τὸ κινησόμενον), namely the passive part of the motion;
the cause of a dyad in opposition, and thus of motion, is *inequality* (*αἰτία* δὲ ἀνισότης αὖ τῆς ἀνωμάλου φύσεως, 58a1).

Beginning with Cornford, scholars have seriously misunderstood the meaning of inequality, of state of opposition, and of inequality being the cause of opposition, in this fundamental passage.[10] Because of this misunderstanding, scholars failed to realize that passage 57d7-58a2 provides an understanding of the Third Kind, in particular identifies the Receptacle and the primary bodies. We now examine this passage.

[10] Cornford 1937, p. 240: "Now, however, the bodies have received their shapes, and these shapes are unequal in size in two ways: each kind of body has a particle of a different size from the rest, and there are several grades of each kind. This may be what is meant by inequality as the cause of heterogeneity."

Gregory and Waterfield 2008, p. 146: "presumably Plato is thinking of the 'inequality' between all the different elementary triangles."

As the term "inequality" does not appear at all in 57c7-d6, these are arbitrary interpretations, making it impossible to appreciate the fundamental role of incommensurability in Plato's Third Kind.

8.4.2. *Inequality, in passage 57d7-58a2, interpreted as anthyphairetic infinity and incommensurability*

Essential help comes from Aristotle and Simplicius.

Aristotle in *Physics* 201b18-26 states that the cause of platonic motion is *inequality* (ἀνισότητα), adding the *Not-Being* (μὴ ὄν, 201b20) and the *indefinite* (ἀόριστόν, 201b24) (ἑτερότητα καὶ *ἀνισότητα* καὶ *τὸ μὴ ὂν* φάσκοντες εἶναι τὴν κίνησιν; *ἀόριστόν* τι δοκεῖ εἶναι ἡ κίνησις);

Eudemus, quoted by Simplicius *In Physics* 431, 6–12, adds to the platonic cause of motion, next to inequality, not-Being, and opposition, also the indefinite dyad *the great and small* (τὸ μέγα καὶ μικρὸν) ("Πλάτων δὲ τὸ μέγα καὶ μικρὸν καὶ τὸ μὴ ὂν καὶ τὸ ἀνώμαλον καὶ ὅσα τούτοις ἐπὶ ταὐτὸ φέρει τὴν κίνησιν λέγει").

Simplicius, *In Physics* 430, 11–12, clarifying Aristotle, points out that Plato identifies the *Not-Being* (μὴ ὄν) as the cause of motion in the *Sophistes* (καὶ δέ, φησί, τὴν κίνησιν ὁ Πλάτων λέγει ἐν Σοφιστῇ); in 453, 25–455, 14 he connects the great and the small (453, 27) with *the indefinite dyad* (ἀόριστος δυάς, 453, 26), and *the infinite* (ἄπειρον, 453, 27–28). It becomes clear from these passages that inequality, the cause of platonic motion, is a *Philebean infinite* (*Philebus* 23b5-25e2).

But now, the Philebean Infinite has been interpreted by Negrepontis 2005, (B), (D), (E), as being the *anthyphairetic infinity*, an interpretation fully supported by Proclus.[11]

We note that Aristotle and commentators on his work, even if strongly critical of Plato (e.g., Simplicius, *In Physics* 431, 6–12, cites Eudemus stating that it is preposterous (ἄτοπον) and claiming that it is ridiculous (γελοῖον) for Plato to maintain that the *cause* of motion is inequality), nevertheless fully understand and describe faithfully Plato's views.

In sum, according to the crucial comments by Aristotle, Eudemus, and Simplicius, and our interpretation, the platonic cause of motion, identified as *inequality*, is *anthyphairetic infinity*, and thus (by Proposition X.2 of the *Elements*) *incommensurability*.

8.4.3. *The inequality in passage 57d7-58a2 refers to the inequality of the sides of Triangle B; it is an Incommensurability in relation to, and completing, the Equality of Triangle A, as shown in Sect. 7.*

In order to clarify the meaning of "inequality," we must examine all the earlier (i.e., before 58a2 in the *Timaeus*) occurrences, where Plato, as he himself states, has dealt with inequality. The prior occurrences of words with the root *ἀνισ-* are in 53d3,

[11] "and if [Platonic, Philebean] infinite did not exist (καὶ τῆς μὲν ἀπειρίας οὐκ οὔσης),
all magnitudes would be commensurable with each other (τά τε μεγέθη πάντα σύμμετρα ἂν ἦν), and
there would be nothing irrational (οὐδὲν ἄρρητον) nor "alogon" (οὐδὲ ἄλογον)"
Proclus, *Comments to the First Book of Euclid's Elements* 6, 19–21 [our translation]

53d4, 54c2, and 55e5, and they all have to do with the inequality of the sides in the right angle of the Triangle B.

Now inequality, on the one hand, has been interpreted in 8.4.2 as an incommensurability and an anthyphairetic infinity and, on the other hand, as the inequality of the sides in Triangle B. But incommensurability makes sense only for a dyad, not for a single entity, such as Triangle B. The parallel that Aristotle and his commentators draw of inequality with *Sophistes*' Not-Being (8.4.2) is helpful. The Not-Being *completes* the Being, and together they form an intelligible dyad, hence, according to Sect. 1, a dyad in philosophic infinite anthyphairesis.

Thus the "inequality" of Triangle B completes the "equality" of Triangle A, and together they form an incommensurable dyad,

as Not-Being completes Being in the *Sophistes*, and

as Being completes One in the second hypothesis of the *Parmenides*.

Thus, the claim in passage 57d7-58a2 that

Inequality is a state of continued (anthyphairetic) opposition (ἀνωμαλότης)

is equivalent to the statement that

the dyad Triangle A and Triangle B is incommensurable.

Indeed, inequality of Triangle B produces anthyphairetic infinity and *incommensurability* when (anthyphairetically) opposed to Triangle A. We recall that the incommensurability of Triangle A, B was hinted by Plato, and we established it in Sect. 7; now, we see that it reappears in 57d7-58a2 with fundamental importance, as the cause of sensible generation and motion.

8.4.4. *Plato's choice for the dyad Equality, inequality, the cause and generator of all sensible bodies in 57d-58a, is the dyad primary earth PE to primary water PW.*

The statement in 57d-58a that *the cause* of every motion is inequality, namely as explained in the previous Sect. 8.4.3, the Incommensurability of Triangle A to Triangle B, according to our interpretation, is remarkable (This is exactly what the Aristotelian Eudemus calls preposterous and ridiculous).

But how can the incommensurability of the dyad Triangle A to Triangle B be the cause of all sensible motion?

The cause for all motion that Plato has in mind is not necessarily the dyad Triangle A to Triangle B; besides the dyad Triangle A, Triangle B itself, certain dyads of primary bodies exhibit such incommensurability.

Since each of the dyads <PW, PA>, <PW, PF>, and <PA, PF> consists, in any conceivable combination of sizes, of entities that contain *only Triangle B,* it cannot constitute an incommensurable dyad, and hence it must be rejected as the cause of generation or motion.

But according to the Main Proposition of Sect. 7, each of the remaining three dyads of primary bodies

PE, PW; PE, PA; and PE, PF,

is inextricably linked with inequality, is *incommensurable* in the sense of Equality versus inequality, in fact in periodic anthyphairesis, and thus is a possible candidate for the cause for the generation of motion.

Thus, in total there are four conceivable dyads for which the opposition inequality versus Equality materializes, namely

$$< \text{Triangle A}, \text{Triangle B} > ; < \text{PE}, \text{PF} > ; < \text{PE}, \text{PA} > ; \text{and}, < \text{PE}, \text{PW} > .$$

The first dyad, the geometrical dyad Triangle A, Triangle B itself, should be excluded. We might base this exclusion on the fact that we expect the cause of sensible motion to be some dyad with fundamental physical significance; and also because the dyad of triangles includes many other instances of incommensurabilities beyond those produced by the "physically meaningful" primary bodies; but the main reason for excluding this dyad is the fact that, as we shall see in Sect. 10, *only* the primary bodies *participate* in the intelligible, *not* these triangles.

We are thus naturally directed to seek the all-important dyad/cause among the three dyads that may be formed from the four *primary bodies*.

At this point however we must take into account the restriction that Plato sets early: primary earth cannot interact directly with primary fire (and with primary air) but only through the intermediary of primary water.

> Hence, in beginning to construct the body of the World,
> God was making it of fire and earth (ἐκ πυρὸς καὶ γῆς).
> But it is not possible that two kinds alone should be conjoined (καλῶς συνίστασθαι)
> without a third; for there must needs be some intermediary bond (δεσμὸν *ἐν μέσῳ*)
> to connect (συναγωγὸν) the two (ἀμφοῖν) 31b6-c2

> what conjoins (συναρμόττουσιν) solids is never one intermediary term alone but always two
> *(μία μὲν οὐδέποτε, δύο δὲ ἀεὶ μεσότητες).*
> Thus it was that in the midst (ἐν μέσῳ) between fire and earth God set water and air 32b2-4

Passage 60e4-61a1, in Sect. 9, Appendix 2, is far more explicit, confirming that there is no direct interaction of PE with PA or PF, but only through PW.

Thus primary earth PE is in direct contact and interaction *only* with primary water PW, and *not* with PA or PF. Interaction between PE and PA or PF is possible only indirectly as a *composition* of the interaction of PE with PW followed, by the use in some form of the commensurable ratios PW/PA = 5/2 and PW/PF = 5/1. It follows that the dyad materializing the incommensurability of Equality versus inequality that is the cause for all motion, according to passage 57d-58a, can only be the dyad PE, PW of primary bodies.

8.4.5. *The sense in which* ***the dyad PE, PW*** *is* ***the "cause of motion"*** *in passage 57d-58a*

In order to understand the use of this term "cause for motion," let us note that,

on the one hand, as we saw in Sect. 2, in pre-*Timaeus* theory of sensibles, a sensible is fully described and known by an initial *finite* segment of the *infinite*

anthyphairesis of the intelligible dyad <IE, IF> (so that a sensible body is the end result of an initial *finite* segment of the *infinite* intelligible anthyphairesis of IE to IF, namely *a sensible body is generated* by an initial finite segment of the intelligible anthyphairesis of IE to IF);

on the other hand, according to traditional, pre-*Timaeus* Plato's dialectics, e.g., in *Phaedo* 100c, the intelligible Being is *the cause* of every sensible body participating in it (Sect. 2.4).

But an initial finite segment of an (intelligible) anthyphairesis describes a motion, exactly as stipulated in the 57d-58a passage:

The first step of the motion is realized by the initial dyad <IE, IF>, the part IF as the active part (that which moves) of the motion, and the part IE as the passive (that which is moved); the anthyphairesis of this dyad generates the part IE_1, and the second step of the motion is realized by the dyad <IF, IE_1>, with IE_1 as the active (moving) part and IF the passive (moved), and so on.

After a finite, say n, number of steps, we will have completed a motion, whose end product is *the generation* of a sensible body, according to Sect. 2; and furthermore a change from one sensible, say generated at stage n, to another sensible, say generated at a later stage m, is also described by a motion.

Thus, the intelligible Being, in fact the intelligible dyad with its *infinite* anthyphairesis, is *the cause* of all sensible bodies that participate in it, and of the motions that generate them and transform one of them into the other, all completed in an initial *finite* number of steps of the infinite intelligible anthyphairesis.

We expect that the meaning of *the cause* of all motion that is claimed in 57d-58a for the inequality, but eventually, following our argument in 8.4.4, for *the dyad PE to PW*, is analogous; in other words, we expect that every sensible body will be fully determined and known by an initial *finite* segment of the *infinite* anthyphairesis of the dyad PE to PW.

8.4.6. *It follows from 57d-58a that*

the generation of the sensible bodies in the Receptacle (49a-57c) is ***caused*** *by the incommensurability of the primary bodies and is equivalent to the generation of the sensibles by means of "diakena" and "hedrai" (58bff);*

the ***Third Kind*** *consists of the incommensurable dyad primary Earth PE, primary Water PW, assisted with the commensurable ratios PW/PA = 5/2 and PW/PF = 5/1.*

It follows, from passage 57d7-58a2, that every motion, hence the motion of the Receptacle (by which all sensible bodies are generated), is governed by the incommensurability of the primary bodies. Indeed:

First, the motion of the Receptacle mentioned in 57c2-6, and in fact every motion of the Receptacle mentioned in 49a-57c, *before* the crucial passage 57d-58a, e.g., 52e1-5, is really a motion involving not just the Receptacle, since every motion is possible only if there is a dyad of opposites, but a dyad, and this dyad can only be the Receptacle & its Content.

Second, *the cause* of the motion of the Receptacle & Content is, as is of every motion, the inequality, which, according to the arguments developed just above, takes the form of the *incommensurable* dyad of primary bodies PE, PW.

Third, by our interpretation of the term "cause," the sensible generated by the motion of the Receptacle & Content is fully described by a *finite* initial segment of the *infinite* anthyphairesis of PE to PW.

On the other hand, every motion described *after* the passage 57d7-58a2 never refers to the Receptacle, but to the "diakena" of the greater primary body, and is in fact described as the motion of a dyad of primary bodies; this motion has also the same cause, namely the *incommensurable* dyad PE, PW, and is thus fully described by a *finite* initial segment of the *infinite* anthyphairesis of PE to PW.

These remarks, depending on the fundamental passage 57d-58a and our interpretation of inequality and cause, lead to the identification of the Receptacle, before passage 57d7-58a2, and the "diakena" of the primary bodies, after passage 57d7-58a2.

Passage 57d7-58a2 thus serves as the connecting link between the description of the Third Kind as Receptacle (49a-57c), mentioned at 57c3 for the last time, and as "diakena" and "hedrai" (58b -62a ff), mentioned at 58a4 for the first time, of the greater primary body (For the relation between "diakena" and "hedrai," see Sect. 9, Appendix A).

Thus the correct interpretation of passage 57d7-58a2, something not done before, shows that the mysterious Receptacle is fully accounted by the action of the primary bodies.

Receptacle	*"diakena" of primary body*	*"hedrai" of primary body*
49a6, 50b6, d3, e3, 51a5, 52a2, d5, d6 ff, 57c3		
57d7-58a2		
	58b3, b5	
		59d6, 60c4
	60e5, 61a5, b1, b4	
		62a8
		[79b4, b6, 80c5]

But our interpretation of the fundamental passage 57d-58a leads us close to an understanding of the Third Kind: Indeed, what would be a better way to describe the Third Kind than by the requirement that it substitutes (for) the intelligible as *the cause* and generator of all sensible bodies.

Thus *the Third Kind* consists of the four primary bodies, in the following arrangement: The principle constituent of the Third Kind is the incommensurable dyad of primary earth PE to primary water PW, substituting (for) the intelligible dyad IE, IF, aided by the auxiliary commensurable dyads PW/PA = 5/2 and PW/PF = 5/1.

The dyad PE, PW generates all waterlike or earthlike sensibles, as initial anthyphairetic approximations of the anthyphairesis of PE to PW. But the airlike and firelike sensibles are in fact completely determined and generated as well by the

dyad PE to PW, if we take into account the auxiliary ratios PW/PA = 5/2 and PW/PF = 5/1.

We will confirm and understand in depth this description of the generation of all sensibles from the primary bodies in Sect. 11 (The exact description of the Third Kind is given in Sects. 10 and 11).

Section 9 *The description of the generation of the Timaeus sensibles in terms of the "diakena" of the primary bodies in 58a4-c2, 61a7-b3 is equivalent to the corresponding description in terms of the Receptacle in 50a-51a, 52d4-53a6, and 57a7-c6, revealing the equivalent descriptions of the Third Kind by the "diakena" of the primary bodies and by the Receptacle, and the mathematical origin of the term Receptacle.*

From our interpretation of passage 57d-58a, in Sect. 8, we have concluded that the Content is a mass of smaller primary bodies that enter the Receptacle, a mass of greater primary bodies, thus showing that the primary bodies and the Receptacle are equivalent descriptions of the Third Kind.

In the present section, we provide a second, more detailed, argument showing the equivalence of the Receptacle and the primary bodies. This argument can be sketched as follows:

In 58a4-c2, Plato describes the generation of the sensibles and their motions in terms of the "diakena" of primary bodies (Sect. 9.1).

This process consists in four steps, as follows:

[1]. A mass of the small primary body, say PW, enters the "diakena" of a mass of the great primary body, say PE.
[2]. In fact not just enters, as in [1], but fills, leaving not even one part empty.
[3]. The result of [2] is either that the great consolidate the small, making the defeated small dissimilar to themselves and similar to the victorious great, or that the small disperse the great, making the defeated great dissimilar to themselves and similar to the victorious small, forming in both cases a sensible body, in which the defeated stays together with the victorious in an entity both one and two (or "confounded" dyad, cf. 11.4.1 below).
[4]. The sensible thus formed by [3] moves to the proper place of the victorious part.

This elaborate process of four steps is essentially repeated in with the "diakena" of the primary bodies replaced by the Receptacle (Sect. 9.2). The four steps, by means of which the generation of the sensibles and their motion takes place in and by the Receptacle, are so uncannily similar to the four steps, by means of which the generation of the sensibles and their motion takes place in and by the primary bodies, that the conclusion that the Receptacle and the primary bodies are equivalent descriptions of the Third Kind is inescapable. We thus identify the Content with the smaller primary body and the Receptacle with the "diakena" of the greater primary body (In the *Appendix*, we show that the "diakena" of a primary body further coincide and are equal to the faces/hedrai of that primary body).

9.1. *The generation of a sensible body as a dyad of a small primary body entering and filling the "diakena" of the greater primary body, in such a way that*

either the great body consolidates the small body, the great body is *minimum for the given small body, the small body becomes similar to the great body and dissimilar to itself, and*

the dyad as a unit moves to the place of the great body

or the small body disperses the great body, the small body is maximum for the given great body, the great body becomes similar to the small body and dissimilar to itself, and

the dyad as a unit moves to the place of the small body.

9.1.1. [1] *A mass of the smaller* ***primary*** *body enters the* ***"diakena"*** *of a mass of the greater* ***primary*** *body* (58a7-b6).

Wherefore (διὸ δὴ),
[primary] *fire* most of all has permeated all things (εἰς ἅπαντα *διελήλυθε* μάλιστα),
and in a second degree [primary] *air*, as it is by nature second in fineness,
(ἀὴρ δὲ δεύτερον, ὡς λεπτότητι δεύτερον ἔφυ);
and so with the rest (καὶ τἆλλα ταύτῃ)
for those [primary bodies] that have *the greatest constituent parts*
(τὰ γὰρ *ἐκ μεγίστων μερῶν* γεγονότα)
have the *largest empty place* left in their construction
(μεγίστην *κενότητα,* 58b3, ἐν τῇ συστάσει παραλέλοιπεν),
and those that have *the smallest* (τὰ δὲ σμικρότατα) [constituent parts]
the least (ἐλαχίστην) [empty place].
58a7-b4
. . .
Therefore, when *small* [primary] bodies are placed beside *great* [primary bodies]
(σμικρῶν οὖν παρὰ μεγάλα τιθεμένων) 58b6

Thanks to 58a5, according to which the passage in question refers to the Kinds (τὰ γένη, 58a5) which were *comprehended (συμπεριέλαβεν)* by the sphere of the World, these Kinds are uniquely determined as the four *primary bodies* (cf. *σφαίρᾳ περιλαβεῖν*, Book XIII. of the *Elements*), and thanks to the fundamental passage 57d7-58a2, analyzed in Sect. 8, the masses in the present passages are masses of *primary bodies*.

We must *not* conclude from passage 58a7-b6 that the particles of primary fire PF or primary air PA can enter and fill the "diakena" of primary earth PE; in fact, this is explicitly excluded in passage 60e4-61a1.

Masses of *[primary] earth* (γῆς ὄγκους) are
not dissolved (οὐ τήκει) by *[primary] fire or [primary] air*;
because their particles are smaller (σμικρομερέστερα πεφυκότα),
than the [primary] earth's "diakena",
so that, having plenty of room to pass through (διὰ πολλῆς εὐρυχωρίας ἰόντα),
not violently (οὐ *βιαζόμενα*),
they do not dissolve (ἄλυτον) [primary] earth but leave it undissolved (ἄτηκτον)
whereas the particles of *[primary] water*, having greater parts (μείζω. . . μέρη),

make their exit through violently (*βίαιον* ποιούμενα τὴν διέξοδον)
and so break (λύοντα) the [primary] earth and dissolve (τήκει) it.
[Primary] earth, then,
namely earth that has not been formed violently into sensible earth (**ἀσύστατον** ὑπὸ **βίας**),
is dissolved in this way by [primary] water only, 60e4-61a1

Thus, the typical case is when the great body is primary earth PE and the small body is primary water PW (The other possibilities will be examined in Sect. 11). We express mathematically the passage by an inequality as follows:

A mass M.PW of M particles of primary water enters into the "diakena" of a mass N.PE of the greater primary earth (where M and N are some natural numbers), so that

$$\text{M.PW} \leq \text{N.PE}.$$

Now we note that we cannot possibly have Equality, since, according to Sect. 7, PE and PW are incommensurable; thus, in fact

$$\text{M.PW} < \text{N.PE}.$$

for some natural numbers M, N.

9.1.2. [2] *The **smaller primary body** enters, fills quite fully, and tightens **the "diakena" of the greater primary body** in such a way that either the smaller primary body is Maximum, or the greater primary body is Minimum* (58b4-5,a7, 617b3).

Thus the process of *tightening* (ἡ δὴ τῆς *πιλήσεως* σύνοδος)
compresses together (συνωθεῖ)
the small [primary] bodies (*τὰ σμικρὰ*) into (*εἰς*)
the empty places (τὰ διάκενα, 58b5) of the great (τῶν μεγάλων) [primary bodies]. 58b4-5

In passage 58b4-5, the process of tightening and compressing sets the mass of small primary body in the "diakena" of the great primary body.

We now come to an important and misread passage.

and thus it allows not even one part (οὐδεμίαν)
of empty place (κενὴν χώραν) to be left (ἐᾷ λείπεσθαι) 58a7.

The key expression here is

"οὐδεμίαν κενὴν χώραν" 58a7.

We claim that it should be rendered,
not as "no empty space," as it is customary,
but more literally and separately, as two words,
as "not even (οὐδὲ) one (μίαν) empty space."

There is an analogous passage with the terms *οὐδὲν* and *μηδὲν* in the *Parmenides* 144c2-6:

—What about this: Is there any among them which
is part of being (ἔστι μὲν μέρος τῆς οὐσίας), and yet not one **(οὐδὲν)** part?
— ᾽How could there be?
—Rather, I suppose, since it is, so long as it is
it must always be some *one (ἕν) thing*;
it cannot be *not one (μηδὲν δὲ ἀδύνατον)*. *Parmenides* 144c2-6
(Allen 1997, p. 28, modified by the author)

Here (οὐδὲν) and μηδὲν make sense only if they are rendered more literally and separately, as two words, as not (οὐδ) one (ἐν), and as not (μηδ) one (ἐν).

It is emphatically clear that both terms (οὐδὲν, μηδὲν) cannot be rendered as "nothing," but only as "not even one" (cf. Negrepontis TA).

According to our interpretation of "οὐδεμίαν κενὴν χώραν," "not even one empty space is left out." The passage, according to this interpretation, states that

namely, not even one part of the smaller primary body PW can be added,

namely, we cannot do with a greater mass of small body than M.PW,

namely, adding even one part PW to the mass of smaller body M.PW, and obtaining thus the mass (M+1).PW, this mass cannot be entered into the greater mass.

With this interpretation of *"οὐδεμίαν κενὴν χώραν,"* we can now connect in a satisfying way with the tight double inequality corresponding to Dispersion, in Sect. 4.

Indeed, let us assume, as in 9.1.1, that a mass M.PW of primary water enters into the "diakena" of a mass of the greater primary earth N.PE, so that

$$\text{M.PW} < \text{N.PE}.$$

The mathematical inequality can now be completed, in accordance with our present interpretation,

not even one part of *the smaller primary body*, namely of M.PW, can be added;

to obtain the following double mathematical inequality:

$$\text{M.PW} < \text{N.PE} < (\text{M} + 1).\text{PW}.$$

We can express this situation by saying that the mass of the small primary body is *Maximum* for the given mass of the great primary body.

The filling and forcible tightening **(***βίᾳ συμπεπιλημένα***)** of the "diakena" of a mass of particles of primary earth PE with a mass of particles of primary water PW in the 61a7-b3 passage confirms the interpretation we have just given.

As regards the classes of bodies (τὰ δὴ τῶν σωμάτων),
which are *compounds (συμμείκτων)* of earth and water (ἐκ γῆς τε καὶ ὕδατος)
so long as the water (ὕδωρ) occupies (κατέχῃ)
the interspaces of earth (τὰ τῆς γῆς διάκενα) which are
forcibly tightened (βίᾳ συμπεπιλημένα),

the portions of water (τὰ μὲν ὕδατος μέρη) which approach from without (ἐπιόντα ἔξωθεν)
find no entrance (εἴσοδον οὐκ ἔχοντα),
but flow round the whole mass (περιρρέοντα τὸν ὅλον ὄγκον)
and leave it undissolved (ἄτηκτον εἴασεν). 61a7-b3

But we now cannot fail to observe, with satisfied surprise, that this is the exact same double inequality we have obtained for the odd stage anthyphairesis of the dyad PE, PW.

We may now complete the picture by considering the dual case (it is not stated in the Timaeus for the "diakena," but it is hinted for the Receptacle, cf. 9.2.2 below): The mass of the great body is being filled quite fully by the mass of the small body in the sense that

not even one part of *earth*, N.PE, can be excluded;
this can only mean that

$$(\mathrm{N}-1).\mathrm{PE} < \mathrm{M.PW} < \mathrm{N.PE},$$

exactly the double tight inequality that corresponds to the even stages of the infinite anthyphairesis of PE to PW that we have found in Sect. 4.

We can express this situation by saying that the mass of the great primary body is *Minimum* for the given mass of the small primary body.

Thus, a mass of great primary bodies is filled quite fully by the mass of smaller primary bodies. We obtain the double tight inequalities

$(\mathrm{N}-1).\mathrm{PE} < \mathrm{M.PW} < \mathrm{N.PE}$, or $\mathrm{M.PW} < \mathrm{N.PE} < (\mathrm{M}+1).\mathrm{PW}$, respectively.

(equivalently, in convergent form,

$\mathrm{M/N} < \mathrm{PE/PW} < \mathrm{M}/(\mathrm{N}-1)$, or $\mathrm{M/N} < \mathrm{PE/PW} < (\mathrm{M}+1)/\mathrm{N}$, respectively).

These inequalities coincide with those obtained in the Basic Proposition of Sect. 4, where they were shown to be equivalent to a finite initial system of the infinite anthyphairesis of PE to PW of even or odd length, respectively, thus to a description of the *Timaeus* (waterlike or earthlike) sensibles bodies.

9.1.3. [3] *Interpretation of Dispersion and Consolidation in terms of the double inequalities of the basic Proposition in Sect. 4: the generation of the sensible bodies in the form of a dyad*

either in Dispersion (the smaller primary body is active, victorious)
or in Consolidation (the greater primary body is active, victorious),
the defeated taking the form of the victorious in both cases,
the sensible body consisting of the ***smaller and greater primary body****, as one body, moves to the proper place of the victorious body*
(58b6-7, 58d4-e2)

and the smaller (τῶν ἐλαττόνων) *disperse (διακρινόντων)* the greater (τὰ μείζονα)
while the greater (ἐκεῖνα) *consolidate (συγκρινόντων)* the smaller (τῶν μειζόνων) 58b6-7

It is clear from the very form of the double tight inequality

$$(\mathrm{N}-1).\mathrm{PE} < \mathrm{M}.\mathrm{PW} < \mathrm{N}.\mathrm{PE}$$

in Sect. 9.1.2 that

a mass of the great primary body of earth PE is the *active* and the *victorious* part,

the defeated mass of the small primary body of water PW becomes approximately *assimilated* to PE,

and thus we witness the *generation* of an *earthlike sensible body.*

The process, *Consolidation (σύγκρισις), is*, as we saw in Sect. 4, the remainder free representation of every first initial segment of an even 2n number of steps of the infinite anthyphairesis of PE to PW in terms of the generalized side and diameter numbers, namely the modern "convergents":

$$(\mathrm{N}_{2\mathrm{n}}-1).\mathrm{PE} < \mathrm{M}_{2\mathrm{n}}.\mathrm{PW} < \mathrm{N}_{2\mathrm{n}}.\mathrm{PE}.$$

It is clear from the very form of the double tight inequality

$$\mathrm{M}.\mathrm{PW} < \mathrm{N}.\mathrm{PE} < (\mathrm{M}+1).\mathrm{PW}$$

in Sect. 9.1.2 that

a mass of the small primary body of water PW is the *active* and the *victorious* part,

the defeated mass of the small primary body of earth PE becomes approximately *assimilated* to PW,

and thus we witness the *generation* of a *waterlike sensible body.*

The process, *Dispersion (διάκρισις), is*, as we saw in Sect. 4, the remainder free representation of every first initial segment of an odd 2n − 1 number of steps of the infinite anthyphairesis of PE to PW in terms of the generalized side and diameter numbers, namely the modern "convergents":

$$\mathrm{M}_{2\mathrm{n}-1}.\mathrm{PW} < \mathrm{N}_{2\mathrm{n}-1}.\mathrm{PE} < (\mathrm{M}_{2\mathrm{n}-1}-1).\mathrm{PW}.$$

We introduce the following notation:
$\boxed{\mathbf{PW}} < \mathbf{PE}$ to mean a dyad $\mathrm{M}.\mathrm{PW} < \mathrm{N}.\mathrm{PE}$ in Dispersion
$\mathrm{M}.\mathrm{PW} < \mathrm{N}.\mathrm{PE} < (\mathrm{M}+1).\mathrm{PW}$, generating a waterlike sensible
$\mathbf{PW} < \boxed{\mathbf{PE}}$ to mean a dyad $\mathrm{M}.\mathrm{PW} < \mathrm{N}.\mathrm{PE}$ in Consolidation
$(\mathrm{N}-1).\mathrm{PE} < \mathrm{M}.\mathrm{PW} < \mathrm{N}.\mathrm{PE}$, generating an earthlike sensible.

The kinds of [[*sensible*]] water ***(τὰ δὲ ὕδατος)*** are, primarily, two *(διχῇ),*
the one being the *liquid* (τὸ μὲν ὑγρόν), the other the *fusible* (τὸ δὲ χυτὸν) kind.
Now *the liquid (ὑγρὸν)* kind, inasmuch as it partakes of
those small (σμικρά) *unequal* (*ἀνίσων*) particles of [primary] water,
is *mobile* both in itself and by external force
owing to its *state of opposition* (διὰ τὴν ἀνωμαλότητα)
and the shape of its *figure* (τὴν τοῦ σχήματος ἰδέαν).
But the other *(τὸ δὲ)* [[*fusible (χυτὸν)*]] kind, which is composed of
large and *unopposed* (ἐκ μεγάλων καὶ ὁμαλῶν) particles [of primary eart

is more *stable* than the first and is *heavy*,
being solidified (πεπηγὸς) by its *state of non-opposition (ὁμαλότητός)* 58d4-e2

It is clear that
the "liquid" waterlike kind consists of a dyad $\boxed{\text{PW}} < \text{PE}$ in Dispersion, while
the "fusible" earthlike kind consists of a dyad $\text{PW} < \boxed{\text{PE}}$ in Consolidation.

We note that most scholars believe that both liquid and fusible waters are just mixtures of masses of only primary water PW in various sizes. The list of examples of the "fusible waters" (χυτὰ ὕδατα) is revealing: gold (χρυσὸς), adamant (ἀδάμας,) bronze (χαλκὸς), rust (ἰὸς) (59a8-c5), and the species of stone (λίθων εἴδη) (61b7-c1); it is simply impossible to accept that Plato believes gold, stones, etc. to consist only of water!

9.1.4. [4] *Motion of sensibles to their proper spaces* (58b8-c2)

they all shift (μεταφέρεται) up and down (πάντ' ἄνω κάτω)
towards their own proper *spaces* (πρὸς τοὺς ἑαυτῶν *τόπους*);
for the change in their several sizes (μεταβάλλον γὰρ τὸ μέγεθος ἕκαστον)
causes their position in *space* also to change (καὶ τὴν *τόπων* μεταβάλλει στάσιν) 58b8-c2

9.2. *The generation of a sensible body as a dyad of a Content entering and filling the Receptacle, in such a way that*

either the Receptacle consolidates the Content, the Receptacle is *minimum for the given Content, the Content becomes similar to the Receptacle and heavy, and*
the dyad as a unit moves to the place of the Receptacle
or the Content disperses the Receptacle, the Content is maximum for the given Receptacle, the Receptacle becomes similar to the Content and light, and
the dyad as a unit moves to the place of the Content.

9.2.1. *A mass of the Content enters, in fact filling quite fully and tightening, the mass of the* ***Receptacle*** (52d4-6, 52e1-3, 57a7-b1).

and that the Nurse of Generation (τὴν δὲ δὴ γενέσεως τιθήνην),
being liquefied (ὑγραινομένην) and ignified (πυρουμένην) and
receiving (δεχομένην) also the forms of earth and of air (τὰς γῆς τε καὶ ἀέρος μορφὰς),
52d 4-6

The *Nurse of Generation,* another name for the Receptacle, *receives (δεχομένην)* its Content, which is identified as forms of earth, water, air, and fire. From the interpretation of 57d-58a in Sect. 8, the Content is some primary body, as well. The fact that from the wording it appears that the Receptacle becomes similar to the Content is not clear at this stage, and we will come to this point later.

but owing to [the Receptacle] being *filled quite fully* (τὸ *ἐμπίμπλασθαι*) with powers (δυνάμεων) that are neither similar nor balanced (μήθ' ὁμοίων μήτε ἰσορρόπων) **52e1-3**

The Receptacle does not only receive the Content, but also is *filled quite full (ἐμπίμπλασθαι)* by the parts of the Content that enter the Receptacle. This is a central feature of the process of generation.

> the *smaller* particles, when *enveloped* (περιλαμβανόμενα) in the many *greater* ones'
> [the small enter into the large] **57a7-b1**

The mysterious Nurse of Generation/*Receptacle* of earlier passages, such as 50a-51a, 52d4-6, receives in general various bodies, while in the present passage 57a7-b1 which also deals with the Receptacle (it is a part of the longer passage 57a7-c6, and the Receptacle is mentioned in 57c3), it is clarified that the great body envelopes/receives the smaller: The smaller (*τά τε αὖ σμικρότερα*) parts, *the Content,* are comprehended *(περιλαμβανόμενα)* in the greater *(ἐν τοῖς μείζοσιν), the Receptacle.*

We may express mathematically by an inequality the passage as follows: A mass M.C of M particles of the Content enters into a mass N.R of the Receptacle (where M and N are some natural numbers), so that

$$M.C \leq N.R.$$

9.2.2. *The **Content** enters and fills quite fully **the Receptacle**, in such a way that (either the Content is Maximum for the given Receptacle or) the Receptacle is Minimum for the given Content* (52e2-3).

> in *not even one part of herself* (*κατ' οὐδὲν αὐτῆς*) [=the Receptacle] is
> she equally balanced (ἰσορροπεῖν) 52e2-3

We claim that

> "κατ' οὐδὲν αὐτῆς" in 52e2-3

should be translated,
not as "no part of herself," as it is customary,
but more literally and separately, as two words, as "not even (οὐδὲ) one (ἐν) part of herself" (as in 9.1.2 above).

According to this interpretation, the phrase states that
not even one part of the Receptacle (*κατ' οὐδὲν αὐτῆς*, 52e2-3) can stay balanced,
namely, not even one part of the Receptacle can be spared,
namely, we cannot do with a smaller Receptacle,
namely, subtracting even one part from the Receptacle, it cannot be Receptacle any more. Thus the Receptacle is *minimum for the given Content.*

The Receptacle is being filled quite fully by the Content, in accordance with 9.2.1, but now tells us in addition that the manner of filling is such that
not even one part of *the Receptacle*, namely of N.R, can be excluded;
this can only mean that

$$(\mathrm{N}-1).\mathrm{R} < \mathrm{M}.\mathrm{C} \leq \mathrm{N}.\mathrm{R},$$

exactly the double tight inequality that we have found in Sect. 4.

9.2.3. *Dispersion and Consolidation* (57c3-4)

> because of the motion of the Receptacle (διὰ τὴν τῆς δεχομένης κίνησιν), become *dissimilar to themselves* (τὰ δὲ ἀνομοιούμενα ἑκάστοτε ἑαυτοῖς) and similar to others (ἄλλοις δὲ ὁμοιούμενα) 57c3-4

The motion of the Receptacle, more precisely the motion of the dyad of opposites Receptacle & Content results

either in the Content becoming dissimilar to itself and similar to the Receptacle, and the Receptacle consolidates the Content, so that the Receptacle is the victorious part and the Content the defeated,

or in the Receptacle becoming dissimilar to itself and (approximately) similar to the Content, and the Content disperses the Receptacle, so that the Content is the victorious part and the Receptacle the defeated.

The defeated becomes approximately similar to the victorious body and stays with it**.**

9.2.4. *The motion of the generated sensible, Content & Receptacle, to the space proper to the victorious part*

> Moreover, it is owing to these affections (παθήματα) that they all interchange (διαμείβεται) their places (τὰς χώρας); for while the mass of each Kind keeps separate (διέστηκεν) in its proper space (κατὰ τόπον ἴδιον) because of the motion of the Receptacle *(διὰ τὴν τῆς δεχομένης κίνησιν)* yet those masses which from time to time become dissimilar to themselves and similar to others are moved (φέρεται), because of the shaking *(διὰ τὸν σεισμὸν)*, towards the space which is proper to those masses whereto they have been assimilated (τὸν ἐκείνων οἷς ἂν ὁμοιωθῇ τόπον). **57b7-c6**

9.3. *The equivalence of the Third Kind as Receptacle and as Primary bodies/ diakena,*

The mathematical origin of the term "Receptacle."

Relations of the descriptions of Receptacle as mother, nurse, with ancient similar descriptions of Earth (γαία).

9.3.1. *The Receptacle description of the generation of sensibles is strictly equivalent to the corresponding "diakena" of primary bodies description.*

Diakena of primary bodies First occurrence 58a (Sect. 9.1)	*Content & receptacle* Last occurrence 57c (Sect. 9.2)
[1] The small primary body enters the diakena of the great primary body 58a7-b6	[1] The content enters the receptacle 52d4-6, 52e1-3, 57a7-b1
The process of *tightening compresses* the small into the diakena of the great. Not even one part of the diakena is left empty 58b4-5,a7, 617b3	[2] Not only enters but fills fully, not even one part of the receptacle is left unbalanced 52e2-3
[3] The great consolidate the small; the small disperse the great 58b6-7, 58d4-e2	[3] *Dissimilar to themselves* and similar to others 57c3-4
[4] *Motion of sensibles to the proper spaces of the victorious part* 58b8-c2	[4] *The motion of the generated sensible, to the space proper to the victorious part* 57b7-c6

The four steps, by means of which the generation of the sensibles and their motion takes place in and by the Receptacle, are so uncannily similar to the four steps, by means of which the generation of the sensibles and their motion takes place in and by the "diakena" of the primary bodies, that the conclusion, that the Receptacle and the primary bodies are equivalent descriptions of the Third Kind, is inescapable.

9.3.2. ***Mathematical*** *origin of the term* ***Receptacle:*** *it reflects the* ***tightness*** *of the double inequalities.*

On the other hand, the steps in the process expressing the finite initial stages of the incommensurability of PE to PW in remainder-free expressions lead to the tight double anthyphairetic inequalities

$M.PW < N.PE < (M + 1).PW,$
[N.PE holds M.PW, but N.PE can hold nothing more]
$(N - 1).PE < M.PW < N.PE,$
[N.PE holds M.PW, but nothing less can hold M.PW]

that are eminently suitable, exactly because of their impressive tightness, to be termed in vivid and suggestive nonmathematical language as Content PW to Receptacle PE, in accordance with Aristotle's description of Receptacle (in Sect. 4). It

should be noted that the *tightness* is the predominant feature of these inequalities, a property expressed, besides the imagery Receptacle to Content, by a variety of expressions: "*tightens* (*σφίγγει*)" "the process of *tightening* (ἡ δὴ τῆς *πιλήσεως* σύνοδος)," "*compresses together (συνωθεῖ)*" *58a4-5*, and "allows not even one part (οὐδεμίαν) of empty place (κενὴν χώραν) to be left (ἐᾷ λείπεσθαι)" 58a7.

9.3.3. *The reason the Receptacle is called Mother (50d3) and Nurse (*τιθήνη *49a6, 52d5,* τροφὸς καὶ τιθήνη *88d6) is that the ultimate Receptacle of every sensible entity is primary Earth.*

According to our interpretation in 9.3.2, and our description of all sensibles in Sect. 11, *the ultimate Receptacle* of every sensible entity is primary earth PE. Primary water PW serves as final or intermediary Content in relation to PE, and as intermediary Receptacle in relation to primary air PA or primary fire PF.

But Earth, *the only female* of the four "elements" (water and fire are neutral, air masculine), is traditionally called *mother* and *nurse*

not only by various ancient authors (here is a small sample):

Γαῖα μάτηρ. Pindar, *Olumpia* 7,38
γῆ πάντων μήτηρ, Hesiodus, *Opera et Dies* 563
γαῖα μαῖα, Aeschylus, *Choephoroi* 43

In *Orphic Hymn* 63,5-16, earth is described as mother and as nurse/τιθήνη.

And all the living creatures (καὶ ζώιων πάντων), that earth (*γαῖα*), the godess *(θεὰ)*, mother (*μήτηρ*) nurses (*τιθηνεῖ*) ἐν κόλποισι *Orphic Hymn* 63,15-16

Cf. also γαῖα θεὰ μήτηρ *Orphic Hymn* 15,4; Γαῖα θεά, μῆτερ 26,1
but in fact also by Plato himself, who is repeatedly referring to Earth in exactly these terms:

ἡ ἡμετέρα γῆ τε καὶ *μήτηρ*, *Menexenus* 237e5-6
καὶ ἡ γῆ αὐτοὺς *μήτηρ* οὖσα ἀνῆκεν,
καὶ νῦν δεῖ ὡς περὶ *μητρὸς καὶ τροφοῦ τῆς χώρας Politeia* 414e2-6

In *Kratulos* 410b8-c3, Plato finds the etymology of earth as the place of generation, likening it to "mother":

The word γῆ (earth) shows the meaning better in the form γαῖα; for γαῖα is a correct word for *γεννήτειρα "mother,"* as Homer says, for he uses γεγάασιν to mean γεγενῆσθαι (be born). *Kratulos* 410b8-c3
[translated by Harold N. Fowler, modified by the authors]

Finally, in the *Timaeus* itself:

γῆν δὲ *τροφὸν* μὲν ἡμετέραν, *Timaeus* 40b8

These repeated occurrences suggest that Plato called the Receptacle "mother" and "nurse" precisely because the principal Receptacle, according to our interpretation, is primary Earth, and Earth was long connected with these terms. Thus these references of the Receptacle to "mother" and "nurse" in fact constitute one more argument in favor of our interpretation of the *Timaeus* Receptacle.

Appendix 9A *The identification of the "diakena" of the greater primary body with the "hedrai" and thus with the area of the faces of the greater primary body*

9A.1. In 53c-55c, where Plato introduces the canonical polyhedral (described in Sect. 6), the suffix -εδρον is never used to refer to the 4hedron (pyramid), 6hedron (cube), 8hedron, or 20hedron, contrary to the practice in the Theaetetean Book XIII of the *Elements* (where there are 77 references with this suffix). Furthermore, the face of a polyhedron is called βάσις, basis in the *Timaeus*; thus the square, consisting of two triangles A, declared a basis more stable (ἀσφαλεστέρα 55e3, 4; στασιμωτέρως 55e7) than the equilateral triangle consisting of two triangles B. However, the root hedr-/*εδρ-* does come up rather unexpectedly in 59d6-7, when the bases/faces of primary water, namely the equilateral triangles/triangles B, are declared less "hedral" (*ἑδραίους*) than those of earth, called "hedral" (*ἑδραῖον* 64b7) once more in 64b6-c6, thus introducing an indirect link between the faces and the term "hedra"

(rendered by Lamb again as less "stable," and thereby losing the connection).

> Also it [primary water] is soft owing to the fact that its bases (*βάσεις*), being less *stable* (*ἑδραίους*, 59d6) than those of earth, 59d6-7

In Greek, the text makes sense only if Plato has in the back of his mind the word "hedra" in place of the word "basis" he is using:

> the *hedra/basis* of the 20*hedron* is less *hedral* (*ἑδραία*) than the *hedra*/ basis of the 6*hedron*/ cube.

9A.2. In 52b1, Plato refers to the Receptacle, as

> providing *"hedra" (ἕδραν)* [room, Lamb] for all things that are generated

a statement that might simply describe the Receptacle as the seat or place of generation, with no association at all of the Receptacle with the face of a polyhedron (Same thing with 53a2). However, it is a fact that more "hedral" hedrai/faces, those of the 6hedron, make the best "hedra," in the sense of Receptacle, since it is the "hedra" of the greater primary body.

9A.3. *Identification of "hedrai" with "diakena"; identification of the "hedrai" of a primary body with the surface of the canonical polyhedron corresponding to that primary body*

But there are three occurrences of "hedrai" (59a3, 60c4, 62a8) *in the plural.*

It is tempting to consider that the empty places "diakena" of the great primary bodies are precisely the sum of the areas of the triangles constituting the great primary body, and that these are empty spaces in the sense that they wait to be filled by the corresponding triangles constituting the small primary dyad. Since the fundamental interaction between primary bodies is, as we have seen, the anthyphairesis of their surfaces, this interpretation is the only one that makes sense.

As we have seen, in Sect. 10, a mass of the small primary body tightens (σφίγγει, πιλήσεως) and compresses (συνωθεῖ) into the diakena of the great primary body:

> "tightens (σφίγγει) . . . all and allows no empty space (κενὴν χώραν) to be left", 58a7;
>
> "Thus the process of tightening (ἡ δὴ τῆς πιλήσεως σύνοδος) *compresses together (συνωθεῖ)* the small [primary] bodies (τὰ σμικρὰ) into (εἰς) *the empty places (τὰ διάκενα)* of the great (τῶν μεγάλων) [primary bodies]", 58b4-5;
>
> "primary water is *compressed* in the *"diakena"* of primary earth", 61a8;
>
> "the water (ὕδωρ) occupies (κατέχῃ) the "diakena" of earth (τὰ τῆς γῆς διάκενα), which are *forcibly tightened (βίᾳ συμπεπιλημένα)",* 61b1-2;
>
> "primary fire enters in the *"diakena"* of primary water", 61b3-4

Plato in several passages in the *Timaeus*, which we now examine, is exchanging the somewhat opaque term "diakena" with the term "hedrai," in plural; it should be obvious that "hedrai" cannot refer but to the "faces" of the polyhedron to which the "hedrai" belong.

> Again, since the [primary] fire, on exiting from the [*"hedrai"/faces* of the primary] water,
> does not exit into a void,
> *the* adjacent *[primary] air* (ὁ πλησίον ἀὴρ),
> being pressed (ὠθούμενος) into the *"hedrai"/faces* [[*places, Gregory*; abodes, Lamb]]
> previously occupied by fire (εἰς *τὰς τοῦ πυρὸς ἕδρας*, 59a3),
> *compresses (συνωθῶν)* the, still mobile, mass of [primary] water [with primary earth],
> and combines it (αὐτὸν) [the water and earth mass] with itself (αὑτῷ) [primary air].59a1-4

It is clear in passage 59a1-4 that "hedrai" is used instead of "diakena."

Primary air exits from the "diakena"/"hedrai" of primary water, and adjacent external primary air enters into and compresses these "diakena"/"hedrai" of primary water, previously occupied by primary fire.

> [primary air] being heavy,
> when pressed (*ὠσθεὶς*) and poured round
> the ["hedrai" of the] mass of [primary water mixed with] primary earth,
> it [the primary air] tightened forcibly (*σφόδρα ἔθλιψεν*) and compressed *(συνέωσέν*)
> it [the mass of primary earth and primary water] into the *"hedrai"/faces (εἰς τὰς ἕδρας)*
> [of primary water] from which [*the "hedrai"* of primary water] the new air was ascending;

and the [primary] earth compressed (συνωσθεῖσα) into [primary] water (ὕδατι) by [primary] air (ὑπὸ ἀέρος) so as to be indissoluble stone (πέτρα) is generated 60c2-6

It is clear in passage 60c2-6 that "hedrai" and "diakena" are used interchangeably.

Primary air enters, tightens, and compresses forcibly into the "hedrai," exactly as into the "diakena" of primary water, with the result that primary earth is compressed into primary water.

When [*mass of primary*] *water* with greater particles (μεγαλομερέστερα),
which surround the [primary earth] body,
enter (εἰσιόντα) into it [the *"hedrai"* of the primary earth],
they press out (ἐξωθοῦντα) the smaller [primary water] particles (τὰ σμικρότερα);
but as they [the greater particles of primary water] cannot enter into the *"hedrai (ἕδρας, 62a8)* [[room, Lamb; places, Gregory]] [of the smaller primary water],
they compress (συνωθοῦντα) the [primary] water (τὸ νοτερόν) within us,
so that instead of state of opposition *(ἐξ ἀνωμάλου)* and motion *(κεκινημένου)*
they produce immobility *(ἀκίνητον)*, as a result of the non-opposition (δι' ὁμαλότητα),
and *compression (σύνωσιν)* 62a6-b2

We interpret this difficult passage as follows:

There is a change from a Dispersion (at even stage 2n) waterlike dyad PW < PE, in which the primary water particles dominate the primary earth particles, and thus appear as small:

$$M_{2n-1}.PW < N_{2n-1}.PE < (M_{2n-1} + 1).PW$$

to a Consolidation (at next odd stage 2n + 1) earthlike dyad PW < PE, in which the primary water particles are dominated by the primary earth particles and thus appear great:

$$(N_{2n} - 1)\,.PE < M_{2n}.PW < N_{2n}.PE.$$

The primary water particles that appear great cannot enter and compress into the "hedrai"/ἕδρας of the primary water particles that appear small, but they enter and compress into the "hedrai" of the primary earth particles.

It is clear in passage 62a6-b2 that "hedrai" is used instead of "diakena." A mass of primary water enters and compresses a mass of primary earth in consolidation and presses out.

There are two more occurrences of "hedra,"

where [primary] air presses *(ὠθεῖ)* the adjacent body from its "hedra (*ἐκ τῆς ἕδρας,*79b4),
the body thus pressed (τὸ δ' ὠθούμενον) drives out in turn the next

and a situation of "antiperistasis" "εἰς *τὴν ἕδραν*" in 79b6, in the singular, but related to those above in the interchangeability with "diakena."

In all these passages, the smaller primary body enters into the "hedrai" (*in plural!*) of the greater primary body and "*sunothei*" (compresses) it. This is precisely the action of the smaller primary entity air described in 58a5, where, by the process of

tightening (ἡ τῆς πιλήσεως σύνοδος), the small (τὰ σμικρὰ) enter into the "diakena" (διάκενα) of the great (τῶν μεγάλων) and compress (συνωθεῖ) them (58b4-5; Sect. 10,1-2).

Furthermore, Simplicius' (*In Physics* 694,19-21) comment on the *Timaeus* 59a1-4 passage is revealing:

> and it is clear δῆλον that the "hedra" (ἕδρα) of [primary] fire was some chora (χώρα) in which the [primary] fire was. And that this is an area (διάστημα) of equal measure (ἰσόμετρον) in every body (τῷ ἐν αὐτῷ παντὶ σώματι). [Our rendering]

We have no choice, on the one hand, but to identify the "diakena" of the great with the "hedrai" of the great, and we also hardly have any choice but to identify *the "hedrai"* of the primary water, namely of the surface of the canonical *20hedron*, with the faces, the bases of the *20hedron.*

Note that it then makes perfect sense that the greatest primary bodies have the greatest empty spaces, as stated in 58a7-b6, something that scholars of *Timaeus* have difficulty explaining by misinterpreting, as they do, "diakena" as external empty space between particles of primary bodies, e.g., Cornford 1937, p. 245:

> But the coming-together involved in the condensing process (ἡ τῆς πιλήσεως σύνοδος) is always driving the smaller particles into the spaces left between the larger.

Section 10 ***The comprehension of the four primary bodies in the sphere, the geometric and that of the soul of the kosmos, has two fundamental consequences:***

(a) ***[Intelligence] The Third Kind participates in the Intelligible Being indirectly, by means of the Soul of Kosmos, thus forming a bastard intelligible (33a6-b5,* 36d8-e5, 52d2-4, 53a7-*b7, 53d6-7, 54a1-b1).***

(b) ***[Necessity] The dyad PE, PW, trapped in the sphere, and thus not able to generate anthyphairetic remainders and infinite kosmoi, is forced to use the participation in the intelligible for the generation of sensibles in Receptacle form leading to the One Kosmos (58a4-7).***

The Third Kind, consisting of the four primary bodies, has the power to generate all sensible bodies, as shown in Sect. 8, on the basis of passage 57d-58a. But two fundamental issues are still unresolved, leaving the whole endeavor in doubt.

One is that so far we have no argument in favor of the participation of the Third Kind in the intelligible, even a most confusing participation, as claimed in 51a7-b1, turning it into a "bastard" intelligible. This is *Plato's fundamental*, sine qua non *tenet:* Without participation of the sensibles in the intelligible, the intelligible loses its main "reason d' etre."

The second open issue is whether the type of sensibles that the Third Kind generates leads to the continued ad infinitum generation of anthyphairetic remainders and to the undesired infinite kosmoi, or to a Receptacle type sensibles, in terms of tight

double remainder-free inequalities, leading to the desired One Kosmos. The fact that we have switched from the intelligible cause to a geometric cause of sensibles does open the possibility for the desired second alternative but does not impose it, and as we saw in Sect. 8, the passage 57d-58a is neutral on this question.

It turns out that the answer to both problems, given in the short but fundamental passage 58a4-7, is based on the very special and unique property of the canonical solids, the *comprehension* of the canonical solids, and hence of the primary bodies, *in a sphere*, as shown by highly nontrivial proofs in the second half of Propositions 13–16 of the Theaetetean Book X of the *Elements*. It turns out that the setting of order, the *"kosmesis,"* of the World from its previous state of chaos and disorder coincides with the start of *participation* of the primary bodies in the intelligible by means of the Soul of the World.

10.1. *The participation of the Third Kind in the intelligible (51a7-b1, 52b1-2)*

Participation of the Sensibles in the Intelligible is absolutely necessary if we are to keep within Plato's philosophical system, in which the intelligible Beings dominate and govern the sensibles. Since the sensibles are generated by the Third Kind, it becomes absolutely necessary that the Third Kind *participates* in the intelligible, and that the participation of the sensibles to the intelligible will be assured exactly because of this participation of the Third Kind to the intelligible. On the other hand, such participation, turning the Third Kind into a "bastard" intelligible (52b1-2), will be difficult to explain, as Plato warns us in no uncertain terms in 51a7-b1, calling it "most difficult to conquer" (δυσαλωτότατον), since it will be a participation that no other geometric dyad shares, making the primary bodies unique among geometrical entities possessing such participation.

> we describe her as a Kind (εἶδός) invisible (ἀνόρατον) and unshaped (ἄμορφον), all-receptive (πανδεχές), and in some way most perplexing (ἀπορώτατά) and most difficult to conquer (δυσαλωτότατον) *partaking (μεταλαμβάνον) of the intelligible (νοητοῦ)* 51a7-b1

The statements in 53d6-7, asking whether the two triangles A and B and the sensibles originate from a higher principle (τὰς δ' ἔτι… ἀρχὰς ἄνωθεν), and in 54a2-b1 declaring that Triangle B forms, together with Triangle A, the most beautiful (τὸ κάλλιστον) geometric dyad, thus making them closest to the intelligible,[12] are preliminary toward the flat statement of participation.

10.2. *The setting of order ("kosmesis") of the World consists exactly in the institution of the* ***participation*** *of the Third Kind in the intelligible (51a7-b1, 52d2-4, 53a2-b7).*

[12] We know from *Sumposion* 211c8-d1, e1, that *every* intelligible true Being, and not just the Idea of the Beautiful, is a most beautiful entity; it is beauty itself. This is the method that Plato describes by means of which the idea of the beauty is most beautiful, in short "ephexes" and "exaiphnes," and is exactly the method of (anthyphairetic) Division and Collection (Logos Criterion), respectively, that holds for every intelligible Being.

There was a time that the World was not yet set in order. The intelligible Beings, the sensible bodies, and the Receptacle, and even the *separation* into four different *places* for the four primary bodies, existed *before* order "kosmesis" was set in the World:

> Let this, then, be, according to my verdict, a reasoned account of the matter summarily stated,—that Being (**ὄν**) and Place (χώραν) and Generation (**γένεσιν**) were existing, three distinct things, *even before the Heaven (οὐρανὸν) came into existence* 52d2-4

> wherefore also these [four Kinds] [τὰ τέτταρα γένη]
> occupied different places (διὸ δὴ καὶ χώραν ταῦτα ἄλλα ἄλλην ἴσχειν),
> even before that the Universe (*πρὶν* καὶ τὸ πᾶν) was *set in order (διακοσμηθὲν)*
> and *generated out of them* (ἐξ αὐτῶν γενσθαι).
> *Before that time* (τὸ μὲν δὴ *πρὸ τούτου*), in truth, all these things were in a state
> without "logos" or measure (ἀλόγως καὶ ἀμέτρως) 53a2-8

It is clear that *if* the sensibles *participate* in the intelligible, then the sensibles *cannot* be in a state without logos and measure, without the presence of God. It is clear from this description that before order was set there could be *no participation* of the World and of the Receptacle in the intelligible.

> but when the work of setting in order this World (κοσμεῖσθαι τὸ πᾶν) was being undertaken, fire and water and earth and air, although possessing some traces of their own nature, were yet so disposed as everything is imaged (εἰκὸς) in the absence of God; and inasmuch as this was then their natural condition, God began by first *marking them out into shapes (διεσχηματίσατο) by means of forms and numbers (ε δεσί τε καὶ ἀριθμοῖς)*. And that God constructed them, so far as He could, to be as *fair (κάλλιστα) and good (ἄριστά)* as possible, whereas *they* had been otherwise (*ἐξ οὐχ οὕτως ἐχόντων*) —this above all else must always be postulated in our account. 53b1-7

Again it is clear that when order was instituted in the World, then the sensibles were shaped by kinds and numbers, namely by intelligible materials, something that would be impossible without participation of the sensibles in the intelligibles. So the time when order of the kosmos was instituted must be *exactly* the time when the sensible bodies started to participate in the intelligible. What did not exist then, before "kosmesis," was *participation* of the sensible bodies in the intelligible.

We have shown already (a) that the sensible bodies are generated from the primary bodies, and also (b) that the Third Kind has been identified with the four primary bodies. Plato expressly states (10.1) that the Third Kind participates in the intelligible; on the other hand, a priori there is no reason to believe that the geometrical primary bodies *qua* geometrical bodies participate in the intelligible, and indeed Plato prepares us for a mode of participation most unusual and far from obvious. It follows that we are urgently in need of an indication that the Third Kind, namely the primary bodies, participates in the intelligible.

10.3. *The participation of the Soul of the World in the intelligible dyad <IE, IF> (33a6-b5, 36d8-e5)*

The World is spherical (σφαιροειδές) and circle-like (κυκλοτερὲς, 33b5, same as in 58a5), *because* it is constructed in similarity to, in imitation of the paradigm, the *intelligible Being* with dyad <IE,IF> (*Timaeus* 33a6-b5), which has *periodic* anthyphairesis;

the periodic motion, the revolution of the World, is due to its Soul and is controlled and powered by the intelligible periodic anthyphairesis (*Timaeus* 36d8-e5). It is clear that periodicity is the property that makes the World *ageless* (ἀγήρων, as in *Politicus* 273e4, *Philebus*, 15d8).

10.4. *The comprehension of the primary bodies in the spherical Soul of the Kosmos results in the "δυσαλωτότατον" participation of the Third Kind, the secondary, bastard cause,* via *the spherical Soul of the World, in the Intelligible, the real, and ultimate cause and in sensibles is described by double tight remainder-free inequalities, ensuring the One Kosmos (58a4-7).*

We now come to the brief but fundamental passage 58a4-7.

> The revolution of the World (ἡ τοῦ παντὸς περίοδος),
> [1] since it comprehended (συμπεριέλαβεν) the [primary] Kinds (τὰ γένη),
> [2] seeing that it is circle-like (κυκλοτερὴς) and
> wishing (βούλεσθαι) naturally (πεφυκυῖα) to come together (συνιέναι) to itself (πρὸς αὑτὴν),
> [3] tightens (σφίγγει) everything (πάντα) and thus it allows not even one part (οὐδεμίαν) of empty place (κενὴν χώραν) to be left (ἐᾷ λείπεσθαι). 58a4-7

[1]. This is the only passage that connects the four primary bodies (τὰ γένη), and thus the Third Kind, with the Revolution of the World, and thus establishes the participation of the Third Kind in the intelligible. The key sentence is "since [the revolution of the World] *comprehended* (συμπεριέλαβεν) the primary bodies."

What does this mean exactly?

The first thing to realize is that the word "*comprehended/συμπεριέλαβεν*" that Plato is using in 58a4-7 essentially coincides with the word "*comprehend/περιλαβεῖν*" used in the Theaetetean Book XIII of Euclid's *Elements,* where in Propositions XIII.13–16 it is shown, with great effort, that each of these regular solids can be comprehended in the sphere (σφαίρᾳ περιλαβεῖν). It is clear that Plato is using here Theaetetus' mathematical Propositions and terminology.

We distinguish two consequences of this comprehension.

The first is that *the real and ultimate cause* of the sensibles, namely the anthyphairetically periodic *intelligible Being*, is able to transmit its intelligible anthyphairetic power of motion, by means of the intermediary, namely the sphericity of the World Soul (κυκλοτερὴς, 33b5, 58a5),

to the *Third Kind*,

namely to the incommensurable dyad PE, PW, aided by the auxiliary commensurable dyads PW/PA = 5/2 and PW/PF = 5/1,

exactly because the constituents of the Third Kind, the primary bodies, are *comprehended* in the intermediary sphere and revolve with it;

as a result, the Third Kind is provided with power of motion and generation of the sensibles and becomes a *secondary cause* for the sensibles, as described on the basis of 57d-58a, in Sect. 8.

[2]. But we are next told that the motion of the Soul of the Kosmos is cyclical, periodic, tending to return to itself. Yes indeed, this is characteristic of every circular, periodic motion, and so it applies to the motion of the Soul of the Kosmos. But what is new or interesting?

[3]. Well, as a consequence the revolution of the Soul of the World

> tightens (σφίγγει) everything (πάντα) [the primary bodies]
> and thus it allows not even one part (οὐδεμίαν) of empty place (κενὴν χώραν)
> to be left (ἐᾷ λείπεσθαι).

We believe this consequence, essential for Plato, has not been understood properly by scholars thus far. We attempt to enlighten it. An explanation that tries to explain this situation of "tightening," "compressing," and "not leaving even one empty space" mechanically, solely by means of the nature of the circular motion in general, has no chance of success.

Here we are dealing with the Third Kind, consisting of the incommensurable dyad PE, PW (plus the two auxiliary ones), already having assumed, by passage 57d-58a, the role of the cause of the sensibles, substituting in that role the intelligible. Thus, the dyad PE, PW generates initial segments of its infinite anthyphairesis.

Now it is crucial to see what form will these products assume. A priori the dyad could generate either anthyphairetic remainders or double tight remainder-free inequalities. As we have already noted (Sect. 8), passage 57d-58a does not take sides on this issue. The evidence presented in Sect. 9 in favor of the second alternative (remainder-free inequalities, leading to the elimination of infinite kosmoi) does not offer a reason for this choice.

This reason is offered precisely now with 58a4-7. Indeed, the primary bodies are caged, trapped *within the confines of the One spherical World/Kosmos,* and hence cannot employ the power of motion granted them by the intelligible to generate new, potentially infinitely many anthyphairetic remainders and hence potentially infinitely many kosmoi. Thus the primary bodies have no other choice and *are forced by necessity* to turn the transmitted intelligible power of motion internally, according to the mathematics of generalized side and diameter numbers, tight, double remainder-free inequalities, toward one kosmos. Thus we have finally the answer to the question why all generations of sensibles examined (in Sect. 9) were always toward preserving the one kosmos.

This reasoning is lost among scholars because they do not realize that what the Third Kind is set to accomplish is the elimination of infinite kosmoi.

This is a very important passage, because it seems to be the only one that connects the four primary bodies, and their *geometric* infinite, in fact periodic, anthyphairesis of the triangles A and B, with the *intelligible, by means* of the period of the soul of the *world*. At the same time, the primary bodies are the only geometric entities that are endowed with participation in the intelligible by means of the World Soul.

This turns the Third Kind, on the one hand, *purely geometric downward* as it generates the sensible bodies as syllables of the form Content & Receptacle, and

on the other hand, it keeps contact with the intelligible through the soul of the World and substitutes for the intelligible as the cause of the sensibles, thus in fact a *"bastard intelligible" upward.*

10.5. *The kosmos is a mixture of intelligence and necessity*

Now we are ready to make sense of the passage 47e-48a and of the exact meaning of necessity in relation to kosmos and the Sensibles:

> The foregoing part of our discourse, save for a small portion, has been an exposition of the operations of *Intelligence* (διὰ νοῦ, rendered by Lamb Reason); but we must also furnish an account of what comes into existence through *Necessity* (δι' ἀνάγκης). For, in truth, this Cosmos in its origin was generated as a compound (μεμειγμένη), from the combination of Necessity and Intelligence.
> And inasmuch as Intelligence was *controlling* Necessity by persuading (τῷ πείθειν) her to conduct to the best end the most part of the things coming into existence, thus and thereby it came about, through Necessity yielding (ἡττωμένης) to intelligent persuasion (πειθοῦς ἔμφρονος), that this Universe of ours was being in this wise constructed at the beginning.
> Wherefore if one is to declare how it actually came into being on this wise, he must include also the form of *the Errant (or Wandering) Cause* (τὸ τῆς πλανωμένης εἶδος αἰτίας), and how it is in its nature to cause movement and change (φέρειν πέφυκεν). 47e3-48a7
> [translation by WRM Lamb, modified by the authors]

"Nous" is rendered by most translators (Lamb, Cornford, and Waterfield) as Reason, but what is meant is precisely intelligence. Thus, kosmos is a mixture of intelligence and necessity.

The role of intelligence in the kosmos is well-defined: It is precisely the participation of the sensibles in the Intelligible (10.1-3).

Since the Receptacle/Chora is considered *necessary* (*εἰσαναγκάζειν*, 49a3; *ἀναγκαῖον ε ναί*, 52b3-5) for the adequate description of kosmos, it is clear that necessity is related to the Receptacle. But the Receptacle, the Third Kind, as we have seen, was necessary to be introduced solely for ensuring that there is only one kosmos with four kinds while, on the other hand, as we have seen in Sect. 3, the oneness of kosmos is NOT a consequence of its participation in the intelligible; the intelligible is neutral and makes it equally possible for one kosmos or for infinitely many kosmoi.

The One Kosmos has been imposed essentially by the nondialectical choice of introducing the Third Kind in terms of the primary bodies. The fundamental fact that

imposed One Kosmos was *the comprehension of the primary bodies, the constituents of the Third Kind, in a sphere*.

Thus, the primary bodies, and specifically the dyad <PE, PW>, on the one hand, because of its participation to the intelligible, are endowed with the valuable transmitted knowledge of how to generate sensibles and, on the other hand, are trapped, caged, and confined in the cosmical sphere, and thus unable to generate infinite remainders and infinite kosmoi, since these remainders would have to go, as lighter above and beyond the sphere (as explained in Sect. 3); they are forced then by *sheer necessity* to turn to the only other possible manner of generating sensibles, that of Receptacle and tight Inequalities, thus staying within the confines of one kosmos.

Thus, the kosmos is indeed a mixture of intelligence and necessity, with intelligence in control of necessity, by means of True Opinion.

11. *The nature and the structure of the sensible bodies in the Timaeus: The sensibles participate in the intelligible* via *the Third Kind; every sensible is a syllable of the form Receptacle & Content; and there are four kinds of sensibles in one kosmos*

We are now, at long last, able to describe fully the nature and the structure of the *Timaeus* sensibles.

Every sensible body contains the dyad
a mass N.PE as Receptacle, and a mass M.PW as Content,
thus forming the sensible

$$\text{M.PW} < \text{N.PE}$$

either in Dispersion or in Consolidation.
To this dyad M.PW < N.PE as Receptacle,
a mass K.PA or a mass K.PF may be added as auxiliary Content,
thus forming the sensible

$$\text{K.PA} < (\text{M.PW} < \text{N.PE}) \text{ or } \text{K.PF} < (\text{M.PW} < \text{N.PE}),$$

either in Dispersion or in Consolidation. These are essentially the only possible sensible bodies. The complete list of sensibles is given in Sect. 11.2 below.

We conclude that the sensibles *participate in the intelligible* via *the Third Kind, and* are syllables of the form Content & Receptacle, four kinds of them in one kosmos.

11.1. *The sensible bodies as the class of entities whose* ***cause*** *is the primary dyad PE, PW*

It will turn out that every sensible body is a syllable with basic dyad PW < PE.

(a) *The dyad PE, PW is* ***the cause*** *of the generation and motion of the earthlike and waterlike sensibles.*

This is precisely the content of Sects. 9.1 and 9.2.

(b) *The dyad PE, PW is* ***the cause*** *of the generation and motion of the airlike and firelike sensibles, as well.*

But the class of earthlike and waterlike sensibles does not exhaust all sensibles.

The fact that Plato has rejected the possibility of direct interaction of PE with either PA or PF does not mean that there is no interaction between these primary bodies.

How can sensible air be generated?

We will need a *syllable* that contains primary air PA.

Can the dyad $\boxed{\text{PA}} \leq \text{PW}$, presumably in Dispersion, meaning

$\text{K.PA} \leq \text{L.PW} < (\text{K} + 1).\text{PA}$ for some numbers K, L, be such a syllable? Such a dyad *by itself* is certainly rejected as a possible sensible air, exactly because this dyad has *no cause* for generation or motion (or equivalently, it *does not participate* in a periodic anthyphairesis of the primary bodies, since the dyad PW, PA is commensurable with finite anthyphairesis).

Since a direct interaction between PE and PA has been excluded (Sects. 8.4 and Sect. 10), it is clear that the only possibility is to *form*

the triad $\text{PA} \leq (\text{PW} < \text{PE})$,

symbolizing that

the particles of the primary water PW enter and fill fully, not leaving even one empty space, into the "diakena" of the primary earth PE, and

the particles of the primary air PA enter and fill quite fully into the already existing "diakena" of the water PW;

furthermore, in order that the triad in question is *sensible air,* the two dyads $\text{PW} < \text{PE}$ and $\text{PA} \leq \text{PW}$ must be in Dispersion, symbolically

$\boxed{\text{PA}} \leq \left(\boxed{\text{PW}} < \text{PE}\right)$.

Thus, the interaction of PA or PF with PE does materialize but in an *indirect* way, namely as a triad $\text{PA} \leq \text{PW} < \text{PE}$, *defined* as the composition of

the principal incommensurable interaction dyad $\boxed{\text{PW}} < \text{PE}$, meaning, as we have said, that there are natural numbers M, N, such that $\text{M.PW} < \text{N.PE} < (\text{M} + 1).\text{PW}$,

followed by the auxiliary commensurable perturbation

dyad $\boxed{\text{PA}} \leq \text{PW}$,

meaning that there is a natural number K, such that $\text{K.PA} \leq \text{M.PW} < (\text{K} + 1).\text{PA}$,

or dyad $\boxed{\text{PF}} \leq \text{PW}$,

meaning that there is a natural number K, such that $\text{K.PF} \leq \text{M.PW} < (\text{K} + 1).\text{PF}$.

Let us make sure that the primary dyad PE, PW deserves to be considered *the cause* not only for every sensible dyad $\text{PW} < \text{PE}$,

but also for all triads $\text{PA} \leq \text{PW} < \text{PE}$, or $\text{PF} \leq \text{PW} < \text{PE}$.

Assume, e.g., $\boxed{\text{PA}} \leq \left(\boxed{\text{PW}} < \text{PE}\right)$.

Since $\boxed{\text{PW}} \leq \text{PE}$ is in Dispersion, there are natural numbers M, N, such that $\text{M.PW} < \text{N.PE} < (\text{M} + 1).\text{PW}$.

Since $\boxed{\text{PA}} \leq \text{PW}$ is in Dispersion, there is a natural number K, such that
$\text{K.PA} \leq \text{M.PW} < (\text{K} + 1).\text{PA}$.
But since *PW/PA = 5/2*, it is immediate that
K is just the integer part of the rational (5/2).M.
Thus, K is determined by the number M, itself determined by the cause PE, PW!
Thus, we can say that the triad $\boxed{\text{PA}} \leq (\boxed{\text{PW}} < \text{PE})$ is caused by the dyad PE, PW.
Thus, sensible air is by necessity a triad, a syllable with three primary elements/letters,
and similarly a sensible fire is a triad $\text{PF} \leq \text{PW} < \text{PE}$, as described in 60e-61a
(or possibly even a tetrad $\text{PF} \leq \text{PA} \leq \text{PW} < \text{PE}$, although Plato does not appear to mention any such example).

Examples (58d8-e3, 61a2-4)
Passage 58d8-e3 describes the most violent change from earthlike sensible to firelike.

> But the other kind, which is composed of great and *unopposed* (ὁμαλῶν) particles,
> is more stable than the first and is heavy,
> being solidified by its state of *non-opposition* (ὑπὸ *ὁμαλότητός*) 58d8-e2

The initial state is clearly the fundamental sensible dyad in consolidation, generating an earthlike sensible dyad: $(\text{N}_{2n} - 1).\text{PE} < \text{M}_{2n}.\text{PW} < \text{N}_{2n}.\text{PE}$; thus, $\text{PW} < \boxed{\text{PE}}$.

> but when fire enters (ὑπὸ δὲ *πυρὸς* εἰσιόντος, 58e2) and dissolves it,
> this causes it to abandon its state of *non-opposition (ὁμαλότητα*, 58e3)

Fire enters into the diakena of water, and finite Dispersion takes place between PW and PF, symbolically $\boxed{\text{PF}} \leq \text{PW}$. This Dispersion unbalances the internal situation in the initial dyad PE and PW, and, by going to the next step in the infinite anthyphairesis, a change occurs *from Consolidation* $\text{PW} < \boxed{\text{PE}}$ *to Dispersion* $\boxed{\text{PW}} < \text{PE}$ by going to the next stage of the anthyphairesis of PE to PW:
$\text{M}_{2n+1}.\text{PW} < \text{N}_{2n+1}.\text{PE} < \text{M}_{2n+1}.\text{PW}$.
There results the following firelike sensible triad: $\boxed{\text{PF}} \leq (\boxed{\text{PW}} < \text{PE})$, a *firelike* sensible body: $\text{K}_{2n+1}\text{PF} \leq \text{M}_{2n+1}.\text{PW} < (\text{K}_{2n+1} + 1)\text{PF}$. Since $\text{PW/PF} = 1/5$, K_{2n+1} is immediately seen to be equal to $\text{M}_{2n+1}.\text{PW/PF} = 5$. M_{2n+1}, determined by M_{2n+1}.
A similar situation is described in passage 61a2-4:

> If mixed [primary Earth] (*συνεστηκυῖαν* δὲ [[γῆν]]) [with primary water] is generated
> only [primary] fire (πλὴν πυρὸς οὐδέν) can disperse it,
> for no entrance (εἴσοδος) is left (λέλειπται) for anything but fire *(οὐδενὶ πλὴν πυρὶ)*. 61a2-3

Mixed *(συνεστηκυῖαν)* primary earth necessarily means that primary earth is mixed with primary water, the only possible mixture for primary earth, in the sense that masses of primary water enter into the "diakena" of primary earth, and consolidation takes place generating an earthlike sensible body, namely the dyad PW < $\boxed{\text{PE}}$. Only primary fire *(πλὴν πυρὸς οὐδέν)* can disperse this earthlike sensible body, as follows: Primary fire PF enters and fills the diakena of primary water PW, in Dispersion $\boxed{\text{PF}}$ < PW; this Dispersion forces the infinite anthyphairesis of PE to PW to go to the next stage, *from Consolidation* PW < $\boxed{\text{PE}}$ *to Dispersion* $\boxed{\text{PW}}$ < PE, and generating a sensible firelike triad ($\boxed{\text{PF}}$ < ($\boxed{\text{PW}}$ < PE).

Primary air PA is presumably too weak for such a drastic action, which is the prerogative of PF.

> *[primary] water* when most forcibly mixed (βιαιοτάτην ὕδατος σύνοδον) *[with primary earth]* is dispersed *by [primary] fire only* (*πῦρ μόνον* [διαχεῖτον]) 61a3-4

Primary water *mixed* with primary earth is PW < PE; primary water *most forcibly* mixed with primary earth can only mean that PE is victorious, namely PW < $\boxed{\text{PE}}$, thus forming an earthlike sensible body, a "chuton." Only primary fire (πῦρ μόνον) can disperse this sensible body, changing as above *from internal Consolidation to Dispersion* by proceeding one step in the infinite anthyphairetic division, and thus forming the sensible *firelike* triad $\boxed{\text{PF}}$ < ($\boxed{\text{PW}}$ < PE) (61a3-4).

Thus the case "συνεστηκυῖαν [γῆν]" in 61a2-3 coincides with the case "βιαιοτάτην. . . ὕδατος σύνοδον" in 61a3-4.

11.2. *The sensible bodies are **not elements** and principles, but at least syllables (48b3-c2)*

This passage points out that the traditional four elements theory of sensibles will be modified by Plato, so that there are indeed four kinds of bodies but they are not elements *but at least syllables*, with elements other entities more basic and elemental.

> We must gain a view of the real nature of fire and water, air and earth,
> as it was before the generation of Heaven (πρὸ τῆς οὐρανοῦ γενέσεως),
> and the properties they had *before that time* (καὶ *τὰ πρὸ τούτου πάθη*);
> for *at present* (*νῦν* γὰρ) no one has as yet declared their generation (**γένεσιν**),
> but we assume that men know what fire is, and each of these entities,
> and we call them *principles (ἀρχὰς)* and
> presume that they are *elements (στοιχεῖα)* of the Universe,
> although in truth they do not so much as deserve to be likened (ἀπεικασθῆναι)
> with any likelihood, by the man who has a grain of sense,
> to the class of *syllables only* (ἐν συλλαβῆς εἴδεσιν *μόνον*) 48b3-c2

Indeed, according to Sect. 11.1, every sensible has one of the following forms:

Sensible bodies			
Earthlike	Waterlike	Airlike	Firelike
PW < [PE]			
	[PW] < PE		
PA < (PW < [PE])			
	PA < ([PW] < PE)		
		[PA] < ([PW] < PE)	
PF < (PW < [PE])			
			[PF] < ([PW] < PE)

It is thus understood in what sense every sensible body must be "at least a *syllable.*"

It follows that a primary body by itself cannot be a sensible body, as believed by most scholars. Let us make sure what dyads or even long "syllables"[13] consisting of the same primary body *cannot* constitute a sensible body.

Every sensible entity
either consists of the dyad PW < PE, in Consolidation or in Dispersion,
or of the dyad PW < PE extended by PA or PF.
For every sensible body, the ultimate Receptacle is primary Earth.

11.3. *The participation of the Third Kind in the intelligible and the incommensurability of the primary earth to primary water cause incessant motion and change of the sensibles into each other.*

We are finally in a position to explain why in Plato's model the motion of the world is incessant.

> but we have not declared
> how it is that it is *not* the case (*πῶς δέ ποτε οὐ*) that
> these [primary] bodies are separated according to their several Kinds (κατὰ γένη διαχωρισθέντα), and cease from their motion and passage one through another
> (*πέπαυται* τῆς δι' ἀλλήλων κινήσεως καὶ φορᾶς)
> Wherefore, we shall once more expound the matter as follows. 58a2-4

Plato has managed to deal with the "infinite kosmoi" infinity, and the result is a One Kosmos, with four Kinds theory, participating in the intelligible. But it should not be thought that infinity has disappeared altogether from Plato's dialectics on Sensibles. The potential infinity of the collection of sensibles is still present, but in a

[13] Plato speaks of syllables in various instances. Typical syllables he mentions are ΣΩ (*Theaetetus* 203a8), ΘΕ, ΤΕ (207e8-208a1); certainly, he would never regard something like ΘΘΘΘΘ, and so on, as a syllable.

different manner. In fact, it manifests itself in a "positive" manner, by guaranteeing incessant, *infinite motion* of "kosmos."

[1]. Let us imagine for the moment that all primary bodies were placed in their natural places/regions, with no interaction between them.
[2]. Before the "kosmesis" and the participation to the intelligible, there would be no reason for the primary bodies to move so as to come close to each other and to interact, and in fact this is the basic difference of the universe before "kosmesis" and after. With "kosmesis" and participation of the Third Kind in the intelligible, the primary bodies do not stay for long by themselves; they try to fill their empty "diakena"/"hedrai"/faces, and generate sensibles.

This may be thought of as the will (βούλεσθαι) of "kosmos," whose Soul is like a living creature, whose purpose is to become as similar as possible to the intelligible paradigm. This means that the soul of the kosmos will "compress" and "tighten" to proceed as far as possible in the stages of anthyphairesis, and obtain a better approximation of the paradigm. But the number of available particles of, say, primary water and of primary earth is in any particular case finite, hence the process of approximation and convergence will necessarily stop at some finite stage, exactly because the number of such parts is finite. Thus, the result will be an imperfect sensible approximation of the intelligible (Cf. 10.1, 10.2).

[3]. In both cases, *motion* to the place/region of the victorious results upward in Dispersion, downward in consolidation (Cf. 10.3, 10.4, 10.5).
[4]. Whatever is generated, namely a sensible, might be transformed into another sensible but will eventually be corrupted, because it is not perfect, but a highly imperfect approximation of something perfect but infinite. The world, pressed by the intelligible, strives for something better. Thus, every sensible body is eventually corrupted, dismantled, and the constituting materials, the primary bodies, go to their own places/regions. Thus, we return to stage [1].

> And since in this way and for these reasons
> the generation (γένεσις) of irregularity (τῆς ἀνωμαλότητος)
> is perpetually *(ἀεὶ)* maintained *(διασῳζομένη),*
> it brings about unceasingly, both now and for the future,
> the perpetual *(ἀεὶ) motion (κίνησιν)* of these bodies. 58c2-4

Infinite motion of the "kosmos" is an eventual consequence both of the built-in anthyphairetic infinity, the same built-in infinity that caused the problem of "infinite kosmoi," and of the participation of the Third Kind in the intelligible.

Cornford's 1937, p. 245, interpretation, finding only mechanical causes, and not connecting the infinity of motion with the infinity of incommensurability, and eventually with the participation to the intelligible, is far from adequate.

11.4. *Sensibles differentiated from the intelligible*

11.4.1. *The difference between*
a sensible dyad
(*by assimilation,* ***simultaneously both one and two*** *(52c7-d1), a* ***confounded dyad*** *(Politeia 524c3-9)) and*
an intelligible dyad
(by equalization, separate with ***equalized units*** *(Politeia 524b10-c1))*

The equalization of the two parts of the dyad, in an intelligible Being, has been described in the second hypothesis of the *Parmenides* and analyzed in Negrepontis TA. Thus, the unequal parts One, Being, since part One is greater than part Being, are nevertheless *equalized* by the condition:

the *number* of parts of the One is finite and equal to the *number* of parts of the Being
Parmenides 144e1-3,

a consequence of the philosophical analogue of the Logos Criterion and the anthyphairetic periodicity of the intelligible Being. The same criterion obtains for every intelligible Being, for example, the Being and not-Being in the *Sophistes*, and also for every geometric periodic anthyphairesis. Intelligible equalization is an equivalence, thus symmetric, relation. This equalization of the two parts allows for the platonic definition of the intelligible number Two.

> If, then, each is one (ἐν ἑκάτερον) and both two (ἀμφότερα δὲ δύο),
> the very meaning of 'two' is that
> the soul (νοήσει) will conceive them as separate (κεχωρισμένα) *Politeia* 524b10-c1

By contrast, the sensible two does not consist of parts each of which separately is a unit equal to the other. The description of the sensible dyad as confounded, but not separated as the intelligible, in the *Politeia* 524c is not entirely understandable by what is said in the *Politeia*.

> Sight (ὄψις) too saw (ἑώρα) the great and the small *(Μέγα. . .καὶ σμικρὸν)*, we say,
> not separated (οὐ κεχωρισμένον) but confounded (ἀλλὰ *συγκεχυμένον*). *Politeia* 524c3-5

Plato clarifies this difference as follows:

> *the intelligence* is compelled to contemplate *the great and small*,
> *not thus confounded (οὐ συγκεχυμένα) but as distinct entities (ἀλλὰ διωρισμένα)*,
> in the opposite way from *sensation*. *Politeia* 524c6-9

But it becomes quite clear when completed with the *Timaeus* 52c7-d1 description as "simultaneously both one and two."

One of the two parts of the sensible dyad PW < PE, the Content, is a mass of primary water, and the Receptacle is a mass of primary earth; one of them is the active/victorious, the other the passive/defeated, the passive becomes approximately similar to the active, and the whole dyad becomes active-like (*Timaeus* 51e6-52a4, 52c5-d1). There is thus no equalization in the *sensible,* of the Receptacle & Content parts; the initial opposition results in a fight, a battle, and there is a victorious and a defeated part, the defeated assuming, in an approximate way, the form of the victorious. Here the assimilation is NOT symmetrical, only approximate (Sect. 10.4).

whereas to the aid of the intelligible Being (τῷ δὲ ὄντως ὄντι)
there comes the accurately true logos (ἀληθὴς λόγος), that so long as
one thing will be *generated in the other thing* (οὐδέτερον *ἐν οὐδετέρῳ γενόμενον*)
so that they *become* (γενήσεσθον) *the same thing (ταὐτὸν)*
simultaneously *(ἅμα) both one and two (ἓν καὶ δύο). Timaeus* 52c5-d1

Finally, according to 52c5-d1, the dyad of the intelligible Being does not admit a description of the form Content & Receptacle, with a mass of Content particles inside a mass of Receptacle particles; this passage is more informative than 51e6-52a4, in that it is the intelligible Logos of periodicity and equalization (hinted in the 51c-52a passage by the reference to the self-identical kind) that obstructs the generation of one part in another. The *intelligible* equalization is a consequence of the Logos Criterion of anthyphairetic periodicity, while the *sensible* assimilation is a consequence, in the absence of the Logos and periodicity, of a far weaker condition of only an initial finite approximation of the infinite anthyphairesis, in the form of Receptacle & Content.

11.4.2. *A sensible body participating in intelligible F is sometimes F, sometimes not F (Sumposion 210e-211a, Politeia 479a-b)*

This is a persistent and prominent pre-*Timaeus* feature of the sensibles (*Sumposion* 210e-211, *Politeia* 475e-480a); the latter passage was described in Sect. 2.2. It is clear that Plato regards it as fundamental, but students of Plato have never explained it in a satisfactory manner. With our present analysis of the structure of the sensibles bodies, this property becomes transparent.

For example, the same sensible pair PW < PE is

sometimes heavy, at every even stage of the anthyphairesis, namely when the dyad Receptacle & Content is in *consolidation***,** and then is a *sensible earthlike*, PW < $\boxed{\text{PE}}$, and

sometimes light, at every odd stage of the anthyphairesis, namely when the dyad Receptacle & Content is in *Dispersion*, and then is a *sensible waterlike*, $\boxed{\text{PW}}$ < PE. It is thus clear that the origin and nature of this property is anthyphairetic.

An excellent example of PW < PE is the generation of flesh (74c5-d2), in Dispersion, not E, and the generation of bone (73e1-5), in consolidation, E.

11.5 *Sensibles differentiated from the Third Kind:*

Only the ***nonsensible, geometrical*** *triangles A and B and the primary bodies that are the elements of the Receptacle-Content inequalities can be called* ***"this,"***

while a ***sensible body****, represented by the Receptacle-Content inequalities, must be called* ***only "such" and not "this"*** *(Timaeus 49d4-50a2).*

> Whatsoever object we perceive to be constantly changing from one state to another (ἄλλοτε ἄλλῃ γιγνόμενον), like [sensible] fire, that object, be it [sensible] fire,
> we must never describe as "this" but as "suchlike," (μὴ τοῦτο ἀλλὰ τὸ τοιοῦτον ἑκάστοτε προσαγορεύειν πῦρ) nor should we ever call [sensible] water "this" but "suchlike" (μηδὲ ὕδωρ τοῦτο ἀλλὰ τὸ τοιοῦτον ἀεί),
> nor should we describe any other element, as though it possessed stability (βεβαιότητα), of all those which we indicate by using the terms "this" and "that" (τῷ τόδε καὶ τοῦτο) and suppose ourselves to refer to a definite object.
> For such an object shuns (φεύγει) and eludes (οὐχ ὑπομένον) the names "this" and "that" (τὴν τοῦ τόδε καὶ τοῦτο καὶ τὴν τῷδε) and every name which indicates that they are stable (μόνιμα). Thus we must not call the several elements "these,"
> but in regard to each of them and all together
> we must apply the term "suchlike" (τοιοῦτον) to represent what is always circling round: thus we shall call that which is constantly "suchlike" (τοιοῦτον) by the name of fire,
> and so with everything else that is generated (ἅπαν ὅσονπερ ἂν ἔχῃ γένεσιν). 49d4-e7
> But that "wherein" (ἐνᾧ) they are always, in appearance, coming severally into existence,
> and "wherefrom" (ἐκεῖθεν) in turn they perish,
> in describing that and that alone should we employ the terms "this" and "that"
> (μόνον ἐκεῖνο αὖ προσαγορεύειν τῷ τε τοῦτο καὶ τῷ τόδε προσχρωμένους ὀνόματι)
> 49e7-50a2

According to the "much mis-read" passage

- the sensible entities, namely the entities generated by Dispersion and consolidation, the syllables of primary bodies, seen as fire, air, water, earth, should be always called
  "such" and not "this,"
- while the elements of the receptacle, the primary bodies, and their surfaces should be called "this" and not "such."

H. Cherniss attempted to interpret this passage thus:

> '*ἀεὶ* περιφερόμενον ὅμοιον' here must have the same meaning:
> "always recurring alike," i. e. "always self-identical in its recurrences."
> Cherniss 1954, p. 121

But in fact Plato has explained elsewhere when an entity is called "such":

> 'being *such* (ὁ τοιοῦτοςὢν) he will be *like* **to** *such* (καὶ ἐοικέναι τοῖς τοιούτοις)'
> 'Then each is *such* (τοιοῦτος ἄρα ἐστὶν) as that to which he is *like* (οἷσπερ ἔοικεν.)'
> *Politeia* 349d8-10
> 'as that to which he is *similar* (ᾧγε ὅμοιος ἑκάτερος εἴη), *such* is each (τοιοῦτον καὶ ἑκάτερον εἶναι.)' 350c7-9

As we have seen, the generation of a sensible by dispersion (diakrisis) or consolidation (sugkrisis) makes the sensible to be dissimilar to itself and *similar to others*. Cf. Sections 9.1 and 9.2. So it is *by being similar to others* that a sensible becomes "*such*," contrary to Cherniss' interpretation. If "such" were going to be a fitting name to something that always stays *similar to itself*, then the receptacle, even the intelligible, would be the perfect entity to be called "such." But the receptacle is called "this" and not "such," precisely, because it is not assimilated to something different.

Thus, Cherniss is right in that "such" must refer to some likeness, but is mistaken to believe that it is *self-similarity*, and not *becoming similar to the others*; his explanation is thus totally off the mark.

Theaetetus 201c-202c passage bears some verbal resemblance to the *Timaeus* 49d-50a passage. But there, the differentiation is between

"name only," namely anthyphairetic division only, true opinion/sensible, and

"name and logos," division and Logos criterion of periodicity, intelligible.

Section 12 ***Additional comments on earlier interpretations of the Receptacle***

[1]. Scholars have described the Receptacle as difficult to understand and perhaps self-contradictory, e.g.,

> "The account that Plato gives of the Receptacle itself is at best obscure (as he admits, 49a3-4) and at worst incoherent." Zeyl 2000, p. lxi
> "The receptacle is probably the hardest and most philosophically challenging concept in *Timaeus*." A. Gregory, R. Waterfield, 2008, p. xlix

But in fact, as we have seen, the Receptacle makes perfect sense as a vivid and imaginative description of the double tight remainder-free anthyphairetic inequalities.

[2]. Scholars either ignore or undermine the fundamental "infinite kosmoi" question and problem as simply a puzzling peripheral problem (Cornford 1935; Gregory 2001) or suggest rather exotic interpretations (Rashed 2013).

> This passage [55c] is extremely puzzling. There is nothing in the previous context to suggest a plurality of worlds; p. 221
> the ancients, who seem to have found the passage as obscure as it is to us
> Cornford 1937, p. 220

But, as we saw in Sect. 3, Plutarch had a quite good notion of the problem of infinite kosmoi, besetting Plato's pre-*Timaeus* theory of sensibles.

> The passage at Timaeus 55c has long been considered a rather puzzling one.
> Gregory 2001, pp. 42–43

Both Cornford and Gregory express puzzlement that Plato returns to the question at 55c, right before identifying the primary bodies with the corresponding canonical solids. But, as we have shown, the problem of infinite kosmoi is the sole problem that

necessitated the introduction of the Receptacle. Treating the problem of infinite kosmoi as peripheral makes it impossible to explain the necessity in introducing the Third Kind.

[3]. Most scholars (including Cornford 1937; Algra 1995; Zeyl 2000; Gregory 2001; Miller 2003; Gregory and Waterfield 2008; Broadie 2011) do not suspect, or even mention, the fundamental nature of the incommensurability between the two fundamental triangles, and therefore between the primary earth and each of the other three primary bodies.

Scholars believe that earth does not in any way interact with the other three kinds of bodies and are unable to understand why Plato appears to isolate the earth from the other three kinds. Thus, e.g., Cornford and Zeyl try to justify the situation, presenting it as a necessary evil following Plato's decision to introduce the correspondence of the traditional "elements" with the four canonical solids:

> The *exclusion of earth* from the cycle of transformation is simply a consequence of the decision to assign the cube to earth. Cornford 1937, p. 215–6
> The *exclusion* is driven by *his theoretical commitments*... Plato was *eager* to assign a distinct solid to each of the four kinds. Zeyl 2000, p. xix

Pohle's 1971, p. 46, even timid realization of the fundamental role of incommensurability between the two triangles and therefore between the primary earth and each of the other three primary bodies is definitely the exception:

> The fact that Plato appears at 54A to leave open the possibility that some mathematical conjecture might be disproved with the concomitant destruction of his "too long" a reason why his elements are best, certainly does not rule out considerations of areal *incommensurability*. Issues having to do with the nonexistence of transformation reactions between earth particles and the other three kinds supply further indirect evidence that Plato was aware of the problems raised by areal *incommensurability*, as does the Theaetetus.

As a result, authors in general follow Cornford 1937 (p. 162), in erroneously believing that the primary bodies are themselves sensible bodies, while in fact the primary bodies replace and substitute for the intelligible Being as the cause of the sensibles.

[4]. There is no realization or agreement among scholars (a) that the Receptacle coincides with the "diakena" of the greater primary body, and (b) that the "diakena" in turn coincide with the "hedrai"/faces of the greater primary body.

Thus, Broadie 2011, p. 238, differentiates the Receptacle from the primary bodies:

> Timaeus provides *two accounts* of the genesis of the four materials: one in which the Receptacle is presented as the 'third kind' along with the two metaphysical kinds familiar from the first beginning; and one in which god constructs the elements geometrically. ...
> the Receptacle and its content are whatever they need to be, in order to ensure that the elements are free of any metaphysically primordial self-sufficient agenda that could interfere with their subservience to the intelligence that made the cosmos.

In fact, both accounts are vivid and imaginative descriptions of the same geometric anthyphairetic inequalities, and the Receptacle account is no more suggestive of participation to the intelligible than the openly geometric one. Participation is assured via the Soul of the World, a claim independent of both accounts.

Gregory 2001, p. 36, correctly differentiates the Receptacle from geometric atomism:

> Geometrical atomism (GA) is introduced at 53c. By geometrical atomism I mean the theory that the elements (earth, water, air, fire) can be analysed into three-dimensional particles of definite shape (cubes, octahedra, eikosahedra, tetrahedra, which I shall call 'atoms', in the modern sense), and that these particles can be further subdivided into planes, and that these planes can be further subdivided into one of two types of triangle. GA does not sit entirely easily with the receptacle passage. It may develop or modify the receptacle theory, and certainly it has a considerable bearing on the nature of the receptacle. At the very least we need to think carefully about how the entities proposed by GA relate to the receptacle. What is undeniable is that the rest of the Timaeus (53c *ad fin*) discusses the phenomena in terms of GA and not the receptacle.

However, the role of the primary bodies is erroneously described as Geometric Atomism GA by Gregory, because it is not realized that incommensurability and its approximation by "convergents" are at work. As we have shown, the precise relation between Receptacle and primary bodies is the following: The Receptacle coincides with the faces/"diakena"/"hedrai" of the greater primary body, its Content being the smaller primary body, the generated "confounded" dyad the sensible body.

[5]. Brisson and Ofman 2020. The authors correctly reject interpretations of *Timaeus' Receptacle* in terms of "atomism," citing the existence of incommensurable magnitudes, but they fail to realize the fundamental importance of incommensurability and of periodic anthyphairesis for understanding (all of Plato, and in particular) the Receptacle. The authors spend quite a bit of time attempting to dissociate the *Timaeus* triangles from the "indivisible lines," not realizing that Plato in the *Parmenides*, as Aristotle correctly states, strongly supports their existence, whose paradigm is the One of the second hypothesis, rejecting at the same time the geometric point, whose paradigm is the One of the first hypothesis, without in the least compromising the infinite divisibility of geometric magnitudes (cf. Negrepontis 2019, Sect. 3).

[6]. Harte 2010, p. 140

> On this view, the reality of something being fi re comes precisely from the formal mathematical shaping the demiurge imposes, a fact that goes hand in hand with the view that the being of Fire—its formal nature—is constituted by something **essentially intelligible**, and not perceptible.

The geometrical/mathematical content of the four primary bodies does not endow them with some participation to the intelligible. This participation of the sensibles to

the intelligible is assured only by the passage 58a4-b8 (in conjunction with the participation of the Soul of the world in the intelligible), which Harte does not invoke.

Epilogue: The Debt of *Timaeus*' Receptacle to Theaetetus

The dedication of a work on the Receptacle in Plato's philosophic dialogue *Timaeus* to the mathematician Theaetetus calls for some justification. Throughout our present work, our thesis is that behind *Timaeus*' Receptacle one senses the mathematical achievements of the great mathematician Theaetetus.

The canonical comparison of Triangle A and Triangle B (Sect. 7), on the basis of which alone it is possible to show that these triangles are incommensurable, needs fine computations relating the sides of the faces of a canonical solid to the diameter of the sphere in which it is comprehended, found in the second half of Propositions 13, 15, 18 of the Theaetetean Book XIII of the *Elements*.

The mathematical Proposition in Sect. 4 on the generalized side and diameter numbers (modern "convergents" of continued fractions), originated by the Pythagoreans, must owe a great deal to Theaetetus, as the Proposition is needed for a reconstruction of the proof of his great Theorem (essentially in Sects. 7, 8, and 9) stating the following: If a,b are line segments, such that $a > b$, a,b incommensurable, and a^2, b^2 commensurable, then the anthyphairesis of a to b is palindromically periodic. In Negrepontis 2018, it is shown (a) how the study of the *Politicus* dialogue leads to the certain conclusion that Plato had perfect knowledge of this Theorem, and (b) how the Theaetetean Book X of the *Elements* contains all the tools necessary for the proof of this Theorem, in particular the notions of the *apotome* line and of the line *of two names*, and principally their conjugacy in Propositions X.112–114. A complete proof of the palindromic periodicity theorem, using only the tools of Book X of the *Elements*, is given in Negrepontis (Notes by V. Kleftaki), 2003.

For the definition of the primary bodies and the Third Kind (Sects. 6 and 7), Plato needs the laborious constructions of the first half of Propositions 13–16 of the Theaetetean Book XIII of the *Elements*.

Plato claims that the Third Kind participates in the intelligible and generates sensibles so as to eliminate the infinite kosmoi (Sects. 8, 9, and 10); Plato needs again the very essential Propositions of Book XIII.13–18 on the comprehensibility in the sphere of the canonical solids and on their uniqueness.

References

Algra K (1995) Concepts of space in Greek thought. E.J. Brill, Leyden

Allen RE (1997) Plato's Parmenides, translated with comment. University of Minnesota Press, Minneapolis, 1983; Revised edition, Yale University Press, New Haven/London, 1997

Annas J (1981) An introduction to Plato's Republic. Clarendon Press, Oxford

Babbitt FC (1936/1978) Translation of Plutarch, The Obsolescence of Oracles. Moralia V:347–501. Loeb Classical Library. [Translation of texts in paper, with modifications by the authors]

Bostock D, Waterfield R (1996) Aristotle physics, translated by R.W, introduction and notes by D.B. Oxford University Press, Oxford. [Translation of texts in paper, with modifications by the authors]
Brisson L, Ofman S (2020) The mathematical anti-atomism of Plato's Timaeus. ffhal02923266f
Broadie S (2011) Nature and divinity in Plato's Timaeus. Cambridge University Press
Burnyeat MF (2005) EIKÔS MYTHOS. Rhizai II:2, 143–165
Cherniss H (1954) A much misread passage of the Timaeus (Timaeus 49C7–50B5). Am J Philol 75: 113–130
Cornford FM (1937) Plato's cosmology: the Timaeus of Plato. Routledge
Farmaki V, Negrepontis S (2021) The paradoxical nature of mathematics. In: Papadopoulos A (ed) Topology and geometry: a collection of papers dedicated to Vladimir G. Turaev. European Mathematical Society Publishing House, Berlin, p. 599–642
Gregory A (2001) Plato's philosophy of science. Bloomsbury Academic, London
Gregory A, Waterfield R (2008) Plato's Timaeus and Critias, translated by R. Waterfield, with an introduction and notes by A. Gregory, Oxford world's classics. Oxford University Press
Harte V (2010) The receptacle and the primary bodies: something from nothing? In: Mohr RD, Sattler BM (eds) One book the whole universe Plato's Timaeus today. Parmenides, Las Vegas, pp 131–140
Miller D (2003) The third kind in Plato's Timaeus. Vandenhoeck & Ruprecht, Goettingen
Negrepontis S (Notes by V Kleftaki) (2003) A reconstruction of the proof of the palindromic periodicity of the anthyphairesis of commensurable in power only lines, employing only tools from Book X. of Euclid's elements, 2003. Manuscript (22 pages)
Negrepontis S (2005) The anthyphairetic nature of the platonic principles of infinite and finite. In: Proceedings of the 4th Mediterranean conference on mathematics education, 28–30 January 2005, Palermo, Italy, pp 3–27
Negrepontis S (2012) Plato's theory of knowledge of Forms by Division and Collection in the Sophistes is a philosophic analogue of periodic anthyphairesis (and modern continued fractions), arXiv:1207.2950, 12 Jul 2012
Negrepontis S (2018) The anthyphairetic revolutions of the Platonic ideas. In: Sialaros M (ed) Revolutions and continuity in Greek mathematics. Walter de Gruyter, pp 335–381
Negrepontis S (2019) Plato on geometry and the geometers. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, pp 1–88
Negrepontis S (2024) The anthyphairetic interpretation of knowledge as recollection in Plato's Meno 80d-86e and 97a-98b. In: Sriraman B (ed) Handbook of history and philosophy of mathematical practice. Springer Nature, Cham
Negrepontis S From musical intervals to Hapseis and Logoi: the periodic anthyphairetic nature of the one of the second hypothesis in Plato's Parmenides, proceedings of the conference Mathematiques et musique : des Grecs a Euler, 10–11 September, 2015, Strasbourg, eds Hascher X, Papadopoulos A, Hermann, Paris, to appear
Negrepontis S, Pappa Birba S True opinion of an intelligible being and knowledge of sensible bodies in the Politeia 475e–480a and the divided line, under preparation
Plato, Translation of Plato in twelve volumes. Vol. 1 (Phaedo) translated by H. N. Fowler, 1966; Vols. 5 & 6 (Republic) translated by P. Shorey, 1969; Vol. 9 (Timaeus) translated by W.R.M. Lamb, 1925; Vol. 12 (Kratulos, Theaetetus) translated by H, N. Fowler, 1921. Harvard University Press/William Heinemann, London/Cambridge, MA. [Translation of texts in paper, with modifications by the authors]
Pohle W (1971) The mathematical foundations of Plato's atomic physics. Isis 62(1971):36–46
Rashed M (2013) Plato's five worlds hypothesis (Ti. 55cd), mathematics and universals. In: Chiaradonna R, Galluzzo G (eds) Universals in ancient philosophy. Edizioni della Normale, Pisa, pp 87–112
Vlastos G (1954) The third man argument in the Parmenides. Philos Rev 64(1954):319–349
Wigner EP (1960) The unreasonable effectiveness of mathematics in natural sciences. Commun Pure Appl Math 13(1960):1–14
Zeyl DJ (2000) Translation, Plato's Timaeus. Hackett Publishing, Indianapolis/Cambridge, MA

The Indivisible Line Coincides with the Platonic Intelligible Being

Stelios Negrepontis

But Plato, for the sake of concealment,
employed mathematical names,
as screens of the truth of the things,
in the same manner as theologists employed fables,
and the Pythagoreans symbols.
Proclus, *Commentary to Timaeus* 2,246,4-7

Contents

1 Introduction 602
2 Periodic Anthyphairesis and Theaetetus' Logos Criterion 604
2.1 The Concept of Anthyphairesis 604
2.2 Finite and Infinite Anthyphairesis 604
2.3 Periodic Anthyphairesis 605
2.4 The Theaetetean Theory of Ratios of Magnitudes 605
2.5 Proposition. The [Theaetetean Preservation of] Logos Criterion for Anthyphairetic Periodicity 606
2.6 The Already Infinite (Hede Apeiron) in the *Philebus* 16e1-2 Coincides with the Theaetetean Logos Criterion 607
3 Plato's and Xenocrates' *Indivisible Lines* Are Considered Mysterious by Modern Scholars 607
4 The *Indivisibility* of Plato's Intelligible Being in the *Sophist* 226b1-231b8 and the *Phaedrus* 277b5-c6 608
4.1 The *Indivisibility* of the Intelligible Being Noble Sophistry in the *Sophist* 226b1-231b8 608
4.2 The *Indivisible Line* and Logos in the *Phaedrus* 277b5-c6 610
5 Proclus' and Simplicius' Explanation of the Sense in Which an Intelligible Being Is *Indivisible* 611
5.1 Proclus, *Commentary to Timaeus* 2,245,23-246,7 611
5.2 Simplicius, *Commentary to Aristotle's Physics* 140,6-13, and 142,16-18 612

S. Negrepontis (✉)
Department of Mathematics, Athens University, Athens, Greece

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_139

6 Our Interpretation of the Meaning of *Indivisibility* of an Intelligible Being 613
6.1 Every Intelligible Being Is a Dyad in the Philosophical Analogue of Infinite Anthyphairesis, Satisfying the Logos Criterion for Periodicity 613
6.2 The Dialectical Number defined by Means of the Plus One Rule (*Parmenides* 148d-149d) in an Intelligible Being 614
6.3 The Paradox of an Infinite *Multitude* of Parts Versus a Finite *Number* of Parts: The Multitude of the Parts in the Intelligible Being Is Infinite, but the Dialectical Number of These Parts Is Finite 617
6.4 The Whole Part B Is *Indivisible* in the Sense that It Is Infinitely Divisible but nevertheless *Unable to Generate New Units* for Dialectical Numbers After the Completion of a Period 618
6.5 *The Units of the Dialectical Numbers* Are the intelligible *Kinds/Eide* 618
6.6 The Whole Part B Is *Indivisible* in the Sense that It Is Infinitely Divisible in Parts but nevertheless *Unable to Generate New Kinds* After the Completion of a Period 618
6.7 Examples 619
6.8 Confirming Evidence for Our Interpretation 619
7 We Are Now in a Position to Obtain Our Interpretation of the Heretofore Mysterious Aristotle's *Metaphysics* 992a19-22 Passage, Since the One of the First Hypothesis in the *Parmenides* Is Identified with the Geometric Point and the One of the Second Hypothesis in the *Parmenides* Is Identified with the Intelligible, Hence Indivisible, Being 620
7.1 Aristotle, *Metaphysics* 992a19-22 on Plato's Indivisible Lines 620
7.2 Aristotle's Passage Is Understood as a Comment on the Comparison Between the First and the Second Hypothesis in Plato's *Parmenides*, Fully Confirming that the Indivisible Line Coincides with the Paradigmatical One of the Second Hypothesis in the *Parmenides* 621
8 The Hypothesis-Free/Anhupotheton in the *Republic* 624
8.1 The Passages *Republic* [D1] 510b6-9, [D2] 511b3-7, [D3] 533c7-d3, [D4] 534b3-d2 Identify the Anhupotheton/Hypotheses-Free with the Intelligible Being 625
8.2 The Term "Anhupotheton" in the *Republic* Suggests the Indivisibility of the Intelligible Being 627
8.3 According to the Passages *Republic* [G1] 510b4-6, [G2] 511b3-7, [G3] 510c2-d3, the True Foundation of Geometry Is Not the Hypotheses of the Geometers but the Hypothesis-Free/Anhupotheton 627
9 The Geometric Analysis Is Connected with the Ascent from Hypotheses to the Anhupotheton in Plato's Intelligible Being, and the Geometric Synthesis and Deduction with Plato's Descent from the Anhupotheton 629
9.1 Geometric Analysis Is Connected with Plato's Division (and Logos) by Proclus *In Euclidem* 211,18-212,4 629
9.2 The Ascent and Descent in Plato's *Republic* Are Connected with the Analysis and Synthesis, Respectively, in Pappus' *Collection 634,11–636,14* (Negrepontis and Lamprinidis 2008) 630
10 The Two Infinites, the Dialectical and the Eristic Infinite, in the Platonic Dialogue *Philebus* 16c5-17a5, Essentially Coincide with the Two Infinites, the Actual and the Potential Infinite, Respectively, in Aristotle's *Physics* 204a-207b 631
10.1 Plato Accepts the *Dialectical, "Already" (Hede Apeiron) Infinite,* Namely the Anthyphairetically Periodic Infinite Induced by the Logos Criterion, and Rejects the *Eristic Infinite,* Namely the Infinite Established Without the Logos Criterion *Philebus* 16c5-17a5 631
10.2 Aristotle in *Physics 204a-207b* Accepts the Potential Infinite but Rejects the Actual, Completed Infinite 632

10.3 The Aristotelian Potential Infinite Essentially Coincides with the Platonic Eristic Infinite, and the Platonic Dialectical Already Infinite Essentially Coincides with the Aristotelian Actual Infinite 633
11 The Self-Motion of the Soul in Plato's *Phaedrus* 245c5-9 and the Second Version of Zeno's Dichotomy Paradox in Aristotle's *Phaedrus*, Are Each a Dialectical/Actual Infinite (*Physics 257a34-b20* and *263a4-264b9*), a Platonic Intelligible Being, but Nonexistent According to Aristotle 635
11.1 The Self-Motion of the Soul in Plato's *Phaedrus* 245c5-9 Is a Platonic Intelligible Being, a Dialectical/Actual Infinite, but Nonexistent According to Aristotle's *Physics* 257a34-b20 635
11.2 The Second Version of Zeno's Dichotomy Paradox Is a Dialectical/Actual Infinite (*Physics 263a4-264b9*), Identical with a Platonic Intelligible Being, but Nonexistent According to Aristotle 638
12 The Application of Zeno's Basic Argument, as Interpreted in Negrepontis (2023a), Is Employed, in the Pseudo-Aristotelian Work *On Indivisible Lines* 968a18-b4, with the Almost Contradictory Property of the Second Version of Zeno's Dichotomy Paradox, of Not Only Touching but Numbering the Infinites as Well, Shows that Zeno's True Being Coincides with the Indivisible Line (Atomos Gramme, 968b4), Hence, According to Sects. 2, 3, 4, 5, 6, 7, 8 and 9 Above, with Plato's Intelligible Being 640
12.1 The Pseudo-Aristotelian Passage *On Indivisible Lines* 968a18-b4 640
12.2 An Outline of Our Anthyphairetic Interpretation of the Pseudo-Aristotelian Passage *On Indivisible Lines* 968a18-b4 as a Special Case of Zeno's Basic Argument 641
12.3 Of Interest Are Two Sentences Appearing in the Passage 641
13 Conclusion for Sections 10, 11, and 12 642
Translations 642
References 643

Abstract

In the present chapter, we provide a novel interpretation of the concept of Plato's and Xenocrates' *indivisible line*, in fact, we show that indivisibility is just another description of the Platonic intelligible true Being. Our claim and arguments are based on our earlier interpretation of Plato's intelligible Being as the philosophic analogue of a dyad of opposite kinds in periodic anthyphairesis (as revealed primarily in the *Meno*, the second hypothesis of the One in the *Parmenides*, and the *Sophist*), and take into account the ancient sources on the indivisible by Plato (*Sophist, Phaedrus, Republic, Philebus*), Aristotle (*Metaphysics*), Proclus, and Simplicius. Additional arguments, based on Zeno (Dichotomy Paradox), and Aristotle (*Physics,* and the pseudo-Aristotelian *On Indivisible Lines*), establish that Zeno's true Being is also indivisible. Earlier attempts by modern Platonists were not successful, mainly because these have not taken into account the geometric/anthyphairetic nature of a true Being.

Keywords

Plato · Zeno · Theaetetus · Aristotle · Proclus · Simplicius · Intelligible being · Indivisible line · Anthyphairesis · Logos Criterion · Already infinite · Dialectical infinite · Eristic infinite · Active infinite · Potential infinite · Self-moving Soul · Dichotomy Paradox

AMS Codes

00A30 · 01A20

1 Introduction

As was acknowledged by A. E. Taylor in 1960 and more recently by M. Rashed (2013), Plato's and Xenocrates' *indivisible line* remains a mystery for the scholars on Plato (Sect. 3). The chapter aims to resolve the mystery and provide a satisfying interpretation of Platonic indivisibility, based on our fundamental interpretation of an intelligible Being as the philosophic analogue of a dyad in periodic anthyphairesis.

The Division of the intelligible Being Noble Sophistry in the *Sophist* 226b1-231b8 and the description of the Division of every intelligible Being in the *Phaedrus* 277b5-c6 passage, in conjunction with our anthyphairetic interpretation of the Divisions in the *Sophist* (Negrepontis 2012), quickly lead us to the conclusion that an intelligible Being becomes "indivisible" exactly at the moment that the Logos Criterion establishing anthyphairetic periodicity is instituted (Sects. 2 and 4).

The exact sense in which an intelligible Being becomes indivisible is lucidly described by Proclus and Simplicius (Sect. 5). For our interpretation of indivisibility (given in Sect. 6), we take into account our earlier interpretation (in Negrepontis 2024) of the definition of dialectic number, namely the number with equalized units, in an intelligible Being, in the second hypothesis of the *Parmenides* and in the *Sophist,* in terms of the "plus one rule," and its finiteness, resulting from the periodicity of the set of ratios of successive anthyphairetic remainders. The identification of the kinds (eide) with the units of the dialectical number leads to the indivisibility of the intelligible Beings, and of course, does not affect the divisibility *ad infinitum qua* magnitudes.

Aristotle's mysterious *Metaphysics* 992a19-22 passage, on Plato's rejection of the point in favor of the indivisible line as the correct foundation of Geometry, is now understood in terms of the rejection of the One of the first hypothesis, partless as a geometric point, versus the acceptance of the self-similar One of the second hypothesis in the *Parmenides* by Plato, both as a true Being and as the foundation of Geometry and Arithmetic (Sect. 7).

In conclusion, the *indivisibility* of an intelligible Being is simply a form of poetic license (a "screen of concealment," as Proclus has it) that allows the transfer, using the *plus one rule* enunciated in the *Parmenides* 148d-149d, of the periodicity of the sequence of the ratios of the successive remainders of the anthyphairesis of the intelligible Being into the finiteness of the number of parts.

The fact that the true foundation of Geometry and Arithmetic is by means of the intelligible Being is expanded in the *Republic* 510b4-511c2 and 533b6-d3, 534b3-d2, where the kinds (eide) generated by the philosophical anthyphairetic Division are

identified with the hypotheses (definitions and postulates) of the geometers, from which the geometers derive, moving downwards/synthetically, the whole body of mathematical propositions. However, the geometers, keep their hypotheses unmoved and undivided, as entities known and true, although they are not, and are content to move downwards, synthetically and deductively to a large extent because of their use of sensible geometric diagrams; thus, the geometers are unable to move upwards, analytically, by a philosophical analogue of the anthyphairetic division of the hypotheses and Logos anthyphairetic periodicity, achieving the status of anhupotheton/hypothesis-free, namely, the status that no new hypotheses/kinds (eide) are generated anymore, exactly the status of indivisibility. Thus the *Republic's* criticism of the geometers fully confirms the identification of the true, intelligible Being with the indivisible line (Sect. 8). The close relation of the upward movement with anthyphairetic division, Logos, and Analysis, while Synthesis and Deduction is alloted to the downward movement, but also with the generation of Geometry and Arithmetic from the intelligible Being, analyzed by Negrepontis and Lamprinidis (2008), is outlined in Sect. 9.

Sections 10, 11 and 12 examine the connections of the indivisible line, not only with Plato's intelligible Being, but, going backward in time, with the second version of Zeno's Dichotomy Paradox (stated in Aristotle's *Physics* 263a-b), both in Aristotle's *Physics* and in the pseudo-Aristotelian work *On Indivisible Lines*.

The passage *On Indivisible Lines* 968a18-b4 on the second version of Zeno's Dichotomy Argument, traditionally considered as hopelessly obscure, is recognized as an instance of Zeno's Basic Argument (presented in *Parmenides* 127d6-128a3), where the almost contradictory Compresence of Opposites property of the true Beings is the numbering of the infinites; the general interpretation of Zeno's Basic Argument, given in Negrepontis (2023a), applied to this instance, shows that Zeno's true Being is identified with the indivisible line (Sect. 12).

In Sect. 10, we obtain the identification of the two Platonic infinites, the dialectical (already infinite, based on the Logos Criterion) versus the eristic, in the *Philebus* 16c5-17a5, with the two Aristotelian infinites, the actual and the potential, respectively, in Aristotle's *Physics*. One basic argument for the identification of the Philebean dialectical infinite with the Aristotelian active infinite is that both are closely associated with the equalization of opposites, fundamental for Plato's philosophy and rejected as impossible by Aristotle.

In Sect. 11, we employ the identification obtained in Sect. 8 to establish the identical nature of Zeno's true Being (in the form of the second version of Zeno's Dichotomy Paradox) and Plato's intelligible true Being (in the form of the definition of the Soul as an entity in self-motion in the *Phaedrus* 245c5-d7). The reason for this identification is that both are deemed impossible to exist because each is an actual infinite (Plato's Soul in *Physics* 257a34-b20, Zeno's true Being in Physics *263-264*).

It follows from Sects. 10, 11 and 12 that the indivisible line is identified not only with Plato's intelligible Being, but also with Zeno's true Being.

2 Periodic Anthyphairesis and Theaetetus' Logos Criterion

2.1 The Concept of Anthyphairesis

Let a, b be two natural numbers or magnitudes (e.g., line segments), with $a > b$.

We measure the great a with the small b and find how many small bs are needed to cover the large a. There are then a natural number k_0, the quotient, and c_1, the remainder, such that

$$a = k_0 b + c_1, \text{with } c_1 < b.$$

This is the first step of the anhyphairesis of a to b.

If c_1 is zero, this is the end of the process. Otherwise, a new pair of great and small is formed, namely the pair $b > c_1$. Notice that the role of b is reversed: in the first step b is small, and in the second step b is great.

We repeat the process. There are a natural number k_1 and c_2, such that

$$b = k_1 c_1 + c_2, \text{with } c_2 < c_1.$$

We continue in this way:

$$c_1 = k_2 c_2 + c_3, \text{with } c_3 < c_2,$$
$$\ldots$$
$$c_n = k_{n+1} c_{n+1} + c_{n+2}, \text{with } c_{n+2} < c_{n+1}$$
$$\ldots$$

2.2 Finite and Infinite Anthyphairesis

- If this process ends at some point, namely for some index n, the remainder c_n divides the immediately previous remainder c_{n-1} ($c_{n-1} = k_n c_n$), then we say that the anthyphairesis of a to b is *finite*.

  This is always the case if a and b are natural numbers, and the last remainder c_n is then the *greatest common measure/divisor* of a and b (Propositions VII.1 & 2 of the *Elements*).
- If this process does not end, then we say that the anthyphairesis of a to b is *infinite*. This can happen only if a and b are not natural numbers. It is then proved that a and b do not have any common measure, namely that a and b are *incommensurable*:

Proposition X.2 of the *Elements*
If a and b are magnitudes, $a > b$, and the anthyphairesis of a to b is infinite, then a and b are incommensurable.

Note that in this case there is an infinite, strictly decreasing sequence of remainders:

$$a > b > c_0 > c_1 > c_2 > \ldots > c_n > c_{n+1} > \ldots.$$

2.3 Periodic Anthyphairesis

The most important case of infinite anthyphairesis is the periodic one.

An anthyphairesis is periodic if the sequence of quotients

$$k_0, k_1, k_2, \ldots, k_n, k_{n+1}, \ldots$$

is a periodic sequence of natural numbers.

2.4 The Theaetetean Theory of Ratios of Magnitudes

The Pythagoreans had an excellent theory of proportion for natural numbers (Book VII of the *Elements*), and, probably, for commensurable magnitudes, but they definitely did not have a theory of proportion for magnitudes.

Aristotle's *Topics* 158b24-159a2, a passage important for the history and reconstruction of Greek mathematics, informs us that before the Eudoxian theory of ratios of magnitudes presented in Book V of the *Elements*, there was an earlier theory based on the following definition.

Definition (Def. V.3 in Euclid's Elements) Two magnitudes a and b possess ratio a/b only if a and b are *homogeneous* (homogene) (namely only if a and b are both straight lines, or both circular lines, or both surfaces, or both volumes).

Definition If a, b are homogeneous magnitudes and c, d are homogeneous magnitudes, then

$$a/b = c/d \text{ if and only if } \mathrm{Anth}(a, b) = \mathrm{Anth}(c, d)$$

(where Anth(a, b) denotes the sequence of successive *quotients* of the anthyphairesis of a to b).

Although Aristotle does not mention the creator of this theory, it is generally accepted that the theory of ratios based on anthyphairesis is due to Theaetetus. A theory of ratios based on this pre-Eudoxian definition, if it were meant to be a theory as general as the Eudoxian theory, would need *Eudoxus' condition* (Def. 4 in Book V. of the *Elements*), and so would be historically untenable. Knorr (Knorr 1975) reconstructs such a theory, but, as carefully noted by Knorr, his crucial Theorem

6, p. 338, the analogue of the fundamental Eudoxian Proposition V.8 of the *Elements*, makes essential use of Eudoxus' condition (V, Def. 4).

There is however evidence that this pre-Eudoxian theory applied only to a limited class of pairs of magnitudes, chiefly the lines that are rational with respect to an assumed line, in the sense of Definition 3 of Book X of the *Elements* – hence the term "alogos" ("ἄλογος") for those ratios whose square is not commensurable. One anonymous comment in *Scholia in Euclidem* V.22 (eds: Heiberg and Stamatis 1977) alludes to a restricted, in fact, quadratic, pre-Eudoxian theory of magnitudes, and another, *Scholia in Euclidem* X.2 (eds: Heiberg and Stamatis 1977), criticizes the general Eudoxian theory, clearly from a Platonic point of view and in favor of the earlier restricted quadratic theory of ratios of magnitudes, which was quite adequate for the purposes of dealing with quadratic incommensurabilities, and for which, as can be shown, there is no need to appeal to Eudoxus' condition.

The primitive *Pythagorean method of preservation of Gnomons/Application of Areas* for finitizing and obtaining full knowledge of the infinite anthyphairesis of diameter to side, evolved, in Plato's time, into the more sophisticated and general *Theaetetean Criterion of [the preservation of] Logos* for establishing periodic anthyphairesis. The formulation of such a Criterion presupposes a theory of ratios and proportions of magnitudes. With this definition of proportion, the Logos Criterion takes the following form:

2.5 Proposition. The [Theaetetean Preservation of] Logos Criterion for Anthyphairetic Periodicity

Let a, b be two line segments, with a > b, their anthyphairesis given as follows:

$$
\begin{aligned}
a &= k_0 b + c_1, & b &> c_1,\\
b &= k_1 c_1 + c_2, & c_1 &> c_2,\\
&\cdots\\
c_{n-1} &= k_n c_n + c_{n+1}, & c_n &> c_{n+1},\\
c_n &= k_{n+1} c_{n+1} + c_{n+2}, & c_{n+1} &> c_{n+2},\\
&\cdots\\
c_{m-1} &= k_m c_m + c_{m+1}, & c_m &> c_{m+1},\\
c_m &= k_{m+1} c_{m+1} + c_{m+2}, & c_{m+1} &> c_{m+2},\\
&\cdots
\end{aligned}
$$

and assume that there are indices n < m such that $c_n/c_{n+1} = c_m/c_{m+1}$

[namely, $\mathrm{Anth}(c_n, c_{n+1}) = \mathrm{Anth}(c_m, c_{m+1})$] ("Logos Criterion").

(in particular, assume that there are indices n<m, such that $c_n c_{m+1} = c_m c_{n+1}$).

Then the anthyphairesis of a to b is eventually periodic, in fact Anth(a, b) = [k_0, k_1, . . ., k_n, period (k_{n+1}, k_{n+2}, . . ., k_m)].

The *Proof* is an application of the Theaetetean definition of proportion. Thus,

$$\begin{aligned}
&\mathrm{Anth}(a,b) = \\
&[k_0, k_1, \ldots, k_n, \mathrm{Anth}(c_n, c_{n+1})] = \\
&[k_0, k_1, \ldots, k_n, k_{n+1}, k_{n+2}, \ldots, k_m, \mathrm{Anth}(c_m, c_{m+1})] = \text{(by Logos Criterion)} \\
&[k_0, k_1, \ldots, k_n, k_{n+1}, k_{n+2}, \ldots, k_m, \mathrm{Anth}(c_n, c_{n+1})] = \\
&[k_0, k_1, \ldots, k_n, k_{n+1}, k_{n+2}, \ldots, k_m, k_{n+1}, k_{n+2}, \ldots, k_m, \mathrm{Anth}(c_m, c_{m+1})] = \text{(by Logos Criterion)} \\
&\ldots \\
&\mathrm{Anth}(a,b) = [k_0, k_1, \ldots, k_n, \text{period } (k_{n+1}, k_{n+2}, \ldots, k_m)].
\end{aligned}$$

2.6 The Already Infinite (Hede Apeiron) in the *Philebus* 16e1-2 Coincides with the Theaetetean Logos Criterion

The Theaetetean Logos Criterion for anthyphairetic periodicity plays a pivotal role in Plato's philosophy, as seen in the *Meno*, in the second hypothesis of the *Parmenides*, and in the *Sophist's* Divisions. The Logos Criterion can be understood as follows: once the Logos Criterion has been established at some *finite* stage of the process of anthyphairesis, we then have established periodicity and thus infinity of the process, and thus at some *finite* stage we have *already arrived at infinity.* In fact, this is exactly the way, as *already infinite* (*hede apeiron*), in which Plato describes the Logos Criterion for anthyphairetic periodicity in the *Philebus* 16e1-2 (as results from the comparison of the second hypothesis in the *Parmenides* with the *Philebus* passage 16c-17a (given in Negrepontis 2024, Sect. 5).

3 Plato's and Xenocrates' *Indivisible Lines* Are Considered Mysterious by Modern Scholars

The *indivisible line*, formally introduced by Xenocrates, but mentioned informally by Plato is something of a mystery for modern scholars of Plato.

A. E. Taylor (1960, p.505–506), writes:

> Plato stated that the "point" was a "fiction of the geometers," and spoke, instead, of the "starting point of the line" (Met. A 9920 20) . . ..
>
> What Plato may have meant by the expression *[[indivisible line]]* we can only conjecture. As a conjecture, I offer the suggestion that his intention is precisely to deny the conception attributed to some Academic, apparently Xenocrates, by the Peripatetic tract. A line, however short, is *"indivisible"* in the sense that you cannot divide it into elements which are not themselves lines in other words, it is a "continuum."

This is certainly not a satisfactory explanation. It is unsupported by ancient sources, it would imply that *every* line is an indivisible line, and it is not explained how such an interpretation would be of any help to Plato in finding a better foundation for Geometry.

Taylor's failure to recognize, in his 1960 work, the anthyphairetic periodicity as the underlying cause of the indivisibility of every intelligible Being (as I will explain in the sequel) is certainly disappointing, because Taylor (1926) was, as far as we know, the first modern scholar of Plato to understand, albeit in an incomplete manner, the connection of platonic philosophy with infinite anthyphairesis.

M. Rashed (2013, p. 106–107) writes

> These remarks may account for Aristotle's *mysterious attribution* to Plato of a theory of *indivisible lines.* According to Aristotle, Plato considered the geometrical point as a doctrine (δόγμα) put forward by geometers, i.e., by geometers only. As such, the geometrical point is deprived of any real existence. Moreover, Plato held some lines to be *indivisible*. To my knowledge, *commentators have never really explained this doctrine.*

4 The *Indivisibility* of Plato's Intelligible Being in the *Sophist* 226b1-231b8 and the *Phaedrus* 277b5-c6

Although Plato does not speak explicitly about the indivisible property of an intelligible Being, there are at least two instances in which he definitely and revealingly alludes to it.

4.1 The *Indivisibility* of the Intelligible Being Noble Sophistry in the *Sophist* 226b1-231b8

We have examined elsewhere (Negrepontis 2023b) Plato's definition of intelligible Beings by the method Name and Logos in the *Sophist*, and concluded that Name stands for the philosophical analogue of the anthyphairetic division of a dyad, and Logos for the philosophical analogue of Theaetetus' Logos Criterion for periodic anthyphairesis, thus exhibiting the intelligible Being as the philosophic analogue of a dyad in periodic anthyphairesis. The examples Plato uses in the *Sophist* are the intelligible Beings of the Angler, the Noble Sophistry, and the Sophist. In the definition of the Noble Sophistry Plato, rather unexpectedly introduces a fleeting, but quite revealing, question about the Indivisibility of the intelligible Being to be defined.

We start with *the art of discrimination* A (diakritiken [technen], 226c8).

The art of discrimination A is divided into two kinds (duo eide, 226c10-11), *the discrimination of like from like* B1, and the discrimination of worse from better A1 (226d1-3), named *purification* (katharmos, 226d10).

Purification A1 is divided into purification of the body B2 and *purification of the soul* A2 (227c7-9).

Two different kinds of vices (kakias, 227d9) of the soul call for two different kinds of purification of the soul A2:

- a kind of disease (noson, 228d7, e3) of the soul, including cowardice (deilian, 228e2, 229a3), insolence (hubrin, 229a3), injustice (adikian, 228e3, 229a3), calls for *[harsh] purification by the imposition of Justice* B3 (Dike, 229a5), akin to medicine (iatrike, 229a1) required for the disease of the body, and
- a kind of deformity (aischos, 228e5) of the soul, ignorance (agnoia, 228d10), calls for *[soft] purification by Instruction* A3 (Didaskalike, 229a9), akin to Gymnastics (gumnastike, 229a1) required for the deformity of the body. 227d-229a.

The difference between the imposition of Justice B3 and the provision of Instruction A3 is certainly the striking difference between a *harsh purification* and a *soft purification*, respectively, (as we have added in brackets).

Next, Instruction A3 is divided into

- instruction in [lowly] *handicraft* B4 (demiourgikas didaskaleias, 229d1-2), and
- [exalted] *education* A4 (paideia, 229d2-3; paideutike 231b5), as provided mainly in Athens (enthade, 229d2), but also among nearly all the Hellenes (229d4).

At this point, the Eleatic Stranger is asking the *crucial question*:

> [Stranger] But we must examine further and see whether it is *already (hede*, 229d5*)* an *indivisible* (*atomon*, 229d5) whole or still admits of division important enough to have a *name.*229d5-6

did the Division scheme *already* (hede) reach the status of an *indivisible* whole (*atomon hede esti pan*, 229d5-6),

or does it still admit a division worthy of a name (tina echon diairesin axian eponumias, 229d6)?

This question suggests some concept of finiteness for division. But finiteness appears to run counter to our interpretation of the Division as the analogue of periodic, hence infinite, anthyphairesis, and in favor of the standard interpretation treating these definitions as some sort of Linneus biological strictly finite divisions. So this question on indivisibility is crucial and must be answered convincingly, if we are to obtain the correct interpretation of the nature of these divisions. We postpone our answer after we consider the way the division finally comes to an end.

The immediate answer given by the Eleatic Stranger is that there is still one more division step (eti schizesthai, 229d8), as follows:

> Education A4 is next divided into a *rougher* method (*trachutera hodos*, 229e1-2), the traditional method of *admonition* B5 (*nouthetetike*, 230a3), and another *smoother* section (*morion leioteron*, 229e2), based on cross-questioning (elegchon, 230d7), the cross-questioning (elegchon) of empty conceit of wisdom (peri ten mataion doxosophian, 231b5-6), *the art of noble sophistry* A5 (genei gennaia sophistike, 231b7-8).

At this point, it becomes clear that the *Logos Criterion* of periodicity has been reached; indeed,

Justice B3 (rougher purification of the soul)/Instruction A3 (smoother purification) = Admonition B5 (rougher Education)/Noble Sophistry A5 (softer Education).

<table>
<tr><td colspan="6">*Sophistes* 231b3-8</td></tr>
<tr><td colspan="6">*Art of Discrimination* A</td></tr>
<tr><td>Discrimination *of like from like* B1</td><td colspan="5">Discrimination of worse from better, *purification* A1</td></tr>
<tr><td></td><td>Purification *of the body* B2</td><td colspan="4">Purification *of the soul* A2</td></tr>
<tr><td></td><td></td><td>*Justice* B3</td><td colspan="3">*Instruction* A3</td></tr>
<tr><td></td><td></td><td></td><td>*Handicraft* B4</td><td colspan="2">*Education* A4</td></tr>
<tr><td></td><td></td><td></td><td></td><td>*Admonition* B5</td><td>*Noble Sophistry* A5</td></tr>
</table>

$$\textit{Logos criterion}\ \ \mathrm{B3/A3} = \mathrm{B5/A5} = \mathrm{harsh/soft}$$

[Plato. Plato in Twelve Volumes, Vol. 12 translated by Harold N. Fowler. Cambridge, MA, Harvard University Press; London, William Heinemann Ltd. 1921, including the *Sophist*, modifications by the author]

According to our interpretation, the definition of Noble Sophistry is complete, and we have attained full knowledge of the infinite periodic anthyphairesis of the intelligible Being *Noble Sophistry*. It is also tacitly clear that, together with the full knowledge obtained by the Logos Criterion, the status of Indivisibility is meant to have been achieved. The question is: in what sense is the intelligible Being indivisible, once the Logos Criterion is achieved? Before answering this question, we will look at one more occurrence where Plato calls an intelligible Being indivisible once the Logos Criterion has been reached.

4.2 The *Indivisible Line* and Logos in the *Phaedrus* 277b5-c6

Socrates. Before someone obtains *knowledge* (eidei) of the truth about each of those he *sets logoi* (legei, 277b6) [speaks] (or writes, he must be able to define (horizesthai, 277b6) everything; and having defined (horisamenos, 277b7), he must obtain *knowledge* (epistethei, 277b8) by *dividing (temnein*, 277b7*) according to kinds (kat' eide*, 277b7*) till the indivisible (mechri tou atmetou, 277b7)*; in the same way (kata tauta) understanding about the nature of the soul, by finding out the kind adapted to each such nature, this is how he sets and adorn the *Logos (ton logon, 277c2)* [*discourse*], giving elaborate and harmonious *logoi* (*logous*, 277c3) [*discourses*] for the complex soul, and to the simple [*logoi, 277c3*] [talks] for the simple soul. Previous to that, it is not possible to use artfully (technei) the genus of *Logoi*

> *(logon, 277c5)* [*speak*], either for purposes of instruction or of persuasion, as the preceding *Logos* (logon, 277c6) [*discussion*] has made clear.

[Plato. Plato in Twelve Volumes, Vol. 9 translated by Harold N. Fowler. Cambridge, MA, Harvard University Press; London, William Heinemann Ltd. 1925, including the *Phaedrus*, with modifications by the author]

The *Phaedrus* 277b5-c6 is similar to the *Sophistes* passage, in that again the indivisible status is obtained in the Division according to kinds, as in the *Sophistes* situation. But now there is an explicit and insistent reference to the *logoi*, which, according to our interpretation, are clearly the *Logoi*, namely the Theaetetus' *Logos Criterion* instituting anthyphairetic periodicity.

We must take note that H.N. Fowler's classical translation of the passage has mistranslated, in our view, all six occurrences of *log-*, [kept in brackets in the above-translated passage] *legei*, 277b6 by *speak, ton logon*, 277c2 by *discourse, logous*, 277c3 by *discourses, [implied] logoi*, 277c3 by *talks, logon*, 277c5 by *speak, logon*, 277c6 by *discussion*, wholly losing the meaning of the passage. Cf. our remark in Sect. 8.2 for corresponding renderings of the root log- in the *Republic*.

The *Phaedrus* passage makes clear that indivisibility has been attained exactly at the point when the Logos Criterion is achieved, and thus our interpretation must provide a satisfying and convincing state of *indivisibility* of the intelligible Being, keeping at the same time its infinite divisibility, since the intelligible Being is an entity in periodic hence infinite anthyphairetic division!

5 Proclus' and Simplicius' Explanation of the Sense in Which an Intelligible Being Is *Indivisible*

The commentators Proclus and Simplicius provide essential help and guidance for the correct interpretation of the indivisible line.

5.1 Proclus, *Commentary to Timaeus* 2,245,23-246,7

> Why, therefore, should we any longer fear those skillful Peripatetics, who ask us, what kind of *line (grammen)* Plato here assumes? Is it a physical line? But this would be absurd: for this is the end of bodies. Is it then a mathematical line? But this is *not self-motive (akinetos)* and is *not a true Being (ouk ousia)*. Plato however says that the soul is a *true Being* and is separate from bodies. We say, therefore, that they in vain make these inquiries. For long before this, we have not ceased asserting that this line is of true Being *(grammen ousiode)*. And before us, Xenocrates calls *a line of this kind indivisible (atomon grammen)*. For it would be ridiculous for anyone to think that there is an *indivisible magnitude (adiaireton megethos)*. It is evident, however, that Xenocrates thought it requisite to call it *the logos of true Being (ton logon tes grammes ton ousiode)* of a line. But Plato, *for the sake of concealment (di'epikrupsin)*, employed *mathematical names (onomaton)*, as *screens (para-petasmasin)* of the truth of the things, in the same manner as theologists employed fables, and the Pythagoreans symbols (tois sumbolois). Proclus, *Commentary to Timaeus* 2,245,23-246,7

[[*The Commentaries of Proclus of the Timaeus of Plato*, translated from the Greek by Thomas Taylor, London, 1820, vol. II, p. 116–117, With modifications by the author]]

According to Proclus, *Commentary to Timaeus* 2,245,23-246,7, Xenocrates by the term *indivisible line* did not mean the literal indivisibility of a mathematical/geometrical line, something that would be ridiculous and impossible for someone who knew geometry very well and therefore knew the infinitely divisible nature of the geometric line. Proclus clearly refers to the *"self-motive" line*, by means of which Plato defines the soul in the *Phaedrus* 245c5-9, a definition that is practically identical to the definition of the intelligible Being in the *Sophist* 247d-e, and is possible only under the Logos Criterion ("ton logon tes grammes ton ousiode") and anthyphairetic periodicity, resulting in the equalization of the moving/active and the being-moved/passive part, and thus "self-motive." Thus, Proclus' description of the indivisible line leads to the intelligible Being, in which the infinite divisibility is not an issue, and that periodic anthyphairesis does not preclude describing in some philosophical sense the intelligible being as indivisible.

5.2 Simplicius, *Commentary to Aristotle's Physics* 140,6-13, and 142,16-18

Simplicius does provide the exact sense according to which an intelligible Being is indivisible; and what is more, this is in complete agreement with the structure of intelligible Beings as given by the One of the second hypothesis of the Parmenides, and by the Being in the Sophist.

According to Simplicius, *in Physics* 140, 6-13, 142, 16-18, Xenocrates, being a geometric and wise man, could not contradict the geometric principle and did not ignore the nature of magnitude [namely the infinite divisibility of magnitudes, ["ten ep' apeiron tomen"]; by "invisible line" he meant a line that, even though it was divisible *ad infinitum*, nevertheless was not divisible in kind ["eidei diairetton"], namely *the number of kinds was finite*.

5.2.1 Simplicius, *Commentary in Aristotle's Physics* 140,6-13

> Those in the circle of Xenocrates agreed (sunechoroun)
> *only* with the first *part of the statement (akolouthian)* that
> the intelligible Being is One and undivided (adiaireton)
> [namely, that the true Being is One],
> *but* did not agree with the second, that Being is undivided (undivided).
> [translation of the Simplicius passages is by Huby PM & Taylor CCW, 2011]

This sentence, with a rather complex structure, has nevertheless quite a clear meaning. Xenocrates, like Plato, accepted that the true intelligible Being is One, but he rejected that the true, intelligible Being is partless; in other words, Xenocrates, like Plato rejected that the intelligible Being is the partless One of the first hypothesis in the *Parmenides*. That Plato rejected the partless One of the first hypothesis is stated most explicitly in the *Parmenides* 141e. A similar rejection is mentioned by

Simplicius, on the authority of Eudemus, already by Zeno, the precursor of Plato, in his *Commentary to Aristotle's Physics* 97, 9-16 and 138,3-139,3.

> Hence again Being is *not only one* (mede hen monon) *but many* (alla pleio).

Thus Plato's true Intelligible Being is not only One, but One and Many. That an intelligible Being is Many, in fact infinitely many, is a consequence of the fact that the intelligible Being is a dyad (the dyad <One, Being> in the *Parmenides* second hypothesis, the dyad <Being, not-Being> in the *Sophist*) in infinite anthyphairesis. That an intelligible Being is One is a consequence of the periodicity of the anthyphairesis of the dyad, effected by the Logos Criterion, by means of which the two parts of the dyad, and in fact each of the infinite multitude of the remainders/parts is equalized to the whole.

> However, while divided (diaireton), the intelligible Being
> *does not go on to infinity* (me ep' apeiron ienai),
> *but stops* at some *indivisibles* (all' eis atoma katalegein).
> These however are
> ***not*** indivisibles (atoma) in the sense of being *without parts (amere)* and smallest,
> ***but*** divisible (tmeta) and having parts (mere echonta) with regard to the multitude (to poison) and magnitude (ten hulen), while *indivisible (atoma) and primary (prota) with regard to kinds (kinds)*, assuming that there are certain primary indivisible lines (grammas atomous) Simplicius, *in Physics* 140,9-13

5.2.2 Simplicius, *Commentary to Aristotle's Physics* 142, 16–18

> But since *Xenocrates* was also a wise man (sophos aner),
> how did he come to suggest *the indivisible lines (tas atomous grammas)*?
> For he was not ignorant of the [[infinitely divisible]] nature (τὴν [[ep' apeiron diaireten]] phusin) of the *magnitude (the megethous), but he did say that it was not divisible (diairetin) in kinds (τῷ εἴδει).* Simplicius, *in Physics* 142, 16–18

This second Simplicius passage is similar to the first.
We will provide our interpretation of indivisibility in the next section.

6 Our Interpretation of the Meaning of *Indivisibility* of an Intelligible Being

We now explain the precise sense in which Plato's intelligible being is *indivisible*.

6.1 Every Intelligible Being Is a Dyad in the Philosophical Analogue of Infinite Anthyphairesis, Satisfying the Logos Criterion for Periodicity

According to our interpretation of Plato's intelligible Being, as it results from the *Meno*, the second hypothesis of the *Parmenides*, and the *Sophist*, an intelligible Being is the philosophical analogue of a dyad of opposites in periodic

anthyphairesis. We start with our interpretation of In the *Meno* the dyad is, in fact, geometric, it is the dyad of the diameter to the side of a square; in the paradigmatical intelligible Being of the second hypothesis in the *Parmenides* the dyad is <One, Being>; in the *Sophist*, it assumes the general form <Being, not-Being>258a-b (including as special cases the intelligible dyads <Great, not Great>, 257 d-e, <Beautiful, not-Beautiful> 257e, <Just, not-Just> 258a).

Adapting the *Sophist* terminology, an intelligible Being is a dyad <B, N>, where B stands for Being, N for not-Being, such that the dyad <B, N> satisfies the philosophical analogue of periodic anthyphairesis. Thus

(a) The dyad <B,N> satisfies the condition of *infinite anthyphairesis*

$$\begin{aligned}
&B = N + N_1, \text{with } N_1 \text{ equal to a part of B, for short } N_1 < B,\\
&N = N_1 + N_2, \text{with } N_2 \text{ equal to a part of N1, for short } N_2 < N_1,\\
&N_1 = N_2 + N_3, \text{with } N_3 \text{ equal to a part of } N_2, \text{for short } N_3 < N_2,\\
&\ldots\\
&\textit{ad infinitum},
\end{aligned}$$

and thus an *infinite multitude* of parts/anthyphairetic remainders

$$B, N, N_1, N_2, N_3, \ldots$$

is generated.

(b) The *Logos Criterion* for anthyphairetic periodicity

$$B/N = N_{k\text{-}1}/N_k$$

is satisfied for some index k.

6.2 The Dialectical Number defined by Means of the Plus One Rule (*Parmenides* 148d-149d) in an Intelligible Being

Plato wishes to generate Numbers in every intelligible Being. The classical and practically universal way in which numbers are defined in Greek Mathematics is that a number is a finite collection of units. This is the way in which numbers are introduced in Euclid's *Elements*, definition VII.2. The first impulse would be to generate numbers by taking as units the infinite sequence of anthyphairetic remainders

$$B > N > N_1 > N_2 > N_3, > \ldots,$$

and this is in fact the abortive attempt at the beginning of the second hypothesis in the *Parmenides*, leading to failure, because as explained in *Philebus* 56d-e, the units for the numbers as used by the many (we might say the *eristic* numbers) are unequal, while the numbers used by the philosophers (we might say the dialectical numbers) must be equal. Thus Plato realized that the dialectic numbers cannot be introduced by taking into account only the anthyphairetic division but there is a need for the anthyphairetic periodicity, as well, if *equalization* of the anthyphairetic remainders is to be achieved.

The manner in which the *dialectic numbers* are defined is crucial for understanding the sense in which an intelligible Being is indivisible.

Definition of the dialectical number by means of the "plus one rule," Parmenides 148d-149d.

The *dialectical number* of a sequence of terms in succession

$$x_1, x_2, x_3, \ldots, x_n$$

is equal to

$$(\text{the } \textit{number} \text{ of the set of (different) successive ratios } x_1/x_2, x_2/x_3, \ldots, x_{n-1}/x_n) \text{ plus one.}$$

Examples If the sequence consists of one term only, x_1, then no ratio can be formed, and thus the number of the terms is 1.

If the sequence consists of two terms, x_1, x_2, then only one ratio x_1/x_2 can be formed, so that the number of the sequence is $1 + 1 = 2$.

If the sequence consists of the terms x_1, x_2, x_3, then we can form the two ratios x_1/x_2, x_2/x_3. There are two cases to consider:

If the ratios x_1/x_2, x_2/x_3 are unequal, then the dialectical number of the sequence x_1, x_2, x_3 is $2 + 1 = 3$; but if the ratios are equal, $x_1/x_2 = x_2/x_3$, then the dialectical number of the sequence will be $1 + 1 = 2$.

As an example, the sequence $x_1 = 2$, $x_2 = 4$, $x_3 = 5$ has dialectical number 3, but the sequence $x_1 = 2$, $x_2 = 4$, $x_3 = 8$ has dialectical number 2.

We apply now the definition of the dialectical number to the sequence of the successive remainders of the intelligible Being with initial dyad <B, N>, which is infinite, but with a periodic sequence of successive ratios of remainders.

Proposition (*Parmenides* 144d4-e3, Negrepontis 2024)
The *dialectical number* of the sequence of parts of *every* part of the infinite sequence

$$B, N, N_1, N_2, , , , , N_n, N_{n+1}, \ldots$$

is $k + 1$.

Proof

(i) It follows from the Logos Criterion 6.1. (b) that the sequence of k successive ratios

$$B/N, N/N_1, N_1/N_2, \ldots, N_{k-2}/N_{k-1}$$

forms a complete period, repeating itself cyclically *ad infinitum*:

$$\begin{aligned} N_{k-1}/N_k &= B/N, \\ N_k/N_{k+1} &= N/N_1, \\ N_{k+1}/N_{k+2} &= N_1/N_2, \\ &\ldots, \\ N_{2k-1}/N_{2k} &= N_{k-2}/N_{k-1}. \end{aligned}$$

Thus the number of the terms in the sequence

$$B/N, N/N_1, N_1/N_2, \ldots, N_{k-2}/N_{k-1}, \ldots, N_n/N_{n+1}, \ldots$$

is equal to the number of the terms in the sequence

$$B/N, N/N_1, N_1/N_2, \ldots, N_{k-2}/N_{k-1},$$

namely is equal to the number k. By the "plus one rule,"

the *dialectical number* of the infinite sequence $B, N, N_1, N_2, , , , , N_n, N_{n+1}, \ldots = k + 1$.

(ii) By periodicity, the number of the different ratios in the sequence

$$N/N_1, N_1/N_2, \ldots, N_{k-2}/N_{k-1}, \mathbf{N_{k-1}/N_k}$$

is clearly equal to the number of different ratios in the sequence

$$\mathbf{B/N}, N/N_1, N_1/N_2, \ldots, N_{k-2}/N_{k-1}$$

namely the number k, hence

the *dialectical number* of the infinite sequence $N, N_1, N_2, \ldots, N_n, N_{n+1}, \ldots = k + 1$

A similar argument shows that for every n

the *dialectical number* of the infinite sequence $N_n, N_{n+1}, \ldots = k + 1$.

Corollary

(i) The parts B, N, N_1, N_2, . . ., N_n, . . .of the intelligible, Being with initial dyad <B, N> are mutually *equalized* in the sense that

the dialectical number of the parts of B =
the dialectical number of the parts of N =
the dialectical number of the parts of N_n for every n.

(ii) The greatest number (o arithmos pas, *Philebus* 16d8) that can be generated in the intelligent Being is equal to the number k+1 of the length of the anthyphairetic period of the intelligible Being +1.
(iii) The intelligible Being satisfies the fundamental property *One and Many*, since the intelligible Being is Many, in fact infinitely many, but it is also One, in the self-similar sense that all parts are equalized to the whole.

6.3 The Paradox of an Infinite *Multitude* of Parts Versus a Finite *Number* of Parts: The Multitude of the Parts in the Intelligible Being Is Infinite, but the Dialectical Number of These Parts Is Finite

Proof Infinite multitude follows from the infinite anthyphairesis of B to N (as follows from Sect. 6.1).

Finiteness in number follows from the periodical nature of the anthyphairesis of a true Being, and the definition of the dialectic number by the plus one rule (as explained in Sect. 6.2).

Note The statement in Sect. 6.3 is certainly *paradoxical*. The fact that Plato expects that we will indeed find the statement paradoxical is made clear in the *Sophist* 257a8-12 (*kai tauta ou duscheranteon*, 257a8). The fundamental tenet of Plato's philosophy that "not-Being is (a true intelligible Being," can be understood only by having first an understanding of the anthyphairetic nature of a true intelligible Being, namely as a mixture of infinite anthyphairesis and the Logos Criterion, establishing periodicity. The reader is referred to Negrepontis (2024) for details. For traditional attempts by Platonic scholars to explain why not-Being is a Being, the reader may consult Heidegger (1924–25), Frede (1992), and Crivelli (2012).

6.4 The Whole Part B Is *Indivisible* in the Sense that It Is Infinitely Divisible but nevertheless *Unable to Generate New Units* for Dialectical Numbers After the Completion of a Period

We can describe the situation reached in Sects. 6.2 and 6.3 by stating that the part B is infinitely divisible by anthyphairesis, but at some finite stage, no new units for the dialectic numbers are generated, so that the part B is *indivisible* in generating new units for numbers after a certain point.

6.5 *The Units of the Dialectical Numbers* Are the intelligible *Kinds/Eide*

The identification of the units for the dialectical numbers, namely numbers with equalized units, generated in an intelligible Being, with the Kinds/eide of the intelligible Being, is most clearly explained by Aristotle. In fact, Aristotle in *Metaphysics*

(a) states repeatedly, that the Platonic numbers are numbers with *equalized* units (isasthenton 1081a25, 1083b24, 31, 1091a25, isasthenai 1091a26), in agreement with Plato's requirement for the dialectical numbers, and
(b) refers to these Platonic dialectical numbers as *eidetic* (1081a21, 1086a5,8, 1088b34, 1090b35), as consisting of eide/Kinds (987b21-22).

6.6 The Whole Part B Is *Indivisible* in the Sense that It Is Infinitely Divisible in Parts but nevertheless *Unable to Generate New Kinds* After the Completion of a Period

Since in the intelligible Being with dyad <B, N> there are, according to Sects. 6.2 and 6.3, and since, according to Sect. 6.5, the Kinds in the intelligible Being are identified with the Kinds (eide), it follows that

$$\text{there are exactly } k+1 \text{ Kinds (eide)},$$

the immediately next anthyphairetic division step

$$N_{k-2} = N_{k-1} + N_k$$

generates a new remainder N_k, since the anthyphairesis is infinite, BUT this remainder N_k *cannot* be *a new kind (eidos)*, because of the stringent "plus one rule" for the generation of monads/kinds. It is precisely in this sense that from this point on

the intelligible Being is indivisible.

In conclusion, the *indivisibility* of an intelligible Being is simply a form of poetic license (a "screen of concealment," as Proclus has it) that allows the transfer, using the *plus one rule* enunciated in the *Parmenides* 148d-149d, of the periodicity of the sequence of the ratios of the successive remainders of the anthyphairesis of the intelligible Being into the finiteness of the number of parts.

Thus in an intelligible Being, there are as many Kinds as units for the whole dialectical number of the intelligible Being.

The line B is infinitely divisible by anthyphairesis, but at some finite stage, no new eide/kinds are generated, so that the line B is *indivisible* in generating new eide/kinds after a certain point.

6.7 Examples

We compute the number of the Kinds in the intelligible Being *Noble Sophistly*, examined in Sect. 4.1. We compute the kinds from the Division, but starting from the point where there is pure periodicity, namely from the Kind B3 (Justice). The Being Noble Sophistry has *five* Kinds/units for dialectic numbers, namely

B3 (Justice), A3 (Instruction), B4 (Handicraft), A4 (Education), B5 (Admonition),

and becomes indivisible, in the sense that, even though the division, being periodic, proceeds *ad infinitum*, it nevertheless does not have the power to generate new Kinds exactly at the point where the Logos Criterion is established.

The *mean and extreme ratio* (Elements, Propositions II.11 and VI.13) is a purely periodic, and the diameter to the side of a square is an eventually periodic, anthyphairesis, each with a whole dialectical number equal to 2.

6.8 Confirming Evidence for Our Interpretation

Our interpretation explains why the indivisibility occurs exactly at the point of establishment of the Logos Criterion, as in Plato's passages in Sect. 4. Furthermore, our interpretation of the indivisible line in Sect. 6.7 is essentially identical to the description given by Proclus and Simplicius in Sect. 5, recall example.

6.8.1 Simplicius, *Commentary to Aristotle's Physics* 142, 16-18

> But since *Xenocrates* was also a wise man (sophos aner),
> how did he come to suggest *the indivisible lines (tas atomous grammas)*?
> For he was not ignorant of the [[infinitely divisible]] nature (ten [[ep' apeiron diaireten]] phusin) of the *magnitude (the megethous), but he did say that it was not divisible (diairetin) in kinds (τῷ εἴδει).* Simplicius, *in Physics* 142, 16-18"

In addition, our interpretation makes sense of the mysterious Aristotle's *Metaphysics* 992a19-22, as we shall see in Sect. 7 below.

On the contrary, the standard interpretation by Platonic scholars has the Divisions in the *Sophist* to be divisions of strictly finite length, not infinite divisions (as becomes manifestly clear in the Division of the Statesman in the *Statesman*), involving the Logos Criterion and anthyphairetic periodicity, and the reference to indivisibility in the *Sophist* and the *Phaedrus* to simply indicate that we have reached the last step of the finite division. Cf, Crivelli (2012 p.18):

> The division goes on until indivisible kinds are reached (cf. Phdr. 277b7–8; Sph. 229d5–6). If it aims at definition, then only one kind counts as indivisible: the one to be defined. If division aims instead at classification, then only kinds count as indivisible that encompass no subordinate kinds.

7 We Are Now in a Position to Obtain Our Interpretation of the Heretofore Mysterious Aristotle's *Metaphysics* 992a19-22 Passage, Since the One of the First Hypothesis in the *Parmenides* Is Identified with the Geometric Point and the One of the Second Hypothesis in the *Parmenides* Is Identified with the Intelligible, Hence Indivisible, Being

7.1 Aristotle, *Metaphysics* 992a19-22 on Plato's Indivisible Lines

> Further, from what principle will the presence of *the points* (*hai stigmai*) in the line be derived? *Plato* even used to object (diemacheto) to this class of objects (toutoi toi genei) as being *a geometrical fiction (geometrikoi dogmati)*. He gave the name of principle of the line- and this he often posited (touto de pollakis etithei) to *the indivisible lines (tas atomous grammas)*. Aristotle, *Metaphysics* 992a19-22

[[Aristotle, *Metaphysics*, *a revised text with introduction and commentary* by William David Ross, Clarendon Press, Oxford, 1924. With modifications by the author]]

The passage states that Plato rejected the concept of the partless geometric point ("stigme") as a foundational concept of geometry, and opted instead for the concept of the "indivisible line" ("atomous grammas").

According to Aristotle's *Metaphysics* 992a19-22 Plato	
rejects the geometric point	
	accepts the indivisible line
as the true foundation of geometry	

The attentive reader of Plato will doubtlessly have noted that there is no mention of geometric point in his dialogues (although he talks about numbers, straight lines, circles, and angles). It is difficult to conceive of geometry without the concept of geometric point, but Plato suggests just that and replaces it with something mysterious, and seemingly contradictory, *the indivisible line*.

However, we will try to make sense of both, quite unexpected and difficult-to-comprehend, statements in Aristotle's passage.

In fact, we shall see that Aristotle wrote his Metaphysics *passage having open in front of him Plato's* Parmenides, *on which Aristotle's passage is essentially a most natural comment.*

7.2 Aristotle's Passage Is Understood as a Comment on the Comparison Between the First and the Second Hypothesis in Plato's *Parmenides*, Fully Confirming that the Indivisible Line Coincides with the Paradigmatical One of the Second Hypothesis in the *Parmenides*

7.2.1 The One of the First Hypothesis Cannot be the Foundation of Geometry, but the One of the Second Hypothesis Can

7.2.1.a *We first note that the One of the first hypothesis in the* Parmenides *137c5-6 is partless, and thus resembles the geometric point*

- Then there can be no parts of it, nor can it be a whole. 137c5-6

Thus the One of the first hypothesis in the *Parmenides* is partless, like a geometric point.

7.2.1.b *The One of the first hypothesis in the* Parmenides *cannot generate the basic figures of Geometry, the straight line, and the circle, but the One of the second hypothesis does*

7.2.1.c *According to the* Parmenides *137d8-138a1, the One of the first hypothesis, a completely partless entity (according to 137c4-d3), having as a model the geometric point does not have the power to generate the basic geometrical figures, straight line and circle*

> -And it [the One of the first hypothesis in the *Parmenides*] is without figure (aneu schematos), for it participates (metechei) neither of the round (stroggulou)) nor of the straight (eutheosς).
>
> -How so?
>
> -The round (stroggulos), of course, is that of which the extremes are everywhere equally distant from the center.
>
> -Yes.
>
> -And the straight (euthu), again, is that of which the middle obstructs the communication (epiprosthen) between the two extremes.
>
> -It is.
>
> -Then the one would have parts and would be many, whether it participated (metechoi) either of straight (eite eutheos) or of round (either peripherous) figure.
>
> -Certainly.
>
> -Then it is neither straight (euthu) nor round (peripheres) since it has no parts.
>
> -Right. *Parmenides* 137d8-138a1

7.2.1.d *According to the* Parmenides 145b3-5, *the One of the second hypothesis does have the power to generate the basic geometrical figures, straight line, and circle*

While on the other hand, *the One of the second hypothesis*, whose model, as we have seen, is a dyad in periodic anthyphairesis, does have the power to generate the two basic geometrical figures, straight line, and circle.

> -And the one (to hen), apparently, being of such a nature (toiouton on),
> will participate (metechoi an) of some figure (schematos),
> whether straight (etoi eutheos) or round (e stroggulou) or a mixture of the two.
> -Yes, it will. *Parmenides* 145b3-5

7.2.1. According to His Dialogue *Parmenides* Plato	
rejects the One of the first hypothesis 137d8-138a1, which is a partless One, like a geometric point 137c5-6	
	accepts the One of the second hypothesis 145b3-5
as the true foundation of geometry	

7.2.2 The One of the First Hypothesis in the *Parmenides* Is Not a True Intelligible Being, the One of the Second Hypothesis Is a True, Intelligible Being

7.2.2.a *The One of the first hypothesis in the* Parmenides *is* ***not*** *a true intelligible Being*

> -Then if the one has no participation in time whatsoever, it neither has become nor became nor was in the past, it has neither become nor is it becoming nor is it in the present, and it will neither become nor be made to become nor will it be in the future.
> -Very true.
> -Can it then participate in being in any other way than in the past, present, or future?
> -It cannot.
> -Then the one has no share in being at all.
> -Apparently not.
> -Then the one is not at all.
> -Evidently not. 141e3-10

The One of the first hypothesis in the *Parmenides* is definitely not a true, intelligible Being.

7.2.2.b *The One of the second hypothesis is a true, intelligible Being*

> -But since the one participates in time and can become older and younger, must it not also participate in the past, the future, and the present?

-It must.

-Then the one was and is and will be and was becoming and is becoming and will become.

-Certainly.

-And there would be and was and is and will be something which is in relation to it and belongs to it?

-Certainly. 155c8-d6

The One of the second hypothesis in the *Parmenides* is definitely a true, intelligible Being.

7.2.2 *According to his dialogue* Parmenides *Plato*	
rejects *the partless One of the first hypothesis 141e3-10*	
	accepts *the One of the second hypothesis 155c8-d6*
as the true intelligible Being	

7.2.3 The Comparison of Aristotle's *Metaphysics* 992a19-22 with Plato's Two Hypotheses in Plato's *Parmenides* Leads to the Identification of the Indivisible with the Paradigmatical Intelligible Being, the One of the Second Hypothesis

Indeed, **the *One of the first hypothesis*** is a completely partless entity and thus is closely related to the **geometric point**, defined (in Euclid's *Elements*) as a geometric entity without parts.

In addition, the *One of the first hypothesis* is not a true Being and **does not have the power** to generate the two basic geometrical entities, the straight line and the circle, according to Plato's *Parmenides*, in agreement with Plato's view, as reported by Aristotle in the *Metaphysics* passage, that the *geometric point* **cannot** serve as the foundation of geometry according to Plato.

On the other hand, **the *One of the second hypothesis*** is a **true intelligible Being**, and **does possess the power** to generate the two fundamental geometrical figures, the straight line and the circle, according to Plato's *Parmenides*, in agreement with Plato's view, as reported by Aristotle in the *Metaphysics* passage, that the so-called *indivisible line* can serve as the foundation of geometry.

Thus, the comparison of Plato's *Parmenides* with Aristotle's *Metaphysics* passage rather strongly suggests that *the indivisible line* is identified with *the One of the second hypothesis, the paradigmatical intelligible Being.*

Comparing the *Metaphysics* passage with the *Parmenides* passages (in Sects. 7.2.1 and 7.2.2), we conclude that

	The indivisible line must be closely connected or even coincide *with the One of the second hypothesis* in the *Parmenides*

7.2.4 The Identification of Plato's Indivisible Line with the One of the Second Hypothesis in the *Parmenides*

According to the passages examined in Sects. 6 and 7, we may regard the One of the second hypothesis, namely the dyad $a > b$ in periodic anthyphairesis, as an indivisible line in the following exact sense:

The dyad $a > b$ is divisible *ad infinitum*, but after the stage n where the Logos Criterion holds, there are no new logoi/ratios of successive parts/remainders, no new units for numbers, and no new kinds (eide). Thus the paradox that occurs in the "indivisible line," namely that the parts of the "indivisible line" are infinite in multitude, while the kinds ("eide") are only finite, practically coincides with the paradox we have observed in the One of the second hypothesis in the *Parmenides*, namely that the parts of the One are *infinite in multitude*, while the parts of the One are *finite in number*.

We have no difficulty in identifying the "indivisible line" with the One of the second hypothesis in the *Parmenides*. The heretofore mysterious *Metaphysics* 992a passage becomes now fully understandable: according to Plato, Geometry must be founded, not on the geometric point (as was done eventually in Euclid's *Elements),* which is clearly the model for the One of the first hypothesis, but on the One of the second hypothesis in the *Parmenides*, Plato's paradigmatical intelligible Being, and which essentially coincides with the indivisible line.

8 The Hypothesis-Free/Anhupotheton in the *Republic*

(a) *Is a True Intelligible Being*
(b) *Is Indivisible*
(c) *Serves as the Foundation for Geometry*

The hint given in the second hypothesis in the *Parmenides* that the intelligible Being is the true foundation of Geometry (Sect. 7) is greatly expanded in the *Republic.* We outline the relation of the intelligible Being, called *anhupotheton/hypothesis-free principle*, as the foundation of Geometry by considering the *Republic* passages listed below.

Dialectical geometry	Geometry as practiced by geometers
	G1 510b4-6
D1 510b6-9	
	G2 510c2-d3
D2 511b3-7	
D2+ 511b7-c2	
	G3 533b6-c6
D3 533c7-d3	
D4 534b3-d2	

8.1 The Passages *Republic* [D1] 510b6-9, [D2] 511b3-7, [D3] 533c7-d3, [D4] 534b3-d2 Identify the Anhupotheton/ Hypotheses-Free with the Intelligible Being

[D1] *In an intelligible Being, there is an intelligible ascent from the hypotheses to the hypothesis-free principle*

> while there is another section in which it proceeds *from a hypothesis ("ex hupotheseos")* to *a hypothesis-free principle ("ep' archen anhupotheton")*, and in which it makes no use of the images employed by the other section, relying on ideas ("autois eidesi") only and progressing systematically through ideas ("di' auton"). 510b6-9

This passage introduces the unfamiliar term *"anhupotheton"/hypothesis-free*. According to passage D1, the intelligible Being proceeds from hypotheses to the hypothesis-free, without any recourse to sensible images, but using kinds.

From the *Meno*, the *Republic*, the Divisions in the *Sophist,* and the second hypothesis in the *Parmenides*, we have reached an interpretation of the structure of an intelligible Being, in terms of anthyphairetic division (described as *True Opinion*, or *Name*) and the Theaetetus' Logos Criterion for anthyphairetic periodicity (described as *Logos*) (cf. Negrepontis 2019, 2023b, 2019).

The parts/kinds/names in the Divisions of the *Sophist*, but also of the *Statesman*, are systematically referred to as *hypotheses*, as seen by the list that can be found in Negrepontis (2019, Sect. 1.5.4).

[D3] *The most crucial and revealing passage concerning the dialectics in the* Republic *is passage 533c7-d3; according to it, the ascent to the anhupotheton is achieved by "anairousa" the hupotheses*

> ". . .Then," said I, "is not dialectics the only process of inquiry that proceeds (poreuetai) in this manner, *"anairousa" the hypotheses (tas hupotheseis anairousa, 533c8)* toward *the principle itself* {auten ten archen) [the anhupotheton] to find confirmation there? And it is literally true that when the eye of the soul is sunk in the barbaric slough, dialectic gently draws it forth and *leads it up* ("anagei ano") 533c7-d3

Furthermore, we have established in Negrepontis (2019, Sect. 1.6.3), that the crucial statement "tas hupotheseis anairousa," in passage D3, refers to the division of the hypotheses, namely the Division according to kinds.

> *Thus, we interpret the term "hypotheses" in passages D1 and D3 to mean the parts/kinds/ names generated by the anthyphairetic/"antanairetic" division of an intelligible Being.*

[D2] The passage introduces the vital role of *Logos* in an intelligible Being; it is the presence of Logos that forces the hypotheses to be not devoid of motion but to act as springboards and footings toward the hypothesis-free.

> ". . .Understand then," said I, "that by the other section of the intelligible I mean that which *Logos itself ("autos o logos," 511b4)* [Shorey's rendering: reason itself] is in contact with

("haptetai") the power of dialectics ("dialegesthai"), treating its *hypotheses* not as principles ("archas") but truly as hypotheses, footings ("epibaseis"), and springboards ("hormas") so to speak, to enable it *to ascend to the hypothesis-free (anhupothetou*) which is the principle ("archen") of the all 511b3-7

[D4] *The fundamental importance of logos for obtaining the Knowledge of an intelligible Being* 534b3-d2

-And do you not also give the name dialectician to the man who is able to exact *the logos (logon*, 534b3) [Shorey's rendering: account] of the essence of each thing? And will you not say that the one who is unable to do this, in so far as he is incapable of rendering *the logos (logon*, 534b4) [Shorey's rendering: account] to himself and others, does not possess intelligence (noun, 534b5) about the matter?

-How could I say that he does? he replied.

-And is not this true of the good likewise – that the man who is unable to define *using logos (toi logoi*, 534b9) [Shorey's rendering: discourse] and *subtract (aphelion*, 534b9) from all other things the idea of the good, and who cannot, as it were in battle, running the gauntlet of all tests, and striving to examine everything by true Being (ousian) and not by opinion (doxan), hold on his way through all this without tripping (aptoti) *using logos (toi logoi*, 534c3) [Shorey's rendering: reasoning] – the man who lacks this power, you will say, does not really have *knowledge (eidenai*) of the good itself or any particular good; but if he apprehends any adumbration of it, his contact with it is by opinion, not *by knowledge (epistemei*); and dreaming and dozing through his present life, before he awakens here 534b3-d2.

[Plato. Plato in Twelve Volumes, Vols. 5 & 6 translated by Paul Shorey. Cambridge, MA, Harvard University Press; London, William Heinemann Ltd. 1969, including *Republic*, with modifications by the author]

We *note* that the classical P. Shorey translation renders *logos* in 510c7 as *account*, in 511b4 as *reason*, in 533c2 as *account*, in 534b3 as *account*, in 534b4 as *account*, in 534b9 as *discourse*, and in 534c3 as *reasoning*. Cf. Sect. 4.2 for corresponding renderings of log- in the *Phaedrus*.

In passages D2 and D4, the Logos plays a prominent, fundamental role. But, as we have seen in the references to our work cited above,

we interpret the term Logos in passages D2 and D4 to mean the Logos Criterion for anthyphairetic periodicity, by which full knowledge of the intelligible Being is obtained.

Our analysis of the *Republic* passages D1, D2, D3, and D4 shows that

- the hypotheses are the *kinds (eide) of the intelligible Being*,
- the Ascent from hypotheses to higher hypotheses is identical with *the anthyphairetic division and motion* of the kinds, while
- by the "Logos provided to the hypotheses" is meant exactly *the Logos Criterion for anthyphairetic periodicity.*

Thus the structure of the anhupotheton in the *Republic* is identical to the structure of the paradigmatical intelligible Being, the One of the second hypothesis in the *Parmenides*.

8.2 The Term "Anhupotheton" in the *Republic* Suggests the Indivisibility of the Intelligible Being

But why does Plato call the intelligible Being "anhupotheton," namely hypotheses-free?

We first note that by passage D3 of the *Republic,* especially the statement *"anairousa" the hypotheses (tas hupotheseis anairousa, 533c8)*, considered in Sect. 8.2, the principal feature of the hypotheses is their division, in fact their dichotomy, into two higher hypotheses, and then into still higher ones. But the only division there is in an intelligible Being, which according to our interpretation is the philosophical analogue of a periodic anthyphairetic dyad, applies to the parts of the anthyphairetic division. So, the hypotheses must be parts/remainders of the intelligible anthyphairetic division.

On the other hand, the Kinds (eide), called also species or genera, which appear in the Divisions in the *Sophist* and the *Statesman*, as the division proceeds toward the Logos Criterion, are systematically referred to as hypotheses, as shown in the extensive list in Sect. 4.4 in Negrepontis (2019).

Thus the hypotheses appearing in the *Republic* description of the structure of the anhupotheton can be safely identified with the Kinds (eide) appearing in the definition of the intelligible Being. The terminology "anhupotheton" is naturally understood as follows: the hypotheses are the Kinds (eide); with the establishment of the Logos Criterion, there are no more new Kinds (eide), and thus no new hypotheses, and so we have reached the state of hypotheses-free. Thus the term anhupotheton has the force of indivisible, according to our interpretation in Sect. 6.

8.3 According to the Passages *Republic* [G1] 510b4-6, [G2] 511b3-7, [G3] 510c2-d3, the True Foundation of Geometry Is Not the Hypotheses of the Geometers but the Hypothesis-Free/ Anhupotheton

8.3.a In the *Republic* passages G1 510b4-6, G2 510c2-d3, and G3 533b6-c6, Plato criticizes the way the geometers regard the hypotheses, definitions, and postulates, as foundational starting points for Geometry and Arithmetic, and proposes instead that these hypotheses should themselves be founded on the intelligible Being/anhupotheton

[G1] By the distinction that there is one section of it [the intelligible] which the soul is compelled to investigate by treating as images the things imitated in the former division and *starting from hypotheses (ex hupotheseon)*, it proceeds not up to *a first principle (ep' archen)* but *down to a conclusion (teleuten)* 510b4-6

[G2] For I think you are aware that students of geometry and reckoning and such subjects set as *hypotheses (hupothemenoi)*, the odd and the even and the various figures and the three kinds of angles and other things akin to these in each branch of science, and these regard

them *as known (hos eidotes),* and, treating them as absolute *hypotheses (hupotheseis)*, do not deign to render any further *logos (logon, 510c7)* [Shorey's rendering: account] of them to themselves or others, taking it for granted that they are obvious to everybody. They take their start (archomenoi) from *these [hypotheseis]*, and pursuing the inquiry from this point on consistently, *conclude (teleutosin)* with that for the investigation of which they set out. 510c2-d3

[G3] while the remnant which we said did in some sort lay hold on reality – *geometry* and the studies that accompany it – are, as we see, dreaming about being, but the clear waking vision of it is impossible for them as long as they leave the *hypotheses (hupothesesi)* which they employ *without allowing them to move (akinetous eosi)* and cannot give a *logos (logon, 533c2)* [Shorey's rendering: account] of them. For where the starting-point (arche) is something that the person does not know, and the end (teleute) and all that is between (between) is an *intertwining (sumpeplektai)* of things not known, what possibility is there that assent in such cases can ever become *knowledge (episteme)*?
None, said he. 533b6-c6

We can summarize the *Republic* passages G1, G2, and G3 as follows:

The geometers have some *hypotheses*, such as the *numbers* (the even and the odd), the *two figures* (the straight line and the circle), and the *three kinds of angles* (right, obtuse, and acute), which they use as starting points, as principles, and, with no justification, as something they pretend to know. We must note at this point that Plato identifies the *mathematical hypotheses*, namely the definitions and postulates of Arithmetic and Geometry, with the philosophical hypotheses, which are the parts/kinds/names/predicates in the philosophical analogue of the anthyphairetic Division that describes an intelligible Being.

From these hypotheses, the geometers proceed, *not* upwards toward the hypothesis-free, as the dialectician, *but* in the converse direction, downwards to their logical consequences. Since the geometers never employ their hypotheses in order to move upwards to the Logos of periodicity and the resulting hypothesis-free, they do not possess knowledge of the hypotheses, but are in some sort of dream condition, In particular, they do not attain the intertwining of Logoi that forms the circle of periodicity, and thus they have no knowledge of the *beginning, middle, and end* of that period of the whole, as explained in the second hypothesis in the *Parmenides* (145a4-b1).

8.3.b *An informal reconstruction of the generation within the intelligible Being/ anhupotheton of the mathematical hypotheses mentioned in the* Republic *passage G2 510c2-d3*

Plato does not provide details for the generation of the mathematical hypotheses from the structure of the intelligible Being, but we can reasonably reconstruct such a generation. Since we have already attempted this reconstruction, we will refer the reader to Negrepontis (2019, Sects. 1.10 and 1.11).

The dialectical numbers, namely the numbers with equalized units, are generated in the intelligible Being of the second hypothesis in the *Parmenides* and are based on the anthyphairetic periodicity. The whole number generated in an intelligible Being is finite and equal to the number of ratios of successive remainders of the anthyphairesis of the intelligible being plus one. Dialectic numbers are also discussed in the *Republic* 522b8-526e8, *Philebus* 16e-17a, and *Philebus* 56c-d. For details, consult Sect. 6 of the present chapter and Negrepontis (2024, Sects. 3 and 4); (2023a, b, Sect. 1.10.1).

The *straight line* is generated, employing the abbreviated Zeno-like representation of the Platonic Divisions, based on the infinite anthyphairetic division and the "already infinite" (cf. Sect. 1.10.2), *the circle* is generated from the anthyphairetic periodicity (cf. Sect. 1.10.4), and the first three Postulates are derived in a way similar to that outlined by Proclus in his *Commentary to Euclid* 185, 8-21, where Proclus endows the geometric point with the motion of the intelligible Being (cf. Sects. 1.7.3 and 1.10.3).

The *three kinds of angle* and *Postulate 4* of the *Elements* are generated from the Pell property of the side and diameter numbers, the anthyphairetic approximations of the intelligible diameter to the side of a square (*Republic* 510d7-8) (cf. Sect. 1.11).

The main failure of Plato's dialectics lies in its inability to derive within the intelligible Being the *Fifth Postulate*. Platonists believed that such derivation would be possible, perhaps because both each of the intelligible Being and the Fifth Postulate is a finitization of the Infinite, but in quite different ways.

9 The Geometric Analysis Is Connected with the Ascent from Hypotheses to the Anhupotheton in Plato's Intelligible Being, and the Geometric Synthesis and Deduction with Plato's Descent from the Anhupotheton

9.1 Geometric Analysis Is Connected with Plato's Division (and Logos) by Proclus *In Euclidem* 211,18-212,4

> Nevertheless, there are certain methods that have been handed down, the best being the method of *analysis (analuseos)*, which traces the desired result back to an acknowledged principle (archen). Plato, it is said, taught (paradedoken) this method to Leodamas, who also is reported to have made many discoveries in geometry using it.
> A second is the method of *diairesis ("diairetike")*, *dividing ("diairousa")* the genus ("genos") proposed into its natural parts ("kat' arthra"), and which [diairetike] provides a starting point ("aphormen") for the demonstration ("tei apodeixei") of the construction ("tes kataskeues") of the proposed [genus] by means of the division ("anaireseos") of the other [genera]. This method also Plato praised (exumnesen) as an aid (epikouron) in all the sciences (epistemais). Proclus, *In Eucliden* 211, 23-27

Here Proclus appears to correlate the method of Division of a genus into species, the basic tool in Plato's philosophy, with the method of Analysis, a geometric method of which Plato was quite fond. The method of Division of genera into species is certainly the method that Plato exemplifies with his Divisions in the *Sophist* and the *Statesman*, and according to our interpretation, it is an imitation of the geometric method of periodic anthyphairesis.

9.2 The Ascent and Descent in Plato's *Republic* Are Connected with the Analysis and Synthesis, Respectively, in Pappus' *Collection 634,11–636,14* (Negrepontis and Lamprinidis 2008)

One may wonder what would be the relation of the Platonic Division method with Analysis and Synthesis. The close connection of the Ascent from hypotheses to the anhupotheton and the converse Descent with Pappus' Analysis and Synthesis, respectively, as suggested by Proclus' connection (Sect. 9.1), presented in Negrepontis & Lamprinidis (2008) and in Lamprinidis (2015), are outlined here.

Indeed, the abbreviated Zeno-like form of Plato's Division and Logos method in the *Sophist*, by which Plato defines the Angler, the noble Sophistry, and the Sophist, makes it quite reasonable to think of Division as Analysis, and of the successive genera of the Division as higher hypotheses: starting from the initial genus G, in analysis fashion we look for a higher hypothesis: we divide the initial genus G into a and b, and then decide that of the two a,b, the suitable higher hypothesis from which G follows is b, and so on, in quite the same way that geometrical analysis, such as the one by Pappus in the *Collection* 634,11–636,14, is described as a process proceeding from hypotheses to higher hypotheses.

Thus the abbreviated representation of anthyphairetic *Division* in the *Sophist* is precisely the *Analysis*: starting with G, G follows from b, b follows from b_1, b_1 follows from b_2, . . ., b_{n-1} follows from b_n, ending into something *known* b_n, exactly when the Logos criterion, say $a_n/b_n = a_1/b_1$, and Analysis are achieved; while the direction opposite to anthyphairetic Division corresponds to the Synthesis and the Deduction →, namely

$$(\text{Logos Criterion } a_1/b_1 = a_n/b_n) \rightarrow b_n \rightarrow b_{n-1} \rightarrow \ldots \rightarrow b_2 \rightarrow b_1 \rightarrow G.$$

[D2+] *The dialectic descent from the anhupotheton* Republic *511b7-c2*

> again ("palin") having as parts the parts of it [the beginning], so to *move downward* ("katabainei") to the consequences ("epi teleuten"), making use exclusively of intelligible entities, moving on through kinds ("eidesi autois") to kinds ("eis auta") and ending with kinds ("teleuta eis eide"). 511b7-c2

10 The Two Infinites, the Dialectical and the Eristic Infinite, in the Platonic Dialogue *Philebus* 16c5-17a5, Essentially Coincide with the Two Infinites, the Actual and the Potential Infinite, Respectively, in Aristotle's *Physics* 204a-207b

10.1 Plato Accepts the *Dialectical, "Already" (Hede Apeiron) Infinite,* Namely the Anthyphairetically Periodic Infinite Induced by the Logos Criterion, and Rejects the *Eristic Infinite,* Namely the Infinite Established Without the Logos Criterion *Philebus* 16c5-17a5

The claim that the dialectical, already infinite in the *Philebus* 16c5-17a5 is the anthyphairetic periodic infinite induced by Platonic Logos, the philosophic analogue of the Theaetetus Logos Criterion, has been established in Sect. 5 of Negrepontis (2024). We outline the content of the *Philebus* passage.

[1]. The fundamental two properties of an intelligible Being are:
(a) The Being is a mixture of infinite and finite
(b) The Being is (infinitely) many and a self-similar One

> *Socrates* A gift (dosis) of gods to men, as I believe, was tossed down from some divine source through the agency of a *Prometheus*, together with a gleaming fire; and the ancients, who were better than we and lived nearer the gods, handed down the tradition that all the things which are said to be eternal *Beings* (ton aei legomenon einai) consist of *One and Many* (*ex henos kai pollon*) and have inherent in them *the Finite (peras)* and *the Infinite (apeirian)* 16c5-10

[2]. The "already infinite" (*hede apeiron*) is essentially the Logos Criterion for anthyphairetic periodicity, since the Logos Criterion implies periodicity of the ratios of successive remainders, allowing for the definition of the *"whole number" (arithmon panta)* by the plus one rule:

> This being the way in which these [true Beings] are arranged (diakekosmemenon), we must always assume that there is in every case *one idea* (mian idean) of everything and must look for it – for we shall find that it is there (euresein gar enousan) – and if we get a grasp (metalabomen) of *this*, we must look (skopein) next for two, if there be two, and if not, for three or some other *number (arithmon)*; and again we must treat each One (hen hekaston) [of the true Beings] in the same way, until we obtain the knowledge (idei) *not only* (me monon) that the original One is one and many and infinite, *but exactly how many* (alla kai hoposa) it is.
>
> And we must *not* offer (prospherein) the idea of the infinite to the multitude (plethos) *until* (prin) we obtain knowledge (katidei) of its *whole number* (*arithmon panta*) between (metaxu) *Infinite and One*; then *already (hede)* each One (hen hekaston) of the true Beings may be let go (methenta chairein ean) *into the infinite (eis to apeiron)* 16c10-e2.

[3]. The dialectic method, employing the already infinite, is an evolution due to Theaetetus of the traditional Pythagorean way of proving incommensurabilities, while the more recent methods, of proving quadratic incommensurabilities, without the Logos Criterion, presumably the arithmetical nonanthyphairetic proofs employing the Principle of the Least, by Archytas, are called *eristic*.

> The gods, then, as I said, handed down to us this mode of investigating, learning, and teaching one another; but the wise men of the present day make the one and the many too quickly or too slowly, in haphazard fashion (hopos an tychosi), and they put the infinites (apeira) right after (euthus) One and (ekpheugei) the intermediate ones (ta mesa), and it is this which distinguishes between the *dialectic* and the *eristic* methods of discussion.16e3-17a5

Thus, the crucial features of the dialectic infinite that differentiates it from the eristic infinite are (a) the Logos Criterion of periodicity, (b) the resulting equalization of the parts/Kinds, and finally (c) the generation of dialectical numbers. Because infinity is established by the Logos Criterion, something that occurs at a finite stage of the anthyphairesis, Plato calls the dialectical infinite the "already" infinite (hede infinite).

10.2 Aristotle in *Physics 204a-207b* Accepts the Potential Infinite but Rejects the Actual, Completed Infinite

10.2.1 The Definition of the Potential Infinite by Aristotle in *Physics*

For there will *always (aei)* be something to take *outside (exo). 206b17-18*

> The infinite turns out to be the contrary of what it is said to be.
> It is not what has nothing outside (meden exo) it that is infinite,
> but what always has something outside (aei ti exo) it. 206b33-207a1
> The translation of these and following passages from Aristotle's Physics is from Hardie RP & Gaye PK, 1984

We note that Aristotle considers that his definition is new and goes against the usual notion of infinite at the time; we conclude that the older notion of the infinite was the completed infinite.

10.2.2 Aristotle's Rejection of the Actual Infinite in *Physics*

It is plain, too, that the infinite *cannot be* an actual (energeiai) infinite and a Being (ousia) and principle (archen). 204a20-21

Now, as we have seen, magnitude is **not** actually infinite (kat' energeian apeiron). 206a16 but by division (diairesei) it is infinite. 206a17

The alternative then remains that the infinite is a potential (dunamei) infinite 206a18.

10.2.3 Aristotle Reassures Mathematicians that His Definition of Infinite Exclusively as Potential Infinite Poses No Poblem for Mathematics Proofs. *Physics* 207b27-34

> Our account does not rob (ouk aphaireitai) the mathematicians of their science, by disproving (anairon) the actual infinite (energeiai apeiron) in the direction of increase, in the sense of the untraversable (adiexiteton).
>
> In point of fact they do not need (oude deontai) the infinite and do not use it (ou chrontai). They postulate only that the finite straight line (peperasmenen) may be produced as far as they wish. 207b27-34

10.3 The Aristotelian Potential Infinite Essentially Coincides with the Platonic Eristic Infinite, and the Platonic Dialectical Already Infinite Essentially Coincides with the Aristotelian Actual Infinite

We will now argue that the two infinities, actual and potential, introduced by Aristotle in *Physics*, are not really new but they are a variation of the two infinities, dialectical, and eristic, respectively, introduced by Plato in the *Philebus* for the identification of the Platonic dialectical infinite, with the Aristotelian *actual infinite*.

10.3.1 The Dialectical Infinite/Already Infinite and the Actual Infinite Are Both Described as Infinites Completed at Some Finite Stage

The Platonic *dialectical infinite* is also described as *already infinite (hede apeiron*, *Philebus* 16e1-2), an expression similar to the expression *actual infinite*, which describes the Aristotelian *completed infinite*.

10.3.2 Dialectical Infinite as Actual Infinite is realized in the Incommensurability Proofs by the Pythagoreans, Theodorus, and Theaetetus, and Eristic Infinite as Potential Infinite in the Incommensurability Proofs by Archytas

A central occurrence of the infinite in Greek Mathematics and philosophy occurs in relation to incommensurability. The infinite described by Plato as *dialectical infinite* and as *already infinite* is clearly the infinite produced in the incommensurability proofs given by the older people that Plato regards as wiser, namely the Pythagoreans, Theodorus, and Theaetetus. The Pythagoreans proved the incommensurability of the diameter to the side of a square, and Theodorus proved the quadratic incommensurabilities of line a to line b, in the cases $a^2 = Nb^2$, with N nonsquare number, $2 < N \leq 17$ (*Theaetetus* 147c-d), both employing the *preservation of Gnomons/application of areas*, the precursor of the Logos Criterion for anthyphairetic periodicity; and, of course, Theaetetus was the first to introduce a theory of ratios for magnitudes, and proved the deep quadratic incommensurabilities with the use of his *Logos Criterion*, another name for the *already infinite*.

Plato criticizes and is ironic toward the later geometers and philosophers who do not employ the Logos Criterion, and go directly to the infinite, but in fact, proceed with less speed.

We can have a good guess whom Plato ironically calls the "now wise," who do not use the dialectic infinite, but the eristic infinite: Archytas, Eudoxus, and their students like Menaichmus, who indeed had no use for the Logos Criterion and the dialectical/aleady infinite. Instead, their basic mathematical results depended heavily on the Principle of the Least (every nonempty subset of the natural numbers has a least element), Archytas for the proof of quadratic and solid incommensurabilities, using the basic Propositions VIII.7 & 8, based on Proposition VII.27 of the *Elements*, itself based on Proposition VII.2 employing the Principle of the Least, and Eudoxus for the proof of the fundamental Proposition V.8, employing directly the Principle of the Least, Archytas and his students had at about that time employed the arithmetic nonanthyphairetic methods, appearing the Books VII and VIII of the Elements of quadratic and solid incommensurabilities by Archytas.

At the same time, Aristotle in *Physics*, a work written not long after the *Philebus*, in the interesting passage 207b27-34, reassures the mathematicians that they do not need *actual infinity* to deal with infinity. Which mathematicians does Aristotle attempt to reassure that rejection of the actual infinite poses no problem for Mathematics? Certainly, not Theaetetus (already dead anyway at that point), who based all his great incommensurability proofs on the Logos Criterion. Aristotle here reassures Archytas, Eudoxus, and their students such as Menaichmus, who were instrumental in turning Mathematics away from anthyphairesis.

An arithmetical proof, based on the Principle of the Least, of incommensurability of, say, the diameter to the side of a square, implies, according to the converse of Proposition X.2 of the Elements, that the anthyphairesis of the diameter to the side is infinite. But, lacking the Logos Criterion, the infinite does not become actual, it is potential. So Plato's eristic is in agreement with Aristotle's potential.

10.3.3 On the One Hand, the Dialectical Infinite Has the Power to Equalize the Opposite Parts and thus Produce an Intelligible Being that Is One and an Intelligible Being with the Power to Generate Numbers with Equalized Units; and, on the Other Hand, the Actual Infinite Is Rejected by Aristotle Exactly Because It Claims to Equalize Opposites that Are Deemed Impossible to Be Equalized, Like the Teacher with the Student and the Medical Doctor with the Patient

The main reason that Plato's dialectic infinite must essentially be identical to Aristotle's actual infinite is the fact that both infinites achieve the crucial *equalization* of opposite parts.

On the one hand, the most prominent property of the *Platonic already infinite*, instituted by the *Logos Criterion* and anthyphairetic periodicity, is the equalization of the two initial opposing elements of the anthyphairetic dyad, in fact, the equalization of all the parts/remainders, generated by the periodic anthyphairesis. This property of *equalization* precisely allows Plato's intelligible Being to be a

(self-similar) One, even though it has been divided into an infinite multitude of parts. It is precisely this equalization that allows Plato to introduce in an intelligible Being the dialectical numbers, the numbers with equalized units.

On the other hand, Aristotle's main argument against *actual infinity* is the fact that equalization is impossible. Does not mention explicitly that the crucial property of the actual infinite is the power to *equalize* the parts, and to generate *dialectical numbers*, but these properties are prominent in Aristotle's argument rejecting the Platonic "self-motion" and Zeno's Dichotomy in the *Physics*, and also in the pseudo-Aristotelian treatise *On Indivisible Lines* rejecting again Zeno's Dichotomy.

For Aristotle, this claimed *equalization* makes no sense, and, as he scathingly notes, it is like equalizing the active teacher with the passive student, who is being taught, or the passive sick person who is being healed with the active medical doctor, who is healing. Thus for Aristotle, the rejection of the actual infinite is equivalent to the rejection of the *equalizing* power of the Logos Criterion, which is the *sine qua non* of Plato's nature of intelligible true Beings, namely, that it has the power to produce equalization of parts, since only equalization effected by periodicity justifies the Oneness of the intelligible Being.

As we will see in more detail in Sect. 11, Aristotle rejects

- in *Physics* 258a Plato's self-moving Soul in the *Phaedrus* 245c–246a, and implicitly Plato's practically identical definition of the intelligible Being in the *Sophist* 247d
- in Physics 263a–b, the second version of Zeno's Dichotomy paradox, in which the infinite multitude of parts produced by the (symbolically dichotomic) division and motion is at the end numbered in a dialectical, equalized manner. Thus equalization is the main feature of both the dialectical infinite and the actual infinite.

11 The Self-Motion of the Soul in Plato's *Phaedrus* 245c5-9 and the Second Version of Zeno's Dichotomy Paradox in Aristotle's *Phaedrus*, Are Each a Dialectical/Actual Infinite (*Physics 257a34-b20* and *263a4-264b9*), a Platonic Intelligible Being, but Nonexistent According to Aristotle

11.1 The Self-Motion of the Soul in Plato's *Phaedrus* 245c5-9 Is a Platonic Intelligible Being, a Dialectical/Actual Infinite, but Nonexistent According to Aristotle's *Physics* 257a34-b20

11.1.1 *For Plato*, Every Motion Is Generated by a Dyad Consisting of the Passive Part that Is Being Moved and the Active Part that Is Causing the Motion. *Timaeus* 57a3-6

Plato in the short but clear *Timaeus* 57a3-6 passage defines Motion as the interaction of an active and a passive part

it is difficult (chalepon), or rather impossible (dunaton),
for *that which is to be moved* (to kinesomenon) to exist without *that which is to move* (tou kinesontos),
or *that which is to move* without *that which is to be moved;*
but in the absence of these, there is no motion (kinesis) *Timaeus* 57e3-6

This description of motion makes motion appear to be close to anthyphairesis. Eudemus' comments given just below go much further in this direction.

11.1.2 The Definition of Soul Using Self-Motion in the *Phaedrus* 245c5-9

Every soul is immortal (athanatos), for that which is ever moving (aeikineton) is immortal but that which moves something (kinoun) else or is moved (kinoumenon) by something else, when it ceases to move, ceases to live.
Only that which moves itself (to auto kinoun), since it does not leave itself, never ceases to move (oupote legei kinoumenon), and this is also the source and beginning of motion for all other things which have motion. *Phaedrus* 245c5-9

11.1.3 *Eudemus's Comments, Reproduced Verbatim by* Simplicius, *Commentary to Aristotle's Physics* 1237,10-21, Reveal that the Nature of Plato's Definition of the Soul in the *Phaedrus* as a Self-Moving Entity Is *Anthyphairetic ("Antanairountes")*

Eudemus divides this too in this way.
"because the moving part [b] is either unmoved or is being moved.
But if the moving part [b] is unmoved, then we have proved what we wanted to prove, namely that the self-moving entity consists of a part being moved [a] and a part that moves and is itself unmoved [b].
But if the moving part [b] moves being itself moved, since everything being moved is moved by something. This part [b] too will be moved by something [a_1].
But this something [a_1] cannot be an external entity, since then it would not be a self-moving entity, but it would be said to be moved by some other part. Hence the moving part [b] will be moved by a part of the self-moving entity, either by one of its parts or by a part of the part that is being moved.
And again, the moving part [a1], moves another part either being itself unmoved, and then it will be the unmoved prime mover, or being itself moved by some part [b_1]; and this moving part will again be a part either by one of its parts or by a part of the part that is being moved.
And thus, *anthyphairating (antanairountes, 1237,21)* in this way, as Eudemus says,
Simplicius, Commentary to Aristotle's Physics 1237, 10-21.
[translation by the author]

Of course, *antanairesis* is another term for *anthyphairesis*; cf. Alexander Aphrodisieus, *Commentary to Aristotle's Topics* 545,16-17.

11.1.4 The Identification of the Self-Moving Soul in the *Phaedrus* with the Intelligible Being in the *Parmenides* and the*Sophist* and with the Dialectical Already Infinite of the *Philebus* 16c5-17a5.

According to Eudemus (Sect. 11.1.3)

(a) "Self-motion" in the *Phaedrus* is interpreted as
  (i) *Motion* in the form of anthyphairetic division of a dyad of opposites, in which
  (ii) The two opposing parts of the dyad are *equalized*; hence we can talk about *self-motion*, presumably by anthyphairetic periodicity and the Logos Criterion.

We realize that the anthyphairetic description of the Soul as self-motion in the *Phaedrus* makes Plato's definition of the Soul identical with

(b) The One of the second hypothesis in the *Parmenides*, the paradigmatical intelligible Being, forming the anthyphairetic dyad One Being, as analyzed in Negrepontis (2024), and

(c) The intelligible true Being, as defined in the *Sophist* 247d8-e4, 257b6-8, forming the anthyphairetic dyad Being, not-Being.

Furthermore, the equalizing power of the intelligible Being in (b) or in (c), in consequence of anthyphairetic periodicity and the "plus one rule" of the *Parmenides* 148d-149d, allows for the generation of dialectical numbers within the intelligible Being, so that, each of the entities (a), (b), (c) is equivalent to

(d) The *dialectical, already infinite* in the *Philebus 16c5-17a5*.

11.1.5 Aristotle's Rejection in *Physics* 257a34-b20 of Plato's Self-Moving Soul, on the Ground that Equalization of Parts Is Impossible (*Physics* 202a21-b22, above) Can Be Extended, by Sect. 11.1.4, to the Rejection of Plato's Intelligible Being, on the Ground that It Is an Actual/Dialectic Infinite, Satisfying the Logos Criterion

In *Physics* 202a21-b22, Aristotle rejects the possibility of equalizing the opposite parts of a dyad, of the active element with the passive element, referring repeatedly to the dyad <active teaching, and passive learning>, and concluding that it is absurd (atopon), impossible (adunaton), without Logos (alogon), these two actions to be equalized. Since equalization of parts is crucial for the One and Many properties of an intelligible Being, it is clear that Aristotle aims his criticism against the nature of Plato's intelligible Being.

This correlation becomes explicit later, in the *Physics* 257a34-b20, where Aristotle employs the earlier passage 202a21-b22 to show that self-motion is impossible, namely the equalization of the active moving part with the passive being moved part of the dyad of motion is impossible, namely exactly because the *actual infinite is rejected as impossible*. But, in Platonic philosophy, these active and passive parts *are* equalized, achieving self-motion, by the Logos Criterion, forming an *already infinite*. So Aristotle's rejection of the actual infinite is in fact rejection of the Platonic already infinite (as noted in Sect. 10.3.3).

The completion of the Eudemus passage, reported by Simplicius, in his *Commentary to Aristotle's Physics* 1237, 21-26 (appearing in Sect. 11.1.3), with the rejection of Plato's definition, in 1237, 10-21, makes clear that what is being rejected is *the periodic*, and thus necessarily *infinite, anthyphairesis*, namely the dialectic infinite.

"And thus, *anthyphairating (antanairountes, 1237,21*) in this way," as Eudemus says,
"we eventually arrive at the prime mover (to protos kinounta),
Those were such that were not being moved by some, but all by all.
Because *it is not possible that the anthyphairesis (antanairesin, 1237,23) is infinite;*
for they would not be in motion, if there were no first moving
for it is shown that it is impossible that the whole is being moved by the whole.
Hence it is necessary that the first part that is active in motion is causing motion but itself is not in motion."
Simplicius, *Commentary to Aristotle's Physics* 1237, 21-26
[[translation by the author]]

11.2 The Second Version of Zeno's Dichotomy Paradox Is a Dialectical/Actual Infinite (*Physics 263a4-264b9*), Identical with a Platonic Intelligible Being, but Nonexistent According to Aristotle

The *first version* of Zeno's Dichotomy paradox initially given in Aristotle's *Physics* 233a21-33, 239b11-14, and repeated here in 263a11-22, states that the runner traversing a finite line interval must contact, pass through an infinite multitude of points in finite time, an impossible task. Aristotle's solution is that there is *no paradox or impossibility at all* because the runner will have to contact and pass through an, indeed, infinite multitude of (half-)points, but this space infinity is comparable *not* with the finite total time needed to complete the action, which action after all is the transversal of a finite length of distance, but will be done *not* in an infinite time, namely in a time that is divided infinitely many times, exactly as the length is divided infinitely many times. Thus additively, finite length is transversed in finite time, and division-wise, infinite length is transversed in infinite time.

But things are radically different in the *second version* of Zeno's Dichotomy Paradox, in which the runner not only contacts an infinite multitude of intervals but also *numbers the infinite* (apeiron arithmekenai arithmon, 263a9-10), as well.

What is the difference between the first version (involving only an infinite multitude of subintervals) and the second version (involving the *numbering* of the infinite multitude of subintervals) of Zeno's Dichotomy? Aristotle states that this turns out to be a substantially different paradox, that requires a radically new solution,

(a) *The second version of Zeno's Dichotomy Paradox is solved by Aristotle in a way analogous to his solution of the Plato's Definition of the Soul in the Phaedrus.*

Aristotle states that his solution of the second version of Zeno's Dichotomy paradox will be similar to his solution of Plato's definition of the soul, (ton auton tropon apanteteon, 263a4), and thus the second version will be rejected as impossible (adunaton, 263a11) because it has the form of.an actual infinity.

According to the general remarks in Sect. 10.3 and our analysis in Sect. 11.1 of Plato's definition of the Soul, including Eudemus's passage, reported by

Simplicius, that motion is *anthyphairetic*. We expect that the structure of the second version of Zeno's Dichotomy Paradox is that of periodic anthyphairesis.

But how can we be convinced that the second version of Zeno's Dichotomy Paradox describes with its numbered divisions an infinite periodic anthyphairesis?

(b) *The condition that when the runner reaches the end will have numbered the infinites suggests equalization of the parts generated by the Division, hence actual infinite.*

The *numbering* of the subintervals is possible ONLY if the subintervals are acceptable units for *dialectic number*, but a dialectic number is possible ONLY if these units/subintervals are *equalized*; but as we have seen in the second hypothesis of the *Parmenides* and the *Sophist*, the subintervals will be equalized ONLY if the *Logos Criterion* holds, namely, as we have seen if we have an *already infinite* (in the sense of *Philebus* 16c5-17a5), and thus an *Aristotelian actual infinite*.

If we want to have a numbering of the infinite, with dialectical numbers, namely with equalization of the units, then, as just explained, the infinite must be aristotelianly actual (entelecheiai)/ /platonically dialectic. Indeed, the numbering of the infinite is associated with the actual, not the potential, infinite; cf. Simplicius, *Commentary to Aristotle's Physics* 1290, 23-24, 28-32, 35-37, 1291,5-7.

(c) *The existence of contrary motions in pairs strongly suggests a motion based on anthyphairesis*.

The usual *continuous* way of dividing an interval into repeated halves does produce an infinite multitude of subintervals, and in fact constitutes *the first form of Zeno's Dichotomy Paradox*, but this infinite is Aristotelianly potential (dunamei)/platonically eristic, and thus with no equalization and no numbering.

On the contrary, the motion in a straight line that results in a completed, actual infinite is NOT a continuous motion, as traditionally thought for the motion of Zeno's Dichotomy paradox, BUT a discontinuous one, in successive steps, so that immediately after a step in straight motion there is a *contrary* (enantia) one in the opposite direction, during which contrary motion there is rest in the straight direction (antikeimene heremia, ?) (264a7-b9).

(d) *The structure of the second version of Zeno's Dichotomy Paradox is identical to the structure of the Divisions of the Sophist, which have been shown to be based on periodic anthyphairesis.*

The divisions in the *Sophist* (of the Angler, the Noble Sophistry in Sect. 2.1, the Sophist) have been shown to have the structure of *anthyphairetic division plus the Logos Criterion* (Negrepontis 2012). Each of these divisions has exactly the same structure as the second version of Zeno's Dichotomy argument, namely

(i) An anthyphairetic division, but in abbreviated, continuous form, mentioning only the first, the third,. . ., in general, the odd stage of the anthyphairesis

(ii) Moreover a periodic one, satisfying the Logos Criterion, hence having the power to equalize the parts generated by the division, a fact indicated by its description as a *Dichotomy*

We conclude that Zeno's true Being is essentially identical to Plato's intelligible being and with the indivisible line.

12 The Application of Zeno's Basic Argument, as Interpreted in Negrepontis (2023a), Is Employed, in the Pseudo-Aristotelian Work *On Indivisible Lines* 968a18-b4, with the Almost Contradictory Property of the Second Version of Zeno's Dichotomy Paradox, of Not Only Touching but Numbering the Infinites as Well, Shows that Zeno's True Being Coincides with the Indivisible Line (Atomos Gramme, 968b4), Hence, According to Sects. 2, 3, 4, 5, 6, 7, 8 and 9 Above, with Plato's Intelligible Being

12.1 The Pseudo-Aristotelian Passage *On Indivisible Lines* 968a18-b4

> Again, according to the argument of Zeno [5] there must be some *partless magnitude (megethos ameres*, 968a19*)*,
> since
> [4] it is impossible to contact an *infinite multitude* (apeiron hapsasthai, 968a20)
> [Hett's rendering: number] in a finite time,
> [3] when contacting each of them, and that which moves must first reach half-way, and half clearly belongs to that which is not without parts (tou de me amerous).
> But if *anything moving along a line* (to epi tes grammes pheromenon) contacts an *infinite multitude* (haptetai ton apeiron, 968a23) [Hett's rendering: series] in a finite time,
> secondly if the faster it travels the greater the space it covers in the same time, and
> [2*] lastly, if the intelligible motion (dianoias kinesis, 968a25-26) is the quickest motion,
> [2] then even intelligible thought (dianoia) must contact an *infinite multitude* (ton apeiron ephaptoito, 968a28) [Hett's rendering: series] one by one in a finite time.
> If, then, intelligible thought (dianoian) contacting the parts one by one is *numbering (arithmein, 968b2)* [Hett's rendering: counting], then
> [1, 3] it must be possible *to number (arithmein, 968b2)* [Hett's rendering: count] *the infinites* (ta apeira, 968b3) [Hett's rendering: series] in finite time.
> [4] If this is *impossible (adunaton, 968,b3)*,
> [5] then there must exist an *indivisible line (atomos gramme, 968b4*).

[[Aristotle, *Minor Works (On Colors, On Things Heard, Physiognomics, On Plants, On Marvellous Things Heard, Mechanical Problems, On Indivisible Lines, Situations and Names of Winds, On Melissus, Xenophanes, and Gorgias)*, translation by W. S. Hett, The Loeb Classical Library, Cambridge Massachusetts, Harvard University Press, London, William Heinemann, 1936. With modifications by the author]]

The passage is about the second (and more essential than the first) version of Zeno's Dichotomy Paradox, as described in Aristotle's *Physics* 263a-b (The first version of Zeno's Dichotomy Paradox, described in Aristotle's *Physics* 233a, 239b, is, as Aristotle correctly points out, nonparadoxical).

12.2 An Outline of Our Anthyphairetic Interpretation of the Pseudo-Aristotelian Passage *On Indivisible Lines* 968a18-b4 as a Special Case of Zeno's Basic Argument

We will analyze this passage, based on our anthyphairetic interpretation of Zeno's Basic Argument in Negrepontis (2023a) in five steps, according to the numbering in brackets that we have added.

The first step in Zeno's Basic Argument is

[1]. The real Being coincides with a sensible Being.

This step is not explicit in the pseudo-Aristotelian passage on indivisible lines, but it is clearly implied at the point marked as [1,3] in the passage in Sect. 12.1, where the sensible property of contacting/passing through an infinite multitude of successive midpoints in finite time is turned into the "intelligible" property of numbering the infinites in finite time.

The second step in Zeno's Basic Argument is the introduction of an almost contradictory property satisfied by the true Beings; in this case

[2]. The Real Being possesses the almost contradictory property [P] of numbering the infinites while passing through them.

The third step consists in transferring property [P] from the true Being to the sensible, on account of [1]. Thus

[3]. A sensible entity must satisfy property [P], namely to have the power to number the infinites when it completes the motion from a point A to a point B.

The fourth step consists in noting, as a consequence of our common sensible experience, and definitely NOT as a logical contradiction, that

[4]. It is impossible that property [P] holds for the sensibles.

The fifth step in Zeno's Basic Argument consists in deducing that

[5]. There is a true Being separate from the sensible entities; and thus, given our two arguments (a) identifying Zeno's true Being with Plato's true Being in Negrepontis (2023a), and (b) identifying Plato's true Being with the Indivisible line in Sects. 2, 3, 4, 5, 6, 7 and 8 in the present chapter, [5] there is an *indivisible line* (*atomos gramme*, 968b3), as claimed in the pseudo-Aristotelian passage.

12.3 Of Interest Are Two Sentences Appearing in the Passage

partless magnitude (megethos ameres, 968a19*)*

is the property of Zeno's true Being, stated in Zeno's Fragment B1, according to which a true Being has an infinite multitude of parts, each with magnitude, and at the same time is measureless. The property is a precursor of the fundamental property One and Many of Plato's intelligible Being. In the above pseudo-Aristotelian passage, it serves as a property essentially equivalent to the property of Being an indivisible line; and indeed both are transfers of the anthyphairetic periodicity to the parts generated by the division, using the plus one rule.

[2*] lastly, if the intelligible motion (dianoias kinesis, 968a25-26) is the quickest motion

This seems an ironic reference to the fact that intelligible, dialectic motion becomes already infinite at some finite stage, achieving, with an incredibly "quickest motion," actual infinity, an impossibility for Aristotle.

13 Conclusion for Sections 10, 11, and 12

We conclude that the second version of Zeno's Dichotomy Paradox, Zeno's true Being, is described, by Aristotle in *Physics*, as an intelligible Plato's true Being (Sects. 10 and 11), and by the pseudo-Aristotelian work *On Indivisible Lines*, as an indivisible line (Sect. 12), confirming the conclusions of both Sects. 2, 3, 4, 5, 6, 7, 8 and 9 of this work that Plato's intelligible Being is an indivisible line, and our interpretation of Zeno's Paradoxes and Arguments in Negrepontis (2023a) that Zeno's true Being coincides with Plato's intelligible true Being.

Acknowledgments I express my warm thanks to Athanase Papadopoulos for encouragement and extensive and essential discussions that have led to an essential improvement of this work.

Translations

Aristotle, *Minor Works (On Colors, On Things Heard, Physiognomics, On Plants, On Marvellous Things Heard, Mechanical Problems, On Indivisible Lines, Situations and Names of Winds, On Melissus, Xenophanes, and Gorgias)*, translation by W. S. Hett, The Loeb Classical Library, Cambridge Massachusetts, Harvard University Press, London, William Heinemann, 1936.

R. P. Hardie and R. K. Gaye, *Physics*, 1984, in *The Complete Works of Aristotle: The Revised Oxford Translation* (J. Barnes, ed.) 1, Princeton University Press, Princeton.

G.R. Morrow, 1970. *Proclus: A Commentary on the First Book of Euclid's Elements, Translated, Introduction*, *Notes*. Princeton University Press, Princeton

Plato. Plato in Twelve Volumes,

Vols. 5 and 6, *Republic*, translated by P. Shorey, 1969,

Vol. 9, including the *Phaedrus*, translated by H. N. Fowler, 1925,

Vol. 12, including the *Sophist*, translated by H. N. Fowler, 1921,

Cambridge, MA, Harvard University Press; London, William Heinemann Ltd.

Proclus, *The Commentaries of Proclus of the Timaeus of Plato*, translated from the Greek by Thomas Taylor, London, 1820.

P. M. Huby and C. C. W. Taylor, *Simplicius: On Aristotle Physics 1.3–4*. (Translation) Ancient Commentators on Aristotle (R. Sorabji, general ed.), Bristol Classical Press, London, 2011.

[Translations of ancient passages in the present chapter, unless otherwise indicated, are from the above sources with modifications, including italicization of some words for emphasis, by the author]

References

Aristotle (1936) Minor works (on colours, on things heard, physiognomics, on plants, on marvellous things heard, mechanical problems, on indivisible lines, situations and names of winds, on melissus, xenophanes, and gorgias), translation by Hett WS, The Loeb Classical Library, Cambridge, MA, Harvard University Press, London, William Heinemann.

Crivelli P (2012) Plato's account of falsehood. A study of the Sophist. Cambridge University Press

Frede M ([1992], 2022) Plato's Sophist on false statements, in volume ed. [First edition Kraut R 1992] Second edition Ebrey D, Kraut R, Cambridge companion to Plato. Cambridge University Press, pp 433–463

Hardie RP, Gaye RK (1984) Physics. In: Barnes J (ed) The complete works of Aristotle: the revised Oxford translation, vol 1. Princeton University Press, Princeton

Heidegger M (1924–5). Platon: Sophistes (Schußler I, ed). Frankfurt a.M., 1992

Huby PM, Taylor CCW (2011) Simplicius: on Aristotle physics 1.3–4. (Translation) Ancient Commentators on Aristotle, Sorabji R, general ed.). Bristol Classical Press, London

Knorr WR (1975) The evolution of Euclidean Elements: a study of the theory of incommensurable magnitudes and its significance in early Greek geometry. Reidel, Dordrecht

Lamprinidis D (2015) The anthyphairetic nature of Plato's Mathemata. Doctoral Dissertation, Department of Mathematics, Athens University [in Greek]

Negrepontis S (2012) Plato's theory of knowledge of Forms by Division and Collection in the Sophistes is a philosophic analogue of periodic anthyphairesis (and modern continued fractions) arXiv:1207.2950 [pdf] math. HO

Negrepontis S (2019) Plato on geometry and the geometers. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Cham, pp 1–88

Negrepontis S (2023a) Zeno's arguments and paradoxes are not against motion and multiplicity but for the separation of true Beings from sensibles. In: Papadopoulos A (ed) Essays in geometry – dedicated to Norbert A'Campo. European Mathematical Society, Berlin

Negrepontis S (2023b) The anthyphairetic interpretation of knowledge as recollection in Plato's Meno 80d-86e, 97a-98b

Negrepontis S (2024) The periodic anthyphairetic nature of the one of the second hypothesis in Plato's parmenides. In: Hascher X, Papadopoulos A (eds) Proceedings of the Conference Mathématiques et Musique: des Grecs à Euler, 10–11 September 2015, Strasbourg. Hermann, Paris

Negrepontis S, Lamprinidis D (2008) The platonic anthyphairetic interpretation of Pappus' account of analysis and synthesis, in history and epistemology in mathematics education. In: Barbin E, Stehlikova N, Tzanakis C (eds) Proceedings of the 5th European Summer School ESU5. Vydavatelsky Servis, Plzen, pp 501–511

Plato. Plato in twelve volumes, vols 5 and 6, Republic, translated by Shorey P, 1969; vol 9, including the Phaedrus, translated by Fowler HN, 1925, vol 12, including the Sophist, translated by Fowler HN, 1921. Harvard University Press, Cambridge, MA; William Heinemann Ltd, London

Proclus (1820) The commentaries of Proclus of the Timaeus of Plato, translated from the Greek by Thomas Taylor, London

Rashed M (2013) Plato's five worlds hypothesis (Ti. 55cd), mathematics and universals. In: Chiaradonna R, Galluzzo G (eds) Universals in ancient philosophy, vol 2013. Edizioni della Normale, Pisa, pp 87–112
Taylor AE (1926) Forms and numbers: a study in platonic metaphysics. Mind 35:419–440
Taylor AE (1960) Plato, the man and his work. Methenn, London

Brouwer's Intuitionism

Mathematics in the Being Mode of Existence

Victor Pambuccian

Contents

1 Introduction 646
2 A Very Short Introduction to L. E. J. Brouwer's Life 652
3 East and West, to Have or to Be 661
3.1 Haas on the Main Characteristic of the Mind East and West 661
3.2 Fromm on to Have or to Be 663
4 Language 665
5 The Eastern Origin of Subjective Idealism 673
6 Brouwer's View on Mathematics 675
7 Wisdom and Mysticism: Beyond Mathematics and Time 689
8 Conclusions 692
References 694

Abstract

It is argued that Brouwer's philosophy of mathematics makes perfect sense if viewed from an *Eastern* philosophical perspective, as a mathematics in what Erich Fromm called "the being mode of existence." The difficulty Western philosophers have accepting its validity under Brouwer's own justifications is that mathematics is one of the highest prized treasures of Western philosophy (those footnotes to Plato's dialogues).

Keywords

L.E.J. Brouwer · Intuitionism · Erich Fromm · William S Haas · Being mode of existence · Eastern mind

V. Pambuccian (✉)
School of Mathematical and Natural Sciences, Arizona State University – West Campus, Phoenix, AZ, USA
e-mail: pamb@asu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_103

1 Introduction

L. E. J. Brouwer and *his* intuitionism is perhaps unique in the annals of the history of mathematics and its philosophy by the quality of the hostility encountered from mathematicians and philosophers alike, from his own time all the way to the present day. Brouwer appears to be unique by the unusual fact of the absence of restraint in the verbal abuse, ridicule, and belittling of his philosophy even from mathematicians who have been attracted to intuitionistic mathematics *sans* underlying philosophy.

Craig Smoryński, a mathematical logician, relates in Smoryński (1994) that "some years ago my topology teacher told the class that Brouwer had been a competent topologist who went off the deep end and invented intuitionism" and that "Brouwer is still widely believed to have been a Nazi, and indeed the pre-publication version of one major obituary of him reported that Brouwer had joined the SA and marched in its parades, his hair flowing majestically in the wind."

The latter part of this slanderous piece of disinformation came from a draft of Kreisel and Newman (1969). There is no drop of truth in it. Brouwer never had any kind of sympathies for any nationalist movement, to say nothing of a biologically driven one, as anyone reading (Van Dalen 2013) will learn. It just fits well into the simplistic picture of the world painted by ideologists to claim that counter-modernists (see below) are right-wingers. After sending that draft to Arend Heyting, and receiving the following reply on March 15, 1969 (copy in Heyting Papers, Haarlem; unpublished, courtesy of Mark van Atten), Kreisel removed, thanking Heyting, that passage:

> It is not true that Brouwer supported National-Socialism. I have very good reasons to say this emphatically. Brouwer was a friend of my wife's father, who was a fervent anti-national-socialist. (He was the conductor of the Residentie-orkest, the philharmonic orchestra of The Hague, and he got into trouble by refusing to let the orchestra play the Horst Wessellied in honour of Prince Bernhard, who then, bethroted to Princess Juliana, was still a German subject.) If Brouwer had in the least supported National-Socialism, my wife's father would have broken all links with him, and my wife would have known that.

By 1994, Smoryński had considerably softened his view on Brouwer and pointed out that the above were misrepresentations and that the author of the book he was reviewing had "done a hatchet job on Brouwer" that "van Stigt's psychological analysis of Brouwer is excessive and excessively negative."

In 1977, his take on Brouwer had been more negative, for he had written that (Smoryński 1977, p. 822):

> L.E.J. Brouwer was making the rounds in a bizarre attempt to turn mathematics into a religion. When, in 1920, Hermann Weyl fell prey to Brouwer's lunacy, David Hilbert decided to intervene.

Charles McCarty, an expert on intuitionistic logic, found Brouwer's philosophy to be an "idiosyncratic idealism" (McCarty 2008, p. 39) and argued as follows that one

ought to found intuitionism on philosophical assumptions that bear no resemblance to Brouwer's own (McCarty 1987, p. 542).

> Now, we seem to be left with only one alternative: that constructions are themselves phenomenal objects akin to pains and afterimages. Their intrinsic properties are only those which they seem to have and these are solipsistic: unshared, unsharable and, on some outlooks, incommunicable. As is well known, Brouwer encouraged just this sort of mathematical solipsism. He declared intuitionistic mathematics to be the record of the fruits of the creative activity of a single individual. Constructions are phenomenal objects which, in the mind of the mathematician, make up a kind of Lego set, pieces of which can be assembled and displayed on the inner visual field. Ultimately, the facts of mathematics are just those which he can put together from his fundamental elements.
>
> It is clear that we ought not to follow Brouwer.

We will see later that "single individual," "solipsistic," and "mathematical solipsism" are misunderstandings of a superficial reading of Brouwer's writings.

And Giovanni Sambin, a constructivist mathematician, tells us (Sambin 2008, p. 301):

> We are here to celebrate a master, namely L.E.J. Brouwer, and one hundred years from the birth of his intuitionism. I believe that the best way to keep Brouwer alive is to try and go beyond Brouwer himself. [. . .] To go beyond Brouwer means to learn his lesson, but also try to fix his mistakes and to soften (in ourselves) some hardness in his personality. And in this way to keep him alive. To confine Brouwer into the prison of what he has materially written, would mean to kill him (again).

The philosopher Carl Posy, who devotes a book to mathematical intuitionism (Posy 2020), asks in Posy (1998, p. 308), ironically:

> First, suppose, as I have argued, that Brouwer's position is primarily mathematical. Why does he encumber it with this bugaboo about language in general and about Hilbert's concentration on formal language in particular?

And Georg Kreisel, a mathematical logician who contributed to intuitionistic logic, has mostly derogatory things to say about Brouwer, for example:

> Gödel was utterly bored by Brouwer unlike several logicians and mathematicians who, being dry themselves, were buoyed by Brouwer's probably genuine exuberance. (Kreisel 1987, p. 146)
>
> Apparently, he [Brouwer] did not like to be interrupted anyway; fittingly, for a good solipsist. (Kreisel 1987, p. 147)

With little knowledge of the actual biographical facts that led to the pause in Brouwer's publications, which were not available at that time, Kreisel proposes an entirely implausible explanation, if one takes Brouwer's own philosophical writings seriously (which Kreisel obviously does not, and makes his disdain explicit in Kreisel (1977)):

> Be that as it may, it is a fact that Brouwer himself devoted the ten years after publication of the papers by Heyting and Gödel to non-scientific activities, although, objectively, these papers concerned essential points of his life's work – and although he was only fifty and in good health. It does not seem unreasonable to suppose that he had received an intellectual shock. (Kreisel and Newman 1969, p. 45)
>
> Gödel's immensely natural proofs did not need anything like the heavy distinctions which Brouwer thought essential to his development of constructive mathematics (nor, incidentally, the ingenious constructions that Hilbert had introduced into proof theory). What they did involve was a clear perception of aims and principles, a philosophical, non-doctrinaire analysis, alien to the logical activists of the time. Seeing the incomparable superiority of this kind of analysis Brouwer had to face the question, consciously or unconsciously, to what extent he had even begun to master his own logical ideas. (Kreisel and Newman 1969, p. 44)

The absurdity of Brouwer mastering "his own logical ideas" and the supposed Gödel-envy in Brouwer will become apparent when we get to know Brouwer's views on logic. That Heyting's formalization of intuitionistic logic caused an "intellectual shock" is ludicrous but very much in line with the tone one felt free to use whenever the subject of discussion was Brouwer.

A historian of mathematics found the need to chip in as well while reviewing Van Dalen (1999b): "Part of the mystique around Brouwer is caused by the notorious obscurity (one could say incoherence) of this philosophy. (Grattan-Guinness 1999, p. 529) "Arend Heyting [. . .] became a major protagonist of intuitionism without the supernatural." (Grattan-Guinness 1999, p. 531)

In addition to these squarely hostile remarks about Brouwer's philosophy of mathematics, there is a subtler way of rejecting his philosophy while trying to find firm foundations for a highly valued intuitionistic mathematics by ignoring Brouwer altogether, the way Michael Dummett, a philosopher, did in Dummett (1975) and Dummett (2000).

This is all the more remarkable, given the achievement of the man whose philosophical thought has inspired so little respect in the mathematical and philosophical world. As Gerrit Mannoury summarized one of Brouwer's achievements, in the retelling of Dirk van Dalen (Van Dalen 2013, p. 747):

> In 1947 it was forty years ago that Brouwer got his doctorate. Friends, students and colleagues who had not deserted the old master, organised a jubilee symposium on 19 February. The ceremony took place in the lecture hall of the Geological Institute of Brouwer's brother Aldert. Mannoury addressed his famous student and colleague; he sketched how Bolyai, Riemann and Peirce had broken the hegemony of the Euclidean and Archimedean axioms, how Einstein had 'widened the visual field of physical science', but how Aristotelian logic had survived all disturbances until Brouwer 'emancipated human thought from the authority of the logical principles and so ran down that stronghold itself'.

Sure, the discovery of non-Euclidean geometry had not been greeted with enthusiasm either. But 100 years after its discovery, all doubts regarding its validity and significance had vanished. Why, with many more professional mathematicians and philosophers interested in mathematics active in the twentieth and twenty-first centuries, did the passing of time not similarly validate Brouwer's revolutionary

achievement? It had never attracted more than a handful of devotees and survived into the twenty-first century only in the form of a formal logical theory, far removed from Brouwer's own conception of what intuitionist mathematics ought to be. In the words of Meister Eckhart, who was dear to Brouwer, his intuitionistic mathematics was a work connected with him, which vanished with him.

> For as soon as the work is accomplished, it has instantly vanished, as has the time in which it happened and there is nothing left to do with the work. Should it somehow continue to act, then it must do so with other works and also in another time. (Trans. V. P.)
>
> [Denn sobald das Werk war, ist es sogleich zunichte geworden, ebenso die Zeit, in der es geschah, und ist nun mit dem Werke nichts mehr zu tun. Soll es irgendwie weiter wirken, so muss es mit anderen Werken geschehen und auch in einer anderen Zeit.]. (Eckhart 1934, p. 204)

The task that we set for ourselves is to find out the deep reasons for the rejection of Brouwer's intuitionistic mathematics and of his philosophy of mathematics. We refer to the *deep* reasons to distinguish them from reasons connected with his biography, with Hilbert's animosity and inability to withstand professional criticism, which we take to be *superficial* in the sense that, even in their absence, rejection would have occurred.

We could not find any such explanation in the literature on Brouwer's philosophy, which contains works in which Brouwer's intuitionism is taken seriously and is treated with utmost respect. The literature that takes Brouwer seriously tries to establish links between his philosophy and that of other figures in the history of philosophy. The first such attempt is due to Oskar Becker who found in Becker (1927) that Brouwer's intellectual predecessors were Aristotle, Descartes, and Kant and who understands Brouwer's intuitionism through the prism of a thoroughly *verheideggert* phenomenology. Similarities between Brouwer's philosophy and the *Lebensphilosophie* of the 1920s, as well as with Bergson (which is also emphasized in Largeault (1993)), were put forward in Hesseling (2003). Parallels with Husserl (*sans* Heidegger) are the subject of Van Atten (2007), van Atten (2015, pp. 237–281) (where "a close intellectual kinship between [Husserl and Brouwer]" is detected), and Tieszen (2008). We find these analyses very interesting, yet their aim is to bring Brouwer closer to some sort of Western philosophical mainstream and thus do not aim to explain the historical phenomenon of exclusion and derision of Brouwer's philosophy. Aristotle, Descartes, Kant, and Husserl are pillars of Western philosophy, and hardly anyone dares to write about them in the style sampled above for Brouwer's philosophy. Bergson and the *Lebensphilosophie* had their decades of glory. Moreover, Husserl was a close friend of Hilbert, and their philosophies of mathematics, despite differences, "share a 'mathematics first,' non-revisionist approach toward mathematics" (Hartimo 2017, p. 245). Was the *Grundlagenstreit* between Hilbert and Brouwer, as Heyting would have it (as reported in Van Dalen (2013, p. 829)), only "a historical accident, a blatant failure of common sense and of communication"? Would a mediation by Husserl – who wrote in a letter of May 9, 1928, to Heidegger that "the most interesting in Amsterdam were the long conversations with Brouwer, who made a very important impression on me, one of

a thoroughly original, radically sincere, genuine, very modern man" ["das Interessanteste waren in Amsterdam die langen Gespräche mit Brouwer, der einen ganz bedeutenden Eindruck auf mich machte, den eines völlig originellen, radikal aufrichtigen, echten, ganz modernen Menschen" (Husserl 1994, p. 156)] – have avoided its outcome? Such a mediation could have certainly helped with raising the level of civility that Hilbert was lacking, but we will argue that the *deeper* reasons for rejecting Brouwer's philosophy were not of a nature that could be mediated away. Schopenhauer as a possible influence is dealt with in Eggenberger (1976), Largeault (1993), Detlefsen (1998), Koetsier (2005), and, in-depth, Van Belle (2021). Largeault also finds that

> Another feature that relates Schopenhauer to Brouwer: the discredit that struck their work. Should we believe that each type of culture has its own incompatibilities? [Trans. V. P.]
>
> [Autre trait qui apparente Schopenhauer et Brouwer: le discrédit qui a frappé leur œuvre. Faut-il croire que chaque type de culture comporte ses incompatibilités propres?]. (Largeault 1993, p. 166)

Could a similarity with Schopenhauer – which is entirely hypothetical, as Brouwer mentions very few Western philosophers, and, from those of times past, only Kant and Spinoza – explain the outsider status of his philosophy and the animosity it elicits? Schopenhauer's influence in his lifetime and long past his death has been considerable. His philosophy influenced in the nineteenth century some late Romantics (such as Eminescu), Tolstoy, and philosophers (such as Nietzsche) and in the twentieth century Wittgenstein and Cioran, psychologists (such as Freud and Jung), writers and poets (such as Thomas Mann, Hermann Hesse, Samuel Beckett, Jorge Louis Borges, and Rainer Maria Rilke), and even physicists (such as Einstein or Wolfgang Pauli). Nothing remotely similar can be said about Brouwer's philosophy. One should also consider that Brouwer himself refers to a proximity to "perennial philosophy" (Brouwer 1948, p. 487), not to some *dernier cri* in philosophy.

However, we hold onto the observation that "each type of culture has its own incompatibilities" as a particularly insightful remark.

One may think, with Herbert Mehrtens, that Brouwer's philosophy of mathematics and his intuitionism were doomed by being *counter-modernist.* What this means is spelled out succinctly in Mehrtens (1996, pp. 521, 522) (the long version can be found in Mehrtens (1990)):

> The two common traits of the various modernisms that I identify as central are, first, the autonomy of cultural production and, second, the departure from the vision of an immediate representation of the world of experience. The artist, the composer or the scientist define what is good and true in their respective fields, and they define it autonomously, independently of political, religious, or philosophical authority, and largely not only independently from but against the beliefs dictated by common sense.
>
> [. . .]
>
> The counter-modernist attitude arises with modernism. It is part of modernity, of the modern world. That is why I chose the term counter-modernism instead of anti-modernism. It is the counterpart to modernism, insisting on the question whether there is not some natural

> substance to the truth and meaning of mathematics. The counter-modernist attitude claims that there is such substance, mostly called 'intuition'.

However, counter-modernism, on its own, does not doom a philosopher's posterity, as Gottlob Frege's posthumous destiny clearly shows. Frege was not only counter-modern, he failed to even understand what Hilbert was doing, as his correspondence with Hilbert on the latter's *Grundlagen der Geometrie* of 1899 and his outlandish polemic in the pages of the *Jahresbericht der Deutschen Mathematiker-Vereinigung* clearly shows. Frege was not only counter-modern by insisting that axioms ought to be *true* (in an absolute way), and not arbitrarily chosen, and by his rejection of non-Euclidean geometry at a time when the issue had been long resolved among mathematicians (see Tóth 2009), but he was also a nationalist (Mehrtens suggests (Mehrtens 1996, p. 520) that there is a correlation between anti-modernism and nationalism, although he acknowledges the existence of exceptions, such as Hermann Weyl; Brouwer is falsely depicted as an anti-internationalist, which anyone reading (Van Dalen 2013) will conclude is flatly wrong) and held anti-Semitic views. Although ignored during his lifetime, causing him to adopt a strident polemic and satiric writing style in several publications (see Wille 2020a, b), Frege has enjoyed a posterity as the founding figure of the dominant style of philosophy in the Anglo-Saxon world, analytic philosophy. His writings have generated a veritable exegetical Frege industry, in which even attempts to find how he could have been right in the dispute with Hilbert abound (to the disbelief of mathematicians, such as Hans Freudenthal, who wrote already in 1960 that "Frege, rebuking Hilbert like a schoolboy, also joins the Bœotians. (I have never understood why he is so highly esteemed today.)" (Freudenthal 1962, p. 618)).

A major aspect of modernism is a marked relativism regarding truth claims. Truth is generated either individually or by groups as deemed convenient and almost anything goes. There are no universal truths that can be discerned from the human condition, whose very characteristics are considered debatable, to say nothing of an ability to "discern." Contrary claims are considered oppressive. In the field of logic, for instance, this is expressed by Carnap's principle of tolerance (*Toleranzprinzip*):

> *In logic, there are no morals.* Everyone is at liberty to build his own logic, i.e. his own form of language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments. (Carnap 1959, § 17)
>
> [In *der Logik gibt es keine Moral.* Jeder mag seine Logik, d. h. seine Sprachform, aufbauen wie er will. Nur muß er, wenn er mit uns diskutieren will, deutlich angeben, wie er es machen will, syntaktische Bestimmungen geben anstatt philosophlscher Erörterungen. (Carnap 1934, § 17)]

In summarizing the results of his study of Brouwer's philosophy, Michael Detlefsen observes how far from the modernist credo Brouwer's understanding of mathematics is (Detlefsen 1998, p. 331):

> He [Brouwer] was a kind of mathematical 'existentialist' in that he sought a conception of mathematical knowledge that makes it basic to our human existence. For this to be so, he believed, mathematical knowledge must be intimately related to our most basic knowledge of ourselves – that is, to that 'existential' awareness that we have of ourselves as willing, acting beings. Such views are, to be sure, extraordinary when viewed in the light cast by the bland and vapid views that have dominated the philosophy of mathematics of the past few decades. They are not, however, for all that, either silly or out of place.

We find that Largeault's remark on incompatibilities related to "culture" and Detlefsen's remark on the closeness of mathematical knowledge to our human existence, to our *being*, come very close to the two reasons we will put forward for the rejection of Brouwer's philosophy.

The first reason, somewhat similar to that mentioned by Largeault, is connected with the geography of the mind, along the lines of Holenstein (2004), with a mistake made by Central Casting, which placed L. E. J. Brouwer in Europe rather than in "the East" and in mathematics rather than in the humanities. The basics of the characteristic features of thought East and West are taken from Haas (1956) and will be summarized below. We are fully aware that speaking of East and West is not well taken these days, but given that we know only of ideological concerns related to that geography, we will ignore the prohibition to refer to those cardinal points.

The second reason is connected with what Erich Fromm called "the being mode of existence" in Fromm (1976), for we claim that both the philosophy and the mathematics Brouwer envisioned are reflections of a mind in the being mode of existence, and his reason for rejecting classical mathematics and formalism is that he sees them tainted with the having mode of existence.

Given the dominance of Greek philosophy in Western philosophy and mathematics (for, as Whitehead noted "The safest general characterization of the European philosophical tradition is that it consists of a series of footnotes to Plato" (Whitehead 1979, p. 39)) and of the having mode of existence in society, a philosophy and a proposal for a reform in mathematics that go against these two pillars were doomed to fail. Making sense of the connection of the two reasons mentioned above with Brouwer's philosophy and pointing out why that Trojan horse cannot be accepted in the proudest citadel of Western thought, which is mathematics, are, in essence, the aims of this essay.

2 A Very Short Introduction to L. E. J. Brouwer's Life

Luitzen Egbertus Jan Brouwer was born on February 27, 1881, at Overschie, the Netherlands, the first child of a schoolmaster and his wife. He enrolled at the University of Amsterdam as a 16-year-old, passing his final examination in 1904, the year he also got married to Lize de Holl (after being engaged to her since 1902), 11 years his senior, who had a daughter from a previous marriage. His close friends during that time were poets, not mathematicians: Carel Adema van Scheltema, Jan Lockhorst, and Albert Plasschaert. He went three times, on foot, to Italy and back.

The experience as a mathematics student left him wondering whether mathematics would be the field to pursue, whether it would have a life significance.

> As happens so often, I began my academic studies as it were, with a leap in the dark. After two or three years, however full of admiration for my teachers, I still could see the figure of the mathematician only as a servant of natural science or as a collector of truths: – truths fascinating by their immovability, but horrifying by their lifelessness, like stones from barren mountains of disconsolate infinity. And as far as I could see there was room in the mathematical field for talent and devotion, but not for vocation and inspiration. (Brouwer 1946a, p. 474)

The lectures of Gerrit Mannoury, a self-taught mathematician, solidified his interest for mathematics, as he no longer felt in front of a lifeless discipline but rather one intimately connected to life:

> [. . .] The undertone of Mannoury's argument had not whispered: 'Behold, some new acquisitions for our museum of immovable truths', but something like this: Look what I have built for you out of the structured elements of our thinking. – These are the harmonies I desired to realize.[. . .] Behold the harmonies, neither desired nor surmised, which after the completion [of the construction] surprised and delighted me'. (Brouwer 1946b, p. 475)

In a letter written on September 7, 1906, to Adama van Scheltema, Brouwer gives voice to his feelings of estrangement:

> Life is a magic garden. With wondrous, softly shining flowers, but among the flowers the gnomes are walking, and I am so afraid of them. They stand on their head and the worst is that they call out to me that I must also stand on my head; once in a while I try to do so and die with shame; but then sometimes the gnomes cry that I do it very well, and that I am indeed a real gnome too. But on no account will I fall for that. (Van Dalen 2013, p. 196)

He got his doctorate in 1907 with a thesis on both foundational topics and Hilbert's fifth problem. He was appointed a *privaatdocent* at the University of Amsterdam in 1909. In 1912 he was appointed extraordinary professor at that same university and, a year later, ordinary professor on the chair that had been vacated for him by his doctoral advisor Korteweg. He turned down full professorships from Göttingen and Berlin ("I feel far too much a Dutchman and in particular a Friesian Dutchman. And I would rather live here between Dutch friends to enjoy, and Dutch enemies to see through, than far away among strangers!" (Van Dalen 2013, p. 191), from an interview from Wiessing (1960)). Between 1909 and 1913, he published a series of papers that revolutionized topology. One of the results is Brouwer's fixed-point theorem: A continuous map f: $D \to D$, where D is the n-dimensional closed disk, has a fixed point, i.e., there exists an $x_0 \in D$, such that $f(x_0) = x_0$. After 1913, the only major contribution to topology is in 1939, an intuitionistic proof of the triangulation of any differentiable manifold. Starting with 1913, he returned to the subject of his doctoral dissertation: the rebuilding of mathematics on a sound foundation, *intuitionistic mathematics*. Appointed to the editorial board of the leading mathematical journal of the time, the *Mathematische*

Annalen in 1914, at the invitation of Felix Klein, he was removed from that board in 1928 by David Hilbert on fabricated grounds. The editorial board was completely revamped, with Albert Einstein and Constantin Carathéodory, who had resisted Brouwer's removal, removed as well.

Einstein had tried to convince Hilbert that one does not choose one's personality or one's health, regardless of the afflicted organs or systems, and that it is in bad taste to make any fuss about such matters. On October 18, 1928, he wrote to Hilbert

> Mr. Brouwer is an unwilling defender of Lombroso's theory of the close connection between genius and madness. I cannot understand how you can take this man, who in his country not only enjoys the reputation of a great mathematician but also that of a hopeless troublemaker, so bitterly seriously (Trans. V P.)
>
> [. . .]
>
> I consider him [Brouwer], with all due respect for his mind, a psychopath and it is my opinion that it is neither objectively justified nor appropriate to undertake anything against him. I would say: 'Sire, give him the liberty of a jester!' If you cannot bring yourself to this, because his behaviour gets too much on your nerves, for God's sake do what you have to do. I, myself, cannot sign, for the above reasons such a letter. (Van Dalen 2013, p. 556)
>
> [Herr Brouwer ist ein unfreiwilliger Verfechter von Lombrosos Theorie der nahen Verbindung von Genie und Wahnsinn. Ich kann es nicht begreifen, dass Sie diesem Mann, der in seinem Lande neben dem Rufe eines grossen Mathematikers auch den eines hoffnungslosen Querulanten geniesst, so bitter ernst nehmen koennen.
>
> [. . .]
>
> Ich betrachte ihn bei aller Hochachtung vor seinem Geiste als einen Psychopathen und halte es weder fuer objektiv gerechtfertigt noch fuer zweckmaessig etwas gegen ihn zu unternehmen. Ihnen moechte ich sagen:‚Sire, geben Sie ihm Narrenfreiheit'! Wenn Sie dies aber nicht vermoegen, weil sein Verhalten Ihnen zu sehr auf die Nerven gegangen ist, so tun Sie in Gottes Namen, was sie muessen. Ich selbst kann aus besagten Gruenden einen solchen Brief nicht unterzeichnen.]. (Rowe and Felsch 2019, p. 278f)]

A few days later, on October 23, 1928, Einstein complemented the above with a somewhat different evaluation of Brouwer in a letter to Carathéodory:

> Furthermore, I admire him not only as a particularly clairvoyant spirit, but also as a forthright and characterful man. (Trans. V. P.)
>
> [Im uebrigen aber verehre ich ihn nicht nur als ueberaus hellsehenden Geist, sondern auch als geraden und charaktervollen Menschen.]. (Rowe and Felsch 2019, p. 120)

After his removal from the editorial board of the *Mathematische Annalen*, Brouwer worked tirelessly to launch a competing journal with an international editorial board. He became, in 1934, a founding editor of *Compositio Mathematica*, a truly international journal (unlike the *Annalen*, which, even before Brouwer's removal, had only two non-German members on its editorial board) – with all German editors resigning in protest in 1936 due to Brouwer's refusal to purge its Jewish editors – which ceased publication in 1940 due to the German invasion of the Netherlands. When it resumed publication, in 1951, he was removed from the editorial board of his own journal by his colleagues. The effect that episode had on him is expressed in a letter of May 8, 1953, to Mannoury as follows:

> I too have been in bed for months during this winter at a relapse of a heart attack, which I suffered in 1950 being suddenly informed of the theft of my journal and the unbelievable means of deceit and fraud used. It is, I believe, through this event which I on the one hand out of self-preservation must try to forget and to which I can, on the other hand, not resign myself, that I have definitely become misanthropic within our national borders. Remarkable, how little philosophy sometimes protects against psycho-somatic reflexes and reactions. (Van Dalen 2013, p. 789) (Dutch original in Van Dalen (2011b, p. 2789))

In 1918 a group consisting of Brouwer; Frederik Van Eeden, a writer and psychiatrist; Jacob de Haan, a Dutch-Jewish writer; and Gerrit Mannoury, a mathematician, invited many renowned figures, such as Rabindranath Tagore, Martin Buber, Giuseppe Peano, and Henri Borel, a sinologist and journalist, to join it, with the aim of "the creation of new words, which form a code of elementary means of communication for the systematic activity of a new and holier society" (Van Dalen 2013, p. 264). The choice of a sinologist was connected with Brouwer's belief that, unlike Western languages, Eastern languages allow for the expression of a different philosophy (Van Stigt 1990, p. 200).

The International Institute for Philosophy saw its founding that year and was dissolved in 1922 to be replaced by the *Signific Circle*. Buber had declined the invitation, questioning the possibility of a collective creation of words, given that "The creation of a word is for me one of the most mysterious processes of spiritual life" (Van Dalen 2013, p. 262) [Wortschöpfung, Erschaffung des Wortes ist für mich einer der geheimnisvollsten Vorgänge des geistigen Lebens (Van Dalen 2011b, p. 793)].

Brouwer's reply to Buber is of great interest for our thesis of the *Eastern* nature of Brouwer's mindset, as it is one of several places in which he deplores shortcomings in the *Western* cultural setup.

> The word of the Occident does indeed in various cases have in addition to its material, a spiritual value, but the latter is always subjugated to the first, and where the first one has acquired a more certain and permanent orienting influence on the activity of society, in the sense that it induces isolated individuals, in their pursuit of physical security and material comfort, to hinder each other as little as possible and wherever possible to support each other; the latter lacks any influence on the legal state of affairs (except possibly when it is used itself for the devious realisation of injustice); therefore its influences are weak, fleeting and local.
>
> Words which have exclusively a spiritual value and which are suitable for the in- and exhaling of the Weltgeist and for the orientation on the observance of Tao are non-existent in Occidental languages; should they exist, then their influence would be paralysed by the mutual physical hatred, rooted in mutual distrust of the purity of their birth, of people who live too closely together, and which hampers the pursuit of material comfort by the single individual only mildly, but which considerably hampers the in- and exhalation of the Weltgeist. (Van Dalen 2011a, p. 262)
>
> [Das abendländische Wort besitzt zwar in mehreren Fällen neben seinem materiellen einen seelischen Wert, aber letzterer ist immer dem ersteren untergeordnet, und während erstere eine sichere und dauerhaft orientierende Wirkung auf die Aktivität der Gemeinschaft erworben hat, in dem Sinne dass es die einzelnen Individuen dazubringt, im Erstreben körperlicher Sicherheit und materiellen Komforts einander möglichst wenig zu hindern und womöglich zu unterstïutzen, entbehrt letzterer jeden Einflusz auf die Rechtsverhältnisse

(es sei denn insofern er daselbst zur Erschleichung von Unrecht missbraucht wird); demzufolge sind seine Wirkungen schwach, vorübergehend und lokalisiert.

Wörter, welche ausschliesslich seelische Werte besitzen und dazu geeignet sind, die Gemeinschaft auf das Ein- und Ausatmen des Weltgeistes und das Innehalten von Tao hin zu orientieren, gibt es in den abendländischen Sprachen nicht; wenn solche existierten, würde übrigens ihre Wirkung lahmgelegt werden durch den wechselseitigen leiblichen Hass der zu dicht bei einander lebenden Menschen, welcher im gegenseitigen Misstrauen gegen die Reinheit ihrer Geburt wurzelt und das Erstreben materiellen Komforts durch die einzelnen Individuen nur unerheblich, das Ein- und Ausatmen des Weltgeistes aber in hohem Masse hindert. (Van Dalen 2011b, p. 792)]

He expressed himself similarly in 1948:

Eastern devotion has perhaps better expressed this wisdom than any western man could have done. For instance in the following passages of the Bhagavad-Gita which even in translation have conserved their electrifying power. (Brouwer 1948, p. 486)

Another one who declined the invitation to join the project of word creation was Erich Gutkind, the author of a prophetic, mystical book *Siderische Geburt – Seraphische Wanderungen vom Tode der Welt zu Taufe der Tat* (*Sideric birth – Seraphic wanderings from the death of the world to the baptism of the deed*), and a friend of Walter Benjamin und Gershom Scholem. Gutkind "doubted if European languages could simply incorporate or assimilate words of a high spiritual value" (Van Dalen 2013, p. 263) for "words like the Chinese Tao or the Hebrew Torah, so strikingly analogue in meaning and sound, can never be grown on a European foundation and presuppose a radical reform of society, require a higher society" (Van Stigt 1990, p. 200), to which Brouwer replied that "The word cannot wait for the higher form of society, because the higher form of society waits for the word" (Van Dalen 2013, p. 263).

Brouwer became a lifelong friend of Erich and Lucia Gutkind. After Brouwer's first visit of the Gutkinds in Berlin, they wrote to Van Eeden "Brouwer is with us and he will bring you our greetings. We get along wonderfully and we are enormously pleased with him" (Van Dalen 2013, p. 323).

By 1946 Brouwer had come to the conclusion that "Buber was right in his denial of the creative power of collective work in this domain" (Van Stigt 1990, p. 196) of creating a language that would, on the one hand, be "purified of rhetoric, subjective, demagogic and in general of deceptively emotional admixtures" and, on the other hand, contain "words for basic immaterial notions which can serve as elements of a more indicative language, suitable to express general mental values and human emotions of fraternity and solidarity" (Brouwer 1946a, p. 452).

Besides the injustice visited upon him by his removal without any drop of a valid reason from the editorial board of the *Mathematische Annalen*, Brouwer suffered another blow in 1929, which he related in a letter to Hans Hahn of August 9, 1929:

Due to a disturbance in the execution of my travel plan, your letter from Bellagio reached me with a long delay. The disturbance was caused by a great calamity: four days ago my briefcase, which also contained my scientific diary, was stolen from me on the front platform

> of a Brussels tram, by a pickpocket, and both the police and the detectives consider the case as hopeless. Since in this diary my collective scientific thoughts and ideas of the last three years, which have largely disappeared from my memory, and of which only a few have already found a registration elsewhere, had been recorded, this event means for my scientific personality a serious personal mutilation, in a way that is like the decapitation (elimination of the central process) for a pine tree. To my amazement, I remain so far, fairly calm under this blow of fate; I believe, however, from certain phenomena, that I have nonetheless suffered a nervous collapse, the consequences of which will perhaps only later become visible, together with a disorganisation of my scientific thoughts. (Van Dalen 2013, p. 600f)
>
> [Infolge einer Störung in der Abwickelung meines Reiseplanes hat mich Ihr Brief aus Bellagio mit grosser Verzögerung erreicht. Die Störung wurde von einer grossen Kalamität verursacht: vor 4 Tagen wurde mir meine Brieftasche, welche auch mein wissenschaftliches Tagebuch enthielt, auf dem Vorderbalcon der Brüsseler Strassenbahn von Taschendieben gestohlen, und sowohl Polizei wie Detektive betrachten die Angelegenheit als hoffnungslos. Weil in diesem Tagebuch meine sämtlichen wissenschaftlichen Gedanken und Einfälle der letzten drei Jahren archiviert waren, welche grossenteils meinem Gedächtnis entschwunden sind und von denen nur wenige schon anderweitige Festlegung gefunden haben, so bedeutet das Geschehnis für meine wissenschaftliche Persönlichkeit eine schwere Verstümmelung, gewissermassen dasselbe was für eine Tanne die Enthauptung (Entfernung des Zentralsprosses) bedeutet. Zu meiner Verwunderung bleibe ich unter diesem Schicksalsschlag bisher ziemlich ruhig; ich glaube aber nach gewissen Erscheinungen, dass ich trotzdem eine Nervenerschütterung erlitten haben muss, deren Folgen vielleicht erst später, zusammen mit der Desorganisierung meiner wissenschaftlichen Gedankenwelt, ans Licht treten werden. (Van Dalen 2011b, p. 1747)]

Later on, twice, in 1941 and 1944 (Van Dalen 2013, pp. 674, 693), his house was on fire, so that Brouwer would write in a letter to De Loor on June 20, 1945 (Van Dalen 2013, p. 702), that "as a consequence my scientific archive and with it the manuscripts of my unpublished research have been lost." (Van Dalen 2013, p. 702)

If one is looking for the long mathematical silence that ensued, then it is not Gödel's theorems that brought it about, but these events.

Among the seven doctoral students Brouwer had, two developed intuitionistic mathematics further: Maurits Belinfante and Arend Heyting. The former, a Portuguese Jew, Brouwer's first intuitionistic follower, who developed intuitionistic analysis, was deported first to Theresienstadt and then in 1944 to Auschwitz, where he was killed. Heyting wrote, at Brouwer's suggestion, his dissertation on an intuitionistic axiomatic development of projective geometry. Later Heyting developed a formal logical calculus that should capture the essence of intuitionistically valid inferences. Brouwer was "highly pleased" with Heyting's work. However, as related by Heyting, "He [Brouwer] always maintained that formalization is unproductive, a sterile exercise. He never changed his mind about that" (Van Stigt 1990, p. 290). Developments of intuitionistic mathematics after World War II proceeded among formal logic lines.

In the 1930s, he served on the municipal council of Blaricum and was involved in several lawsuits and other time-consuming non-mathematical and unpleasant activities that he felt a compulsion to be involved in out of a very strong sense of justice.

After the war, he was suspended for some months as a suspected collaborationist, reprimanded, and then re-instated by the University. At the end of 1945, Brouwer

contacted his old friend Erich Gutkind, now living in the United States. Gutkind wrote a letter to Einstein, letting him know that Brouwer had expressed a strong wish to emigrate to the United States. He mentioned "that Brouwer had hidden some Jews in his institute and took care of them" as well as that "I recall that Brouwer travelled a couple of times to Germany to get German mathematicians out of German concentration camps" (Van Dalen 2013, p. 716). Gutkind asked Einstein to help Brouwer find a position in the United States. Einstein refused to help Brouwer, given that he was under investigation for collaboration with the occupier.

"There is no doubt that Brouwer suffered under the post-war accusations, – he, one of the leading Dutch scientists, who had always come to the defence of the underdog, was classified as a spineless collaborationist!" (Van Dalen 2013, p. 716).

He published a few more papers, one (Brouwer 1952a) showing that his own fixed-point theorem is intuitionistically false but can be rephrased intuitionistically by stating that, given a natural number n, one can find a point x_0 whose distance to $f(x_0)$ is less than 2^{-n}. He traveled extensively to give lectures on intuitionism. In 1959 his wife Lize died. In 1966, he was run over, crossing the street in front of his house in Blaricum, to deliver some presents for Saint Nicholas' eve to friends across the road, by three cars.

Van Dalen writes that "Somehow he managed to inspire either love and admiration, or hatred and repugnance" (Van Dalen 2013, p. 714). In the first category was the homeopathic physician Ralph Twentyman, who wrote that "the high spiritedness Brouwer brought into any room is one of my most treasured memories. Everything moved up a floor or two when he appeared" (Van Dalen 2013, p. 600f).

We should also add that Brouwer was very far from having adopted the European professorial persona. He could disappear in the middle of a discussion in the garden of his house, with women present, enter his house, and re-appear naked (as he enjoyed nude open-air baths), continuing the previous conversation, with no one commenting (Van Dalen 2013, p. 645). Or, during a visit to Canada, at 72, he would disappear at a picnic to climb a tree (Van Dalen 2013, p. 8f). "When he stayed with [Georg Henrik] von Wright, he asked his host one night, 'would you mind if I slept out on the ground tonight?' Von Wright may have had his doubts about his aged guest, but he saw no reason to refuse the request. And so Brouwer once more was one with nature under the sky, as he loved to be" (Van Dalen 2013, p. 788f).

During an interview with Wiessing, translated from Wiessing (1960, p. 143f) in Van Dalen (2013, p. 190f), Brouwer explained succinctly his intellectual journey:

> Basically my mathematical thinking is non-sensory internal architecture. You may compare these forms of thought to music or poetry. My first inklings of the possibility of such a mathematics emerged, I think, from discussions with my teachers at the time of my HBS [high school] and gymnasium study. But only in my dissertation of 1907 have I started to give these thoughts a definite formulation. Since then this mathematics, nowadays called 'intuitionistic', has developed with interruptions. The recognition in professional circles of this work of mine came only in a rather slow tempo, with many ups and downs. It has by no means found general acceptance! Many view it, even now, as charlatanism. There are also people who say that it may be correct, but that it is totally uninteresting and not even new.

> If I had not, now and then, written about 'ordinary mathematics', I don't think a place at a Dutch University could have been found for me.
>
> [. . .]
>
> [. . .] I don't like mathematics and it basically bores me.
>
> Wiessing: What would you, if it comes to that, rather have done than practice mathematics?
>
> Brouwer: That is hard to say. Let me say: to have no subject and to let my thoughts roam freely. Every attachment to a subject brings, as you will agree, that your realm of thought suffers a certain mutilation. And it is obvious that then one can only have pleasure in such a profession, if one is, as I sometimes observe with some people, supported and driven by ambition or conviction. But that has never been the case with me. Anyhow, life demands that you choose a profession. Well, then I think that science is for a man like me, who is by nature solitary, not such a bad sanctuary. One is less dependent on the public, and one can more easily preserve one's solitude, than if one takes up literature or the visual arts, not to mention music. For no matter how much pleasure and satisfaction art by itself may give to a person, society, I think, demands more violating concessions from artists than science does.

It is worth mentioning that Brouwer, who almost single-handedly developed geometric topology and could have continued on that path to certain glory, decided to return to the concerns of his youth (he had hoped, during his topological years, that his topological work could be made intuitionistically sound (see Dubucs 1988; Johnson 1981), repudiated his topological work, and never lectured on it, although the preoccupation with intuitionism made him appear, in the eyes of many mathematicians, a "charlatan," as he put it. To give a specific example, which also shows that an admiration for the Bhagavad Gita (see Weil 1991) in two European mathematicians can coexist with very different philosophies of mathematics, André Weil attended Brouwer's Berlin lectures on intuitionism and wrote to Fréchet on January 31, 1927, that

> He has declared in his first lecture that the principle of the excluded third is a superstition which is about to disappear. It is a pity that such a remarkable man devotes himself exclusively to such bizarre things. (Van Dalen 2013, p. 595)

Brouwer is one of the very rare cases who not only felt a slight hesitation at the alternative presented in the thought experiment devised by Simone Weil but actually chose to be an object of derision rather than betray the truth. Here is the thought experiment of a manifest counter-modern, one who believed in truth:

> If one were to propose to all those whose profession is to think, priests, pastors, philosophers, writers, scholars, professors of all kinds, the choice, from the present moment on, between two destinies: either to sink immediately and definitively into idiocy, literally, with all the humiliations that such a collapse entails, and keeping only enough lucidity to feel all its bitterness; or a sudden and prodigious development of intellectual faculties, which ensures them an immediate global celebrity and glory after death for millennia, but with the inconvenience that their thought would always stay a little outside of the truth; can one believe that a lot of them would experience for such a choice even a slight hesitation? [Trans. V. P.]
>
> [Si l'on proposait à tous ceux qui ont pour profession de penser, prêtres, pasteurs, philosophes, écrivains, savants, professeurs de toute espèce, le choix, à partir de l'instant

> présent, entre deux destinées: ou sombrer immédiatement et définitivement dans l'idiotie, au sens littéral, avec toutes les humiliations qu'un tel effondrement entraine, et gardant seulement assez de lucidité pour en éprouver toute l'amertume; ou un développement soudain et prodigieux des facultés intellectuelles, qui leur assure une célébrité mondiale immédiate et la gloire après la mort pendant des millénaires, mais avec cet inconvénient que leur pensée séjournerait toujours un peu en dehors de la vérité; peut-on croire que beaucoup d'entre eux éprouveraient pour un tel choix même une légère hésitation?]. (Weil 1949, p. 328)

Needless to say, this very short introduction does not relieve the reader, in particular, the reader who thinks negatively about Brouwer, to consult Van Dalen (2013) or Van Dalen (1999a) and Van Dalen (2005), Van Dalen (2011a), and Van Dalen (2011b) to get whatever picture some footprints on a sandy beach can reveal about a life. Unsurprisingly, both the reviews for *Mathematical Reviews* and for *Zentralblatt für Mathematik* (Dauben 2000; 2007; Soifer 2007) find that Van Dalen was too kind to Brouwer, for they would have seen some facts in a significantly more detrimental light. The much more substantial reviews (Smoryński 2015; Johnson 2014; Kaneko 2002b) are free from that kind of critique. While praising Van Dalen's monumental achievement, Gray (2015, p. 134) finds that "there is a slight feeling of sympathy for Brouwer that the reviewer does not share, being perhaps less enamoured of intuitionism" and that

> there was a solipsistic streak in him that strengthened his inability to get people on his side, that elevated a reasonable but by no means secure philosophical opinion to one of inflexible rectitude, that kept him illiberal and unable to see that his allies were increasingly rightwing in toxic times, and that made him capable of cruelly selfish behaviour. (Gray 2015, p. 134)

Gray is here very selective in the decision regarding who Brouwer's "allies" were, thinking of the National-Socialists Ludwig Bieberbach and Theodor Vahlen, but forgetting his friends with far-left leanings, the communists Henri Wiessing (Van Dalen 2013, p. 281) and Gerrit Mannoury (Van Dalen 2013, p. 513), and devotes significant space in his review to the post-war accusations against Brouwer, which were, to a careful reader of Van Dalen (2013), without any basis in fact, or rather a dispute about the most appropriate resistance strategy. It was by no means clear what that strategy should be, as a reader of Hermans (1949) would come to realize. Equating accusations with guilt would also be ill-advised in an atmosphere in which, even as late as 1963 – not 1945 – Miep Gies would be considered by the police "one of the suspected people" for the betrayal on August 4, 1944, of the Frank family (Gies and Gold 1987, p. 250) on the basis of her having been born in Vienna.

The "illiberal" part is even more mysterious, given that Brouwer was in 1935 and 1939 elected to the council of the town Blaricum, with a significant artist population. What was he advocating for?

> The preservation of the local heath (a nature reserve), [. . .] better access for walkers while keeping out cars [. . .] a raise for the local police [. . .], added improvements to a plan for the construction of a cycle path, [. . .] the improvement of the soccer field. [. . .] When [. . .] the committee of action 'For God' demanded the barring of the periodical of the freethinkers

> from the public library, Brouwer remarked that according to the state's instructions all groups must be represented. (Van Dalen 2013, p. 641)

And do these thoughts sound like those of an "illiberal" person?

> Power over fellow-creatures will be avoided. Firstly because one would get mixed up with limitations of other people's liberty of action. And secondly because those fellow-creatures are part of the reflex image held out to mind from its deepest home, therefore have to be respected, and must not be judged, let alone condemned, despised or rejected, even if they are enemies to be fought against. (Brouwer 1948, p. 486)

If one were to think that van Dalen's biography has cleared the air of the cobwebs of lies that have been spun around Brouwer, one must have forgotten that, as Thomas Francklin wrote:

> Falsehood will fly, as it were, on the wings of the wind, and carry its tales to every corner of the earth; whilst truth lags behind; her steps, though sure, are slow and solemn, and she has neither vigour nor activity enough to pursue and overtake her enemy. (Francklin 1787, Sermon XI (On Vigilance), p. 233)

Defamations, heard through the grapevine, are passed on, regardless of the absence of any mention of anything remotely resembling the accusation in the extensive correspondence related to the *Mathematische Annalen* affair and picked up in AD 2021 in Budianksy (2021, p. 82) (where Brouwer becomes a "committed Aryan nationalist") and Misak (2021) (where Brouwer is mentioned as "a budding Nazi collaborator") as facts, with

> The dispute took on a particular nuance in the sense that the Dutchman Brouwer set himself up as a champion of Aryan Germanness. Consequently Hilbert removed him from the editorial board of the *Mathematische Annalen* after he objected to what he felt were too many Eastern European Jewish ("*Ostjuden*") authors. (Fraenkel 2016, p. 137)

as source.

3 East and West, to Have or to Be

3.1 Haas on the Main Characteristic of the Mind East and West

To substantiate why an approach from an "Eastern" position to mathematics is unacceptable by the philosophical establishment, we need to define what we mean by the distinction between East and West in philosophy. We think that the following quotations from Haas (1956) will give an idea of what is meant by these different fundamental approaches.

> However, the decisive act which constitutes Eastern mind and civilization – the positing or the fixing of the subject – implies the idea that an existence entirely separated from and

independent of the subject cannot be conceived. So what becomes in the Western mind and civilization the autonomous object, is paralleled in the East by the other which is coherent with and held in existence by the subject. (Haas 1956, p. 114)

But there lives on in the world concept and philosophy of the East a definite aversion to those pairs of opposites which dominate the history of Western thought – the living and the inanimate, mind and matter, soul and body, the conscious and the unconscious. And the absence of these antitheses is the expression of the closeness in which the other despite its otherness is held to the subject.[. . .] Being a form of the mind the subjective attitude is not restricted to any field of existence. It is active everywhere [. . .] Its most penetrating consequence is its denial to the external world of an existence in its own right. Such an objective existence, independent of the perception of the subject, seems to the Asiatic mind absurd and paradoxical – a monstrous idea. (Haas 1956, p. 116)

On the other hand, radical idealism with its theory of the illusory character of the external world is not innate in, let alone characteristic of, Asia. It belongs as do so many other extremist theories to the West. (Haas 1956, p. 117)

The East takes opposites as they are immediately experienced from life and nature. Thus while the principle of the contraries is recognized, they are left in their original context and to their original operation. For this reason they have been interpreted by contemplative thought as polarities whose contrariety, which is more apparent than real, conceals a deeper meaning of balance and harmony. (Haas 1956, p. 145)

The structures of the Eastern mind leads it to regard reality as somehow co-ordinated and akin to the subject. The innate tendency is to bridge and diminish the distance between reality and the subject. (Haas 1956, p. 130)

Ignoring the objectifying process of thought, the East dismisses the outer world as the all important object of speculation. Instead it stresses the necessary relationship of the Real with the subject.

This of course, radically changes the position and meaning of the term subject.[. . .] the subject, for the Easterner, is relative not to the object, that is not to an objective world as the essence of reality, but to an entity close to him [. . .]. (Haas 1956, p. 131)

The history of the Western mind in general and of its philosophical thought in particular is characterized by two great conceptions. Neither assumed a prominent role in the East. The one is mathematics – the other the idea of evolution.

Important as the invention and knowledge of mathematics are in China and particularly in India, the East cannot claim anything comparable to the systematic development of mathematical theory and applied mathematics in the West. (Haas 1956, p. 150)

In the history of Western philosophy the importance of mathematical thought on the one hand, and the idea of evolution on the other, is vividly illustrated by the fact that with few exceptions the most important philosophers can be grouped as those who were inspired by mathematical thinking and those who follow the line of the idea of evolution. The first group is represented by men like Pythagoras, Plato, Descartes, Leibniz and Kant while the other can claim the names of Heraclitus, Aristotle and Hegel.

If philosophical mysticism is credited with having vaulted this chasm, it did so at the price of ignoring the structure of the Western mind as well as its premises and aspirations. In this fashion it succeeded in avoiding the issue. It is no mere coincidence that the two outstanding representatives of this are to be found at the periphery of Western civilization – Plotinus, who was born in Egypt and inspired by the currents of Eastern thought, and Spinoza, of Jewish descent. (Haas 1956, p. 151f)

Eastern cognition is interested in consciousness itself. Western cognition is interested in the objects of consciousness. (Haas 1956, p. 167)

In the mind of the East the subject holds the predominant place occupied by the object in the West (Haas 1956, p. 180).

Reduced to its essence Eastern knowledge is a form of being, a state of consciousness, that is lucid and self-sufficient (Haas 1956, p. 182).

> Western knowledge is a form of having (Haas 1956, p. 182).
> At the height of his fulfillment, the man of the East stands utterly alone (Haas 1956, p. 232).

James P. Allen, an Egyptologist, expresses a similar view on the nature of Eastern cognition in (Allen 1988, p. ix):

> We have divorced philosophy, as a discipline, from religion. In the former we appreciate reality *objectively*, as something capable of study; in the latter we understand it *subjectively*, as something that can only be experienced. This dichotomy did not govern ancient Near Eastern thought. To it all appreciation of reality was subjective – "I-Thou" rather than "I-It."

One should not conclude from the above that there is no room in the Eastern mind for forms of realism. The *Nyāya* (one of the six orthodox schools of Hinduism) believed "in the external existence and mind-independence, not just of the ordinary objects of daily experience, but also of the universals or properties which reside in those objects. They combined this 'metaphysical realism' with a correspondingly realist or externalist approach towards cognitive and semantic content" (Ganeri 1996, p. 111f). This is precisely the kind of view repudiated by Brouwer and the Buddha. While several schools of Indian philosophy embrace *ananta* (infinity) – Patañjali recommends meditating on *ananta*, for the Jain "the theory of multiplexity of reality (*anekānta-vāda*) rests on [. . .] the conviction that things relate to each other by an infinite number of relations" (Balcerowicz 2001, p. 379) – both Brouwer and José Ortega y Gasset allow no place for actual infinity in mathematics, the latter, who also rejected the use of the principle of the excluded middle in mathematics, insisting that mathematicians should "focus only on objects that are 'immediately present' for the human intuition" (Rabi 2016, p. 65) and that "the infinite cannot be immediately present for the human intuition" (Rabi 2016, p. 68).

3.2 Fromm on to Have or to Be

Erich Fromm, a radical-humanistic psychoanalyst, has described in Fromm (1976) two basic character orientations, which he calls "the having mode of existence" and "the being mode of existence." The former is the dominant mode of existence in contemporary society. Our claim is that the spirit of Brouwer's philosophy – "the spirit out of which the work takes shape" ["der Geist, aus dem das Werk geschieht" (Eckhart 1934, p. 204)] – is that of the latter. Fromm cites from the works or sayings of Jesus, the Buddha, Meister Eckhart, Spinoza, and Albert Schweitzer, to which he refers as "Masters of Living," to circumscribe the being mode of existence.

Fromm was decidedly counter-modern, for he considered, much like Spinoza – for whom "mental health is, in the last analysis, a manifestation of right living; mental illness, a symptom of the failure to live according to the requirements of human nature" (Fromm 1976, p. 78) – that there exists a "human nature" and wrote about the "pathology of normalcy"; in other words, unlike standard psychiatry, he

found that being adapted to a society that he found insane and acting in its spirit is *not* a sign of sanity. Although he grew up in Frankfurt in the twentieth century, in reality he grew up, as he put it in interviews, in a "medieval atmosphere, in which everything is dedicated to traditional learning" (Friedman 2013, p. 4). Moreover:

> Between 1916 and 1921, Fromm actively studied topics linked to the Hebrew Bible in a small group, led by Rabbi Nehemia Nobel in Frankfurt, that included Martin Buber, Leo Baeck, and Gershom Scholem. (Friedman 2013, p. xxxiii)

Here are a few quotations from Fromm (1976) that refer to themes that occur in Brouwer's philosophy as well.

> Optimum knowledge in the being mode is *to know more deeply.* In the having mode it is *to have more knowledge*. (Fromm 1976, p. 34)
>
> Having refers to *things* and things are fixed and *describable.* Being refers to *experience*, and human experience is in principle not describable. What is fully describable is our *persona* – the mask we each wear, the ego we present – for this persona is in itself a thing. In contrast, the living human being is not a dead image and cannot be described like a thing. In fact, the living human being cannot be described at all. Indeed, much can be said about me, about my character, about my total orientation to life. This insightful knowledge can go very far in understanding and describing my own or another's psychical structure. But the total me, my whole individuality, my suchness that is as unique as my fingerprints are, can never be fully understood, not even by empathy, for no two human beings are entirely alike. Only in the process of mutual alive relatedness can the other and I overcome the barrier of separateness, inasmuch as we both participate in the dance of life. Yet our full identification of each other can never be achieved. (Fromm 1976, p. 71)
>
> *Language* is an important factor in fortifying the having orientation. The name of a person – and we all have names (and maybe numbers if the present-day trend toward depersonalization continues) – creates the illusion that he or she is a final, immortal being. The person and the name become equivalent; the name demonstrates that the person is a lasting, indestructible substance – and not a process. Common nouns have the same function: i.e., love, pride, hate, joy give the appearance of fixed substances, but such nouns have no reality and only obscure the insight that we are dealing with processes going on in a human being. But even nouns that are names of *things*, such as "table" or "lamp," are misleading. The words indicate that we are speaking of fixed substances, although things are nothing but a process of energy that causes certain sensations in our bodily system. But these sensations are not *perceptions* of specific things like table or lamp; these perceptions are the result of a cultural process of learning, a process that makes certain sensations assume the form of specific percepts. We naively believe that things like tables and lamps exist as such, and we fail to see that society teaches us to transform sensations into perceptions that permit us to manipulate the world around us in order to enable us to survive in a given culture. Once we have given such percepts a name, the name seems to guarantee the final and unchangeable reality of the percept. (Fromm 1976, p. 67)
>
> If my *self* is constituted by what I *have*, then I am immortal if the things I have are indestructible. (Fromm 1976, p. 67)

It should be noted that the having mode of existence is by no means restricted to the possession of objects. As part of the Four Noble Truths, *tṛ́ṣṇa* (Sanskrit for desire or thirst), a cause of *dukkha* (unsatisfactoriness, suffering), is not primarily directed to material things.

> It is important to keep in mind that for Buddhism, the problem is not desire for material things; more fundamentally, desire is a thirst for ideas, theories (*dṛṣṭi*), since it is these that give us a sense of identity and define the values we impute to material things, such that we then find them desirable. [. . .] Those ideas, theories *(dṛṣṭi)* are part of the play of a linguistic-cognitive web of closure (*prapañca, āsava, anuśaya, saṃvṛti,* etc.). (Lusthaus 2002, p. 61)

Similarly, for *upādāna* (appropriation, grasping, clinging, attachment), the result of *tṛ́ṣṇa*, we have:

> We appropriate in order to expand, strengthen and affirm our selves. Self-definition, self-identification arises through appropriation. What is other either becomes mine, or else it serves as the boundary marker that circumscribes my limits, and thus what I am. The 'more' I possess, the 'greater' I am. (Lusthaus 2002, p. 65)

4 Language

An important element of Brouwer's philosophy is his position with respect to language, a position that has changed little throughout his life. It is also a major source of misunderstanding by the "bland and vapid views that have dominated the philosophy of mathematics," for it has been read in an absolute, un-nuanced sense. It is that misunderstanding that brands Brouwer with the "solipsist" label. We will see that, even though some might actively dislike it (Posy's "bugaboo"), it is by no means an idiosyncratic idea that sprung out of Brouwer's head and is not encountered in other reasonable minds.

That his views on language are of high relevance to his philosophy *of mathematics* and to his rejection of formalism and of logicism can be best read from his own explanation of the "first act of intuitionism":

> The *first act of intuitionism* completely separates mathematics from mathematical language, in particular from the phenomena of language which are described by theoretical logic, and recognizes that intuitionist mathematics is an essentially languageless activity of the mind having its origin in the perception of *a move of time,* i.e. of the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained by memory. If the two-ity thus born is divested of all quality, there remains the *empty form of the common substratum of all two-ities.* It is this common substratum, this empty form, which is the *basic intuition of mathematics.* (Brouwer 1952b, p. 509f)

His views can be summarized as follows: (i) language originates in will-transmission, in commands given to others, and thus of questionable morality as its original aim is to "break the will of others," making others obey; (ii) it is not, on its own, a means of communication, but in rare instances, it could become an aide in a communication that has been established by non-verbal means; (iii) language is very inadequate at expressing the inner life of a human being; (iv) language does not have the miraculous power of conjuring a life-like world, and it does not have a descriptive function in which it has shaken off the original sin of its birth as will-transmission; (v) even in the very restricted field of mathematics, language cannot

express faithfully the creations of a human subject, as these are living processes. Of these, (i), (iv), and (v) are insights that cannot be found expressed in such stark terms by other thinkers who find language inadequate. There is one mathematician who, after Brouwer, made a (v)-like statement. René Thom, another mathematician-philosopher, coming from an entirely different philosophical perspective, namely, that of mathematical Platonism, wrote in Thom (1970) (see also Papadopoulos 2018) that

> the world of ideas infinitely exceeds our "technical possibilities." It is in the intuition that the *ultima ratio* of our faith in the truth of a theorem resides. And, according to a now-forgotten etymology, a theorem is above all the object of a vision. (Thom 1970, p. 697)
>
> [Car le monde des Idées excède infiniment nos possibilités opératoires, et c'est dans l'intuition que réside *l'ultima ratio* (la raison dernière) de notre foi en la vérité d'un théorème – un théorème étant, selon une étymologie aujourd'hui bien oubliée, l'objet d'une vision.]

Already in Brouwer's earliest writing of 1905, the pamphlet *Life, Art, and Mysticism*, we read:

> The immediate companion of the intellect is language. From life in the intellect follows the impossibility of any form of direct communication with others – instinctively by gesture or looks, or even more spiritually through all separation of distance. People therefore start training themselves and their offspring in some crude sign language, painfully and with little success, for never has anyone been able to communicate with others, soul to soul. Language can only be the accompaniment of an already existing mutual understanding.
>
> [. . .]
>
> Only in the very narrowly restricted domains of the imagination such as in the exclusively intellectual sciences – which are completely separated from the world of perception and therefore touch the least upon the essentially human – only there can mutual understanding be maintained for some time. There is little scope for misunderstanding notions such as "equal" and "triangle," but even then two different people will never feel them in exactly the same way. Even in the case of the most restricted sciences, logic and mathematics – a sharp distinction between these two is hardly possible – no two different people will have the same conception of the fundamental concepts on which these two sciences are constructed; and yet their wills are parallel, and in both there is a small, unimportant part of the brain which forces their attention in the same way. (Brouwer 1996, p. 401)
>
> But ridiculous is the use of language when one tries to express subtle nuances of will which are not part of the living reality of those concerned, when for example so-called philosophers or metaphysicians discuss among themselves morality, God, consciousness, immortality, or the free will. These people do not even love each other, let alone share the same subtle movements of the soul; sometimes they even do not know each other personally. They either talk at cross-purposes or each builds his own little logical system which lacks any connection with reality. For logic is life in the human brain; it may accompany life outside the brain but it can never guide it by virtue of its own power. (Brouwer 1996, p. 401f)
>
> Words are no more than commando-signals [. . .] means of training and controlling
>
> All verbal utterances are more-or-less developed verbal imperatives, i.e. speaking can always be reduced to commands or threats, and understanding to obeying. (Van Stigt 1990, p. 197, unpublished papers)
>
> Language can accompany man's will to dominate the will of others or his will to keep the movements of wills together; for example, the war cry of Red Indians accompanies the will to break the will of others.

> Language by itself has no meaning; any philosophy which searched for a firm foundation based on that presumption has come to grief; lulled into sleep by the assurance of such firm foundation, one was rudely awakened by the appearance later of deficiencies and contradictions. A language which does not derive its certainty from the human will but which claims to live on in the 'pure concept' is an absurdity. (Brouwer 1996, p. 403)
>
> The practical reliability of the logical principles relies on the fact that a significant part of the world of intuition shows much more fidelity and satisfaction with respect to its finite organization than humanity itself. The fact that one has been, from time immemorial, blind to this sober interpretation, has been caused by the fact that the exclusive character of words as a means of will-transmission had not been recognized, and these were seen as a result of a rash superstition as suggestive means of fetish-like 'concepts'. These 'concepts' as well as the links existing between them, are supposed to have an existence independently of the causal stance of man, and the logical principles should represent the concepts and the a priori laws governing their connections. [Trans. V P.]
>
> [Die praktische Zuverlässigkeit der logischen Prinzipien beruht darauf, dass ein großer Teil der Anschauungswelt in bezug auf ihre endliche Organisation viel mehr Treue und Zufriedenheit zeigt als die Menschheit selbst. Dass man von alters her vor dieser nüchternen Interpretation blind war, wurde dadurch verursacht, dass man den ausschließlichen Charakter der Worte als Willensübertragungsmittel nicht erkannte und dieselben infolge eines unbesonnenen Aberglaubens als Andeutungsmittel fetischartiger 'Begriffe' betrachtete. Diese 'Begriffe' sowie die zwischen ihnen bestehenden Verknüpfungen sollten unabhängig von der kausalen Einstellung des Menschen eine Existenz besitzen, und die logischen Prinzipien sollten die Begriffe und ihre Verknüpfungen beherrschende aprioristische Gesetze darstellen. (Brouwer 1929, p. 423)]
>
> In default of a plurality of mind, there is no exchange of thought either. Thoughts are inseparably bound up with the subject. So-called communicating-of-thoughts to somebody, means influencing his actions. Agreeing with somebody, means being contended with his cooperative acts or having entered into an alliance. [. . .] By so-called exchange of thought with another being the subject only touches the outer wall of an automaton. This can hardly be called mutual understanding. Only through the sensation of the other's soul sometimes a deeper approach is experienced. And when wisdom revealed by the beauty of this sensation, finds expression in the antiphony of words exchanged, then there may be mutual understanding. (Brouwer 1948, p. 485)

These quotes encourage analytic philosophers to heap the usual verbal abuse, such as "solipsist," on Brouwer. This is a thorough misunderstanding of what Brouwer is writing *against.* What he writes against is a language that is viewed as possessing self-existence, as having content independently of the human mind, and as capable of transmitting that content. The rather mainstream Western view that language does have these miraculous powers is expressed by Franz Brentano, Gottlob Frege, Nicolai Hartmann, and Karl Popper.

In Hartmann (1964), Hartmann assigns to the spiritual (*das Geistige*) its own "real stratum" in the composition of the world, one which cannot be reduced to subjective consciousness. Subjective is only the carrying out of the act of thinking, the content of the thought itself is the spiritual content that is detachable from the subject and transferable, a fact without which communication between humans would not be possible. And Popper, whose engagement with Brouwer's thought is surveyed in (Naraniecki 2015), finds that

> we may distinguish the following three worlds or universes: first, the world of physical objects or physical states; secondly the world of states of consciousness or of mental states, or perhaps of behavioural dispositions to act; and thirdly the world of *objective contents of thought*, especially of scientific and poetic thoughts and of works of art. (Popper 1967, p. 106)

Popper finds that his "third world" has "more in common still with Bolzano's theory of a universe of propositions in themselves and of truths in themselves" and "resembles most closely the universe of Frege's objective contents of thought" (Popper 1967, p. 106). He argues for "the independent existence of the third world" (Popper 1967, p. 107). "Theories or propositions or statements are the most important third-world linguistic entities" (Popper 1968, p. 157). "The activity of understanding consists, essentially, in operating with third-world objects" (Popper 1968, p. 164).

The Buddha finds that such a separation of the contents of the second and third worlds is impossible:

> Feeling, perception, and consciousness, friend – these states are conjoined, not disjoined, and it is impossible to separate each of these states from the other to describe the difference between them. For what one feels, that one perceives, and what one perceives, that one cognizes.
>
> Wisdom and consciousness, friend – these states are conjoined, not disjoined, and it is impossible to separate each of these states from the other to describe the difference between them. For what one wisely understands, that one cognizes, and what one cognizes, that one wisely understands.
>
> [. . .] wisdom is to be developed, consciousness is to be fully understood. (Nanamoli and Bodhi 1995, Mahāvedalla Sutta, pp. 388–389 (PTS M i 292–293)]

While the Buddha mentions a "theory" that does not depend on human discovery, namely, the doctrine of *dependent arising* (or *dependent origination*) (see Kalupahana 1975), its content, when clothed in language, is subject to the vicissitudes that visit any language-composed doctrine practiced by humans. It is due to precisely the historical destiny of all such doctrines that he prophesized the demise of the *dharma* in India (see Lamotte (1976, p. 210f) for more details).

> And what, bhikkhus, is dependent origination? "With birth as condition, aging-and-death [comes to be]": whether there is an arising of Tathāgatas or no arising of Tathāgatas, that element still persists, the stableness of the Dhamma, the fixed course of the Dhamma, specific conditionality. A Tathāgata awakens to this and breaks through to it. (Boddhi 2000, II.25, Nidānasaṃyutta, p. 551)

As pointed out by David Kalupahana, language, as understood by the Buddha, does not work on its own, in the absence of what Brouwer calls "an already existing mutual understanding."

> Language cannot be a means of communication unless the language users agree upon the various elements of a language as being expressive of their experiences. Thus, agreement or concurrence *(sammuti)*, whether it be of a small group or of the world *(loka)* at large, is an

> important characteristic of language. The use of the term *sammuti*, which implies agreement in the sense of "thinking together," makes language not merely a behavioural phenomenon but a psychological one as well. (Kalupahana 1999, p. 52)

The "propositions in themselves," the "truths in themselves," the "objective contents of thought," and, in general, "third-world objects" claim a *svabhāva* existence (*self-nature* in Kalupahana's translation, *own being* in Robinson's translation), which is not only found to be contradictory in Chap. XV (Examination of Self-Nature) of Nāgārjuna's *Mūlamadhyamakakārikā* (for the extent to which mādhyamika can be seen as a critique of language (see Napper (1989, pp. 75–100)):

> *Svabhāva* is by definition the subject of contradictory ascriptions. If it exists, it must belong to an existent entity, which means that it must be conditioned, dependent on other entities, and possessed of causes. But a *svabhāva* is by definition unconditioned, not dependent on other entities, and not caused. Thus the existence of a *svabhāva* is impossible. (Robinson 1957, p. 299)

but is also considered by the Buddha and his followers to be the origin of all mental confusion:

> The root of all the troubles of cyclic existence and solitary peace is the ignorance conceiving true existence, as well as its predispositions. (Gyatso 1990, p. 78)

If I think in terms of "contents of thought," this objectifying process allows ownership of that "content." If thoughts are fleeting by their very nature, are inseparably bound up with the subject, and disappear with the subject, then they are of no value in the having mode of existence.

Similar thoughts on the inadequacy of language were expressed not just by Erich Fromm but also by a wide variety of twentieth-century thinkers. Here is what the founder of "general semantics" had to say on the matter:

> Words *are not* the things they represent. (7)
>
> The above [. . .] considerations lead to unexpectedly far-reaching consequences.
>
> From (7) – it follows that the objective levels which include the events, ordinary objects, objective actions, processes, immediate feelings, 'instincts', 'ideas', [. . .] represent un-speakable levels, *are not words*. (9)
>
> From (9) – that the use of the 'is' of *identity*, as applied to objective, un-speakable levels, appears invariably structurally false to facts and must be entirely abandoned. Whatever we might say a happening 'is', it *is not;* (10)
>
> From (10) – *structure* appears as the only possible link between the objective, unspeakable, and the verbal levels. (11)
>
> From (11) – the only possible 'content of knowledge' becomes exclusively *structural.* (12)
>
> From (12) – the only aim of 'knowledge' and science appears as the empirical search for, and verbal formulation of, structure.
>
> The only method for acquiring 'knowledge' is found in an empirical investigation of the potentially unknown structure of the world, ourselves included, only afterwards adjusting the structure of languages so that they would be similar, and so of maximum usefulness;

> instead of the delusional reversed order of ascribing to the world the structure of an inherited primitive language. (Korzybski 1931, p. 751f)

Although Krishnamurti does not explicitly refer to the will-transmission character of language, it is implicit when he admonishes his audiences repeatedly that the speaker, although on a podium, tries to figure out the answer to a question *with his audience*, that all following of someone else's advice on matters of wisdom, life, love, ethics, or meditation is counter-productive, for only oneself can discover for oneself the truth of those, that all else is "conditioning." "There is no right *guru*. There is only wrong *guru*, because nobody can teach you anything except for yourself. [. . .] nobody can teach you what you are" (Krishnamurti and Bohm 1983). He also sees the will-transmission component when language is not even used between two people but rather in a monologue, as in the traditional approaches to "meditation":

> Who is meditating? In most systems of meditation, Japanese, Indian, Tibetan, and so on, there is the controller and the controlled. The controller tries to control thought, quieten thought, shape it according to a purposeful direction.
>
> [. . .]
>
> The speaker has gone into this for the last sixty years or more. He has discussed this question with Zen pundits, with Hindus, Tibetans and all the rest of the gang, and he refutes that kind of meditation because their idea of meditation is to achieve an end. The end being complete control of the brain so that there is no movement of thought. Because when the brain is still, deliberately disciplined, deliberately sought after, it is not silent. It is like achieving something, which is the action of desire. (Krishnamurti 1984)

Krishnamurti is also very clear about the inadequacy and the limitations of language:

> Understanding is not verbal, nor is there such a thing as intellectual understanding. Intellectual understanding is only on the verbal level, and so no understanding at all. Understanding does not come as a result of thought, for thought after all is verbal. (Krishnamurti 1956, p. 194)
>
> Sir, love is not a word; the word is not the thing, is it? God is not the word "god," love is not the word "love." But you are satisfied with the word, because the word gives you a sensation. When somebody says "God," you are psychologically or nervously affected, and that response you call the understanding of God. So, the word affects you nervously and sensuously, and that produces certain action. But the word is not the thing, the word "god" is not God; you have merely been fed on words, on nervous, sensuous responses. Please see the significance of this. How can you act if you have been fed on empty words? For words are empty, are they not? They can only produce a nervous response, but that is not action. Action can take place only when there is no imitative response, which means the mind must enquire into the whole process of verbal life. (Krishnamurti 1991, Bangalore fourth public talk July 25, 1948)
>
> Can we say the word is not the thing? Whatever the description, it is not the real, not the truth, however much you embellish or diminish it. We recognize that the word is not that, then what is there beyond all this? Can my mind be so desireless that it won't create an illusion, something beyond? (Krishnamurti and Bohm 1999, p. 121)
>
> Naming only strengthens and gives continuity to the experiencer, to the desire for permanency, to the characteristic of particularizing memory. There must be silent awareness

> of naming, and so the understanding of it. We name not only to communicate, but also to give continuity and substance to an experience, to revive it and to repeat its sensations. This naming process must cease, not only on the superficial levels of the mind, but throughout its entire structure. This is an arduous task, not to be easily understood or lightly experienced; for our whole consciousness is a process of naming or terming experience, and then storing or recording it. It is this process that gives nourishment and strength to the illusory entity, the experiencer as distinct and separate from the experience. Without thoughts there is no thinker. Thoughts create the thinker, who isolates himself to give himself permanency; for thoughts are always impermanent. (Krishnamurti 1956, p. 69)

Here is what Albert Schweitzer wrote on the possibility of understanding one another:

> None of us can claim that we really know another person, even if we have lived together with that person every day for years. Of what makes up our inner experience, we cannot communicate even to our closest others more than fragments. We can neither give the entirety of that experience nor could the others grasp it. We walk in a half darkness, in which neither can recognize the other's features accurately. Only from time to time, through an experience we had with our companion, or through a word that falls between us, does the other stand next to us for a moment, as if illuminated by a lightning bolt. That's when we see him as he is. Afterwards we'll go again, perhaps for a long time, side by side in the dark and try in vain to imagine the other's features. (Trans. V. P.)
>
> [Keiner von uns darf behaupten, daß er einen andern wirklich kenne und wenn er seit Jahren täglich mit ihm zusammen lebt. Von dem, was unser inneres Erleben ausmacht, können wir auch unseren Vertrautesten nur Bruchstücke mitteilen. Das Ganze vermögen wir weder von uns zu geben, noch wären sie imstande, es zu fassen. Wir wandeln in einem Halbdunkel, in dem keiner die Züge des anderen genau erkennen kann. Nur von Zeit zu Zeit, durch ein Erlebnis, das wir mit dem Weggenossen haben, oder durch ein Wort, das zwischen uns fällt, steht er für einen Augenblick neben uns, wie von einem Blitze beleuchtet. Da sehen wir ihn, wie er ist. Nachher gehen wir wieder, vielleicht für lange, im Dunkel nebeneinander her und suchen vergeblich, uns die Züge des andern vorzustellen. (Schweitzer 1974, p. 385f)]

And here is what two neurobiologists wrote about communication by means of language (one of them, Francisco Varela, was a co-founder of the *Mind and Life Institute* to promote dialog between science and Buddhism, so guilty by association of accepting the validity of the *Eastern* approach):

> So long as language is considered to be denotative it will be necessary to look at it as a means for the transmission of information, as if something were transmitted from organism to organism, in a manner such that the domain of uncertainties of the 'receiver' should be reduced according to the specifications of the 'sender'. However, when it is recognized that language is connotative and not denotative, and that its function is to orient the orientee within his cognitive domain without regard for the cognitive domain of the orienter, it becomes apparent that there is no transmission of information through language. It behooves the orientee, as a result of an independent internal operation upon his own state, to choose where to orient his cognitive domain; the choice is caused by the 'message', but the orientation thus produced is independent of what the 'message' represents for the orienter. In a strict sense then, there is no transfer of thought from the speaker to his interlocutor; the listener creates information by reducing his uncertainty through his interactions in his cognitive domain. Consensus arises only through cooperative interactions in which the

> resulting behavior of each organism becomes subservient to the maintenance of both. (Maturana and Varela 1980, p. 32)

And here is a quote from Henri Poincaré about language, quoted at the end of Brouwer (1913):

> People do not understand each other because they do not speak the same language and because there are languages that cannot be learned. (Trans. V. P.)
>
> [Les hommes ne s'entendent pas, parce qu'ils ne parlent pas la même langue et qu'il y a des langues qui ne s'apprennent pas. (Poincaré 1913, p. 161)]

All of Brouwer's utterances on language are taken literally and absolutely by Brouwer's detractors, who deliberately refuse to take notice of the fact that the same person who wrote dismissively about the capabilities of language to provide *exact* communication by means of "objective contents of thought" and wrote about its inherent limitations, was involved in the significs project and was a master of a highly evocative language, as related in a recollection of Mannoury's daughter.

> Mrs. Vuysje, the daughter of Mannoury, described Brouwer's conversation as an almost artistic process: Brouwer could in the middle of a sentence pause to savour a particular expression, and to replace it after consideration by a more suitable one. She remembered him, standing in the garden stretching out his hand, tasting a particular word and contemplating fitting equivalents. The result was not a faltering flow of words; Brouwer had mastered the technique of refinement to such a degree that he managed to produce gentle, flowing sentences of incredible length without discernible interruptions. (Van Dalen 2013, p. 263)

Not even in this area is Brouwer given his due. Reviewing Van Stigt (1990), Smoryński mentions that he could not make any sense of Van Stigt's remark that "Brouwer lacked the ability to express his feelings artistically," for it "certainly runs counter to what every Dutchman I've spoken to has said of Brouwer's correspondence with Adama van Scheltema: Brouwer's prose is uniformly favourably compared with that of the poet" (Smoryński 1994, p. 802). There are, however, mathematicians and philosophers who did not fall into the trap of concluding "solipsism" or "bugaboo" from Brouwer's writings on language. In addition to the works already mentioned by Largeault and Detlefsen, Van Dalen (1999a), Placek (1999), and Kaneko (2002a) treat this aspect of Brouwer's philosophy in earnest, and none come to the negative conclusions mentioned in the Introduction. Placek (1999) disproves that Brouwer was a solipsist and shows that intersubjectivity is possible in Brouwer's philosophy.

There is one more restricted use Brouwer finds for language; it is that of a memory aide.

> Even if language in its origin and in the first place is a function of the activity of social man it plays a significant role also in the reflective thinking and mnemotechnical processes of the solitude of singular man. (Brouwer 1933; Van Stigt 1990, p. 422f)

We can say that, while Brouwer realizes that, unfortunately, language is not a means for mutual understanding, is endowed with an original sin that renders its use morally questionable, we have no choice but to try the impossible. As Verena Huber-Dyson put it:

> Formalization and Poetry are the two most direct means we have of externalizing our thoughts. (Huber-Dyson 1996, p. 67)

Brouwer chose poetry, for at least it has no pretension to convey anything *exactly*, and opts for hinting at a state of mind, a feeling, or an atmosphere. This he found, in his own experience as a student at the University of Amsterdam, listening to Gerrit Mannoury, more helpful than the precision of strings of symbols. Brouwer finds that one is, by the very nature of language, in the Sisyphean, absurd hero position of making, while being thrown upon the use of language, a desperate use of a tool ill-suited for the task, much like a character in Franz Kafka's world, as seen by Albert Camus:

> Kafka's world is in truth an indescribable universe in which man allows himself the tormenting luxury of fishing in a bathtub, knowing that nothing will come of it. (Camus 1955)
>
> [Le monde de Kafka est à la vérité un univers indicible où l'homme se donne le luxe torturant de pêcher dans une baignoire, sachant qu'il n'en sortira rien.]

5 The Eastern Origin of Subjective Idealism

Why is subjective idealism so scandalous in Western philosophy concerned with mathematics? After all, Schopenhauer made statements that put him squarely in the subjective idealist, decidedly *Eastern*, camp, such as

> On the other hand, the Kantian teaching, even without the antinomies, leads to the insight that things and their whole mode and manner of existence are inseparably associated with our consciousness of them. Therefore he who has clearly grasped this soon reaches the conviction that the assumption that things exist as such, even outside and independently of our consciousness, is really absurd. [Trans. E. F. J. Payne]
>
> [Hingegen leitet die Kantische Lehre, auch ohne die Antinomien, zu der Einsicht, daß die Dinge und die ganze Art und Weise ihres Daseyns mit unserm Bewußtseyn von ihnen unzertrennlich verknüpft sind; daher wer Dies deutlich begriffen hat, bald zu der Ueberzeugung gelangt, daß die Annahme, die Dinge existierten als solche auch außerhalb unseres Bewußtseyns und unabhängig davon, wirklich absurd ist. (Schopenhauer 1859, Erstes Buch, Kap. 1, p. 708)]

and this did not prevent him from being influential for a while, nor did Emerson's embrace of subjective idealism (Lyttle 1997) harm his position in American letters and philosophy.

That someone presenting general advice on life, ethics, or, in general, what usually goes for wisdom, embraces subjective idealism presents no threat to the

main pride of Western philosophy: mathematics and evolution. What is scandalous and asks for the entire arsenal of derision, belittling, and name-calling in Brouwer's case is the proposal as the only sound basis for mathematics, a jewel of the Western philosophical approach, of an unmistakably Eastern foundation.

As was observed by Aram Frenkian in 1946, there is not a hint of subjective idealism in any Greek philosophical work that precedes the contact with the East brought about by Alexander. He detected in Frenkian (1946) the first text suggestive of subjective idealism in the Memphite Theology, as inscribed on the Shabaka Stone (see Allen (1988), Hoffmeier (1983), and Finnestad (1976) for a detailed study of its contents), which dates from the 25th dynasty of Egypt. There we are told that "Great and important is Ptah,// who gave life to all the gods and their *kas* as well// through this heart and this tongue" (Allen 1988, p. 43). Similar utterances can be found in earlier previous Egyptian texts, such as the Coffin Texts (Allen 1988, p. 45) (late 18th or early 19th dynasty). The conclusions Allen draws in (Allen 1988, pp. 56–61) regarding the subjective idealist nature of Egyptian thought and cosmology, a "construct of opposites in balance" (Allen 1988, p. 57), without citing Frenkian (1946) or Haas (1956), can be taken to be an added, independent confirmation of Frenkian's thesis.

The appearance of something akin to subjective idealism, in various *mind-only* schools of Mahāyāna Buddhism, has been the subject of several studies, such as Guenther (1989), Suzuki (1998), and Lusthaus (2002). Perhaps the most telling line is one by Nāgārjuna, who states:

> The [material world], in the final analysis, cannot exist. (Inagaki 1998, Chap. 33, p. 124)

The first appearance of a "vacillation" between "the ontic-ontological aspect" and "its mentalist-idealist aspect" (McEvilley 1980, p. 181) in the world of Greek letters can be detected in the philosophy of Plotinus, the first Greek philosopher who is seen similar to, on the one hand, an Indian current of thought of the subjective idealism variety, Vijñānavāda Buddhism, and, on the other hand, to an Indian version of objective idealism, the Upanishadic-Vedāntic philosophy, in the opus magnum comparing Greek and Indian thought (McEvilley 2002, Chaps. 22 and 23) (cf. also Sabo 2017). The dearth of reports of Greek philosophers traveling East before Alexander's conquests, mentioned in Männlein-Robert (2009), indicates that Greek philosophers of pre-Hellenistic times were unaware of the very existence of subjective idealist thought.

Subjective idealism is very rare in the history of Western philosophy. Although Berkeley was labeled by Kant and Hegel as a subjective idealist, that designation is, given the role played by God in his philosophy, certainly not accurate, as pointed out in Callahan (2015).

The varieties of idealism favored in Western philosophy tend to leave the physical world in its place. Husserl emphasized in his *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie* the *European* roots of his transcendental phenomenology not just because the year in which the lectures were held was 1936 and he had been branded non-Aryan but, as pointed out in

(Majer 2010), because the roots of his philosophy are unmistakably Western. In fact, there is a version of Husserl minus idealism, developed by one of his students, Ingarden (1975), which is viable. In that sense, the philosophy favored for comparisons with Brouwer escapes the anathema of an Eastern origin.

6 Brouwer's View on Mathematics

As we have seen, the first act of intuitionism distinguishes mathematics from the language of mathematics; establishes the source of mathematics, which is an intuition, the intuition of time; and establishes the nature of the mathematical activity, which is a "languageless activity of the mind." This is Brouwer's version of Fromm's "Being refers to *experience*, and human experience is in principle not describable."

The second act of intuitionism defines the intuitionistic continuum in terms of choice sequences, which are freely chosen sequences of rational numbers.

We will first look at the wide-ranging consequences of the first act. The first part, the statement that mathematics is not to be confused with the language of mathematics, might seem the most shocking. Brouwer provides the following comparison to explain why the two are distinct: "the formal language which accompanies mathematics as the weather-map accompanies the atmospheric processes" (Brouwer 1937, p. 451). We have seen that René Thom expressed similar views on the inability of our *possibilités opératoires* of expressing the world of our ideas in Thom (1970, p. 697). Needless to say, the support comes from a counter-modern, who looked to Aristotle for guidance, followed none of the intellectual fashions of his time, and asked his daughter, then just a child, excited about the events of May 1968: "Mais où prends-tu qu'après une révolution les choses iront mieux?" (Papadopoulos 2018) ["But where did you get it that after a revolution things will get better?" (Trans. V. P.)]

Brouwer is very explicit on the difference between mathematics and the language of mathematics, both in the sense that there is language-specific rubbish that he does not count as mathematics and that there exists ineffable mathematics, making its communication impossible.

> [. . .] I do not recognize as true, hence as mathematics, everything that can be written down in symbols according to certain rules, and conversely I can conceive mathematical truth which can never be fixed down in any system of formulas. (Brouwer 1937, p. 452)
>
> Thus for a human mind equipped with an unlimited memory, pure mathematics, practised in solitude and without using linguistic signs, would be exact, but the exactness would be lost in mathematical communication between human beings with an unlimited memory, because they would still be thrown upon language as their means of understanding. (Brouwer 1933, p. 443)

Although we have seen that other thinkers deny language the ability to express certain processes, it is mostly our inner lives that are deemed ineffable by Fromm or Schweitzer. On the other hand, even something that has been clothed in language cannot be decoded by a naive reading, as pointed out by Dilthey, who asks for

special faculties, such as *Verstehen* (understanding) for the very comprehension of literary textual material, for which the reader needs the abilities of *Hineinversetzen*, *Nachbilden*, *Nacherleben* (Dilthey 1965, pp. 213–216) [putting oneself in somebody else's position, replicating, reliving]. In other words, texts do not communicate by means of a simple transmission of sequences of words. Yet one would perhaps think that mathematics is far removed from these texts belonging to the humanities, that comprehension is straightforward, requires no *Verstehen* in Dilthey's sense, because mathematics is not about human life. If, however, mathematics is a "languageless activity of the mind," then it *is* subject to the same communication difficulties a literary text encounters. Things are further complicated by the fact that

> Mutual understanding between two people is a matter of degree, but in mathematics their mutual understanding is either total or it does not exist, like being-asleep, one either is or is not. (Van Stigt 1990, p. 208, unpublished manuscript]

On an empirical level, anyone who has tried to teach mathematics has been confronted with this issue, as have university administrators who have thought that recording the teaching of a course by a luminary would enable tens of thousands to learn the content of that course. The same holds for the purported proof of the *abc*-conjecture or for the highly complex mathematics of our time, in which certain subjects are understood by no more than a dozen people, and it is practically impossible to learn those subjects from texts. One needs to be present in the centers where there are experts in them. This is a version of Polanyi's "we can know more than we can tell" (Polanyi 2009, p. 4).

When Brouwer refers to "everything that can be written down in symbols according to certain rules," he means what he called in Brouwer (1913) *formalism.* The name refers to Hilbert's understanding of the foundations of mathematics, although Hilbert never used the term. Hilbert understood, for the purposes of proof theory, a mathematical theory as consisting of a set of axioms expressed in a formal language, together with a set of rules of inference, which allow the deduction of new sentences from the axioms. All sentences are just strings of meaningless symbols, satisfying certain syntactic rules. The name "formalism" turned out to be a very popular one and has been in use by everyone except Hilbert and members of his inner circle up to the present day. Hilbert would say that he does not mean that mathematics is concerned with deriving meaningless strings of symbols from other meaningless strings of symbols, that this is just a tool for reaching certain meta-mathematical goals. However, formalism is, nowadays, the official foundation of mathematics. The other competing foundation, logicism, which claimed that mathematics is a part of logic, has faded away, as it was not able to account for modern algebra or even for geometry, as noticed by Bertrand Russell himself, a major proponent of logicism (Gandon 2012; Musgrave 1977).

Now, even this highly controlled and unambiguous language is, in Brouwer's view, fraught with the shortcomings of any language.

> There is neither exactness nor certainty for will-transmission, especially for one mediated through language. And this state of affairs remains undiminished if will-transmission refers to the construction of purely mathematical systems. *So even for pure mathematics there is no certain language*, i.e., no language that excludes misunderstandings in conversation and protects from mistakes in the act of memory support (i.e. from the confusion of different mathematical entities). This circumstance cannot be remedied, in the manner done by *the formalistic school*, that one subjects the mathematical language (i.e. the system of signs that is used to evoke purely mathematical constructions in other people) itself to a mathematical examination, confers it through reworking the accuracy and stability of a material instrument or a phenomenon of exact science and in so doing communicates in a second-order language or metalanguage about it. (Trans V. P.)
>
> [Nun gibt es aber für Willensübertragung, insbesondere für durch die Sprache vermittelte Willensübertragung, weder Exaktheit, noch Sicherheit. Und diese Sachlage bleibt ungeschmälert bestehen, wenn die Willensübertragung sich auf die Konstruktion reinmathematischer Systeme bezieht. *Es gibt also auch für die reine Mathematik keine sichere Sprache*, d. h. keine Sprache, welche in der Unterhaltung Mißverstandnisse ausschließt und bei der Gedächtnisunterstützung vor Fehlern (d. h. vor Verwechslungen verschiedener mathematischer Entitäten) schützt. Diesem Umstande ist nicht dadurch abzuhelfen, daß man, wie es die *formalistische Schule* macht, die mathematische Sprache (d. h. das zur Hervorrufung reinmathematischer Konstruktionen bei anderen Menschen dienende Zeichensystem) selber einer mathematischen Betrachtung unterzieht, ihr durch Umarbeitung die Genauigkeit und Stabilität eines materiellen Instrumentes oder eines Phänomens der exakten Wissenschaft verleiht und sich dabei in einer Sprache zweiter Ordnung oder Übersprache über sie verständigt. (Brouwer 1929, p. 421f)

Now, what could be the ineffable mathematical truths Brouwer referred to? I thought it was some truth belonging to geometric topology, given that I always felt that the idea that geometric topology could be formalized is a pious wish. The fact that Thom, another geometric topologist, stated the same belief in ineffable mathematical truths seems to provide additional support. Van Dalen (Van Dalen 1999a) thinks that choice sequences, which are unfinished objects, were what Brouwer had in mind.

Also part of the first act is the realization that mathematics originates in intuition, in particular in a *single* intuition, that of time, "a single aprioristic basic intuition, which may be called invariance in change as well as unity in multitude" (Brouwer 1907, p. 97). That arithmetic originates in [the intuition of] time was stated by Meister Eckhart as well.

> The one is eternity, which always holds itself alone and is without change. However, the two, that is time, which changes and multiplies itself. (Trans. V. P.)
>
> [Das Eine ist die Ewigkeit, die sich allezeit allein hält und wandellos ist. Die Zwei aber, das ist die Zeit, die sich wandelt und vermannigfaltigt. (Eckhart 1934, p. 240)]

It may appear at first sight, perhaps even after reading Brouwer (1909), that Brouwer dropped the intuition of space, present in Kant's philosophy, given the discovery of non-Euclidean geometries and the untenability of an intuition of Euclidean space (there would still be a possibility to assume the existence of an intuition of absolute geometry, the common part of Euclidean and hyperbolic space). In his thesis, even though there is a unique "basic intuition of mathematics," there is

an acknowledgment of a distinction between the discrete and the continuous intuition of time which does not occur in his later writings:

> Since continuity and discreteness occur as inseparable complements, both having equal rights and being equally clear, it is impossible to avoid one of them as a primitive entity, trying to construe it from the other one, the latter being put forward as self-sufficient; in fact it is impossible to consider it as self-sufficient. Having recognised that the intuition of continuity, of 'fluidity', is as primitive as that of several things conceived as forming together a unit, he latter being at the basis of every mathematical construction, we are able to state properties of the continuum as a 'matrix of points to be thought of as a whole'. (Brouwer 1907, p. 17)

To justify the untenability of an intuition of space, Brouwer brings arguments taken from "mechanics," such as a reference to the "relativity postulate" due to Lorentz (Brouwer 1909, p. 115) (he was probably unaware of Einstein's work at that time) or the impossibility of separating time and space, against both Frege's and Russell's attempted defense of Euclidean geometry in the form that "our faculty of cognition can assimilate the world of experience only in the form of Euclidean geometry" (Brouwer 1909, p. 109). This is surprising, given that time and space are supposed to be those *of intuition*, not those of physics. He mentions this explicitly for time: "Of course we mean here *intuitive time* which must be distinguished from *scientific time*" (Brouwer 1907, p. 61).

It is apparent that he favored a view of geometry that is only about the continuum, and the continuum can be developed from the intuition of time. As he put it in Brouwer (1913, p. 128): "For since Descartes we have learned to reduce all these geometries to arithmetic by means of the calculus of coordinates."

Brouwer's insistence on a single basic intuition might also have been motivated by an aversion, felt by many of those whom Fromm calls Masters of Living, for dualism. We know he was not just in awe of the writings of Meister Eckhart and Jakob Böhme, of the Bhagavad-Gita (all of which he mentioned already in 1905 in Brouwer (1996)), but also of Spinoza, for he wrote, as a case in point for his "logic is not a reliable instrument to discover truths and cannot deduce truths which would not be accessible in another way as well" (Brouwer 1948, p. 488) that "we feel the wisdom that blooms in Spinoza's work as completely independent of his logical system" (Brouwer 1908, p. 107). We think it is the same "definite aversion to those pairs of opposites" found by Haas in "the world concept and philosophy of the East," that Nina Berberova confessed to, that made the mystically inclined Brouwer avoid treating time and space as generating distinct intuitions.

> All dualism is painful for me, all splitting or bisecting contrary to my nature. When Lenin speaks about matter in opposition to energy, when Berdyaev speaks about material principle (reaction) and spiritual principle (revolution), when idealist philosophers speak about spirit and flesh, I am jarred, as by a false note. My commandment was the truth that matter is energy. My whole life has been the reconciliation within myself of the old dichotomy. (Berberova 1992, pp. 23–24)

This monistic view is not generally accepted. It is clear that the intuition of time allows the creation of natural numbers and thus a foundation for number theory. While, in Atiyah's view, modern algebra is also dependent on time, he finds geometry to be distinctly *not* related to time.

> Algebra, on the other hand (and you may not have thought about it like this), is concerned essentially with time. Whatever kind of algebra you are doing, a sequence of operations is performed one after the other and 'one after the other' means you have got to have time. In a static universe you cannot imagine algebra, but geometry is essentially static. I can just sit here and see, and nothing may change, but I can still see. Algebra, however, is concerned with time, because you have operations which are performed sequentially and, when I say 'algebra', I do not just mean modem algebra. Any algorithm, any process for calculation, is a sequence of steps performed one after the other; the modern computer makes that quite clear. The modern computer takes its information in a stream of zeros and ones, and it gives the answer.
>
> Algebra is concerned with manipulation in time and geometry is concerned with space. These are two orthogonal aspects of the world, and they represent two different points of view in mathematics. (Atiyah 2002, p. 6f)

However, the state of mind Atiyah describes for geometry, "I can just sit here and see, and nothing may change, but I can still see," sounds awfully close to that undifferentiated state that one is graced with in meditation, not the kind called up by the I who has decided to "meditate," but the one that establishes its dominion by itself, reminds one of what Krishnamurti describes as "the state of experiencing."

> In the state of experiencing, there is neither the experiencer nor the experienced. The tree, the dog and the evening star are not to be experienced by the experiencer; they are the very movement of experiencing. There is no gap between the observer and the observed; there is no time, no spatial interval for thought to identify itself. Thought is utterly absent, but there is being. This state of being cannot be thought of or meditated upon, it is not a thing to be achieved. The experiencer must cease to experience, and only then is there being. In the tranquillity of its movement is the timeless. (Krishnamurti 1956, p. 32)

That is not the state in which geometry happens. In geometry itself, one has to separate two things "one of which gives way to the other, but is retained by memory," as Brouwer put it, to discover any geometric truth. The mind must still be traveling from one entity to the other, its stillness may give rise to awe, but will never generate an insight, languageless or not, about the *structure* of the observed.

We have here identified the intuition that allows to "mentally differentiate between objects" with the intuition of time. Placek is not sure whether that identification is allowed, as he finds no confirmation for it in Brouwer's writings:

> although one may introduce a number of other intuitions as well, say those allowing the subject to separate a mathematical entity, or mentally differentiate between objects, Brouwer does not identify the intuition of two-ity with any of them. Even more, he does not even mention any such faculty in his writings. Whether these faculties are needed in his doctrine is [a] separate question. (Placek 1999, p. 42)

Only if Brouwer understood geometry strictly as the science of the continuum could he have considered that the intuition of time can serve to construct space by means of the construction of the coordinate space. However, he was too well-informed about the axiomatic foundations of geometry and certainly knew that ordered geometry, as axiomatized by the axioms of incidence and order, in the two-dimensional case, without the addition of any configuration theorem, does not allow for an algebraization. He had read the entire relevant literature, had a regular course on the subject, and was the advisor of a thesis in classical axiomatic geometry.

There is a striking similarity with the confluence of time and space in Buddha's language, which was pointed out by Monier Williams.

> Another interesting feature of the Buddha's language of temporality is his use of the term for 'space' to refer to time as well. In the traditional literature, the term *adhvan* is used exclusively in the sense of road, way, orbit, journey or course. In fact, Monier Williams has noted that Buddhism and Jainism used it in the sense of time. The "Discourse on Convocation" (*Sangīti-suttanta*) enumerates the three *addhā* [paths] as past, present and future. Elsewhere it is said:
>
> A person, wandering for a long stretch of time with craving as companion, does not overcome the continued cycle of wandering *(saṁsāra)* in this or other modes of life. (Kalupahana 1999, p. 59)

Now, leaving aside the fact of a *single* origin, what about the fact that mathematics originates in an *intuition* at all? This is Brouwer's middle way between nominalism (for which mathematical objects, relations, and structures do not exist – as physical objects or as entities having causal powers) and Platonism (for which there are abstract mathematical objects whose existence is independent of human beings, of the language or thought processes of those beings – Plato was, incidentally, not a Platonist in this sense, as pointed out in Pritchard (2010, p. 111)), shying away from both extremes. One can also see it as a middle way between a naturalism that declares "all mathematics" to be "ultimately a product of the human nervous system, the best product produced at each stage of our development" (Korzybski 1931, p. 276), a position defended by the neurobiologist Jean-Pierre Changeux, and a Platonism, backed in conversations with Changeux (Changeux and Connes 1995) by Alain Connes. It is Brouwer's version of denying mathematics a *svabhāva*, a self-existence, it is his version of "dependent origination" (*paṭiccasamuppāda*), his insistence on the interrelatedness of all there is.

So how is this intrusion of *paṭiccasamuppāda* on the altar of Western philosophy, with its relentless search for the object, received? We present two quotes, from critics of intuitionism (which Hartmann calls "Intuitivismus") by Nicolai Hartmann and by Felix Kaufmann, that read as if the authors had read Haas's book, have internalized how Western philosophy has to proceed, and came up with the predictable answers.

> 'Intuition' is from the start a mode of cognition, i.e. a 'grasping', a transcendent act; in that it differs essentially from the 'setting', and it is just the theory that misjudges that. The truth of intuition is that it is not a giving act, but a receiving (receptive) act, and that the giving authority behind it ought to be searched for at the object. It is the one who determines the intuition to the extent that it 'presents' (appears) itself to it, and that happens as an object that

> is indifferent to the very act of intuition. It is thus already assumed as an existent-in-itself. In case such an object were not to present itself, in case no existent were available, that would have its particular suchness already as its own, then the act itself would not be a seeing apprehension either. (Trans. V. P.)
>
> [‚Intuition' ist nämlich von vornherein ein Modus des Erkennens, also ein‚Erfassen', ein transzendenter Akt; sie ist darin von Grund aus anders gestellt als die‚Setzung', und es ist nur die Theorie, die das verkennt. Die Wahrheit der Intuition ist, dass sie nicht gebender, sondern aufnehmender (rezeptiver) Akt ist, und dass die gebende Instanz hinter ihr beim Gegenstande zu suchen ist. Dieser bestimmt die Anschauung, insoweit er sich ihr‚darbietet' (erscheint), und zwar als ein gegen den Anschauungsakt selbst indifferenter. Er ist also schon als ansichseiender vorausgesetzt. Bietet sich ein solcher Gegenstand nicht dar, liegt also kein Seiendes vor, das sein bestimmtes Sosein schon an sich hätte, so ist auch der Akt kein schauendes Erfassen. (Hartmann 1965, p. 236)]
>
> For Brouwer's view that mathematical facts change with mathematical knowledge implies that there is something cognizable that is created only by being cognized, which runs counter to the nature of cognition. For [. . .] all cognition presupposes an object that must be thought of as existing independently of its being cognized. (Kaufmann 1978, p. 56f)
>
> [Denn Brouwers Auffassung, daß sich die mathematischen Tatsachen mit der mathematischen Erkenntnis wandeln, schließt ein, daß es ein zu Erkennendes gebe, welches erst durch die darauf bezogene Erkenntnis geschaffen wurde, was dem Wesen der Erkenntnis zuwiderläuft. Denn jede Erkenntnis setzt [. . .] einen unabhängig von seinem Erkanntwerden als bestehend zu denkenden Gegenstand voraus.]. (Kaufmann 1930, p. 65)

Here is what the theoretical physicist David Bohm had to say about this objectifying tendency of Western philosophy:

> The weakness of thought is that thought inevitably separates itself from what it thinks about. It creates an imaginary other, which it calls the object – which is still really thought. Let's say I'm thinking about the image of a tree. Now that which I am thinking about seems to be separated from me. It seems the image is over there somewhere and I am here. Therefore it seems that I have created two images, one is the tree and the other is me. (Krishnamurti and Bohm 1999, p. 49)

Going even deeper into the process of object-creation, as well as into the phenomenon of *tṛ́ṣṇa*, we find in Brouwer (1948, p. 480) that

> As mind it [consciousness] takes the function of a subject experiencing the present as well as the past sensation as object. And by reiteration of this two-ity-phenomenon, the object can extend to a world of sensations of motley plurality.
>
> In measure of the irreversibility with which the subject has receded from an element of the object, this element loses its egoicity, i.e. gets estranged from the subject, and in measure of this estrangement, mind becomes disposed to desire and apprehension, and consequently to positive or negative conative activity with respect to the element in question.

There is a striking emphasis on the *yang*, on the object, on the external, the visible, which goes hand in hand with a de-emphasis of the *yin*, the internal, the invisible in these critiques of Brouwer's intuitionism. The same emphasis on external features is prominent in formalism, in which the origin of the strings of meaningless symbols, that are its subject matter, is irrelevant. Formalism is not specific to mathematics as a modernist doctrine. Its counterpart in philosophy is analytic

philosophy, which focuses on the language of philosophy, given that, what is philosophy other than a philosophical text, and what other purpose could philosophy serve than producing exact sequences of words? Its counterpart in psychology is behaviorism, which ignores the inner life of a human being and focuses on the observable part, namely, behavior. To put in in Hannah Arendt's terms, "behavior has replaced action as the foremost mode of human relationship" (Arendt 1998, p. 41). In legal theory, the modernist approach is that of the "Pure Theory of Law" (*Reine Rechtslehre*) of Hans Kelsen, against the counter-modernist legal theory of Natural Law (*Naturrecht*), with Leo Strauss as its prominent twentieth-century proponent.

This rejection of mathematics' dependency on intuition, of its birth from intuition, the preference for thinking of mathematics as of a thing – for in its formalist understanding it is a thing, namely, the system consisting of strings of meaningless symbols, together with the formal rules that allow for the generation of other strings of meaningless symbols – is another case of what Günther Anders called "the shame of not being a thing," in which he sees "a second level of alienation," that on which "man has recognized the superiority of things, brings himself in conformity with them, approves of his own having-become-a-thing, respectively condemns his not-having-become-a-thing as a shortcoming" (Anders 1980, p. 30) (Trans. V. P.). The shame of being born – an unavoidable fact in any philosophy of complete interdependency – has been already noticed by Schelling (as mentioned in Anders 1980, p. 325), who writes that "Der Eigendünkel des Menschen sträubt sich gegen diesen Ursprung aus dem Grunde [. . .]" (Schelling 1809, p. 360) ["The self-conceitedness of man rejects this deep origin [. . .]" (Trans. V. P.)].

How does formalism avoid having to admit that time is playing a decisive role? Having created a "material instrument" (Brouwer 1929, p. 422), namely, the strings of meaningless symbols, that one can imagine as being written on paper or stone, it has transformed mathematics into the physical object that is represented by the strings of symbols. The deceptive part in this procedure is that one cannot thus consider that this spatial artifact is free of time, for the strings of symbols are concatenations, so time in the sense of *this after that* and a direction of reading are required. Without a direction of reading and the notion of a successor, the notion of string itself falls apart.

On a very concrete level, philosophically inclined mathematicians fear that, if one were to accept Brouwer's origin story, mathematics would – from being capable of bringing forth the envy of a Christian theologian (who is afraid that a Platonic understanding "precludes God's being the sole ultimate reality," the existence of "uncreatable" abstract objects being "incompatible" with God's aseity (Craig 2017, p. 13)) – become part of psychology. This fear is unjustified (and has been dealt with in both Placek (1999) and Van Atten (2004)) and has been addressed in a very condensed form by Largeault:

> intuitionism has something *aristocratic* about it which is connected with the exceptional nature of intuition. (Despite this fact, intuition, pure and disinterested knowledge that it is, is

> *universal* rather than individual, which refutes the legend of solipsism which has been attributed to the Dutch geometer.) (Trans. V. P.)
>
> [l'intuitionisme a quelque chose d'*aristocratique* qui tient au caractère exceptionnel de l'intuition. (Malgré cela, l'intuition, connaissance pure et désintéressée, est *universelle* plutôt qu'individuelle, ce qui réfute la légende du solipsisme attribué au géomètre hollandais.). (Largeault 1993, p. 224)]

This universal nature of mind, as well as the rejection of what Brouwer refers to as "the plurality of mind," is emphasized by Krishnamurti in many of his talks. One of these is a talk with David Bohm:

DB: You are using the word mind, it means it is not my mind.
JK: Oh, mind, mind, it is not mine.
DB: Universal or general.
JK: Yes. It is not my brain either.
DB: No, but there is a particular brain, this brain or that brain. Would you say there is a particular mind?
JK: No.
DB: Now, you see that is an important difference. You are saying mind is really universal.
JK: Mind is universal – if you can use that word, ugly word.
DB: Unlimited and undivided.
JK: It is unpolluted, not polluted by thought.
DB: But I think for most people there will be a difficulty in saying how do we know anything about this mind. I only know of my mind is the first feeling – right?
JK: You cannot call it your mind. You only have your brain which is conditioned. You can't say, 'It is my mind'.
DB: Yes, well whatever is going on inside I feel is mine and it is very different from what is going on inside somebody else.
JK: No, I question whether it is different.
DB: At least it seems different.
JK: Yes. I question whether it is different, what is going on inside me as a human being and you as another human being, we both go through all kinds of problems, suffering, fear, anxiety, loneliness, suffer, and so on and so on. We have our dogmas, beliefs, superstitions, and everybody has this.
DB: Well we'll say it is all very similar but still it seems each one of us is isolated from the other.
JK: By thought. My thought has created that I am different from you, because my body is different from you, my face is different from you, so we carry that same – we extend that same thing into the psychological area. (Krishnamurti and Bohm 1983)

That universal nature of introspectively discovered truth is also stated in the following "reflection" of Goethe, which appeared posthumously in 1833 (Goethe 1953; Maximen und Reflexionen 1060, p. 514):

> In case I know my relation to myself and to the outside world, I call it truth. And in that way everyone can have their own truth, and it is nevertheless always the same. (Trans. V. P.)
>
> [Kenne ich mein Verhältnis zu mir selbst und zur Außenwelt, so heiß ich's Wahrheit. Und so kann jeder seine eigene Wahrheit haben, und es ist doch immer dieselbige.]

Brouwer himself wrote to Van Dantzig on August 24, 1949, that he sees a psychological interpretation of intuitionism as a misunderstanding:

> I am glad to see that these developments make the essentially negative properties meaningful also to those who do not recognize the intuitionistic creating subject, because with respect to mathematics they hold either a psychologistic point of view, or in any case stick to the 'plurality of mind'.
>
> As I told you in conversation, my example in question is for fundamental intuitionism so much more unassailable than for those of a different persuasion, because the intuitionistic creating subject can certainly, and from the outset put restrictions (or prohibitions of restrictions) on a specific growing mathematical entity, but not on his own possibilities of creation.
>
> My belief that psychological pictures of intuitionistic mathematics, however interesting they may be, never can be adequate, has, if possible, been even strengthened by your comment. (Van Dalen 2011a, p. 439)

Now why and how does this philosophical stance of Brouwer alter concrete mathematics? Why is his own fixed-point theorem becoming false under this understanding? Brouwer had proved that the assumption of a continuous self-map of the closed disc without a fixed point leads to a contradiction. So, if we denote by A the fixed-point theorem, then Brouwer proved the truth of $\neg\neg A$, or, put differently, "the absurdity of the absurdity of A." In classical logic, the double negation of a sentence is equivalent to it. Intuitionistically, that is not valid. Why? In classical mathematics, one thinks of the closed disk as a set of points, so if a continuous self-map f without a fixed point is an absurdity, then there must exist a point in that closed disk that is not moved by f. We might never be able to find that point for a specific f, but its existence is assured. This thinking – even if only encoded in the axiom system chosen, and thus purely formal and not making any overt claim of existence – is a tributary to Platonism, to the belief that the points of the closed disk exist independently of our ability to construct them. However, already in his doctoral dissertation, Brouwer has a section heading that reads "Mathematics can deal with no other matter than that which it has itself constructed" (Brouwer 1907, p. 51) and we find out that "to exist in mathematics means: to be constructed by intuition; and the question whether a corresponding language is consistent, is not only unimportant in itself, it is also not a test for mathematical existence" (Brouwer 1907, p. 96). In short, there are no "non-experienced truths" (Brouwer 1948, p. 488) in mathematics.

This is best illustrated by Brouwer's account of the intuitionistic continuum, whose very nature is that of an unfinished entity, one that is *becoming*. Brouwer found the need to introduce a new language to deal with intuitionistic entities. For example, he does not use the word "set." This has similarities with the Buddha's tendencies in matters of language.

> The Buddha needed a language that avoided [. . .] substantialist and nihilist implications. This was the language of becoming *(bhava)* that more accurately reflected his experience of change and continuity. It is the language of 'dependent arising' *(paṭiccasamuppàda)* that steered clear of the extremes of permanent existence and nihilistic non-existence. Thus the Buddha can be credited with revolution in linguistic philosophy when he consistently utilized the 'language of becoming' to take the venom off the 'language of existence.' (Kalupahana 1999, pp. ii–iii)

Again, Largeault, with his keen sensitivity for the cultural aspect of philosophy, had noticed this fact, for he writes that

> Does this mean that our mathematics is based on a substance-attribute ontology, reflected in our language (subject-copula-attribute sentences), inherited from the ancient Greeks, while intuitionist mathematics presupposes an ontology of processes, which is repelled by our thinking habits? (Trans. V. P.)
>
> [Cela signifie-t-il que nos mathématiques reposent sur une ontologie substance-attribut, reflétée dans notre langue (phrases suject-copule-attribut), héritée des Grecs anciens, tandis que les mathématiques intuitionistes supposent une ontologie de processus, qui répugne à nos habitudes de pensée? (Largeault 1993, p. 175, footnote 1)]

One could wonder why it was and still is so hard for mathematicians and philosophers of mathematics to allow the observer inside the halls of mathematics. Did not physicists do just that? True, they did, but they were forced to do so by the very nature of experiments, not out of an abundance of interest in the Eastern approach. And even physics tends to exclude those who overstep the boundaries of what is, due to experimental outcomes, strictly necessary to accept, such as David Bohm, with his Spinoza-like embrace of the interconnectedness of mind and matter (see also Espinosa Rubio 2000), and his statement, reminiscent of Teilhard de Chardin's "within of things" that

> Mind and matter are inseparable, in the sense that everything is permeated with meaning. The whole idea of the somasignificant or signasomatic is that at no stage are mind and matter ever separated. There are different levels of mind. Even the electron is informed with a certain level of mind. (Bohm and Weber 1987, p. 443)

But, one might ask, what is the value of the fact that Brouwer's philosophy displays characteristics similar to that of an Indian who lived more than 2500 years ago, when science or mathematics were as good as non-existent? And how is the thesis of Brouwer's Eastern disposition helped by having the support of Krishnamurti, to whom Walpola Rahula said in Rahula et al. (1996, p. 19) "there is hardly any difference between your teaching and the Buddha's"? Is that same Eastern mind still present in the twentieth century, in Indians who had extensive contact with the Western mind, and who did not discover "the liberation of insight" (Krishnamurti and Bohm 1999, p. 101)? Pupul Jayakar, in conversation with Krishnamurti, referred to an Indian tendency of "delving into the self, delving into the within, insights into things" (Jayakar and Krishnamurti 1996, p. 154). And then there is Rabindranath Tagore – who, just like Krishnamurti, whose talks Brouwer

frequented (see Van Dalen (2013, p. 364) and (van Atten 2015, p. 179), where we find out that "When visiting Krishnamurti, Brouwer said to a friend: 'Oh my, this is the baby room of philosophy'"), played a role in Brouwer's life – in conversation with Einstein "in the afternoon of July 14, 1930, at the latter's residence in Kaputh":

E.: There are two different conceptions about the nature of the universe: (1) The world as a unity dependent on humanity. (2) The world as a reality independent of the human factor.
T.: When our universe is in harmony with Man, the eternal, we know it as truth, we feel it as beauty.
E.: This is a purely human conception of the universe.
T.: There can be no other conception. This world is a human world – the scientific view of it is also that of the scientific man. There is some standard of reason and enjoyment which gives it truth, the standard of the Eternal Man whose experiences are through our experiences.
E.: This is a realization of the human entity.
T.: Yes, one eternal entity. We have to realize it through our emotions and activities. We realize the Supreme Man who has no individual limitations through our limitations. Science is concerned with that which is not confined to individuals; it is the impersonal human world of truths. Religion realizes these truths and links them up with our deeper needs; our individual consciousness of truth gains universal significance. Religion applies values to truth, and we know truth as good through our own harmony with it.
E: Truth, then, or Beauty, is not independent of Man?
T: No.
E: If there would be no human beings any more, the Apollo of Belvedere would no longer be beautiful?
T: No.
E: I agree with regard to this conception of Beauty, but not with regard to Truth.
T: Why not? Truth is realized through man.
E: I cannot prove that my conception is right, but that is my religion.
T: Beauty is in the ideal of perfect harmony which is in the Universal Being; truth the perfect comprehension of the Universal Mind. We individuals approach it through our own mistakes and blunders, through our accumulated experience, through our illumined consciousness – how, otherwise, can we know Truth?
E: I cannot prove scientifically that truth must be conceived as a truth that is valid independent of humanity; but I believe it firmly. I believe, for instance that the Pythagorean theorem in geometry states something that is approximately true, independent of the existence of man. Anyway, if there is a reality independent of man there is also a truth relative to this reality; and in the same way the negation of the first engenders a negation of the existence of the latter.
T: Truth, which is one with the Universal Being, must essentially be human, otherwise whatever we individuals realize as true can never be called truth – at least the truth which is described as scientific and can only be reached through the process of logic, in other words, by an organ of thoughts which is human.

According to Indian Philosophy there is Brahman the absolute Truth, which cannot be conceived by the isolation of the individual mind or described by words, but can only be realized by completely merging the individual in its infinity. But such a truth cannot belong to Science. The nature of truth which we are discussing is an appearance – that is to say what appears to be true to the human mind and therefore is human, and may be called māyā or illusion.

E: So according to your conception, which may be the Indian conception, it is not the illusion of the individual, but of humanity as a whole.

T: In science we go through the discipline of eliminating the personal limitations of our individual minds and thus reach that comprehension of truth which is in the mind of the Universal Man.

E: The problem begins whether Truth is independent of our consciousness.

T: What we call truth lies in the rational harmony between the subjective and objective aspects of reality, both of which belong to the super-personal man.

E: Even in our everyday life we feel compelled to ascribe a reality independent of man to the objects we use. We do this to connect the experiences of our senses in a reasonable way. For instance, if nobody is in this house, yet that table remains where it is.

T: Yes, it remains outside the individual mind, but not outside the universal mind. The table which I perceive is perceptible by the same kind of consciousness which I possess.

E: Our natural point of view in regard to the existence of truth apart from humanity cannot be explained or proved, but it is a belief which nobody can lack – no primitive beings even. We attribute to Truth a super-human objectivity; it is indispensable for us, the reality which is independent of our existence and our experience and our mind though we cannot say what it means.

T: Science has proved that the table as a solid object is an appearance, and therefore that which the human mind perceives as a table would not exist if that mind were naught. At the same time it must be admitted that the fact, that the ultimate physical reality of the table is nothing but a multitude of separate revolving centres of electric forces, also belongs to the human mind. In the apprehension of truth there is an eternal conflict between the universal human mind and the same mind confined in the individual. The perpetual process of reconciliation is being carried on in our science and philosophy, and in our ethics. In any case, if there be any truth absolutely unrelated to humanity then for us it is absolutely non-existing. It is not difficult to imagine a mind to which the sequence of things happens not in space, but only in time like the sequence of notes in music. For such a mind its conception of reality is akin to the musical reality in which Pythagorean geometry can have no meaning. There is the reality of paper, infinitely different from the reality of literature. For the kind of mind possessed by the moth, which eats that paper, literature is absolutely non-existent, yet for Man's mind literature has a greater value of truth than the paper itself. In a similar manner, if there be some truth which has no sensuous or rational relation to the human mind it will ever remain as nothing so long as we remain human beings.

E: Then I am more religious than you are!

T: My religion is in the reconciliation of the Super-personal Man, the Universal human spirit, in my own individual being. This has been the subject of my Hibbert Lectures, which I have called 'The Religion of Man'. (Gosling 2007, p. 161f)

For the mathematician who is not philosophically inclined, and who could not care less about East or West, the rejection of intuitionistic mathematics is connected with the fact that it is considered to be a mutilation of classical mathematics, given that many of the results of classical mathematics become false in intuitionistic mathematics. The fact that, on the other hand, there are results in intuitionistic mathematics that are flatly wrong in classical mathematics, such as the theorem proved by Brouwer in 1924 that "every full function is uniformly continuous" does not endear it to mathematicians either, given that such results remove from the body of mathematics all results about discontinuous functions. Brouwer's answer to this criticism was:

> In general intuitionism brings about a complete recasting of mathematics, with the result, to our regret, that in many places its supple and elegant character is lost, and it has to assume much harsher, more tortuous and more complicated forms. Alas, the spheres of truth are less transparent than those of illusion. (Brouwer 1933, p. 444)

This displeasure with "fewer" results is indicative of the approach to knowledge in the having mode of existence. *More knowledge* is preferred to knowing *more deeply*, which is what Brouwer proposes as "optimum knowledge in the being mode." As Meister Eckhart put it:

> That one is much more blessed, who can dispense with all things and does not need them, then he, who has taken possession of all things, because he needs them. (Trans. V. P.)
>
> [Der ist viel seliger, der alle Dinge entbehren kann und ihrer nicht bedarf, als der, welcher alle Dinge in Besitz genommen hat, weil er sie braucht. (Eckhart 1934, p. 109)]

Brouwer never claimed that classical mathematics or formalistic mathematics is *false.* He just claimed that just the fact that something is free of error does not make it worthwhile pursuing, that in a finite life, one should pursue "the structured elements of our thinking" (Brouwer 1946b, p. 475), the structures that arise from some lived intuition, not just any possible game of symbols. It is, after all, a *moral* argument for intuitionistic mathematics. He was opposed, also on moral grounds, to the cooperation of mathematicians with society:

> Of course art and philosophy continually illustrating such wisdom cannot participate in cooperation, in particular should not communicate with the state. Supported by the state, they will lose their independence and degenerate. (Brouwer 1948, p. 487)
>
> The two games [logic and mathematics], by dint of their origin, influence each other. By their very nature, they should not interfere with social life. Since the latter has nevertheless claimed them, they undergo the influence of the pragmatic sciences by cooperating, against their nature, to the transformations of social life that are called progress. Luckily, their most

> beautiful developments will probably never have any connection with technological, economical, or political matters.
>
> [Les deux jeux [logique et mathématique], en vertu de leur origine, s'influencent mutuellement. De par leur nature, ils ne devraient pas s'immiscer dans la vie sociale. Celle-ci les ayant néanmoins réclamés, ils subissent l'influence des sciences pragmatiques tout en coopérant, contre leur nature, aux transformations de la vie sociale qu'on appelle le progrès. Heureusement, leurs plus beaux développements n'auront probablement jamais aucun rapport avec les questions techniques, économiques ou politiques. (Brouwer 1950)

The same disdain for applications and the corrupting influence of money was present in René Thom's thinking (Papadopoulos 2018) and is certainly more widespread than meets the eye, given that many mathematicians will likely keep such thoughts private or never get the opportunity to make them public.

7 Wisdom and Mysticism: Beyond Mathematics and Time

While mathematics originates in the intuition of time, Meister Eckhart tells us that

> No thing is as despised by God as time. Not only time, the clinging to time as well. And not only the clinging, already the touch of time. And not only the touch, already the simple smell or taste of time. (Trans. V. P.)
>
> [Gott ist kein Ding so sehr zu wider als die Zeit. Nicht allein die Zeit, auch das Haften an Zeit. Und nicht allein das Haften, schon das Berühren von Zeit. Und nicht allein das Berühren: schon der bloße Geruch und Geschmack von Zeit. (Pfeiffer 1857, 308, 5)]
>
> If you are neither this nor that, then you are all things. – What has number or is number, from that he is far. (Trans. V. P.)
>
> [Bistu weder dies noch das, so bistu alle Dinge. – Was Zahl hat oder Zahl ist, von dem ist er fern.] (Pfeiffer 1857, 162, 7)

And Krishnamurti tells us that "a human being can see for himself that compassion is out of time, truth is out of time, and the depth from which that compassion comes is out of time. And therefore it is not cultivable" (Krishnamurti and Bohm 1999, p. 97).

Brouwer writes already in his dissertation (although his doctoral advisor crosses it out, as not relevant for mathematics):

> Mathematics is a free creation; it exists in developing a primordial intuition which can be called 'permanence in change' or 'the discrete in the continuous'.
>
> Applying it to the exterior world is creating the objective world; this is characteristic of the human strategy in the general struggle for life. As such it is clearly inferior and has nothing to do with religion or with wisdom. (Van Stigt 1990, p. 414f, The Rejected Parts of Brouwer's Dissertation)
>
> Perhaps the greatest merit of *mysticism* is its use of language independent of mathematical systems of human collusion, independent also of the direct animal emotions of fear and desire. If it expresses itself in such a way that these two kinds of representations cannot be detected, then the contemplative thoughts – whose mathematical restriction appears as the only live element in the mathematical system – may perhaps again come through without obscurity, since there is no mathematical system that can replace them. When the mystical

> author joins representations of this kind into more central affections, of which they were one-sided restrictions, he can, with the most ordinary words, gradually break down the barriers round the contemplative sphere and guide us back to the 'all-embracing' which every poet seeks to approach. His language will therefore appear as meaningless to those who expect to find in words only the communication of mathematical systems or a stimulus to mathematical activity. (Van Stigt 1990, p. 409, The Rejected Parts of Brouwer's Dissertation)
>
> In religious truth, i.e. in wisdom, which abolishes the discernment between the subject and something different, and where the perception of time is no longer admitted, there is no mathematical understanding, let alone reliability of logic. (L. E. J. Brouwer, 1908, p. 108)
>
> Searching for wisdom, we may find it in knowing that causal thinking and acting is non-beautiful and hard to justify, and that in the long run it brings disappointment.
>
> And in knowing that the exterior world with its innumerable individuals and with its hypertrophied cooperation is wedded to mind, its disharmonies reflecting mind's free-will-guilt.
>
> As a consequence of this knowing the exterior world and one's own position in it are accepted as they are, so that towards the exterior world generally only acts as reversible as possible aiming at maintenance, but no acts let alone causal acts aiming at change, are undertaken of one's free will. Repair of disadjustments, averting of danger and relief of need, all this negative intervening in human society is justified in itself and sometimes prescribed. But positive activity to change the structure of human society governed by so many unknown forces, will always be checked by the self-admonition: 'not to improve her works has Providence placed thee in this world', and only vocation and inspiartion tested in detachment-concentration will be stronger than this admonition. (Brouwer 1948, p. 485f)

This understanding leads to a "sceptical prognosis" for humanity, one that was the central concern of Fromm or Krishnamurti as well. In times past, this would have been dismissed as "pessimism," perhaps no longer as out of hand. Brouwer had expressed concerns regarding the equilibrium-upsetting interventions of humans in nature, already in 1905, in the opening lines of Brouwer (1996). It is important to notice that Brouwer offers not only a somber prognosis but also the means to escape it by abandoning the delusion of being able to steer and control life in desired directions by means of "causal acts," by "cunning."

> The categorical imperative prescribing the aforesaid attitude towards life has its counterpart in a sceptical prognosis that mankind, possessed by the delusion of causality, will slide away in a deteriorative process of overpopulation, industrialization, serfdom, and devastation of nature, and that when hereby first its spiritual and then its physiological conditions of life will have been destroyed, it will come to its end like a colony of bacteria in the earth crust having fulfilled its task.
>
> All this though timeless art and perennial philosophy continually suggest that the unknown forces governing the destiny of individual and community, are not subject to causality; that in particular the ways of fate cannot be paved with causality, and that security is as unattainable as it is unworthy; that intensification of organization increases vulnerability, that new vulnerability asks for protection through new organization, and that thus for organization which is believed in, there is no end of growth; finally that if the delusion of causality could be thrown off, nature, gradually resuming her rights, would be (except for her bondage to destiny) generous and forgiving to a mankind decausalized and subsiding to more modest and more harmonious proportions. (Brouwer 1948, p. 487)

And we can find already in 1905, in its last sentence, his categorical rejection of the having mode of existence:

> Only he who recognizes that he has nothing, that he cannot possess anything, that absolute certainty is unattainable, who completely resigns himself and sacrifices all, who gives everything, who does not know anything, does not want anything and does not want to know anything, who abandons and neglects everything, he will receive all; to him the world of freedom opens, the world of painless contemplation and of – nothing. (Brouwer 1996, p. 429)

While it might seem strange that Brouwer searches the ultimate answers whether for the source of mathematical intuition or for the source of "wisdom" in himself, in introspection, he is, again, in good company in doing so: Albert Schweitzer and Salman Baruch Rabinkow – a Chassidic Jew of the Chabad Chassidism branch, which originated in Lithuania, and Talmud teacher – with whom Fromm studied for 6 years the Talmud, and about whom he writes (Fromm 1987, p. 103) that he 'influenced my life more than any other man."

> The great secret is to go through life as an unconsumed person. Such an outcome is possible to the one who, instead of reckoning with people and facts, is thrown back on oneself and seeks the ultimate reason for things in oneself. (Trans. V. P.)
>
> [Das große Geheimnis ist, als unverbrauchter Mensch durchs Leben zu gehen. Solches vermag, wer nicht mit den Menschen und Tatsachen rechnet, sondern in allen Erlebnissen auf sich selbst zurückgeworfen wird und den letzten Grund der Dinge in sich sucht. (Schweitzer 1974, p. 312)]
>
> Judaism ascribes absolute value in its teaching only to this autonomous human being, who connects nature and spirit in a continuous synthesis. To the extent that the condition of autonomy is satisfied, Jewish doctrine does not recognize any difference in rank between humans. Everyone is entitled and obliged to say "The world was created for my own sake" (Sanhedrin, Mishnah, chap. 4), because every human being is an end in itself and is, as it were, burdened with the responsibility for the whole of Creation. (Trans. V. P.)
>
> [Allein diesem autonomen, Natur und Geist zur durchgängigen Synthese verbindenden Menschen spricht das Judentum in seiner Lehre absoluten Wert zu. Soweit aber die Voraussetzungen der Autonomie gegeben sind, kennt die jüdische Lehre keinerlei Rangunterschiede zwischen Mensch und Mensch. Jedermann ist berechtigt und verpflichtet zu sagen „Meinetwegen ist die Welt erschaffen" (Sanhedrin, Mischna, Kap. 4), denn jedes menschliche Wesen ist Selbstzweck und ist gleichsam mit der Verantwortung für die gesamte Schöpfung belastet. (Rabinkow 1929, p. 808f)]

Brouwer's emphasis on the inadequacy of language for communication and the overcoming of even the languageless activity that is mathematics with mysticism is not shared by all mystics. While mystics of the negative theological persuasion, such as Meister Eckhart, agree with Brouwer's take on the ineffability of the mystical experience, not all mystics regard language as inadequate. As Rudolf Otto pointed out, mysticism "preserves its special character, which explains itself clearly from the ground over which it domes up" ["wahrt ihren Sondercharakter, der sich deutlich aus dem Grunde erklärt, über dem sie sich wölbt"] (Otto 1926, p. 231). Given the high regard for language in Judaism ("Thy word (or: essence) is true from the beginning" (Psalm 119: 160)), Jewish mysticism highly values and is heavily reliant on

language, as pointed out in Scholem (1972), Ben-Sasson (2018), and Kilcher (1998, pp. 31–94).

> The point of departure of all mystical linguistic theories, among which we should also number those of the Kabbalists, is constituted by the conviction that the language – the medium – in which the spiritual life of man is accomplished, or consummated, includes an inner property, an aspect which does not altogether merge or disappear in the relationships of communication between men. (Scholem 1972, p. 60)

None of this diminished in any way the deep friendship, based to a great extent on a common interest for the mystic experience, between Brouwer and Erich Gutkind.

Questions regarding "the common core thesis on mysticism" (which "holds that all forms are in the end different ways to express the same thing" (Barendregt 2008, p. 131)) were addressed in Van Atten and Tragesser (2015) and in Barendregt (2008).

8 Conclusions

What we have attempted to show in this study is the fact that Brouwer's philosophy of mathematics is coherent and does not stand alone in the arena of philosophies. It is perceived as a foreign body by philosophers trained in the Western tradition, who usually do not even confer the title "philosophy" to its Eastern branch, let alone allow it to have anything to say about Western philosophy's oldest and most valued possession: mathematics.

We have seen that Brouwer's philosophy converses well with that of the Buddha, Erich Fromm, Krishnamurti, Meister Eckhart, Albert Schweitzer, David Bohm, Humberto Maturana, Francisco Varela, or Spinoza. Critics will say, fine, all you have shown is that he hung out with the wrong crowd, with the wrong side of C. P. Snow's "two cultures." What is a modern mind to do with the Buddha, Krishnamurti, Bohm, and Spinoza, who maintain that, although the world is highly causal and thus that there is a great amount of determinism, *insight liberates* (see (Bittner 1994) for the train of thought that led Spinoza to this conclusion)? All of them defend counter-modernist positions (certainly an anachronistic statement in the case of the Buddha, Spinoza, or Meister Eckhart); no contemporary ethics study in the English language, written by an analytic philosopher, even mentions Schweitzer's name. Both the being mode of existence and the Eastern approach are considered to be outside of the area of acceptable academic discourse. Incidentally, whoever thinks that counter-moderns must lie politically on the right spectrum should consider that Bohm was a communist until the Hungarian Revolution of 1956. This is also not the first instance in which Brouwer and Bohm have been brought together in an essay: a proposal for a model of indeterminism in physics (Gisin 2021) has called for the intuitionistic continuum as "an alternative mathematical language that is both powerful enough to

allow scientists to compute predictions and compatible with indeterminism and the passage of time" and mentions Bohmian mechanics.

What we have *not* attempted to do is to take sides. It is clear to us that Platonism, intuitionism, and formalism are all coherently expressed philosophies, and we are not at all interested in poking holes in one of them, with the aim of demolishing it or finding out inner contradictions. The conversation between Jean-Pierre Changeux and Alain Connes, in which none is able to convince the other of the wrong nature of one's own position, shows that such conversations make sense as exercises in informing the other, not in convincing the other of some mistake in one's own thinking (Changeux and Connes 1995).

Although we clearly stated this in a previous, much shorter essay on this topic (Pambuccian 1992), every attempt to defend Brouwer's position as coherent will likely bring on the same reaction that Brouwer's philosophy elicits, as can be read from the following review:

> This essay exhibits the efforts of a mathematician to articulate a philosophy justifying doing mathematics under the constraints of the outlook of the intuitionist L. E. J. Brouwer. But the philosophy does not concern having the proper concept of mathematical reality and, then, having the proper techniques for having truths about it. The word "having" is repeated in the preceding sentence to emphasize that a more fundamental philosophy, based on insights from (Fromm 1976), is pursuit of detachment from the way of existing which is dependent upon having things: the having mode of existence. This includes detachment from having the proper metaphysics and epistemology. The fruit of such detachment is an independent way of being: the being mode of existence. Presumably, those who are seeking this detached way of existing will do mathematics in Brouwer's way because they are detached from having access to a mathematical reality. (Kielkopf 1993)

The writer never attempted to justify his own activity, as he has written all his papers in a strict formalistic tradition and has used a (formal) intuitionistic result only once. The result used was arrived at by Michael Beeson in Beeson (2018). The writer's own use of it in solving an open problem going back to 1852 was straightforward (Pambuccian 2018). This essay is also not written to justify the writer's own or anyone else's activity. That would be pathetic, for, as Schweitzer put it:

> We are in truth when we experience conflicts increasingly deeper. The good conscience is an invention of the devil. (Trans. V. P.)
>
> [In der Wahrheit sind wir, wenn wir die Konflikte immer tiefer erleben. Das gute Gewissen ist eine Erfindung des Teufels.] (Schweitzer 1923, p. 249)

What we hope to have shown is that, if a formalist, whose papers are – unlike those of the majority of mathematicians, who are formalists only on Sundays – filled with actual strings of meaningless symbols, can understand what Brouwer was intending and why, then anyone can. There is no expectation that anyone will change their mind on the validity of Brouwer's philosophy after having read this essay.

References

Allen JP (1988) Genesis in Egypt. The philosophy of ancient Egyptian creation accounts. Yale Egyptological Seminar, New Haven
Anders G (1980) Die Antiquiertheit des Menschen, vol I. C. H. Beck, München
Arendt H (1998) The human condition, 2nd edn. The University of Chicago Press, Chicago
Atiyah MF (2002) Mathematics in the 20th century. Bull Lond Math Soc 34:1–15
Balcerowicz P (2001) The logical structure of the *naya* method of the Jainas. J Indian Philos 29: 379–403
Barendregt H (2008) Buddhist models of the mind and the common core thesis on mysticism. In: van Atten M, Boldini P, Bourdeau M, Heinzmann G (eds) One hundred years of intuitionism (1907–2007). The Cerisy conference. Birkhäuser, Basel, pp. 131–145
Becker O (1927) Mathematische Existenz. Max Niemeyer Verlag, Halle a.S
Beeson M (2018) Brouwer and Euclid. Indag Math (NS) 29:483–533
Ben-Sasson HH (2018) "The name of God and the linguistic theory of the Kabbalah" revisited. J Relig 98:1–28
Berberova N (1992) The italics are mine. Alfred A Knopf, New York
Bittner R (1994) Spinozas Gedanke, daß Einsicht befreit. Deutsche Zeitschrift für Philosophie Berlin 42:963–871
Boddhi B (2000) The connected discourses of the Buddha. A translation of the Saṃyutta Nikāya. Wisdom, Boston
Bohm D, Weber R (1987) Meaning as being in the implicate order philosophy of David Bohm: a conversation. In: Hiley BJ, Peat FD (eds) Quantum implications. Essays in honour of David Bohm. Routledge, London/New York
Brouwer LEJ (1907) Over de grondslagen der wiskunde. Academisch proefschrift. Maas & van Suchtelen, Amsterdam; [On the foundations of mathematics] [Brouwer 1975, pp. 11–101]
Brouwer LEJ (1908) De onbetrouwbaarheid der logische principes. Tijdschrift voor Wijsbegeerte 2: 152–158. [The unreliability of the logical principles] [Brouwer 1975, pp. 107–111]
Brouwer LEJ (1909) Het wezen der meetkunde. Openbare Les. Clausen, Amsterdam; [The Nature of Geometry. Public Fecture.] [Brouwer 1975, pp. 112–120]
Brouwer LEJ (1913) Intuitionism and formalism. Bull Am Math Soc 20:81–86. [Brouwer 1975, pp. 123–138]
Brouwer LEJ (1929) Mathematik, Wissenschaft und Sprache. Monatshefte für Mathematik 36: 153–164. [Brouwer 1975, pp. 417–428]
Brouwer LEJ (1933) Willen, Weten, Spreken. Euclides 9:177–193. [Brouwer 1975, pp. 443–446] [English translation in [Van Stigt 1990, Appendix 5]]
Brouwer LEJ (1937) Signifische dialogen [Signific dialoges]. Synthese 2:168–174. 316–324; [Brouwer 1975, pp. 447–456]
Brouwer LEJ (1946a) Synopsis of the signific movement in The Netherlands. Synthese 5:201–208. [Brouwer 1975, pp. 465–471]
Brouwer LEJ (1946b) Address delivered on September 16th, 1946, on the conferment upon Professor G. Mannoury of the honorary degree of Doctor of Science. [Brouwer 1975, pp. 472–476]
Brouwer LEJ (1948) Consciousness, philosophy and mathematics. In: Proceedings of the 10th international congress of philosophy, Amsterdam, pp. 1235–1249; [Brouwer 1975, pp. 480–494]
Brouwer LEJ (1950) Discours final. In: Les méthodes formelles en axiomatique. Centre National de la Recherche Scientifique, Paris, p. 75. [Brouwer 1975, p. 503]
Brouwer LEJ (1952a) An intuitionist correction of the fixed-point theorem of the sphere. Proc R Soc Lond A 213:1–2. [Brouwer 1975, pp. 506–507]
Brouwer LEJ (1952b) Historical background, principles and methods of intuitionism. S Afr J Sci 49:139–146. [Brouwer 1975, pp. 508–515]
Brouwer LEJ (1975) In: Heyting A (ed) Collected works, vol 1. North-Holland, Amsterdam

Brouwer LEJ (1996) Life, art, and mysticism. Van Stigt WP (trans). Notre Dame J Form Log 37: 389–429
Budianksy S (2021) Journey to the edge of reason. The life of Kurt Gödel. Oxford University Press, Oxford
Callahan G (2015) Was Berkeley a subjective idealist? Collingwood and British Idealism Studies 21:157–184
Camus A (1955) The myth of Sisyphus, and other essays (trans: O'Brien J). Alfred A Knopf, New York
Carnap R (1934) Logische Syntax der Sprache. Springer Verlag, Wien
Carnap R (1959) The logical syntax of language (trans: Smeaton A). Adams and Co, Littlefield
Changeux JP, Connes AF (1995) Conversations on mind, matter, and mathematics. DeBevoise MB (ed & trans). Princeton University Press, Princeton
Craig WL (2017) God and abstract objects. The coherence of theism: aseity. Springer, Cham
Dauben JW (2000) Review of [Van Dalen 1999b]. MR1669279 (2000g,01037)
Dauben JW (2007) Review of [Van Dalen 2005]. MR2171883 (2007i,01004)
Detlefsen M (1998) Constructive existence claims. In: Schim M (ed) The philosophy of mathematics today. Oxford University Press, New York, pp. 307–335
Dilthey W (1965) Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften. Gesammelte Schriften, VII. Band. B. G. Teubner/Vandenhoeck & Ruprecht, Stuttgart/Göttingen
Dubucs J-P (1988) L. E. J. Brouwer: topologie et constructivisme. Revue d'Histoire des Sciences 41:133–155
Dummett M (1975) The philosophical basis of intuitionistic logic. In: Rose HE, Shepherdson JC (eds) Logic colloquium '73 (Bristol, 1973). North-Holland, Amsterdam, pp. 5–40
Dummett M (2000) Elements of intuitionism, 2nd edn. Oxford University Press, Oxford
Eckhart M (1934) Deutsche Predigten und Traktate. Insel, Leipzig
Eggenberger PB (1976) The philosophical background of L.E.J. Brouwer's intuitionistic mathematics. Ph.D. Thesis. University of California, Berkeley
Espinosa Rubio L (2000) Spinoza y David Bohm: un diálogo sobre filosofía natural. Estudios Filosóficos 49:309–328
Finnestad RB (1976) Ptah, creator of the gods: reconsideration of the Ptah sections of the *Denkmai*. Numen 23:81–113
Fraenkel AA (2016) Recollections of a Jewish mathematician in Germany. Edited by Jiska Cohen-Mansfield. Translated by Allison Brown. Birkhäuser, Cham
Francklin T (1787) Sermons on various subjects, and preached on several occasions by Thomas Francklin, vol 3. T. Cadell, in the Strand, London
Frenkian AM (1946) L'Orient et les origines de l'idéalisme subjectif dans la pensée européenne. T. I: La doctrine théologique de Memphis (L'inscription du roi Shabaka). Librairie Orientaliste Paul Geuthner, Paris
Freudenthal H (1962) The main trends in the foundations of geometry in the 19th century. In: Nagel E, Suppes P, Tarski A (eds) Logic, methodology and philosophy of science. Proceedings of the 1960 international congress. Stanford University Press, pp. 613–621
Friedman LJ (2013) The lives of Erich Fromm. Love's prophet. Columbia University Press, New York
Fromm E (1976) To have or to be? Harper & Row, New York
Fromm E (1987) Reminiscenses of Shlomo Barukh Rabinkow. In: Jung L (ed) Sages and saints. Ktav Publishing House, Hoboken, pp. 99–105
Gandon S (2012) Russell's unknown logicism. Palgrave Macmillan, New York
Ganeri J (1996) Numbers as properties of objects: Frege and the Nyāya. Studies in Humanities and Social Sciences 3:111–121
Gies M, Gold AL (1987) Anne Frank remembered. The story of the woman who helped to hide the Frank family. Simon and Schuster, New York
Gisin N (2021) Indeterminism in physics and intuitionistic mathematics. Synthese 199: 13345–13371. https://doi.org/10.1007/s11229-021-03378-z

Goethe JW (1953) Goethes Werke. Band XII. Schriften zur Kunst, Schriften zur Literatur, Maximen und Reflexionen. Christian Wegner Verlag, Hamburg
Gosling DL (2007) Science and the Indian tradition. When Einstein met Tagore. Routledge, London
Grattan-Guinness I (1999) Review of [Van Dalen 1999b]. Bull Am Math Soc 36:529–532
Gray J (2015) Brouwer's certainties: mysticism, mathematics, and the ego. Essay review of [Van Dalen 2013]. Metascience 24:127–134
Guenther HV (1989) Mentalism and beyond in Buddhist philosophy. In: Guenther HV (ed) Tibetan Buddhism in Western perspective. Dharma Publishing, Berkeley, pp. 162–177
Gyatso T (the Fourteenth Dalai Lama) (1990) Opening the eye of new awareness. Wisdom, Boston
Haas WS (1956) Destiny of the mind. East and West. Macmillan, New York
Hartimo M (2017) Husserl and Hilbert. In: Centrone S (ed) Essays on Husserl's logic and philosophy of mathematics. Springer, Dordrecht, pp. 245–263
Hartmann N (1964) Der Aufbau der realen Welt. Walter de Gruyter, Berlin
Hartmann N (1965) Zur Grundlegung der Ontotogie. Walter de Gruyter, Berlin
Herman WF (1949) De tranen der acacia's. Uitgeverij G. A. van Oorschot, Amsterdam
Hesseling DE (2003) Gnomes in the fog. The reception of Brouwer's intuitionism in the 1920s. Birkhäuser, Basel
Hoffmeier JK (1983) Some thoughts on Genesis 1 & 2 and Egyptian cosmology. J Anc Near East Soc 15:39–49
Holenstein E (2004) Philosophie-Atlas. Orte und Wege des Denkens. Ammann Verlag, Zürich
Huber-Dyson V (1996) Thoughts on the occasion of Georg Kreisel's 70th birthday. In: Odifreddi P (ed) Kreiseliana. About and around Georg Kreisel. A. K. Peters, Wellesley, pp. 51–73
Husserl E (1994) Brief von Husserl an Heidegger, 9. V. 1928 (1). In: Schuhmann K (ed) Husserliana. Dokumente. Band III. Briefwechsel. Teil 4: Die Freiburger Schüler. Springer-Science, Dordrecht
Inagaki H (1998) Nāgārjuna's discourse on the ten stages (Daśabhumika-vibhasa). Ryukoku Gakkai, Kyoto
Ingarden R (1975) On the motives which led Husserl to transcendental idealism. Martinus Nijhoff, Den Haag
Jayakar P, Krishnamurti J (1996) Is there an Eastern mind and a Western mind? In: Questioning Krishnamurti. J. Krishnamurti in dialogue. Thorsons, San Francisco, pp. 153–169
Johnson DM (1981) The problem of the invariance of dimension in the growth of modern topology II. Arch Hist Exact Sci 25:85–267
Johnson DM (2014) Review of [Van Dalen 2013]. Not Am Math Soc 61:607–610
Kalupahana DJ (1975) Causality. The central philosophy of Buddhism. The University Press of Hawaii, Honolulu
Kalupahana DJ (1999) The Buddha's philosophy of language. Sarvodaya Viishna Lekha, Moratuwa
Kaneko H (2002a) Brouwer's conception of language, mind and mathematics. Ann Jpn Assoc Philos Sci 11:35–49
Kaneko H (2002b) Review of [Van Dalen 1999b]. Ann Jpn Assoc Philos Sci 11:51–56
Kaufmann F (1930) Das Unendliche in der Mathematik und seine Ausschaltung. Eine Untersuchung über die Grundlagen der Mathematik. F. Deuticke, Wien
Kaufmann F (1978) The infinite in matematics. McGuinness B (ed), Foulkes P (trans). D. Reidel, Dordrecht
Kielkopf CF (1993) Review of [Pambuccian 1992]. MR1174963 (93f:00005)
Kilcher A (1998) Die Sprachtheorie der Kabbala als ästhetisches Paradigma. J. B. Metzler, Stuttgart
Koetsier T (2005) Arthur Schopenhauer and L. E. J. Brouwer: a comparison. In: Koetsier T, Bergmans L (eds) Mathematics and the divine. A historical study. Elsevier, Amsterdam, pp. 569–593
Korzybski A (1931) A non-Aristotelian system and its necessity for rigour in mathematics and physics. In: Korzybski A (ed) Science and sanity, 5th edn. Institute of General Semantics, Brooklyn, pp. 747–761

Kreisel G (1977) Book review of L. E. J. Brouwer. Collected works, volume I, philosophy and foundations of mathematics. Bull Am Math Soc 83:86–93
Kreisel G (1987) Gödel's excursion into intuitionistic logic. In: Weingartner P, Schmetterer L (eds) Gödel remembered. Bibliopolis, Napoli, pp. 67–186
Kreisel G, Newman MHA (1969) Luitzen Egbertus Jan Brouwer 1881–1966. Biogr Mem Fellows R Soc 15:39–68
Krishnamurti J (1956) Commentaries on living, first series. The Theosophical Publishing House, Wheaton
Krishnamurti J (1984) Conversation 1588, Los Alamos, New Mexico, USA, 21 March 1984. http://legacy.jkrishnamurti.org/krishnamurti-teachings/view-text.php?tid=1588&chid=1285
Krishnamurti J (1991) Collected works 5. Choiceless awareness. Kendall Hunt, Dubuque
Krishnamurti J, Bohm D (1983) Is there evolution of consciousness? Dialogue 2 Brockwood Park, England – 20 June 1983. https://jkrishnamurti.org/content/there-evolution-consciousness
Krishnamurti J, Bohm D (1999) The limits of thought. Routledge, London
Lamotte É (1976) Historie du bouddhisme Indien. Des origines à l'ère Śaka. Institut Orientaliste, Louvain-la-Neuve
Largeault J (1993) Intuition et intuitionisme. Librairie Philosophique J. Vrin, Paris
Lusthaus D (2002) Buddhist phenomenology. RoutledgeCurzon, London
Lyttle D (1997) "The world is a divine dream": Emerson's subjective idealism. The Concord Saunterer, New Series 5:92–110
Majer U (2010) The origin and significance of Husserl's notion of the *Lebenswelt*. In: Hyder D, Rheinberger H-J (eds) Science and the life-world. Essays on Husserl's 'crisis of European sciences'. Stanford University Press, Stanford, CA
Männlein-Robert I (2009) Griechische Philosophie in Indien? Reisewege zur Weisheit. Gymnaisum 116:331–357
Maturana HR, Varela FJ (1980) Autopoiesis and cognition. The realization of the living. D. Reidel, Dordrecht
McCarty C (1987) Variations on a thesis: intuitionism and computability. Notre Dame J Form Log 28:536–580
McCarty C (2008) The new intuitionism. In: van Atten M, Boldini P, Bourdeau M, Heinzmann G (eds) One hundred years of intuitionism (1907–2007), the Cerisy conference. Birkhäuser, Basel, pp. 37–49
McEvilley T (1980) Plotinus and Vijñānavāda Buddhism. Philos East West 30:181–193
McEvilley T (2002) The shape of ancient thought. Allworth Press, New York
Mehrtens H (1990) Moderne – Sprache – Mathematik. Eine Geschichte des Streits um die Grundlagen der Disziplin und des Subjekts formaler Systeme. Suhrkamp Verlag, Frankfurt am Main
Mehrtens H (1996) Modernism vs. counter-modernism, nationalism vs. internationalism: style and politics in mathematics 1900–1950. In: Goldstein C, Gray J, Ritter J (eds) L'Europe mathématique/mathematical Europe. Histoires, mythes, identités/history, myth, identity. Éditions de la Maison des Sciences de l'Homme, Paris
Misak C (2021) What are the limits of logic? How a groundbreaking logician lost control. Review of (Budianksy 2021). Times Literary Supplement, November 5, 2021
Musgrave A (1977) Logicism revisited. Br J Philos Sci 28:99–127
Nanamoli B, Bodhi B (transl) (1995) The middle length discourses of the Buddha. A new translation of the Majjhíma Níkāya. Wisdom Publications, Boston
Napper E (1989) Dependent arising and emptiness. A Tibetan-Buddhist interpretation of Mādhyamika philosophy emphasizing the compatibility of emptiness and conventional phenomena. Wisdom Publications, Boston
Naraniecki A (2015) L. E. J. Brouwer and Karl Popper: two perspectives on mathematics. Cosm Hist 11:239–255
Otto R (1926) West-östliche Mystik. Leopold Klotz Verlag, Gotha

Pambuccian V (1992) Mathematik, Intuition und die Existenzweise des Seins. Wissenschaft vom Menschen 3:87–119
Pambuccian V (2018) Negation-free and contradiction-free proof of the Steiner-Lehmus theorem. Notre Dame J Form Log 59:75–90
Papadopoulos A (ed) (2018) René Thom. Portrait mathématique et philosophique. CNRS Éditions, Paris
Pfeiffer F (1857) Meister Eckehart. J.G. Goschen, Leipzig
Placek T (1999) Mathematical intuitionism and intersubjectivity. A critical exposition of arguments for intuitionism. Kluwer Academic Publishers, Dordrecht
Poincaré H (1913) Dernières pensées. Flammarion, Paris
Polanyi M (2009) Tacit knowledge. The University of Chicago Press, Chicago
Popper KR (1967) Epistemology without a knowing subject. In: Popper KR (ed) Objective knowledge. Oxford University Press, Oxford, pp. 106–152
Popper KR (1968) On the theory of the objective mind. In: Popper KR (ed) Objective knowledge. Oxford University Press, Oxford, pp. 153–190
Posy C (1998) Brouwer versus Hilbert: 1907–1928. Sci Context 11:291–325
Posy CJ (2020) Mathematical intuitionism. Cambridge University Press, Cambridge
Pritchard P (2010) Plato's philosophy of mathematics, 2nd edn. Academia Verlag, Sankt Augustin
Rabi L (2016) Ortega y Gasset on Georg Cantor's theory of transfinite numbers. Kairos J Philos Sci 15:46–70
Rabinkow SB (1929) Individuum und Gemeinschaft im Judentum. In: Brugsch T, Lewy FH (eds) Die Biologie des Menschen. Band 4: Soziologie der Person. Urban & Scwarzenberg, Berlin/Wien, pp. 799–824
Rahula W et al (1996) Are you not saying what the Buddha said? In: Questioning Krishnamurti. J. Krishnamurti in dialogue. Thorsons, San Francisco, pp. 18–38
Robinson RH (1957) Some logical aspects of Nāgārjuna's system. Philos East West 6:291–308
Rowe DE, Felsch V (2019) Otto Blumenthal: Ausgewählte Briefe und Schriften II. 1919 – 1944. Springer Spektrum, Berlin
Sabo T (2017) Plotinus and Buddhism. Philos East West 67:494–505
Sambin G (2008) Two applications of dynamic constructivism: Brouwer's continuity principle and choice sequences in formal topology. In: van Atten M, Boldini P, Bourdeau M, Heinzmann G (eds) One hundred years of intuitionism (1907–2007). The Cerisy conference. Birkhäuser, Basel, pp. 301–315
Schelling FWJ (1809) Philosophische Untersuchungen über das Wesen der menschlichen Freiheit und die damit zusammenhängende Gegenstände. In: Schelling FWJ, Sämtliche Werke, I. Abteilung, Bd VII
Scholem G (1972) The name of God and the linguistic theory of the Kabbalah I and II. Diogenes 79:59–80. 80: 164–194
Schopenhauer A (1859) Die Welt als Wille und Vorstellung. Grossherzog Wilhelm Ernst Ausgabe. II. Teil. Insel Verlag, Leipzig
Schweitzer A (1923) Kultur und Ethik. C. H. Beck'sche Verlagsbuchhandlung, München
Schweitzer A (1974) Aus meiner Kindheit und Jugendzeit. In: Schweitzer A (ed) Gesammelte Werke, vol I. C.H. Beck, München, pp. 253–313
Smoryński C (1977) The incompleteness theorem. In: Barwise J (ed) Handbook of mathematical logic. North-Holland, Amsterdam, pp. 821–865
Smoryński C (1994) Review of [Van Stigt 1990]. Am Math Mon 101:799–802
Smoryński C (2015) Review of [Van Dalen 2013]. Math Intell 37:103–107
Soifer (2007) Review of [Van Dalen 1999b] and [Van Dalen 2005]. Zbl 1122.01019
Suzuki DT (1998) Studies in the Laṅkāvatāra sutra. Munshiram Manoharlal Publishers, New Delhi
Thom R (1970) Les mathématiques "modernes": une erreur pédagogique et philosophique? L'Âge de la Science 3:225–236. [English translation: American Scientist 59: 695–699]

Tieszen R (2008) The intersection of intuitionism (Brouwer) and phenomenology (Husserl). In: van Atten M, Boldini P, Bourdeau M, Heinzmann G (eds) One hundred years of intuitionism (1907–2007). The Cerisy conference. Birkhäuser, Basel, pp. 78–95
Tóth I (2009) Liberté et vérité. Pensée mathématique et spéculation philosophique. Éditions de l'Éclat, Paris
van Atten M (2004) On Brouwer. Wadsworth, Belmont
van Atten M (2007) Brouwer meets Husserl. On the phenomenology of choice sequences. Springer, Dordrecht
van Atten M (2015) Essays on Gödel's reception of Leibniz, Husserl, and Brouwer. Springer, Cham
van Atten M, Tragesser R (2015) Mysticism and mathematics: Brouwer, Gödel, and the common core thesis. In: Van Atten M (ed) Essays on Gödel's reception of Leibniz, Husserl, and Brouwer. Springer, Cham, pp. 173–187
van Belle M (2021) Schopenbrouwer. De rehabilitatie van een miskend genie. Marcel van Belle, Tilburg
van Dalen D (1999a) The role of language and logic in Brouwer's work. In: Orowska E (ed) Logic at work. Essays dedicated to the memory of Helena Rasiowa. Physica-Verlag, Heidelberg
van Dalen D (1999b) Mystic, geometer, and intuitionist. The life of L. E. J. Brouwer, 1881–1966. Volume 1: The dawning revolution. Oxford University Press, Oxford
van Dalen D (2005) Mystic, geometer, and intuitionist. The life of L. E. J. Brouwer, 1881–1966. Vol. 2. Hope and disillusion. Oxford University Press, Oxford
van Dalen D (2011a) The selected correspondence of L.E.J. Brouwer. Springer, London
van Dalen D (2011b) Companion to the selected correspondence of L.E.J. Brouwer. Springer, London. extras.springer.com
van Dalen D (2013) L. E. J. Brouwer – topologist, intuitionist, philosopher. How mathematics is rooted in life. Springer, London
van Stigt WP (1990) Brouwer's intuitionism. North-Holland, Amsterdam
Weil S (1949) L'enracinement. Gallimard, Paris
Weil A (1991) Souvenirs d'apprentissage. Birkhäuser, Basel
Whitehead AN (1979) Process and reality. The Free Press, New York
Wiessing H (1960) Bewegend portret. Moussault, Amsterdam
Wille M (2020a) Nach 1879. Über die Anfänge der Polemik im Werk Gottlob Freges. In: Wille M (ed) Fregesche Variationen. Essays zu Ehren von Christian Thiel. Mentis Verlag, Munster, pp. 189–220
Wille M (2020b) >alles in den Wind geschriebene<. Gottlob Frege wider den Zeitgeist. Mentis Verlag, Münster

Mathematical Practice as Philosophy, with Galois, Riemann, and Grothendieck

Arkady Plotnitsky

Contents

1 Introduction 702
2 Philosophy: The Architecture of Concepts and Mathematical Practice as Philosophy 706
3 Algebra: Abel and Galois 716
4 Geometry: Lobachevsky and Riemann 722
5 Algebraic Geometry: Weil and Grothendieck 735
6 Cross-References 745
References 745

Abstract

The primary aim of this chapter is to consider mathematicians' *working* philosophy of mathematics as emerging in and defining actual mathematical thinking and practice, as exemplified in three cases stated in my subtitle: Niels Henrik Abel and Évariste Galois, Nikolai Lobachevsky and Bernhard Riemann, and André Weil and Alexander Grothendieck. I speak of "and" and hence "conjunction," rather than the disjunctive "or," in all three paired cases considered in this chapter because, while my primary concern is on the (historically) second figure – Galois, Riemann, and Grothendieck – in each case, my aim is not merely to juxtapose these figures, especially given the significance of their thinking for transforming mathematics. The accomplishments of Abel, Lobachevsky, and Weil were revolutionary events as well. Instead, while granting the differences between their thinking, my aim is to explore the shared grounding that gives rise to these differences, defining, and defined by, their mathematical practice as philosophy. The chapter's approach to their mathematical practice and to creative mathematical practice in general is parallel to that of Deleuze and Guattari to creative philosophical practice, which or even true philosophy itself is defined by them as

A. Plotnitsky (✉)
Literature, Theory, and Cultural Studies Program, Philosophy and Literature Program, Purdue University, West Lafayette, IN, USA
e-mail: plotnits@purdue.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_97

the practice of the invention of new concepts, with philosophical concepts given a particular definition, in part in juxtaposition to mathematical and scientific concepts. The working philosophy of mathematics this chapter considers, under the heading of "mathematical practice as philosophy," is, analogously, defined as the practice of the invention of new mathematical concepts, defined, against the grain of Deleuze and Guattari's argument, in affinity with (although not identically to) philosophical concepts in their definition.

Keywords

Concepts · Problems · Algebra · Geometry · Non-Euclidean geometry

1 Introduction

It is not uncommon in the historical and philosophical studies of mathematics to discuss mathematicians' own philosophical reflections on mathematics. While such reflections will also be addressed in this chapter, its primary aim is different. It is to consider mathematicians' *working* philosophy of mathematics as emerging in and defining actual mathematical thinking and practice, as exemplified in the three cases stated in my subtitle: Niels Henrik Abel and Évariste Galois, Nikolai Lobachevsky and Bernhard Riemann, and André Weil and Alexander Grothendieck. The chapter's approach to their mathematical practice and to creative mathematical practice in general is parallel to that of Gilles Deleuze and Félix Guattari to creative philosophical practice, as defined by the invention of new concepts, with philosophical concepts given by them a particular definition, in part in juxtaposition to mathematical and scientific concepts (Deleuze and Guattari 1996). The working philosophy of mathematics this chapter considers, under the heading of "mathematical practice as philosophy," is, analogously, that of the invention of new mathematical concepts. Against the grain of Deleuze and Guattari's argument, mathematical concepts will be understood here in affinity with rather than in juxtaposition to philosophical concepts, as the latter are defined by Deleuze and Guattari. This invention can sometimes be collective, either as a product of collaboration, more common in twentieth and twenty-first century mathematics, or of independent work by different mathematicians resulting in the same concepts. Both situations have been less common in philosophy, where the invention of new concepts is rarely collective. No such invention is ever entirely individual, because every concept, no matter how innovative, has a history defined by other concepts from which a new concept emerges and without which it would not be possible. A new concept is always a product of an interplay of continuities with and breaks from preceding concepts, although the balance of continuities and breaks can be different. Galois's and Riemann's concepts, such as that of a group by Galois and that of a manifold [*Manningfaltigkeit*] by Riemann, both especially central to this chapter, were more radically innovative, but still not completely, because they had continuities with earlier concepts, such as

(among others) those of Adrien-Marie Legendre in the case of Galois and those of Karl Friedrich Gauss in the case of Riemann.

This form of mathematical practice becomes especially important at a time when one or another part of mathematics and mathematical practice needs to be transformed, often in response to a crisis on one or another scale, from that defined by a single problem to that affecting a large area of mathematics. Such transformative events, leading to practice as philosophy, are also found in science, such as physics or biology, or of course in philosophy. One might argue that philosophy, especially as enacted in the work of creators of original and powerful concepts, such figures as Plato (ideas), René Descartes (cogito), Gottfried Leibniz (monads), or Immanuel Kant (things-in-themselves), is always practiced at the time of crisis, because each such philosopher always sees philosophy in crisis, to which such a concept responds. Or, to adopt Thomas Kuhn's language (without fully subscribing to his argument, which is difficult to apply in philosophy and which requires qualifications in mathematics and science), philosophy as a creation of concepts only admits a "revolutionary" practice, with a "normal" practice having at most only auxiliary importance for the advancement of philosophy, say, by exploring or applying new concepts (Kuhn 2012). In mathematics or science, both revolutionary and normal practices appear to be equally important for its advancement, as Kuhn recognizes, although they tend to be in a more complex interplay than was argued for by Kuhn. Thus, overtly normal, in *technical* terms, mathematical practice can, by reinterpreting the concepts used, adopt a different form of this practice from that of the revolutionary mathematics it follows and thus introduce a new, "revolutionary," form of mathematical practice.

Things are of course not so straightforward. There is more to a normal practice in philosophy than meets the eye. Institutional philosophical fields, such as analytic philosophy, are based on normal practice, which made important contributions to our understanding of philosophical concepts or theories, and, it might be added, of mathematical and scientific concepts in the analytic philosophy of mathematics or science. The distinction between the normal and the revolutionary, or the importance of each, is not unconditional and requires caution in any domain. There tends to be more continuity than the appearance of a break might suggest, or across a longer history with many revolutions. Thus are the relationships between Cantor's set theory and category theory, especially in Grothendieck's hands, those of a revolutionary break (a more common view) or that of continuity, a view advocated, for example, by Yu. I. Manin, who prefers to see the history of mathematics, as against, that of physics, in terms of continuity (Manin 2002, p. 8, 2019, pp. 129–130)? Perhaps they are equally both, in a complex interplay, in either field. The cases considered here, including that of Grothendieck, could be seen in terms of this interplay. They were revolutionary *events* insofar as they radically transformed the mathematical situation in relation to which they emerged. It may not always be possible, and perhaps is never ultimately possible, to *fully* track how new trajectories emerge from and against previous trajectories of thought. This limitation, however, leaves room for meaningful assessments of this emergence in terms of both continuities and breaks, in the case of breaks, I argue, in particular by virtue of the

invention of new concepts or conglomerates of concepts, some new and some already in place, for example, conglomerates comprising and defining new theories, such as algebraic geometry, as transformed by Grothendieck.

There are other cases that would illustrate my argument in this chapter, for example, that of Richard Dedekind and Leopold Kronecker, a famous example of two opposing (mathematical) philosophical positions. The main reasons for considering the cases selected here is that the first two stand at the origin of what we now see as *modern* mathematics, respectively, modern algebra and geometry – two ways of mathematical thinking (broadly conceived, so as to include arithmetic in the first and topology in the second) and the relationships between them that governed the history of mathematics since Pythagorean mathematics to our own time. Both and their relationships were transformed by modern mathematics that emerged roughly around 1800, although some of its trends can be traced to the mathematics, such as analytic geometry and calculus, contemporaneous with the rise of modernity itself as a cultural formation, defined by a series of transformations – scientific, technological, economic, political, and philosophical – from, roughly, the sixteenth century on. Of course, when it comes to modern mathematics, too, the thinking and work of other figures, such as Gauss in the case of algebra (his geometrical ideas will be considered in conjunction with Riemann's work), was important and some of this thinking and work will be noted as this chapter proceeds. As I argue here, however, two concepts, defining respectively Galois's algebraic and Riemann's geometrical thinking, that of a group and that of a manifold [*Manningfaltigkeit*] respectively, had an unmatched shaping influence on the subsequent developments of algebra and geometry, and in the case of a group, on both, and much of modern mathematics, as well as physics and other sciences, where, especially in physics, the concept of a manifold had played a role as well. Galois also introduced a related concept, that of a group action on another mathematical entity, such as a manifold in geometry or a Hilbert space in quantum physics, a concept that has been important in mathematics and science. Another concept invented by Riemann, that of a Riemann surface, has played a role in the development considered in this chapter, beginning with Riemann's own work, including his invention of the concept of manifold.

While one of the most important and remarkable examples of bringing algebra and geometry together in algebraic geometry and, as such, the main reason for being considered here, the case of Weil and Grothendieck is different. It belongs to twentieth-century mathematics, with the work of Grothendieck having a unique impact on subsequent mathematics, rather than to the rise of modern algebraic geometry in the nineteenth century. This rise followed non-Euclidean geometry (and its effects on the development of projective geometry), on the one hand, and, and on the other, the theory of Abelian integrals, which led Riemann to his concept of a Riemann surface (mathematically, a complex manifold), thus making Riemann a key figure in this development as well. As will be seen, his ideas also had a more direct impact on both Weil's and Grothendieck's work, as did Galois's theory.

I speak of "and" and hence "conjunction," rather than the disjunctive "or," in all three cases considered because, while my primary concern is on the (historically) second figure – Galois, Riemann, and Grothendieck – in each case, I am not merely

juxtaposing, let alone opposing, these figures to each other, especially given the significance of their thinking for transforming mathematics. The accomplishments of Abel, Lobachevsky, and Weil were revolutionary events as well. Instead, while granting the differences between their thinking, my aim is to explore the shared grounding that gives rise to these differences, defining, and defined by, their mathematical practice.

Realizing that mathematics itself or its basic fields cannot be given a single definition, I shall state a general understanding of these fields which is sufficient for my purposes and is, I would contend, in accord with how they are generally viewed. I understand algebra as the mathematical formalization of the relationships between symbols, arithmetic as dealing specifically with numbers, geometry as the mathematical formalization of spatiality, especially in terms of measurement, and topology as the mathematization of the structure of spatial-like objects apart from measurements, through their continuity or discontinuity. The corresponding mathematical fields are algebra, number theory, geometry, and topology. Analysis deals with questions of limit and related concepts, such as continuity, differentiation, and so forth. There are multiple intersections between these fields, and fields that branch off them.

On the other hand, defining algebra as the mathematical formalization of the relationships between symbols makes it part of all modern mathematics. Geometrical and topological mathematical objects always have algebraic components, while algebraic objects may, but need not, have geometrical or topological components. The term algebra has other meanings, such as that (standing at the origins of algebra as a mathematical discipline) of the study of algebraic equations, or that of algebraic structures, such as groups. These forms of algebra are important to both algebra itself and other mathematical disciplines, such as geometry, topology, or analysis. In defining algebra as the mathematical formalization of the relationships between symbols, however, it is not only a matter of having an algebraic component as part of the mathematical structure of geometrical, topological, or analytical objects, but also and especially of defining these objects themselves algebraically.

My argument, then, gives preference to the role of conceptual thinking in mathematics rather than to logical propositions or calculations. Both logic and calculations are essential to mathematics, including conceptual thinking there, because the structure of mathematical concepts is logical, and because calculations require concepts and can lead to new concepts. While, however, fully respecting these features, this chapter views creative mathematical thinking as most essentially defined by the invention of new concepts, a view based on a specific definition or *concept* of concept and of a mathematical theory as an organized assemblage of concepts, as considered in Sect. 2. Mathematics, I argue, advances most by the inventions of new concepts, which defines what I call "mathematical practice as philosophy," a *mathematical* philosophy. For this mathematical philosophy, especially in its emergence, could be analyzed in terms of a more general *philosophy of mathematics*, in part by means of nonmathematical concepts, as it is in this chapter.

The next section outlines a philosophical grounding of this chapter's argument. The subsequent sections are devoted to the three case studies – Abel and Galois,

Lobachevsky and Riemann, and Weil and Grothendieck – as concrete instantiations of this architecture as mathematical practice.

2 Philosophy: The Architecture of Concepts and Mathematical Practice as Philosophy

The concept of a *mathematical concept* to be proposed here is homomorphic to Deleuze and Guattari's concept of a *philosophical concept* proposed in *What is Philosophy?* (Deleuze and Guattari 1996). This homomorphism is not an isomorphism. It preserves the differences (which would be difficult to avoid) between philosophical and mathematical or scientific concepts, and thus the disciplinary specificity of each field, in particular a greater role of logical propositions or formal and quantitative elements in mathematical and scientific concepts vis-à-vis philosophical ones. This role was emphasized by Deleuze and Guattari in their argumentation concerning the difference between and even juxtaposition of philosophical and mathematical or scientific concepts. As discussed below, in the present view (if not strictly in that of Deleuze and Guattari), this difference is also related to the sharpness of the definition of a concept. Deleuze and Guattari's argumentation goes beyond arguing for the difference between philosophical and mathematical or scientific concepts. They see their concept of a philosophical concept as belonging only to philosophy in its *defining* features, while denying these features to mathematics and science, defined by logical propositions and formal mathematical or scientific elements, such as functions. "The concept," they say, "belongs to philosophy and only to philosophy" (Deleuze and Guattari 1996, pp. 11–12, 33–34). They might have argued that the homomorphism proposed here is either too limited or locates (strictly) philosophical subconcepts within mathematical or scientific concepts, something to which they refer as an "interference" (Deleuze and Guattari 1996, pp. 217–218). I argue that this homomorphism is sufficiently rich to be useful and goes beyond such interferences, although the latter are sometimes found in mathematical or scientific concepts. Ultimately, this type of objection would not matter for my argument, insofar as the present definition of a mathematical concept is workable, as I hope it is, for our understanding of what mathematical concepts are and how they work in mathematics. It is this understanding, rather than Deleuze and Guattari's philosophy, that is my concern, helped as my argument may be by theirs or those of other philosophical figures considered here, such as Kant.

In some respects, this chapter returns Deleuze and Guattari's view of concepts to its mathematical origin, given that their concept of a concept was in part modeled on a mathematical concept, that of Riemann's concept of manifold. In addition, as will be seen, part of the genealogy of their argument concerning philosophical concepts is in Abel's and Galois's, in my terms, mathematical practice as philosophy, considered in Deleuze's earlier analysis of "ideas," on which Deleuze and Guattari's argument builds (Deleuze 1995, pp. 180–181). Riemann's concept of manifold is, admittedly, not the only origin of their concept of a concept, which also has a philosophical genealogy. The mathematical genealogy of this concept is, however, important and

paradigmatic historically, beginning with Plato, who, too, both modelled his philosophical concepts on mathematical ones and yet claimed an essential distinction between them (Negrepontis 2019, p. 24, Plotnitsky 2021a, pp. 649–651). Riemann's ideas, as well as those of Galois and Abel, had a major impact on Deleuze and Guattari's philosophy in general, especially, again, Riemann's *concept* of manifold, beyond being a model for their concept of a concept. They saw this concept as heralding a revolutionary philosophical change. This view was helped by the French term "*multiplicité*" as a translation of Riemann's *Mannigfaltigkeit* (which the English translation of Deleuze and Guattari's works often renders, understandably but incorrectly, as "multiplicity"). They say: "It was a decisive event when the mathematician Riemann uprooted the manifold from its predicative state and made it a noun, manifold [*multiplicité*]. It marked the end of dialectics and the beginning of the typology and topology of manifolds" (Deleuze and Guattari 1987, p. 483; translation modified).

Of course, even if mathematical concepts were the only origins of new philosophical ideas or concepts (to the degree that such a single origin is ever possible), this would not challenge claims, such as those of Plato, Hegel, or Deleuze and Guattari, for the fundamental difference between mathematical and philosophical concepts, because these mathematical concepts could be transformed into philosophical concepts in this transition or translation. My argument, however, is that this difference may not be as fundamental as Deleuze and Guattari contend, insofar as arguably the most essential aspect in their definition of a concept, which makes it a composition of its components, a composition specific to each concept, equally defines a mathematical or scientific concept, in the present definition. As will be seen, another key feature, which makes a concept a problem, that they associate strictly with philosophical concepts, can be transferred, even if with qualifications, to mathematical and scientific concepts as well, at least to some of them, such as that of Galois's concept of a group or Riemann's concept of a manifold.

Riemann's thinking and practice provide, arguably, the most famous example of the expressly conceptual approach to mathematics, especially in geometry, where his thinking, leading him beyond all previous approaches to geometry, including those leading to non-Euclidean geometry, abandoned Euclid's axiomatic-theorematic method, previously dominant in foundational thinking concerning geometry. The axiomatic aspects of geometry were at most secondary for Riemann, as against Lobachevsky's and Jànos Bolyia's work, leading each to the discovery of non-Euclidean (hyperbolic) geometry. As I shall argue, however, conceptual thinking plays an essential role in the thinking of all figures considered in this chapter, although the relative balance between conceptual and other forms of mathematical thinking and practice, such as the axiomatic one, is different in each case. Lobachevsky in effect made axioms or postulates of geometry, ultimately all axioms and postulates (rather than only the fifth postulate of Euclid), hypotheses, based in conditional assumptions or conventions, and thus, in the first place, concepts. Intentionally or not, his discovery deprived these axioms and postulates of their status as self-evident or preexisting truths that needed to be assumed or discovered, rather than invented, as concepts and thus geometries always are in the present view.

Ultimately, all axioms, not the least those of Euclid, are conceptual inventions, even when their origins as such are lost. Moreover, as Gödel's theorems tell us, mathematical axioms, assumed to reduce the complex to the simple, are defined by the irreducible complexity that precedes the simple or what appears to be simple. Imre Lakatos, citing Hermann Weyl, referred to Cantor's set-theoretical paradise (as David Hilbert famously dubbed it), as "a bold theoretical construction, and as such the very opposite of analytical self-evidence" (Lakatos 1980, p. 31, n. 3, Weyl 2021, p. 64). This is no less true of Euclid's geometrical paradise.

It is, to reiterate, not my intention to diminish the role of logical or calculational aspects of mathematics or science, or to take anything away from the contribution of those philosophical investigations, such as those in the analytic philosophy of mathematics, that focus primarily on logical and propositional structures in mathematics or science (There are new powerful and rich approaches to logic in mathematics and beyond, such as homotopy type theory (Voevodsky et al. 2013; Corfield 2020). I would contend, however, that concepts still play an essential and unavoidable role in them). My aim instead is to place a proper emphasis on the role of concepts, especially in *creative* thinking in mathematics or science, but not only there, because working with already established concepts is indispensable in all mathematical or scientific practice. Besides, analytic philosophers, too, sometimes, give concepts their due, beginning with one of the founders of analytic philosophy, sometimes even seen as *the* founder of it, Gottlob Frege, who gave priority to the role of concepts in mathematics and philosophy. The unconditional opposition between the logical-propositional and conceptual structures, or between calculations and concepts, is not so easy to maintain, even in considering figures such as Galois and Riemann, or Dedekind. While their thinking was primarily governed by concepts vis-à-vis calculations, they were perfectly capable of and performed difficult calculations. And then, logic and calculations do involve concepts and lead to new concepts, as they did in the case of Abel, or more pointedly (as he insisted on their primacy) Kronecker. Calculations can also function as statements, sometimes even as philosophical statements, thus connected to concepts. (There are still other aspects of mathematical thinking, for example, a narrative one, the role of which, as constitutive rather than merely auxiliary, which has always been recognized, has received a considerable amount of attention during recent decades. For a representative collection of essays on the subject, see Doxiadis and Mazur (2012). For a full disclosure, the present author was among the contributors.) Calculations, too, involve narrative dimensions. My main emphasis in this article is, however, on concepts as, arguably, the primary creative vehicles of mathematics.

In the present understanding, then, extending that of Deleuze and Guattari beyond philosophy, concepts are essential and even primary vehicles of creative thinking in theoretical fields. What is a concept, then? First, a concept is not merely a generalization from particulars (which is commonly assumed to define concepts) or a general or abstract idea, although a concept may contain such generalizations or ideas, specifically abstract mathematical ideas in mathematics or physics, where these abstractions may also become concepts. A concept is a multicomponent structure, defined by the organization, *composition*, of its components, and some

of these components may be concepts in turn. This is, however, only a general, *abstract*, definition. What meaningfully defines a concept is the specific character of its compositional structure, defined by both the nature of each component and the character of their relations within this composition, by how these components relate to each other in the structure of this concept. Such components may be more general or particular, and some of them are concepts in turn. This composition is unique in a new concept, and make this concept unique. The creative essence of concepts is in their compositional individuality, which is only possible by virtue of their generally compositional nature, but is defined by the specificity of this composition. All musical works are compositions, but each gives us different music. A group is composed of its elements, but in accordance to very specific rules, which give it its structure and role in mathematics in group theory and beyond, and then it multiplies further to still more specific concepts of a group, a finite group, a Lie group, an algebraic group, and so forth, and then into further concepts within each of these.

A concept is a multicomponent conglomerate of concepts, figures, metaphors, and so forth, which are conjoint in a composition, a *relational organization* of the components. It is rare for a concept to have only one component, and ultimately impossible to do so. As Deleuze and Guattari say, "there are no simple concepts. Every concept has components and is defined by them. . . . There is no concept with only one component" (Deleuze and Guattari 1996, p. 16). A single-component concept, such as an element of a group, is a product of a provisional suspension or cut-off of its multicomponent organization, such as that defining the concept of an element, usually taken as a primitive concept as it may be. In fact, there are always cut-offs in delineating a concept, cut-offs that result from assuming some of the components of this concept to be primitive entities whose structure is not specified. These components could, however, be specified by alternative delineations, leading to a new concept, containing a new set of more primitive components.

Consider as an example the mathematical concept of space, discussed in detail in connection with Riemann's rethinking of the foundations of geometry in Sect. 4. Emerging historically at the intersection of mathematics, physics, philosophy, and general phenomenal intuition, and defined by such constitutive concept-components as point, line, plane, or distance, as philosophical, physical, and mathematical concepts, this concept and each of these components have a long history of definitions, modifications, transformations, and so forth. As a general phenomenal concept, it is arguably as old as is human consciousness (any concept requires consciousness to be defined as a concept). Kant did not see space and time as concepts, but rather as a priori forms of our phenomenal intuition. He had a point, even if one rejects (as many, including Riemann, have done) his a priori view of space and time, insofar as these phenomenal forms precede any concept of space and time our general phenomenal thinking can form, as Kant realized they can. As a geometrically representable physical concept, space is at least as old as ancient Greek geometry, well preceding Euclid, who does not define, although, arguably, implies space in the *Elements*. Its history as a mathematizable physical and eventually (that happened only in the nineteenth century) purely mathematical concept

extends, to mark some of its junctures, from Euclid to Descartes (a coordinate space) to Riemann (a space defined as a manifold) to Felix Hausdorff (topological space) to Grothendieck (topos). For the first three figures, the concept of space qua space was, while mathematizable, still physical. The concept also has its algebraic aspects and history. After each transformation, such as that of defining it as (representable by) a manifold by Riemann, the new concept of space, became, to stay with Kuhn's idiom, part of the *normal* practice in mathematics, or physics, during which it acquired further features and new definitions.

A concept, then, is an assemblage of its components connected by its composition, which define its specific individual nature. Some of its components may be unique as well but need not be. The components of a concept need not be new, but its composition must be new for a concept to be new. In this respect a concept is akin to a work of art, always defined by its composition, in accord with the ancient Greek meaning of *poiein* (to make), as making something new, in which the term poetry originates. A theory, then, or a disciplinary field, mathematical, scientific, or philosophical, such as geometry, classical physics, or Kant's philosophy (philosophy, again, tends to be more individualized), becomes an organized assemblage of concepts. It is a product of a broader *poiein*, in mathematics or science, more commonly collective, although it may be initiated by a single new concept, such as that of Galois's concept of a group or Riemann's concept of a manifold. In the present view, however, a *creative* conceptual *mimesis* (which may borrow elements of previous concepts but combines them in a new way) is a *technical* mimesis of previous concepts, rather than a Platonist mimesis of some preexisting independent reality, assumed to govern this mimesis.

I am not equating theoretical concepts (or theories) with works of art, either in their compositional nature or in their functioning. For one thing, a work of art, as a composition, may but need not contain concepts, although so-called conceptual art does. In functioning, a concept can be used directly or with minor modifications in, again, in Kuhn's terms, the normal practice in the corresponding theoretical field, especially mathematics and science. As noted, philosophy, at least a certain form of philosophy, may be associated, as it is in Deleuze and Guattari, with the invention of concepts as a revolutionary practice, which thus becomes its normal practice. In this respect, this practice is closer to that of art, which, while it may use a tradition, is compelled to be innovative in each work, but only in this respect, because the nature of the conceptual and artistic composition remains different. The relationships among mathematics, philosophy, and art have been an important and sometimes defining subject in the history of philosophy, from Plato on. It is, accordingly not surprising that art, as defined by artistic composition, versus philosophy, as defined by concepts, or versus mathematics and science, as defined by formal structure and logical propositions (and, in science the concept of reference), was an important part of Deleuze and Guattari's argument for defining philosophy. Art is beyond my scope here. On the other hand, it is difficult to bypass the fact that Deleuze and Guattari deny that concepts in their sense are found in mathematics or science, a denial that can be misleading for understanding the nature of conceptual thinking in either domain.

They even declare that "it is pointless to say that there are concepts in science [including mathematics]," adding that "even when science is concerned with the same 'objects' [as philosophy] it is not from the viewpoint of the concept; it is not by creating concepts" (Deleuze and Guattari 1996, p. 33). This, in my view problematic or at least insufficiently qualified, assessment also includes Riemann's thinking concerning manifolds and spatiality: "Riemannian functions of space [one presumes, those defining the mathematical structure of his manifolds], for example, tell us nothing about a Riemannian concept of space peculiar to philosophy: it is only to the extent that philosophy is able to create it we have a concept of a function" (Deleuze and Guattari 1996, p. 161). That may be true as concerns some philosophical "Riemannian" concepts of space, such as that of Henri Bergson (to whom Deleuze often refers in this connection). Perhaps this is what this statement primarily refers to, insofar as at stake in it is a Riemannian concept of space peculiar to philosophy and not Riemann's concept of space as a mathematical concept. Still, it would be difficult to agree that "it is only to the extent that philosophy is able to create it we have a concept of a function." The opposite appears to be the case: one can only have a function to the extent that mathematics was able to create its concept or the concept of function in the first place. This is certainly so in the case of those functions that appear in Riemann's concept of manifold or in his analysis of functions of a complex variable, which led him to his concept of a Riemann surface. Deleuze and Guattari add irrational numbers as another example, "the concept, on the other hand, refers not to series of [rational] numbers [series defining irrational numbers mathematically] but to a string of ideas that are reconnected over a lacuna (rather than linked by continuation)" (Deleuze and Guattari 1996, p. 161). Is not, however, a new mathematical concept, such as that of an irrational number (any such concept) equally defined by a string of new ideas reconnected over a lacuna left by earlier ideas? Was it not how Dedekind invented his concept of real numbers? One could in principle argue, as Deleuze and Guattari might have, that a mathematical definition of a concept by putting together its components, such as that of a group as a particular organization of its elements, represents a philosophical, rather than mathematical, thinking within mathematics, and as such defines a philosophical concept that underlies this mathematical concept. As noted, Deleuze and Guattari see this situation, or more generally, the relationships between philosophical and mathematical or scientific thinking, or between either thinking and art, in terms of "interference" with each other (Deleuze and Guattari 1996, pp. 217–218). There may be philosophical or, again, artistic dimensions to mathematical or scientific thinking, especially in its intuitive part, involved, for example, in plausible reasoning (Polya 1990; Corfield 2006). It would, however, be difficult to limit this type of creation in this way, even given the *exact* nature of mathematical or scientific concepts and thinking, which distinguishes them from *inexact* but (philosophically) rigorous philosophical concepts and thinking. This difference between the exact and the inexact is invoked by Deleuze and Guattari, in part, it is worth noting, via Edmund Husserl's analysis of the origin of geometry as an exact science vs. inexact, but, again, rigorous, philosophical thinking (Deleuze and Guattari 1987, pp. 407, 483; Husserl 1970). It may ultimately be impossible to strictly maintain this difference, as the exact and inexact

can coexist and invade each other, but it is not possible to disregard it either, these gray areas notwithstanding. One might take advantage here of an astute and far-reaching remark of Heisenberg, aimed especially against Kant, in view of quantum physics, but leading Heisenberg to a broader point, which equally and even more directly applies in mathematics:

> Any concepts or words which have been formed in the past through the interplay between the world and ourselves are not really sharply defined with respect to their meaning; that is to say, we do not know exactly how far they will help us in finding our way in the world. We often know that they can be applied to a wide range of inner or outer experiences, but we practically never know precisely the limits of their applicability. This is true even of the simplest and most general concepts like "existence" and "space and time." Therefore, it will never be possible by pure reason to arrive at some absolute truth [in physics or even mathematics as Kant thought it might be].
>
> The concepts may, however, be sharply defined with regard to their connections. This is actually the fact when the concepts become a part of a system of axioms and definitions which can be expressed consistently by a mathematical scheme. Such a group of connected concepts may be applicable to a wide field of experience and will help us to find our way in this field. But the limits of the applicability will in general not be known, at least not completely. (Heisenberg 1962, p. 92)

Kant is not as far from much of Heisenberg's argument as Heisenberg thinks, and if Kant could have benefited from modern physics, especially quantum theory, or mathematics, beginning with non-Euclidean geometry, he might have agreed with Heisenberg as concerning the role of causality, which he strictly maintained but which was challenged by quantum theory. I shall discuss the question of Kant's claim concerning our capacity to "arrive at absolute truth" in connection with Riemann's views of space (a physical concept for Riemann) and geometry in Sect. 4. Kant's analysis is also characterized by a search for a *sharper* definition and the limits of applicability of concepts, in part through establishing the connections between them, even if not as embedded in "a mathematical scheme." It may well be, however, that, although, by virtue of establishing more rigorous connections between concepts, a definition of an inexact philosophical concept is sharper than in the case of our daily concepts, this definition cannot be as sharp as that of a mathematical concept. Such a mathematical concept can also be used in physics, although a physical concept that uses this mathematical concept may not be defined as sharply as a mathematical concept either.

If one adopts this view, as I do here, the difference between *inexact philosophical* and *exact mathematical* concepts will be maintained as essential and as defining mathematical concepts, on this point in accord with Deleuze and Guattari's view. On the other hand, the compositional character, as equally defining both types of concepts, will be maintained, thus challenging Deleuze and Guattari's unequivocal claim that "the concept belongs to philosophy and only to philosophy," a claim that they appear to extend to the compositional character of concepts. It is conceivable that they would have accepted the view proposed here. Generally, however, they see the relationships between inexact philosophical and exact mathematical or scientific concepts or thinking as either operative alongside each other or in terms

"interference" (in the positive sense of interfering wave fronts rather than in the sense of inhibition), as when "a philosopher attempts to create the concept of . . . a function (for example, a concept peculiar to Riemannian space or to irrational number)" (Deleuze and Guattari 1996, p. 217). By contrast, in the present view, creative mathematical and scientific thinking is defined by the creation of multi-component mathematical or scientific concepts, compositionally, if not *functionally*, analogous or homomorphic to that of philosophical concepts in Deleuze and Guattari's sense, with each concept defined by the specificity of its composition. Just as there are Plato's ideas, Descartes's cogito, and Leibniz's monads, there are Gauss's curvature, Riemann's manifolds, Dedekind's ideals, and Grothendieck's topoi in mathematics. It may be that in mathematics and science such signatures below the concept are erased by history and the concepts themselves more commonly modified or translated by means of new concepts than in philosophy. But, apart from the fact that this is not always so, as in the case of Riemann surfaces or Hilbert spaces (redefined as each has been many times as a concept), this forgetting does not diminish the significance of the creation of concepts. These conceptual compositions are not *isomorphic* because of technical or, again, *exact* aspects of mathematical and scientific concepts, which enables a greater and perhaps maximal possible, even if, as Heisenberg says, never complete, sharpness of their definitions, which is demanded by and has defined the disciplinary nature of mathematics and science throughout their history.

Another difference, noted earlier, is, in Kuhn's terms, in the nature of *normal*, rather than *revolutionary*, practice in each domain. In philosophy, normal practice consists primarily in understanding, interpreting, and commenting on concepts, while in mathematics and science, the normal practice consists primarily (while still reflecting on and interpreting concepts) in working with frames of reference and other mathematical or scientific formations, propositions, and so forth. It is true that for Deleuze and Guattari revolutionary philosophical practice is the primary form of philosophy. This is, however, not in conflict with the view of creative practice in mathematics or science advocated here, insofar as this practice is equally defined by the invention of concepts.

In addition, according to Deleuze and Guattari, a philosophical concept in their sense is also defined as a problem, as positing a problem along with and even *as* this concept. This view has a mathematical genealogy, although the concept of a problem well preceded mathematics and was adopted by it and given a mathematical specificity by ancient Greek mathematics. A problem in this sense is not something that, like a theorem, is derived from assumed axioms by means of logical rules, but is something that is posed, created, along with and as a concept. A theorem could be a problem in this sense, when it arises, from a mathematical concept, such as that of Riemann's concept of manifold or Galois's concept of a group. In Deleuze and Guattari's understanding, however, insofar as it remains a philosophical concept, a problem persists rather than disappears in its solutions, which fact enables such a "concept-problem," as it may be called, to remain "always new," insofar as, as I understand it, its productive significance is retained in the history following its invention (Deleuze and Guattari 1996, p. 5). This is a complex conception, which

is not easy to interpret consistently, let alone definitively. The view of concepts as "concept-problems" that I adopt as applicable to mathematical concepts may only be a *version* or a variation on the theme, rather than an interpretation, of Deleuze and Guattari's understanding of concepts as problems. This view, however, appears to me to be workable in mathematics.

First of all, in this view, not all mathematical or scientific concepts as defined above are concept-problems, while all philosophical concepts in Deleuze and Guattari's sense appear to be for them, which is, however, a secondary matter here. My point here is that some mathematical or scientific concepts are in a way analogous to philosophical concept-problems in Deleuze and Guattari. A new concept-problem, such as that of a group by Galois or a manifold by Riemann, would be introduced, first, just as a concept in order to solve an already existing problem, which defied a solution by means of already existing concepts or perhaps offered such a solution that was unsatisfactory for one or another reason, conceptual, technical, or aesthetic. Finding the algebraic solution to a polynomial equation (or proving the impossibility of doing so) in Galois's case, or that of properly defining any possible geometry in Riemann's case, was such a problem. Their way of thinking already reflects two approaches (predictably with gray areas between them or sometimes combining both) to solving a problem, especially a difficult problem: the first is essentially logical, which aims to solve it by means of logical and technical manipulations of previously existing concepts, and the second is essentially conceptual, which uses a new concept or set of concepts, although technical and logical manipulations remain unavoidable of course. Such a concept may be relatively simple, which is not to say trivial, such as that of a group, or (still more nontrivially) of an action of a group, or it may be more complex, such as that of a manifold, or very complex, such as those of Grothendieck. But, even when relatively simple, such a concept usually arises in response to a complex problem, as in the case of the concept of a group, invented by Galois to find an algebraic solution to a polynomial equation (and proving the impossibility of doing so), a very complex problem at the time. By be means of his concept of a group, Galois redefined this problem by establishing that whether such a solution exists or not is related to the structure of a group of permutations associated with the roots of the polynomial, known as the Galois group of this polynomial. Abel's approach to this problem had elements of this thinking by posing the question whether a solution exists or not, given the structure of the equation, and finding that in general it does not, the finding now known as the Abel-Ruffini theorem or Abel's impossibility theorem. Paolo Ruffini gave an incomplete proof in 1799, and Abel was the first to prove it in 1824.

As a *concept-problem*, however, a new concept does more than merely solving a given initial problem. It defines a new problem-posing field in which new types of actual problems emerge and are being solved, while the concept-problem that initiated and governs this field is not something that is "solved" so as to no longer be necessary for this field to continue to develop, which it does by defining new concepts and concept-problems (some of which are new forms of the original concept-problem). Doing so is a far-reaching form of mathematical practice as

philosophy. Galois's concept of a group not only gave the Abel-Ruffini theorem a more general grounding but also, now as a concept-problem, gave algebra a new problem-posing field, Galois theory. This type of invention of a new concept in response to a concrete problem not only in solving this problem but also making this concept a concept-problem, which leads to a new problem-posing field, can happen as an outcome of a deliberate strategy, as it was arguably the case in Grothendieck or (although it is more difficult be certain) Galois and Riemann, or it might happen contingently, in a complex interplay between chance and necessity. In this sense, such concept-problems give rise to theories, such as Galois theory, Riemann geometry, or Grothendieck's topos theory.

Deleuze and Guattari would have probably seen such concept-problems as philosophical, even when they are operative in mathematics, while seeing mathematics as being about solutions of its problems, a relation that, as discussed in Sect. 3, Deleuze suggested earlier, via Albert Lautman and Jules Vuillemin, in speaking of ideas, rather than concepts, but conceived analogously, still using Abel's and Galois's thinking as a key model of this process (Deleuze 1995, pp. 163–164, 179–182). It is true that mathematics is commonly in pursuit of solutions of its problems, and there is often a philosophical dimension to defining such a new problem-posing field. However, in the present view, as just outlined (which I am, again, not attributing to Deleuze and Guattari), a concept-problem, such as a group, is *mathematical*. It still enables one to solve, or to advance the possibility of solving, concrete problems, beginning, in Galois's case, with a problem that gave rise to the concept of a group – the problem of finding the algebraic solution to a polynomial equation – but it does so by embedding this original problem in a new *mathematical* concept-problem, which defines a new problem-posing mathematical field, or even a new field of mathematics, such as Galois theory in algebra and in effect much more. This concept-problem is, accordingly, mathematical, including, again, by virtue of the sharpness of its formal definition, possible for all mathematical concepts, but difficult to achieve outside mathematics.

That need not and here does not mean that such a concept-problem contains or fully determines a problem-posing field it creates, even if only as a potentiality, in a Platonist fashion. This field emerges in a contingent development, as a complex interplay of causal chains and chance occurrences. In general, in adopting a problem-posing versus axiomatic-theorematic view of mathematics, one is more likely to reject a Euclidean, essentially Platonist, view in which axioms are assumed as pregiven truths rather than as created as concepts, as axioms are from the present perspective, a distinction that, as discussed below, was brought into the foreground by the discovery of non-Euclidean geometry. To sharpen this difference, I shall adopt Lakatos's distinction between a Euclidean theory (paradigmatically based on Euclidean geometry) and a quasi-empirical theory, which is how Lakatos sees mathematics. ("Quasi" reflects the fact that mathematics is "empirical" in dealing with mental rather than material phenomena.) The mathematical practice of figures considered here is quasi-empirical, in parallel with being problem-posing, rather than Euclidean and, hence, axiomatic-theorematic, although some of them, such as Abel, Weil, and Grothendieck might have adopted a more Euclidean ideology of mathematics. This

split is not uncommon; and one might even argue, as Lakatos does, that no Euclidean ideology can avoid quasi-empirical practice, thus making a purely Euclidean practice impossible. According to Lakatos:

> The methodology of a science is heavily dependent on whether it aims at a Euclidean or at a quasi-empirical ideal. The basic rule in a science which adopts the former aim is to search for self-evident axioms—Euclidean methodology is puritanical, antispeculative. The basic rule of the latter is to search for bold, imaginative hypotheses with high explanatory or 'heuristic' power, indeed, it advocates a proliferation of alternative hypothesis to be weeded out by severe criticism—quasi-empirical methodology is uninhibitively speculative.
>
> The development of Euclidean theory consists of three stages: first the naïve prescientific stage of trial and error which constitutes the prehistory of the subject; this is followed by the foundational period which reorganizes the discipline, trims the obscure border, establishes the deductive structure of the safe kernel; all that is then left is the solution of problems inside the systems, mainly constructing proofs or disproofs of interesting conjectures. ([The discovery of] a decision method for theoremhood may abolish this stage altogether and put an end to the development.)
>
> The development of a quasi-empirical theory is very different. It starts with problems followed by daring solutions, then by severe tests, refutations. The vehicle of progress is bold speculations, criticism, controversy between rival theories, problem shifts. Attention is always focused on the obscure border. The slogans are growth and permanent revolution, not foundations and accumulations of eternal truth.
>
> The main pattern of Euclidean criticism is suspicion: Do the proofs really prove? Are the methods used too strong and therefore fallible? The main pattern of quasi-empirical criticism is proliferation of theories and refutations. (Lakatos 1980, pp. 29–30)

Every new concept, especially a concept-problem, is taking a chance (possibly, sacrificing one's own interests to the rigor of thought) on the future of mathematics or science, or philosophy. The invention of new concepts, as "always new" concept-problems, has, according to Deleuze and Guattari, defined the creative practice in philosophy from the pre-Socratics on. Deleuze and Guattari may have been indebted most to Friedrich Nietzsche's view of all true philosophy as "a philosophy of the future," his subtitle to *Beyond Good and Evil: A Prelude to a Philosophy of the Future* (Nietzsche 1966). One can, I argue, take the same view of many new mathematical concepts, such as that of Galois's concept of a group or Riemann's concept of a manifold, although one cannot always know in advance which. Robert Langlands (of the Langlands program fame, a program in part grounded in Galois theory) compared Grothendieck to "Nietzsche's *Philosoph der Zukunft*" (Langlands and Shelstad 2007, p. 486). This is no less true about Galois and Riemann, or Abel, Lobachevsky, and Weil, or Langlands, as the Langlands program proved.

3 Algebra: Abel and Galois

My argument in this section is indebted to that of Jules Vuillemin's analysis, over half a century ago, of Abel and Galois, and their predecessors, such as Legendre and Gauss, and of algebra itself in *La philosophie de l'algèbre* (Vuillemin 1962) and that of Deleuze, building on Vuillemin's analysis (Deleuze 1995). This is not to say that

my argument *follows* that of Vuillemin, which has a strong structuralist orientation shaped by the Bourbaki program (the book is dedicated to, among others, Pierre Samuel, a member of Bourbaki), very different from the conceptual and quasi-empirical one of this chapter. Also, the project of Vuillemin's book is to transpose the methods of algebra, such as those of Abel and Galois, into philosophy, which is not my aim either. My argument absorbs some aspects of Vuillemin's argument within a conceptual orientation, as Deleuze does, inspired by Vuillemin and Lautman, again, in transposing the methods of Abel and Galois into his theory of ideas (a precursor of his and Guattari's theory of concepts) and thus into philosophy. Nor, for the reasons explained in the preceding section in considering concept-problems, does my argument quite follow that of Deleuze either, but instead builds on and yet transforms it. According to Deleuze:

> Abel elaborated a whole method according to which solvability must follow from the form of the problem. Instead of seeking to find out by trial and error whether a given equation is solvable [in radicals] in general, we must determine the conditions of the problem which progressively specify the fields of solvability in such a way that 'the statement contains the seeds of the solution'. This is a radical reversal in the problem-solution relation, a more considerable revolution than the Copernican. It has been said [by Vuillemin] that Abel thereby inaugurated a new Critique of Pure Reason, in particular going beyond Kantian 'extrinsicism'. This same judgement is confirmed in relation to the work of Galois: starting from a basic 'field' (R), successive adjunctions to this field (R', R", R"'...) allow a progressively more precise distinction of the roots of an equation, by the progressive limitation of possible substitutions. There is thus a succession of 'partial resolvents' or an embedding of 'groups' which make the solution follow from the very conditions of the problem: the fact that an equation cannot be solved algebraically, for example, is no longer discovered as a result of empirical research or by trial and error, but as a result of the characteristics of the groups and partial resolvents which constitute the synthesis of the problem and its conditions (an equation is solvable only by algebraic means—in other words, by radicals, when the partial resolvents are binomial equations and the indices of the groups are prime numbers). The theory of problems is completely transformed and at last grounded, since we are no longer in the classic master-pupil situation where the pupil understands and follows a problem only to the extent that the master already knows the solution and provides the necessary adjunctions. For, as Georges Verriest remarks, the group of an equation does not characterize at a given moment what we know about its roots, but the objectivity of what we do not know about them. Conversely, this non-knowledge is no longer a negative or an insufficiency but a rule or something to be learnt which corresponds to a fundamental dimension of the object. The whole pedagogical relation is transformed—a new *Meno*—but many other things along with it, including knowledge and sufficient reason. Galois's progressive discernibility unites in the same continuous movement the processes of reciprocal determination and complete determination (pairs of roots and the distinction between roots within a pair). It constitutes the total figure of sufficient reason. (Deleuze 1995, pp. 179–180)

This is a radical transformation of mathematical practice as philosophy and, on this pattern, of philosophy itself, including the pedagogy of ideas or concepts, which philosophy has always been. The argument also implies that the pedagogy of mathematics, at least as it implicitly functions in mathematical practice (and possibly even in the actual teaching of mathematics, which is, however, a separate subject).

Galois's thinking has a more revolutionary significance for both Vuillemin and Deleuze. Nevertheless, according to Vuillemin, Abel transforms algebra, as the inventor of "a general method":

> In the first place, we analyze, in its most general form, a mathematical relation or a defined set of such relations, which make it possible to determine a property, of which we do not yet know whether or not we can attribute it to a class of beings: by example the character of being algebraically solvable, of being convergent, of being expressible by a defined number of functions of a certain class. In the second place, one considers the class of beings to which it is a question of attributing this property [...]; we analyze these beings from a general point of view, we define the relationships to which their nature allows them to be subjected. Finally, this double examination reveals the cases of incompatibility (demonstrations of impossibility) and, possibly, indicates the way to find the new required relationships in cases of possibility. (Vuillemin 1962, pp. 214; translation mine)

Thus, rather than looking for the solution to the equations of the fifth degree, one asks whether such a solution is possible, in the first place, for any polynomial equation. Abel's new "idea of a general method consists in giving a problem a form such that it is always possible to solve it" (Vuillemin 1962, pp. 209; translation mine). What Abel then shows is that the conditions for any equation to be solvable are not, in general, fulfilled once the degree of the equation is more than four. In this way, by reversing the method of his predecessors, which proceeded from the particular to the general, by generalizing the particular, to proceeding from the general to the particular (the general conditions of solubility of algebraic equations are satisfied for any specific equations of any degree less than five), Abel becomes one of the creators of the modern algebraic method, which Vuillemin sees as, in effect, structuralist in Bourbaki's sense. Against Kant, who limited his construction of the proof of impossibility by positing the possibility of an experience as an exclusive criterion of knowledge, "general demonstrations, in the sense of Abel, change the modality of the proof. Particular demonstrations are actual [*réelles*]: they assume as their grounding principle the possibility of the experience given in the affect of sensation. The general proofs deal with the possible and start from the concept alone, disregarding the restrictive conditions of sensitivity" (Vuillemin 1962, p. 216; translation mine). This initiating concept is, thus, defined not by a generalization from particulars, but as a new structured entity. Mathematics needed Galois in algebra (a group) and then Riemann in geometry (a manifold) to be expressly rethought in this way. Abel thus introduced a new critique (in Kant's sense of an analysis of concepts) of pure reason, at least in mathematics, but, as both Vuillemin and Deleuze argue, extendable beyond mathematics.

While (independently) adopting the same "general method" as that of Abel, Galois brings a new, an even more revolutionary, dimension to this critique by his invention of the concept of a group. Some elements of group-theoretical thinking are found in Abel, or Legendre, earlier, but the concept carries Galois's signature, and it was conceived in accord with a concept of concept as defined here. It was, moreover, accompanied by another, equally innovative and important concept, that of an action

of a group on another structured entity, also defined as a concept, such as an algebraic field or vector space.

For Vuillemin, deeply steeped in Bourbaki's ideology, the emphasis is on the idea of structure, and he sees the concept of a group as the first algebraic structure, introduced and analyzed as such, in the history of mathematics. Galois's thinking, then, is both akin to that of Abel as concerns Abel's philosophy of solving problems, as just considered, and reaches beyond Abel by bringing in the role of a structure and the concept of a group as an entity (we would now say a set) with a structure, a relationally organized multiplicity. Galois, for Vuillemin, becomes the first modern algebraist, who associates with solving a problem the analysis of the structure from which this problem arises, and which makes it possible or not to solve it, such as the problem of finding algebraic solutions of polynomial equations. Vuillemin's structuralist perspective is not out of place. One cannot avoid structures, as all concepts are structures, compositions, and hence, merely dismiss structuralism (Corry 2003). My point here is the emergence of structures as concepts.

Indeed, as in the case of Abel's thinking, one of Vuillemin's main points is more about concepts than structures, keeping in mind that a concept is a structure, too, and conversely any given structure qua structure has to be defined and thus may be a concept, including in the present sense of problem-concepts, as it indeed was for Bourbaki. Vuillemin distinguishes two types of abstraction: the first is conventional, defined by an increasing generalization from the particular to the general, and the second, found in group theory, is conceptual in the present sense. It defines the differences between specific individual cases, because it constructs these cases, from the initial concept, such as that of a group. Vuillemin sees the second abstraction or, more precisely, "formalization" in terms of a double abstraction: the first is that of the formal presentation of the group, as a concept, in terms of algebraic symbols and their relations; the second is that defined by providing a method that makes it possible "to construct individuals, no longer in intuition according to imperfect [more arbitrary] schemes, but [determinately] in the concepts themselves" (Vuillemin 1962, pp. 288–289; translation mine).

Thus, one deals with an algebraically formalized conceptual practice, in accordance with the argument of this chapter, and redefines algebra accordingly, ushering in a new mathematical practice as philosophy. In Galois's and (more extensively) in Abel's case, one also deals with a parallel shift to the conceptual algebraization of analysis, as well as, in effect, geometry (where this shift was already initiated with the analytic geometry of Fermat and Descartes). The subsequent "returns" of geometry still use and are often based on this lateralization rather than dispensing with it (Plotnitsky 2021a). In noting that, in the wake of Galois, "modern mathematics is therefore regarded as based upon the theory of groups or set theory rather than upon differential calculus," Deleuze adds: "Nevertheless, it is no accident that Abel's method consisted in giving a problem a form such that it is always possible to solve it, concerned above all the integration of differential formulae" (Deleuze 1995, pp. 180–181).

Galois's contribution was ultimately more fundamental conceptually, as based in the concept of a group as a multiplicity, in effect as a set, with a structure, an entirely

new type of concept. As a concept-problem, this concept defines a new problem-posing field through the new mathematical formulations of existing problems (such as that of the solution of polynomial equations in radicals) and posing new problems, and the simultaneous establishing of their conditions of solvability. This new thinking changes the future of mathematics as a field of solving problems, because both these problems and the conditions of their solvability change, in particular, by asking under what conditions a given problem (such as that of finding the solution of polynomial equations in radicals) can be solved. Abel does something similar, but there is no concept of the type a group is. Galois, again, also creates the concept of an action of a group on another structural entity, such as a field or a manifold in geometry.

A momentous role of Galois's concept of a group for algebra has been akin to that of Riemann's concept of manifold for geometry. It may not, in retrospect, be surprising that, in the wake of both Galois and Riemann, with the work of, to mention some among the most prominent names, Felix Klein, Sophus Lie, Poincaré, Élie Cartan, and Weyl, the role of group theory – the action of a group on a manifold – became so crucial in geometry and elsewhere in mathematics or physics. Most elementary particle physics is now understood and practiced in terms of group representations and symmetries they define (irreducible representations correspond to elementary particles). The examples are numerous and will undoubtedly continue to accumulate. One of them is especially intriguing because, expectedly, it establishes remarkable connections between Galois theory and quantum field theory (QFT), specifically one of the most important and still controversial aspects of it, namely, the so-called renormalization. I would like, accordingly, to briefly comment on these connections in closing this section.

It was realized by the early 1930s that the computations provided by quantum electrodynamics (QED), the original form of QFT, were reliable only to a first order of perturbation theory. (Perturbation theory is a set of approximation methods using mathematical perturbation for describing a complex quantum system by using a simple one with a known solution and adding to it a "perturbing" Hamiltonian representing a small disturbance of that simple system.) These computations led to the appearance of infinities or divergences when one attempted to use the formalism for calculations that would provide closer approximations matching the experimental data. These difficulties were eventually handled through the renormalization procedure, which became and has remained a crucial part of QFT, because it still contains these divergencies, specifically certain divergent and, hence, mathematically illegitimate integrals. Roughly speaking, the renormalization procedure manipulates such infinite integrals by, at a certain stage of calculations, replacing them with finite integrals through artificial cut-offs. These cut-offs have no proper mathematical justification and are performed by putting in, by hand, experimentally obtained numbers that make these integrals finite, which removes the infinities (which happen to cancel each other) from the final results of calculations (e.g., Teller 1995, pp. 149–168.) At some point of the history of renormalization, a new concept (a concept-problem), that of the renormalization group, was introduced, with the work of Kenneth G. Wilson giving this concept its current form. The renormalization

group is a mathematical technique, part of the so-called effective QFTs, that enables one to properly investigate the changes of a quantum system as represented ("viewed") at different scales. It reflects the changes in the underlying force laws of QFT, changes due to the fact that the energy scale at which physical processes occur varies, with energy-momentum and resolution distance scales obeying the uncertainty relations. Actual computations are difficult and performed by means of various mathematical techniques, massively assisted by, and in fact impossible apart from, the use of very powerful computers. Conceptually, one evaluates the transformations in the formalism in accordance with Feynman's diagrams, which are, however, only heuristic guides to such calculations in most views, including Richard Feynman's own (e.g., Plotnitsky 2021b, pp. 300–302).

Quite unexpectedly, Galois theory can be brought up to recast the mathematics involved in a new and more rigorous form. The technical complexities are, again, formidable, even for most working in QFT, including among other stratospheric techniques, Grothendieck's theory of motives. These complexities may, however, be bypassed here. My point is a remarkable *fact* itself of the role of a Galois group in renormalization theory. The idea was initially suggested by Pierre Cartier, who designated the corresponding Galois group a "cosmic Galois group," in view, one suspects, of its potential significance in the ultimate QFT, possibly including gravity (Cartier 2001). It is sufficient to consider here the basic statement of how a specific ("motivic") Galois group enters, given by Alain Connes and Matilde Marcolli's article, which explains the essential point. The article and their related work also provide necessary formal definitions for a suitably informed reader, as the mathematics is formidable and requires advance levels of training. According to them:

> The divergences of quantum field theory are a highly structured phenomenon [even beyond the renormalization group, which already structures them]. More precisely, they provide data that define an action of a specific "*motivic Galois group*" U^* on the set of physical theories.
>
> In particular, this exhibits the renormalization group as the action of a one parameter subgroup $\mathbb{G}a \subset U^*$ of the above Galois group. . . .
>
> The natural appearance of the "motivic Galois group" U^* in the context of renormalization confirms a suggestion made by Cartier in [Cartier 2001], that in the Connes–Kreimer theory of perturbative renormalization one should find a hidden "cosmic Galois group" closely related in structure to the Grothendieck–Teichmüller group. The question of relations between the work of Connes–Kreimer, motivic Galois theory, and deformation quantization was further emphasized by Kontsevich in [Kontsevich 1999]. . . .
>
> The "motivic Galois group" U^* acts on the set of dimensionless coupling constants of physical theories, through the map of the corresponding group $\mathbb{G}$ to formal diffeomorphisms constructed in [Connes and Kreimer 2001].
>
> This also realizes the hope formulated in [Connes 2003] of relating concretely the renormalization group to a Galois group. . . .
>
> These facts altogether indicate that the divergences of Quantum Field Theory, far from just being an unwanted nuisance, are a clear sign of the presence of totally unexpected symmetries of geometric origin. (Connes and Marcolli 2004, pp. 4073–75) (See also Connes and Marcolli (2007, pp. 95–136) for a comprehensive exposition of these problematics).

The last conclusion is remarkable, because it counters the common sentiments (some of which have been more strongly negative) that have governed the attitude

toward the divergencies of QFT and renormalization from 1930 on. In addition, it brings in yet another, highly complex, instance of the interplay between algebra and geometry into quantum theory. The physical meaning and implications of this mathematics have hardly been explored as yet, even though there were some developments along these lines since. My main point is, however, an unexpected and potentially far reaching role of Galois's group, as a concept-problem in the present definition, in linking algebra, geometry ("symmetries of geometric origins" invoked by Connes and Marcolli), and fundamental physics. Moreover, this role changes our perspective on renormalization in QFT, which has appeared as inhibiting the relationships between mathematics and fundamental physics, but might reflect deeper aspects of these relationships, thanks to the concept-problem of a Galois group, which emerge along with the concept-problems of a group and an action of a group in the work of Galois.

I now move to Riemann and geometry. Galois, however, will make a return in the following section, via the role of Galois theory in algebraic geometry, specifically, Grothendieck's étale cohomologies, another momentous impact of Galois's way of thinking in mathematical practice as philosophy. Grothendieck's theory of étale cohomologies was, actually, one of the *motivations* for his concept, concept-problem, or a motive, and motive theory, which, among its numerous applications, eventually led to the role, just considered, of motivic Galois groups in renormalization in QFT, a theory, again, indebted, on several counts, to Riemann's concepts as well, as is in fact the concept of a cosmic Galois group. It might be added that the so-called anti-de-Sitter spaces, which are the Lorentzian analogies of Lobachevsky spaces, play an important role in developments of QFT theory and its extensions, such as string and brane theories – the holographic principle and Juan Maldacena's AdS/CFT (anti-de-Sitter/conformal field theory). There are also deep connections between these ideas and motivic Galois's groups, connections noted in (Kontsevich 1999), cited above, a key work on the role of motivic Galois theory in QFT.

4 Geometry: Lobachevsky and Riemann

My main focus in this section is Riemann's mathematical practice, as the most explicit exemplification of my main thesis concerning mathematical practice as philosophy, as defined by the invention of concepts. While characterizing all of Riemann's work, this philosophy is especially manifested in his rethinking of the foundations of geometry in his Habilitation lecture, "On the Hypotheses That Lie at the Foundations of Geometry" [*Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*] (Riemann 1854). Lobachevsky's thinking and work, however, not only, obviously, have their own revolutionary place in the history of geometry, but are important in the present context as well. The differences in their thinking and mathematical practice remain important. It would be difficult to argue that, unlike Gauss (expressly mentioned by Riemann in the lecture), Lobachevsky was a precursor of Riemann, and as often noted, it is doubtful that Riemann was familiar with

Lobachevsky's work. On the other hand, it would be difficult to argue either that Riemann was unaware of Lobachevsky's discovery of non-Euclidean geometry (of the hyperbolic type), especially given that Gauss championed Lobachevsky's work, as against that of Bolyai.

Specifically, I shall argue that, while, unlike Riemann, expressly addressing Euclid's fifth postulate, Lobachevsky's thinking implicitly challenges the axiomatic approach, as developed from Euclid on, to geometry at its very core. While not, like Riemann, overtly grounding geometry in hypotheses and concepts, Lobachevsky's discovery in effect showed that the axioms of geometry in general, rather than only the fifth postulate, and, perhaps, all our axioms, including in physics, are hypotheses, based, however imperceptibly, in concepts, such as that of parallel lines in Euclidean geometry. The latter may be a notion arising from our general phenomenal intuition, which may still not be able to conceive how it is possible to have more than one parallel line in the same plane crossing a point outside this line. As a mathematical object, however, either is a concept, and the fact that there is one parallel line to a given one going through a point outside it is a hypothesis, either as concerns a physical space, with which the term geometry was associated at the time, or as concerns a possible mathematical geometry. Lobachevsky, thus, also (in accord with a key trend of modern mathematics), suggested a dissociation of geometry from physics, which is, arguably, one of the reasons why he, initially, referred to his hyperbolic geometry as imaginary geometry. Of course, he also thought that hyperbolic geometry may correspond to that of actual physical space on a broader scale, which, however, still allows for its mathematical independence. He gave his geometry or geometries other names, on which I shall comment below, merely noting now that his final name "pangeometry" may have been a reflection of the view, akin to that of Riemann, that what is mathematically shared is geometrical thinking. This thinking enables us to create new mathematical worlds, "out of the nothing," as Bolyai famously said about his invention of hyperbolic geometry (Stäckel 1913, p. 86). These worlds are creations of thought from its insubstantial *materials*, "such stuff/As dreams are made on," in William Shakespeare's even more famous phrase (*The Tempest*, IV, 1, ll. 156–157). This thinking, then, can lead to either Euclidean or hyperbolic geometry, and possibly still other geometries, eventually an uncontainable multitude of geometries with no connections to physics, although for Lobachevsky, Bolyai, and Riemann the terms space and geometry still referred to physical space and its geometry.

The landscape of geometrical thinking changed even more radically following Riemann's Habilitation lecture, given in 1854 but first published by Dedekind in 1868, 2 years after Riemann's death (On Riemann's subsequent developments of his geometrical ideas and their application to physics, see essays in Ji et al. (2017), where the earlier (longer) version of this discussion of Riemann was published (Plotnitsky 2017). The present version revises the argument offered there on several key points). The lecture offered a radical rethinking of space and geometry, as against the preceding history of both, from before Euclid to the discovery of non-Euclidean geometry. This rethinking was based on the concept of manifold or manifoldness [*Mannigfaltigkeit*], a major mathematical innovation. Establishing

the *possibility* of non-Euclidean geometry by Lobachevsky and Bolyai was a great discovery, with profound implications for mathematics, physics, and philosophy, and indeed culture. Riemann's thinking and argumentation departed even more radically from Euclidean geometrical thinking concerning space and geometry than those of his predecessors, including, Lobachevsky and Bolyai, or even Gauss, Riemann's teacher and, arguably, the main precursor as concerns his geometrical thinking. The question of the fifth postulate of Euclid played no role in Riemann's argument. Riemann pursued a different way of thinking, following Gauss's ideas concerning the curvature of two-dimensional surfaces (Riemann also mentions Johann F. Herbart as an influence or guidance, but, even leaving aside the fact that Riemann was ambivalent and sometimes critical of Herbart's philosophy, this influence does not appear to me ultimately that important, especially in the present context). Riemann's thinking was also, and correlatively, *problem-posing*, or quasi-empirical, rather than *axiomatic-theorematic*, as axioms of geometry were neither Riemann's starting point nor figured significantly, if at all, in his investigation, in the way they did in that of Lobachevsky or Bolyai. As I argue here, however, this juxtaposition is not unconditional. The proximities between their thinking is significant, and the fact that Riemann almost certainly knew about their findings is not insignificant either. Lobachevsky's thinking, I argue, had important conceptual and problem-posing aspects. Riemann's approach, however, through the invention of a new concept, concept-problem, that of manifold, establishing a new problem-posing field (extending beyond geometry), led him beyond a single alternative to Euclidean geometry to an uncontainable multiplicity of geometries and, in principle, to an even greater multiplicity of possible spaces, because some among them would not admit geometry. The latter circumstance, not considered by Riemann, became important for the development of topology, influenced by Riemann otherwise, beginning with his concept of Riemann surfaces, which played a major role in the emergence of topology, as well as algebraic geometry.

The implications for physics, extending those of the discovery of a single non-Euclidean geometry, were dramatic as well. Riemann's contribution to physics in the lecture was nearly as revolutionary as his contribution to mathematics, extending already major changes of thinking concerning space brought about by the discovery of non-Euclidean geometry. Before this discovery, one and only one geometry would be available for a geometrical description of physical space (which is, again, how the term space was understood at the time) – if one assumes that space or, in Riemann's astute phrase, "the reality underlying space" could be mathematically described geometrically (Riemann 1854, p. 33). This had been a grounding assumption of classical physics, from Galileo and Newton on, or of modern philosophy, from Descartes on. Kant's epistemology of phenomena (referring to appearances or representations constructed by our minds) versus noumena or things-in-themselves (referring to how things, material or mental, exist independently of how we perceive or represent them) qualified this assumption (Kant 1997). While important, this qualification did not change the essential import of this assumption for physics. Kant's epistemology does not affect measurements that allow us to ascertain the observable properties of space, specifically whether the observable

space is Euclidean (flat) or non-Euclidean (curved), or is three-dimensional (the only space we can phenomenally imagine) or not. Possible deviations from the flatness of physical space could be established indirectly by using measuring instruments, while our phenomenal experience of space could still be Euclidean. Riemann reflected on this situation in his lecture, leaving the question of the ultimate nature of physical space or of the reality underlying space to future physics. Kant, by contrast, assumed that physical space is Euclidean and three-dimensional, or at least that it was unlikely to be anything else. He allowed that such alternatives, specifically that space may have a dimension higher than three, are logically possible in his early (1747) work, but rejected the idea as improbable, on theological grounds (Kant 2012; de Bianchi and Wells 2015, Pesic 2007, pp. 3–4, 7).

Theological grounds would not have deterred Riemann, who was originally trained in theology, which might have even helped his thinking concerning space and geometry, including his concept of manifold. Although the term *Mannigfaltigkeit* was not uncommon in German philosophical literature, including in Leibniz and Kant, it is worth noting that the German word for the Trinity is "*Dreifaltigkeit*," thus, etymologically, suggesting a kind of "three-folded-ness," which could not have been missed by Riemann, or, for that matter, Leibniz and Kant. Nor would they have been likely to deter Lobachevsky, which is, however, a separate subject, except for the fact that he only considered mathematically two and three dimensional space. It was, however, mathematically crucial that his geometry was three-dimensional, as was that of Bolyai, but it appears not that of Gauss, who thus came short of this discovery, even if he ever reached the establishment of the two-dimensional hyperbolic plane. In Riemann's view, that space was a three-dimensional Euclidean manifold was a hypothesis, reasonably well confirmed by the measurements performed at the time, but a hypothesis nevertheless, the truth of which cannot be ascertained by reason alone, as Kant thought possible.

In any event, the situation dramatically changed, mathematically, physically, and philosophically, and even culturally, with the discovery of non-Euclidean geometry. That actual physical space *may not* be Euclidean (the three-dimensional nature of space was, again, rarely contested) made this discovery a major event, even though no measurements made at the time showed any deviation from Euclidean geometry. The possibility was momentous enough. In reality, this challenge was even more radical, including vis-à-vis Kant, because it went beyond questioning of the nature of physical space and geometry to that of our axiomatic presuppositions in general. Even if Lobachevsky himself only realized this radical nature and implications of this challenge to some degree (which was, I would surmise, likely), his discovery revealed that the fifth postulate and, by implication, all axioms or, in the first place, definitions are conceptual constructions based on our general phenomenal intuition but reaching beyond it as mathematical concepts. Axioms are not preexisting truths but are assumptions based in conceptual constructions. This difference became more radical and pronounced with Cantor's concept of the continuum. A given set of these constructions allows one to define a consistent geometrical system, that of Euclidean geometry. But an alternative set of axioms, such as that without any postulate concerning parallel lines (now known as "neutral geometry" or "absolute geometry")

or that in which the fifth postulate is replaced by the postulate that there are at least lines parallel to a given line crossing a point outside this line (hyperbolic geometry), implies another consistent geometry. Mathematically, it is a matter of decision (a preferable category to that of choice) in which of them or possibly in several one pursues one's mathematics, a decision the nature of which is ultimately not mathematical. Physically, any such geometry becomes a hypothesis concerning the nature of physical space, and as such is subject to testing, verification, qualification, refinement, and so forth. These considerations, although not all of them were pursued by Lobachevsky, give us an insight into his geometrical designations – imaginary geometry, as a construct of mathematical imagination, under the assumption that physical space and geometry are Euclidean, and then pangeometry, which, as Athanase Papadopoulus notes, refers to "general geometry," which included all geometries known at the time, including spherical geometry in two-dimensions (Papadopoulos 2010, pp. 229–231). (There were no known elliptical geometries in three dimensions before Riemann.) Pangeometry may, arguably, also be seen as referring to geometrical thinking itself.

Physically, there is yet another possibility inherent in Lobachevsky's discovery, which does not appear to have been entertained by Lobachevsky himself, but was expressly considered by Poincaré, who rediscovered one of Eugenio Beltrami's models for hyperbolic geometry, which helped to prove that it was equiconsistent with Euclidean geometry. (Lobachevsky might have had some intuitions concerning this equiconsistency as well [Papadopoulos 2010, p. 250].) The decision concerning the geometry of physical space may be arbitrary (not the same as random!), in other words, a convention. Whether it actually is, as Poincaré thought, is a matter of debate, which was given a new impetus with general relativity, discovered in 1915, after Poincaré's death (in 1913). Nevertheless, the possibility, which is inherent in Lobachevsky's discovery, is significant. His discovery might indeed be seen as more radical than that of Copernicus, insofar as it would be parallel to the assumption that the decision between the geocentric and heliocentric view of the solar system is a mathematical convention (something, ironically, the Church wanted Galileo to say, admittedly on different grounds). In view of general relativity, the heliocentric view is a convention, much more convenient mathematically as it is.

Of course, Lobachevsky's (or Bolyai's) decision to consider the fifth postulate as independent of, but consistent with, those of "general geometry" and to consider the hyperbolic alternative as an equiconsistent system, could not emerge without technical mathematical work. And yet, as in the case of Abel's and Galois's rethinking the problem of the solution of the polynomial equations in radicals, it was an entirely new way of posing the problem of geometry (as that of consistency of the set of axioms defining geometry, axioms that are mathematical concepts) and thus also a new concept of geometry. That of Riemann was still more radical, especially as concerns the multiplicity of possible geometries, which required an invention of a new concept-problem and new mathematics, those of manifolds. As a result, Riemann also gave a firmer conceptual grounding and more radical dimensions to the problem of geometry. Before I turn to Riemann, I would like to comment on the

following, in my view, misleading, assessment of Lobachevsky's and Bolyai's discovery by Jeremy Gray and José Ferreirós:

> The new geometry posed a radical challenge to Euclidean geometry, because it denied traditional geometry its best claim to certainty, to wit, that it was the *only* logical system for discussing geometry at all. It also exploited the tension known to experts between the concepts of straightest and shortest. But in other ways it was conventional. It offered no new definitions of familiar concepts such as straightness or distance, it agreed with Euclidean geometry over angles, it merely offered a different behaviour of parallel lines based on a difference in the distant behaviour of straight lines. Its proponents did not offer a skeptical conclusion. Bolyai and Lobachevsk[y] did not say: "See, there are two logical but incompatible geometries, so we can never know what is true." Instead, they held out the hope that experiments and observations would decide. The epistemological price people would have to pay if astronomical observations had come down in favour of the new geometry would, in a sense, have been slight: it would have been necessary to say that straight lines have an unexpected property after all, but one only detectable at very long distances. To be sure, many of the theorems of geometry would then have to be reworked, and their familiar Euclidean counterparts would appear only as very good approximations. But that is broadly comparable to the situation Newtonian mechanics found itself in after the advent of special relativity. (Gray and Ferreirós 2021)

That may be technically correct (although some qualifications would still be necessary), as is the claim that their view was that experiment and observation would decide which describes physical space, a view that Riemann held as well, with deep implications for physics. This assessment, however, misses the radical epistemological and ontological implications of the way of thinking thus brought about concerning our "basic" assumptions, axioms, postulates, and so forth, as no longer basic, but rather as constructions of thought the validity of which can be challenged in physics and in mathematics itself, as for example, when it comes to considering Newtonian physics, as axiomatic. It may be that hyperbolic geometry as such does not challenge physics in the way general relativity (which might have been the main example in this regard for Gray and Ferreirós) and quantum mechanics did as physical theories, but the way of thinking in which hyperbolic geometry was discovered suggests the possibility of questioning the "axioms" of physics. Einstein would have agreed: he did so in questioning our basic assumptions concerning the kinematics of motion in special relativity, or the special status of inertial systems in general relativity. On the other hand, he was unwilling to give up causality and especially realism as basic philosophical "axioms," which led to his refusal to accept quantum theory as a fundamental theory (a theory describing the ultimate constitution of nature).

The boldness and radical nature of Lobachevsky's thinking, or that of Bolyai, and their implications for our thought in physics is missing in Gray and Ferreirós's assessment. If one can create a mathematical "universe out of nothing," it is no longer easy to assume for modern physics, as a mathematical-experimental science, which one of these mathematical worlds corresponds best to the physical world, and perhaps ultimately none does, as quantum mechanics may imply (Plotnitsky 2021b). It is this spirit that matters more than technicalities invoked by Gray and Ferreirós.

They give more credit to Riemann, while, however, still underappreciating the radical nature of his thinking, including as concerns the point of the possible (or impossible) relationships between a given mathematical world and the physical world. Riemann offers a much better way of looking at the connection between radical thinking in mathematics and physics: "Investigations like the one just made here [concerning foundations of geometry], which begin from general concepts, can only serve to insure that this work is not hindered by unduly restricted concepts and that progress in comprehending the connection of things is not obstructed by traditional prejudices" (Riemann 1854, p. 33). Riemann refers to his own mathematics as allowing physics to break with "unduly restricted practice" and "traditional prejudices," such as those of Euclidean geometry or Newtonian mechanics. But the point is more general and concerns all concepts that become "unduly restricted" unquestionable axioms by forgetting their human origins and history.

Riemann's argument, again, *implied* that an infinite number of possible geometries could be associated with physical space or "the reality underlying space." Any such association is a *hypothesis* (this is one, the most direct, of the meanings of the "hypotheses" of Riemann's title) and as such is subject to testing, verification, qualification, refinement, and so forth. These processes can rule out some among possible geometries or require different geometries at different scales, as indeed happens in modern physics. Thus, with the help of Einstein's general relativity, we know reasonably well certain local physical geometries, say, the one, curved, in the vicinity of the solar system, and even more global geometries, say, that (on average flat) in the Milky Way. Or, more rigorously, the corresponding argumentation works well in physics and astronomy as things stand now. It is, however, much more difficult to be sure concerning the ultimate geometry of the Universe, although the current data seems to suggest that it is, *on the average*, flat as far as we can observe it. Locally, space could be curved by gravity or contain a singularity, such as a black hole, in accordance with general relativity. It is also possible that there are other Universes with different geometries. Riemann did not envision this possibility, which arises from quantum considerations concerning the Big-Bang origin of the Universe that we observe. However, his discovery of the infinite number of possible geometries is in accord with this idea, the genealogy of which goes back to Leibniz's concept of compossible worlds, although for Leibniz there existed, created by God, only one world, the best one, in which we live.

The significance and impact of Riemann's lecture for mathematics, physics, and philosophy were immense. This impact was delayed until its publication, in 1868, 2 years after Riemann's death, and 14 years after it was presented in 1854, although some of Riemann's key ideas contained there became known and had their impact earlier. Riemann opens as follows:

> As is well known, geometry presupposes the concept of space, as well as assuming the basic principles for construction in space. It gives only nominal definitions of these things, while their essential specification appears in the form of axioms. The relationship between these presuppositions is left in the dark; we do not see whether, or to what extent, any connection

between them is necessary, or a priori whether any connection between them is even possible.

From Euclid to Legendre . . . this darkness has been dispelled neither by the mathematicians nor the philosophers who have concerned themselves with it. The reason [*Grund*] for this is undoubtedly because the general concept of multiply extended magnitudes [*Grösse*], which includes spatial magnitudes, remains completely unexplored. I have therefore first set myself the task of constructing the concept of a multiply extended magnitude from general notions of magnitude. It will be shown that a multiply extended magnitude is susceptible of various metric relations, so that space constitutes only a special case of a triply extended magnitude. From this, however, it is a necessary consequence that the theorems of geometry cannot be deduced from general notions of magnitude, but that those properties that distinguish space from other conceivable triply extended magnitudes can only be deduced from experience. Thus arises the problem of seeking out the simplest data from which the metric relations of space can be determined, a problem that by its very nature is not completely determined, for there may be several systems of simple data that suffice to determine the metric relations of space; for the present purposes, the most important system is that laid down as a foundation of geometry by Euclid. These data are—like all data—not logically necessary, but only of empirical certainty, they are hypotheses; one can therefore investigate their likelihood, which is certainly very great within the bounds of observation, and afterwards decide on the legitimacy of extending them beyond the bounds of observation, both in the direction of the immeasurably large [*Unmessbargrosse*] and in the direction of the immeasurably small [*Unmessbarkleinen*]. (Riemann 1854, p. 23; translation modified)

These introductory reflections are already profound and far-reaching, and Riemann develops them much further in the lecture. Riemann is, again, not interested in the axiomatic approach to geometry and appears to doubt its effectiveness. In contrast to most previous works on non-Euclidean geometry, the parallel postulate is not the starting point of his investigation. The reason for the ineffectiveness of the axiomatic approach is that the *concepts* of space and of measuring distances in space, necessary for geometry, are not adequately defined. This is what leaves "the relationships between [axioms] . . . in the dark. We do not see whether, or to what extent, any connection between them is necessary, or a priori whether any connection between them is even possible." This questioning of the axioms of geometry was unusual at Riemann's time, and it was much deeper than customary doubts concerning the parallel postulate, in any event no longer assumed as such by the time of Riemann's lecture in view the discovery of hyperbolic geometry by Lobachevsky and Bolyai, of which, if not of their actual work, Riemann, as I suggested, was likely to be aware. As discussed earlier, however, their argumentation was not restricted to merely axiomatic thinking and, in effect, questioned the latter. In any event, only Gauss, rather than either Lobachevsky or Bolyai, was expressly invoked in the lecture (There were other reasons, beginning with the fact that Gauss was Riemann's mentor and the chair of his Habilitation committee. Indeed, Gauss selected this topic among three proposed by Riemann (following the rules). The philosopher R. H. Lotze, a fervent opponent of non-Euclidean geometry, was a member of the philosophy faculty, to which Riemann's Habilitation was presented. Later, Lotze criticized Riemann's approach anyway, as part of his general critique of non-Euclidean geometry. See Laugwitz (1999, pp. 222–226) and Russell (1897, pp. 97–112)).

Gauss does not appear to have been able to establish a possible existence (in the sense of logical consistency) of non-Euclidean geometry, because, unlike Lobachevsky and Bolyai, he did not do this for the three-dimensional case. In spirit, however, Gauss's work on the geometry of two-dimensional surfaces and his *concept* of curvature as intrinsic to a given surface was closer to and must have influenced Riemann's thinking. For Riemann, Gauss was pursuing a trajectory of thought better suited for the foundations of geometry and more fruitful for its conceptual development. In his famous theorem (*theorema egregium*), Gauss proved that the curvature of a surface, which he defined as well, was intrinsic to the surface. More precisely, the theorem states that Gaussian curvature of a surface does not change if one only bends the surface but does not stretch it. This means that curvature can be fully determined by measuring angles and distances on the surface itself, without considering the way it is embedded in the ambient (three-dimensional) Euclidean space, which makes the Gaussian curvature an intrinsic invariant of a given surface. Gauss's theorem thus suggests a two-dimensional conceptual architecture of geometry akin to that of Riemann, who was able to give this architecture a great generality. It suggests that one could see the surface as an *independent* curved manifold, a new concept and a new philosophy emerging in mathematical practice, and then to generalize this concept to higher dimensions via "the concept of a multiply extended magnitude," mentioned in this passage and the concept of manifold, which is what Riemann did. These concepts also helped him to generalize to higher dimensions Gauss's concept of curvature. To do so required yet another new concept, that of the tensor of curvature, and a new form of differential calculus, tensor calculus on manifolds, a generalization of differential calculus, another great invention of Riemann. This calculus was fully developed later (sometimes under the name of absolute calculus), and it played a major role in Einstein's general relativity.

The independence of the curvature or of geometry of a given manifold of any dimension from its embedding also allows one to determine, at least in principle, intrinsically whether the space thus represented is flat or curved, which is crucial for determining the nature of the physical space we inhabit. For example, as noted, space is curved in the immediate vicinity of the solar system, because of the gravity of the Sun, planets, and other material entities there (or locally around other stars), but, given the current experimental data, appears to be on the overage flat on the scale of the observable Universe. It has never been established definitively or, in any event, agreed upon whether such a determination is rigorously possible, at any scale, as opposed to having a practically effective and possibly, within its proper scope, the best available theory or, as Poincaré would have it, convention, without making a truth-claim concerning "the reality underlying space" (Riemann 1854, p. 33).

Riemann's thinking was clearly conceptual and/as problem-posing rather than axiomatic-theorematic, quasi-empirical rather than Euclidean: he grounds mathematics, as well as physics, in hypotheses and concepts, concept-problems, such as the concept of manifold and its subconcepts, or associated concepts (such as distance, curvature, and tensor). The problem of space and geometry was seen by Riemann as that of finding concepts that are necessary to define space and to give it a geometry. A general concept of manifold is assumed to be applicable to any possible

space, while specific manifolds define different spaces or subspaces. Thus, as *foundational* thinking, that of Riemann is different from that conforming to Hilbert's concept of "foundations" [*Grundlagen*] in the sense of axiomatic foundations, an idea that Hilbert also tried to apply to physics. One of the problems of his famous 1900 list (Problem 6) was that of the development of (mathematized) axiomatic foundations, system(s) of axioms, for physics, on the model of his own *Foundations of Geometry* (Hilbert 1979). That such a system is possible is a hypothesis. It became clear subsequently that the existence of such a system is a hypothesis even in mathematics, and in view of Gödel's incompleteness theorems of 1931, any such system cannot be proven to be free of contradiction, once it is large enough to include arithmetic, as most geometries are. While Hilbert's foundational thinking aimed to bring physics closer to mathematics, by giving physics an axiomatic form, that of Riemann brings mathematics closer to physics by grounding it in hypotheses and concepts, rather than in axioms, by making it quasi-empirical.

The parallel postulate is, again, never mentioned in the lecture, although Euclidean geometry is invoked as the geometry of "flat space," merely a particular and very special case of geometry, where metrical relations take an especially simple form, defined by the Pythagorean theorem. Euclidean geometry is no longer the starting point of Riemann's thinking concerning foundations of geometry, which has been noted, although rarely appreciated as a consequence of Riemann's mathematical thinking as defined by concepts. The starting point of geometry is measurement: the concept of the metrical relation and its instantiations characterizing given spaces and their geometry. The character of this relation is a hypothesis on which a geometry could be based, a hypothesis to be tested physically to establish whether it is in accord with actual physical space. Non-Euclidean geometry (Riemann, again, does not use the term) is introduced as a possible case of geometry, that of a curved, rather than flat, space of either negative or positive curvature, defined by Riemann by the corresponding type of quadratic form determining the metric. Riemann, thus, introduced a more general *concept*, concept-problem, of geometry, which also allows for a rigorous conceptual grounding to non-Euclidean geometry of either negative or positive curvature. He also introduced non-Euclidean geometry of positive curvature, elliptical geometry. Eclipsed, as it may have been, by Riemann's overall achievement in the lecture, this was a major mathematical discovery with important cosmological implications, as in anticipating the idea, later considered by Einstein, that the universe may be unbounded and yet finite (As noted, the currently dominant view or hypothesis, which appears to be confirmed by cosmological measurements, is that the universe is on average flat and is expanding). Spherical geometry, known previously, was a two-dimensional example of geometry of positive curvature, and no three or higher dimensional ones were established before Riemann.

Riemann's thinking may, thus, be seen as returning to pre-Euclidean thinking and to the original meaning of the ancient Greek word geometry as geo-*metry* – the way of measuring distances and volumes of figures on the surface of the Earth and extending this measurement to volumes of three- or even higher dimensional figures. In order to do so and to have geometry, two concepts are necessary, that of space and that of measurement of space. Axioms, if one needs them, can come later. Besides, as

Riemann clearly implies from the outset of his lecture, the axioms of geometry require concepts first, an understanding that, as I said, was implicit in Lobachevsky's work. This does not mean that one cannot productively pursue an axiomatic approach to geometry, but only that one cannot do so without concepts. Hilbert, again, did so, and he used and invented concepts, but as part of an axiomatic program. This is, however, not how Riemann thought about geometry. The question for Riemann, as for Galois, is always how to define mathematical entities in terms of their conceptual architecture reflecting the problems they respond to, which led them to defining new problem-posing fields by the corresponding concept-problem, that of a manifold and a group, respectively. Besides mathematical axioms are, first, mathematical concepts, and sometimes concept-problems.

Riemann defined the concept of space, again, understood as physical space (as against the concept of manifold, which is mathematical), as a three-dimensional physical instantiation of the concept of continuous manifold, or as something that can be, in an idealized way, represented mathematically by this concept, in accordance with the hypotheses that he assumed as likely given the experimental data then available. While a manifold, as defined by Riemann, may be either discrete or continuous, the concept of continuous manifoldness has a richer and more complex mathematics, and most of Riemann's lecture is devoted to it. Technically, continuous manifolds considered by Riemann were differentiable manifolds, which means that one can define differential calculus on them. Indeed, they are metrical manifolds, now called Riemannian, which allow for the concept of distance between any two points and thus for geometry (there were no discrete geometries yet. As most of his contemporaries, Riemann did not distinguish continuous and differentiable manifolds. It became eventually clear, however, that not all continuous (also called topological) manifolds are differentiable. There are topological manifolds with no differentiable structure, and some with multiple nondiffeomorphic differentiable structures. Thus, there is a continuum of nondiffeomorphic differentiable structures of $\mathbb{R}^4$. I shall, however, speak of continuous manifolds, following Riemann and his juxtaposition between continuous and discrete manifoldness. In modern use, the term manifold more customarily refers to continuous (but not necessarily differentiable) manifolds, although one also refers to discrete manifolds, which have topological dimension zero. Discrete and continuous manifolds do not appear to have that much in common and may be seen as different concepts. These two concepts could, in modern understanding, be subsumed under the same concept. It is conceivable that Riemann thought along similar lines, which would explain his choice of the term manifold for both discrete and continuous manifolds. It is difficult to be certain on the basis of his lecture or his other writings, because the term had a more general use at the time. Thus, Cantor, initially referred to sets as *Mannigfaltigkeiten* but eventually switched to *Mengen*. In any event, the difference between these two types of manifolds is crucial, in general and for Riemann, especially for his analysis of physical space. Riemann stressed the significance of the relationships between continuity and discontinuity for mathematics, physics, and philosophy (e.g., [Riemann 1991, pp. 515–524; Ferreirós 2006, pp. 77–80]).

Riemann's concept of continuous manifoldness was a new concept of geometrical multiplicity. It is a multiplicity of local subspaces, most specifically those, "neighborhoods," associated with each point, out of which a given space is composed. Riemann defines first the concept of "n-dimensional magnitude," which allows one to determine a position in a manifold by n numerical determinations, in the same way a position is determined by coordinates in the Euclidean space of n dimensions. A crucial and even the single most defining feature of the concept of (continuous) manifold is that it is conceived as infinitesimally Euclidean. This makes a continuous manifold into a conglomerate of local, continuously connected, open neighborhoods around each point. The concept of neighborhood, assumed to be infinitesimally flat (Euclidean), is a component-concept of the concept of manifold. The concept of manifold as composed out of local neighborhoods is extendable to a more general concept of topological space, in which case local neighborhoods need no longer be Euclidean and can be defined with a great degree of generality. It is true that Riemann did not define a concept of topological space as such, although it is conceivable, especially given his concept of a Riemann surface, that he entertained this type of idea, just as he, expressly, entertained the idea that "the reality underlying space" may be discrete at a very small scale. In any event, the conceptual architecture defining topological spaces is a nearly inherent generalization of Riemann's conceptual architecture of a manifold as a "space" composed of neighborhoods. Riemann surfaces, too, were recast in terms of manifolds. Weyl, who clearly considered a Riemann surface to be a concept, was the first to define them as manifolds (Weyl 2013). Riemann, however, must have realized that they were manifolds, and the concept of a Riemann surface was part of the genealogy of the concept of manifold.

In the case of Riemannian manifolds, while each neighborhood is infinitesimally Euclidean, the manifold as a whole is, in general, not, except in the special case of Euclidean manifolds. A manifold may be negatively or positively curved, and, which is another major innovation of Riemann, this curvature can also be variable. Riemann defined the metric form as a quadratic differential function of coordinates

$$1/\left(1+\frac{\alpha}{4}\sqrt{\sum dx^2}\right)$$

which is the only real formula in his lecture (discounting the coordinate expression for the line element). Riemann also assumed that the transition from one local coordinate system to another was differentiable. Thus, he de facto considered differentiable manifolds with positive metrics, Riemannian manifolds. In modern terms, each such a manifold has a differentiable section of positive-definite quadratic forms on the tangent-bundle. While, however, modern technical language can bring out deeper mathematical aspects of Riemann's concepts, it can also obscure and displace how Riemann thought, a displacement sometimes found in twentieth-century English translations of Riemann, including his Habilitation lecture.

Although it was Riemann's theory of continuous or differentiable manifolds that had the greatest impact, the concept of discrete manifolds was important for Riemann's argument, and it is important for the modern understanding of both

spatiality and geometry in mathematics, physics, and philosophy. While a discrete manifold has topological dimension zero, it may still be seen as multiply extended, if defined as forming a very fine lattice with very small intervals between points, which can be "filled," as it were, to form a continuous space of the corresponding topological dimensions. It is also possible to introduce metrical relations for discrete manifolds. This concept is important in the context of the relationships between physical forces in nature and the nature of space or, to return to Riemann's terms, the physical "reality underlying space," although Riemann was cryptic on such metrical relations. Crucially, however, he allows for the possibility that the physical "reality underlying space" might be "a discrete manifold" (Riemann 1854, p. 33). This possibility had been entertained even before Riemann and has been often considered since. Also, mathematically, finite geometries were beginning to be developed, usually in more axiomatic ways, around Riemann's time as well, later on also under the impact of his geometrical thinking (For some of these developments, see Pambuccian et al. (2017) and other articles in Ji et al. (2017)). Riemann saw the relationships between continuity and discontinuity as foundationally central to mathematics, physics, and philosophy (e.g., [Riemann 1991, pp. 515–524; Ferreirós 2010, pp. 77–80]), a view confirmed by the subsequent developments in the foundations of mathematics, from Dedekind and Cantor on, and quantum physics. The latter uses continuous (technically, again, differentiable) mathematics, that of Hilbert spaces over $\mathbb{C}$, to predict, in probabilistic terms, irreducibly discrete phenomena, that is, phenomena that are not, and that possibly cannot be, assumed to be connected to each other by a continuous physical process. The genealogy of this mathematics is still Riemannian, given that Riemann has already envisioned, along with discrete manifolds, infinite-dimensional ones. Tensor calculus, too, is a key aspect of quantum mechanics, reflecting crucial features of quantum phenomena. In sum, beyond Riemann's contribution to physics in his lecture itself, which, as noted, was nearly as momentous as his contribution to mathematics, his mathematical concepts have immense impact on the subsequent development of physics.

The final implication of Riemann's approach I would like to mention, in part, again, as a bridge to the next section, is that it allows one to define a geometrical or topological space (in modern terminology) not as a set of points continuously connected to each other, but in terms of open neighborhoods around each point. As explained above, following this way of thinking, topology describes a given space not only in terms of its points, continuously connected to each other, but also and most essentially in terms of its *open neighborhoods* around each point. These neighborhoods are subspaces of this space, the idea that underlies Riemann's concept of manifold, in this case giving each neighborhood a Euclidean geometry. The approach enabled Riemann, again, following Gauss's theory of curvature of two-dimensional surfaces, to define manifolds of any dimension in terms of its inner properties rather than in relation to the ambient Euclidean space, where a manifold could be placed against the flat Euclidean background. It is true that, if one appeals, as is usual even in considering Riemann, to open sets, this concept retains the concept of a set (of points) as a primitive concept. Thus, Ferreirós's analysis displaces Riemann's thought into the axiomatic and set-theoretical register,

dominant in the wake of Cantor (Ferreirós 2006, pp. 39–80). This displacement reflects Ferreirós's insufficient attention to the structure of Riemann's mathematical concepts, or in the first place, to the *concept* of mathematical concept found in Riemann's practice. Riemann did not think either in terms of axioms or in terms of sets of points, as Ferreirós contends, although his concepts could be translated into these terms (e.g., [Pambuccian et al. 2017]). Riemann thinks in terms of spaces composed of spaces. Riemann's way of thinking concerning manifolds, however, also suggests a possibility of thinking of, and even defining, a space (as a mathematical concept) in terms of its relations to other spaces, in effect categorically, which allows one to use this structure as more primordial by replacing covering a space by *open sets (of points)* with covering it by open *spaces*. A topological space becomes a collection of open spaces as subspaces with algebraic rules for the relationships between them. This way of defining a space in its relation to other spaces (as opposed to their constitution as sets of points) extends to Grothendieck's topos theory, inspired, however, primarily by the idea of a "covering space," originating in Riemann's theory of Riemann surfaces.

5 Algebraic Geometry: Weil and Grothendieck

I begin with a brief summary of the concept of curve, discussed in detail in [Plotnitsky 2019], which this section partly follows within the context and argument defined by the present chapter. By expressly making geometry algebra, or just about (this claim needs qualifications, considered below), the analytic geometry of Fermat and Descartes *in effect* gave mathematics its independence from physics, thus, initiating modern mathematics as a project dealing with ideal objects independent from material objects in nature. I qualify by "in effect" because it took about two centuries for this independence to get its momentum around the time of Gauss, Abel, and Galois, as discussed in Sect. 3. Analytic geometry did so because the equation corresponding to a curve, say, $X^2 - 1 = 0$ for the corresponding parabola, could be studied as an algebraic object, independently of its geometrical representation or its connection to physics, which eventually enabled mathematics to define algebraic varieties over finite fields and thus as discrete objects. A curve becomes, in its composition, defined by its equation, divested from its representational geometrical counterpart. Nor does it any longer geometrically idealize the reality exterior to it. It only represents itself. The equation, as algebra, also defines a separation of a mathematical curve or surface from any curve or surface found in the world, thus also revealing a deeper essence of all mathematics. When one says in mathematics, "consider a curve X," one separates it from every curved object in the world. This separation of algebra from material reality and geometry (as ordinarily understood) also enables it to define a discrete curve, or can make a curve a surface, or a surface a curve, give it an even more complex spatial algebraic architecture, or a yet more complex conceptual composition, such as that of a moduli space, the Teichmüller space (also the Teichmüller *curve*), Grothendieck's scheme and representable-functors, and so forth. We are yet far from exhausting the problem-posing field defined by the concept-problem of a Riemann surface, which is also a curve over

$\mathbb{C}$ (e.g., A'Campo et al. 2016). This view also transforms our view of geometrical objects in the same way: modern geometry, too, separates or liberates its objects and itself from daily phenomenology, even if not as completely as algebra does.

The concept of a Riemann surface is one of Riemann's great contributions to modern mathematics, which shaped several of its areas, most especially here algebraic geometry. In Papadopoulos's account:

> In his doctoral dissertation, Riemann introduced Riemann surfaces as ramified coverings of the complex plane or of the Riemann sphere. He further developed his ideas on this topic in his paper on Abelian functions. This work was motivated in particular by problems posed by multi-valued functions *w(z)* of a complex variable *z* defined by algebraic equations of the form
>
> $$f(w, z) = 0,$$
>
> where *f* is a two-variable polynomial in *w* and *z*.
>
> Cauchy, long before Riemann, dealt with such functions by performing what he called "cuts" in the complex plane, in order to obtain surfaces (the complement of the cuts) on which the various determinations of the multi-valued functions are defined. Instead, Riemann assigned to a multi-valued function a surface which is a ramified covering of the plane and which becomes a domain of definition of the function such that this function, defined on this new domain, becomes single-valued (or "uniform"). Riemann's theory also applies to transcendental functions. He also considered ramified coverings of surfaces that are not the plane. (Papadopoulos 2017, p. 240)

The concept of a Riemann surface gains much additional depth and richness when considered in terms of his concept of a manifold. Riemann, as noted, did not do so himself, although he undoubtedly realized that Riemann surfaces were manifolds, and they have likely been part of the genealogy of the concept of manifold. Riemann's surfaces were first expressly defined as manifolds by Weyl in *The Concept of a Riemann Surface* (Weyl 2013). It is an intriguing question whether Riemann himself thought of Riemann surfaces as curves, but it would not be surprising if he had. Weyl undoubtedly did, although the point does not figure significantly in his book, focused, topologically, on the *surface* nature of Riemann surfaces. This may be surprising. But then, Weyl was not an algebraic geometer. The work of Émile Picard, a key figure in the history of algebraic geometry would be more exemplary in considering this point (Picard 1891–1896; Papadopoulos 2017).

Weyl's argumentation leading him to his definition is an application of a principle akin to Joseph Silverman and John Tate's principle, "Think Geometrically, Prove Algebraically":

> It is also possible to look at polynomial equations and their solutions in rings and fields other than $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$. For example, one might look at polynomials with coefficients in the finite field $\mathbf{F}_p$ with p elements and ask for solutions whose coordinates are also in the field $\mathbf{F}_p$. You may worry about your geometric intuitions in situations like this. How can one visualize points and curves and directions in $\mathbf{A}^2$ when the points of $\mathbf{A}^2$ are pairs (x, y) with $x, y \in \mathbf{F}_p$? There are two answers to this question. The first and most reassuring is that you can continue to think of the usual Euclidean plane, i.e., $\mathbb{R}^2$, and most of your geometric intuitions

> concerning points and curves will still be true when you switch to coordinates in $\mathbf{F}_p$. The second and more practical answer is that the affine and projective planes and affine and projective curves are defined *algebraically* in terms of ordered pairs (r, s) or homogeneous triples $[a, b, c]$ without any reference to geometry. So, in proving things one can work algebraically using coordinates, without worrying at all about geometrical intuitions. We might summarize this general philosophy as: *Think Geometrically*, *Prove Algebraically.*" (Silverman and Tate 2015, p. 277)

Affine and projective planes and curves can be defined without any reference to our phenomenal intuition, which grounds the geometrical thinking Silverman and Tate refer to. Even in these more intuitively accessible cases, we think algebraically, too, by using what may be called "spatial algebra." "Spatial algebra" refers to algebraic structures that *mathematically* define geometrical or topological objects and reflect their proximity to $\mathbb{R}^3$ and mathematical spatial objects there that are close to our phenomenal intuition. This proximity may be, and commonly is, left behind in rigorous mathematical definitions and treatments of such objects, beginning with $\mathbb{R}^3$ itself. The same type of algebra may also be used to define mathematical objects that are not available to our phenomenal intuition at all. Among examples are a projective line (a set of lines through the origin of a vector space, such as $\mathbb{R}^2$, with projective curves defined algebraically, as algebraic varieties) and an infinite-dimensional Hilbert space (the points of which are typically square-integrable functions or infinite series, although a Euclidean space of any dimension is technically a Hilbert space, too). Spatial algebra is an algebraization of spatiality that makes it rigorously mathematical, as opposed to something that is phenomenally intuited spatially, even in the case of more conventional spatial objects in $\mathbb{R}^3$. In addition, as Silverman and Tate state, spatial algebra retains its connection to geometrical thinking because analogies with $\mathbb{R}^3$ continue to remain useful and even indispensable. Weyl's argument concerning "thinking geometrically" about analytic forms is also a manifestation of the spirit of Riemann's mathematical practice, in which geometry and algebra, indeed geometry, topology, algebra, and analysis, come together in a complex mixture of the rigorous and the intuitive, algebraic and spatial-algebraic, mathematical and physical, and mathematical and philosophical. Riemann's work on his ζ-function and number theory reflect this mix as well, as his greatest contribution to algebra, but still not without geometry to do it. According to Weyl:

> It was pointed out . . . that one's intuitive grasp of an analytic form [an analytic function to which a countable number of irregular elements have been added] is greatly enhanced if one represents each element of the form by a point on a surface $\mathcal{F}$ in space in such a way that the representative points cover $\mathcal{F}$ simply and so that every analytic chain of elements of the form becomes a continuous curve on $\mathcal{F}$. To be sure, from a purely objective point of view, the problem of finding a surface to represent the analytic form in this visual way may be rejected as nonpertinent; for in essence, three-dimensional space has nothing to do with analytic forms, and one appeals to it not on logical-mathematical grounds, but because it is closely associated with our sense perception. To satisfy our desire for pictures and analogies in this fashion by forcing inessential representation of objects instead of taking them as they are could be called an anthropomorphism contrary to scientific principles. (Weyl 2013, p. 16)

Weyl resolves this "problem" not only by suggesting, as did Silverman and Tate, that one might think geometrically, while proving algebraically, but also that one can think "spatial-algebraically" (along with intuitively geometrically), by joining Riemann's concepts of Riemann surface and manifolds. Weyl says:

> However, these reproaches of the pure logician are no longer pertinent if we pursue the other approach, already hinted at, in which the *analytic form is a two-dimensional manifold* to which all the ideas of continuity that we meet in ordinary geometry may be applied. To the contrary, not to use this approach is to overlook one of the most essential aspects of the topic. (Weyl 2013, p. 16)

Weyl will proceed, in the spirit of Gauss and Riemann, to his definition of a two-dimensional manifold and then Riemann's surface, *intrinsically*, rather than in relation to its ambient three-dimensional space. This concept enables one to define a Riemann surface as a curve over $\mathbb{C}$. This intrinsic and abstract view of a Riemann surface was often forgotten by Riemann's followers, especially at earlier stages of the history of the concept. According to Papadopoulos, who follows Klein's assessment:

> Riemann not only considered Riemann surfaces as associated with individual multi-valued functions or with meromorphic functions in general, but he also considered them as objects in themselves, on which function theory can be developed in the same way as the classical theory of functions is developed on the complex plane. Riemann's existence theorem for meromorphic functions with specified singularities on a Riemann surface is also an important factor in this setting of abstract Riemann surfaces. Riemann conceived the idea of an abstract Riemann surface, but his immediate followers did not. During several decades after Riemann, mathematicians (analysts and geometers) perceived Riemann surfaces as objects embedded in three-space, with self-intersections, instead of thinking of them abstractly. They tried to build branched covers by gluing together pieces of the complex plane cut along some families of curves, to obtain surfaces with self-intersections embedded in three-space. (Papadopoulos 2017, p. 242)

Hence, Weyl (whom Papadopoulos cites) contends: "Thus, the concept 'two-dimensional manifold' or 'surface' will not be associated with points in three-dimensional space; rather it will be a much more general abstract idea," in effect a spatial-algebraic one. Weyl's position concerning the nature of mathematical reality is a different matter. As is clear from his philosophical writings (e.g., Weyl 2021), Weyl was ultimately a realist in mathematics and physics alike, notwithstanding his major contribution to quantum mechanics (which radically challenged realism), not the least by making group and group representation theories central there. Weyl continues as follows:

> If any set of objects (which will play the role of points) is given and a continuous coherence between them, similar to that in the plane, is defined we shall speak of a two-dimensional manifold. Since all ideas of continuity may be reduced to the concept of neighborhood, two things are necessary to specify a two-dimensional manifold:
> 1. to state what entities are the "points" of the manifold;
> 2. to define the concept of "neighborhood." (Weyl 2013, pp. 16–17)

This means one should define entities, such as points, lines, neighborhoods, by using algebraic symbols and relationships among them, without referring to objects in the world represented by ordinary language, even if still using this language, as the concept of a Riemann surface as a curve and then its avatars such as Gromov's concept of pseudoholomorphic curve (a smooth map from a Riemann surface into an almost complex manifold) exemplify. Its connections to our phenomenal sense of surface are primarily intuitive, as when it comes to the idea of continuity, for example, as defined by Weyl here, in terms of the concept of neighborhood. In any event, a Riemann surface is not a curve in any phenomenal sense.

This multifaceted nature of Riemann surfaces equally and often jointly defined the history of complex analysis, the main initial motivation for Riemann's introduction of the concept of a Riemann surface, and the history of algebraic curves, both building on this concept, and other developments, for example, in abstract algebra and number theory, including Riemann's work on the ζ-function and the distribution of primes. All these developments were unfolding during the period between Riemann and Weyl, whose book initiated the treatment of the concept of Riemann surface that defines the present-day understanding of the concept. This history joins several "Riemanns" – the Riemann of the concept of manifold, the Riemann of the concept of Riemann surfaces, the Riemann of complex analysis. A few more Riemanns could be added. This multiple history shaped algebraic geometry, eventually leading to the work of Weil and Grothendieck, to which I now turn, although Weyl's analysis of the idea of a Riemann surface will make yet another important appearance, via, remarkably, Galois theory.

Weil's and Grothendieck's work on algebraic geometry are among their greatest contributions and are remarkable cases of mathematical practice as philosophy, as the extension of the concept of algebraic variety over a finite field and the study of such objects by the standard tools of algebraic topology, in particular homotopy and cohomology theories. These theories had been previously proven to be very effective for the study of complex algebraic varieties. The project of extending them to finite fields was initiated by Weil by bringing together algebra, geometry, and number theory, in which he was a true heir of Fermat (and he probably saw himself as one), as well as of Kronecker (in this case, Weil expressly saw himself as one). Riemann remains a key presence in the history leading to Weil's work in algebraic geometry, first of all, again, given his concepts of a Riemann surface and a covering space, but also by virtue of introducing the concept of a discrete manifold in his Habilitation lecture. Gino Fano, one of the founders of finite geometry, belonged to the Italian school of geometry (1880s–1930s), contemporary with and an important part of the history of mathematical modernism (Fano was a student of Federigo Enriques), strongly influenced by Riemann. Representatives of the Italian school (quite a few of them, even counting only major figures) made major contributions to many areas of geometry, especially algebraic geometry, which were an important part of a rich and complex history leading to Weil's work on algebraic geometry, culminating in his 1946 *Foundations of Algebraic Geometry*, with the second edition published in 1962 (Weil 1946, 1962). The book is distinguished by many important technical contributions, which would be difficult to properly consider here. My main focus, defined

by the relationships between Weil's and Grothendieck's philosophies of algebraic geometry, is Weil's concept of a "universal domain," assumed by him to be necessary. This concept distinguishes his approach and, thus, his mathematical practice as philosophy from the previous approaches by requiring that in addition to the base field of definition of the objects (algebraic varieties) considered, an assumption of an (algebraically closed) "universal domain" encompassing all fields that may appear in any constructions made over the base field. The Platonist implications of this claim are tempered by his view that such a domain need not be unique. Weil argued that the theorems of algebraic geometry will remain the same regardless of the choice of this domain if it has an infinite transcendence degree and is closed. Grothendieck will abolish this idea, without a return to any previous thinking, on his way to solving a problem posed by Weil, or rather making it a concept-problem. While suggesting a new problem-posing field, Weil did not define the corresponding concept as such, only suggested it as a possibility. Grothendieck's redefinition of this problem as a concept-problem was a result of his mathematical practice as philosophy as different from that of Weil, the difference, arguably, most essentially defined *philosophically* by abandoning Weil's assumption of a universal domain.

As one of his greatest and most visionary contributions, Weil eventually (in 1949) suggested that a cohomology theory for algebraic varieties over finite fields, now known as Weil cohomologies, could be developed by analogy with the corresponding theories for complex algebraic varieties or topological manifolds in general. Weil's motivation was a set of conjectures (these go back to Gauss), known as the Weil conjectures, concerning the so-called local ζ-functions, which are the generating functions derived from counting numbers of points on algebraic varieties over finite fields. These conjectures, Weil thought, could be attacked by means of a proper cohomology theory, although he did not propose such a theory, as a corresponding set of concepts, only posed it as a problem, without such a concept. It was not a concept-problem or set of concept-problems. But the idea led Grothendieck to such a concept-problem and indeed a set of such concept-problems, and a new mathematical philosophy entailed by mathematical practice, a philosophy different, on several counts, from that of Weil that led him to this problem.

To develop such a concept, one needed, first, a proper topology, finding which was nontrivial because the objects in question are topologically discrete. A "native" topology that could be algebraically defined by then, known as Zariski's topology, did not work, because it had too few open sets. The decisive ideas came from Grothendieck, helped by the sheaf-cohomology theory and category theory, known as "cohomological algebra," by then the standard tool of algebraic topology. Using these tools, a hallmark of Grothendieck's thinking throughout his career, and his previous concepts, such as that of "scheme," led him to topos theory and étale cohomology, as a viable candidate for Weil's cohomology, which it had proven to be. Soon thereafter, it also led him to a more general concept-problem of motivic cohomology, mentioned earlier in conjunction with motivic Galois groups, but the subject requires a separate discussion. By using étale cohomology, Grothendieck (with Michael Artin and Jean-Louis Verdier) and Pierre Deligne (Grothendieck's student) were able to prove Weil's conjectures, and then Deligne found and proved

their generalization. Grothendieck's key, extraordinary, insight, also extending what I call spatial algebra in a radically new direction, was to generalize, in terms of category theory, the concept of "open set," beyond a subset of the algebraic variety, which was possible because the concept of sheaf and of the cohomology of sheaves could be defined by any category, rather than only that of open sets of a given space. Étale cohomology is defined by this type of replacement, specifically by using the category of étale mappings of an algebraic variety, which become "open subsets" of the finite unbranched covering spaces of the variety, a vast generalization of Riemann's concept of a covering space (Grothendieck's initial primary area of mathematical research was topological vector spaces, which suggests yet another genealogical link to Riemann, in this case, again, to his concept of manifold). Grothendieck was also building on related ideas of Jean-Pierre Serre. Part of the origin of this generalization was the fact that the fundamental group of a topological space, say, again, a Riemann surface, could be defined in two ways: it can either be defined more geometrically, as a group of the sets of equivalence classes of the sets of all loops at a given point, with the equivalence relation given by homotopy, or it can be defined even more algebraically, as a group of transpositions of covering spaces. In this second, algebraic, definition, it is analogous to the Galois group of the algebraic closure of a field, as Serre was first to consider for finite fields, importantly for Grothendieck's work on étale cohomology, a concept that, thus, has its genealogy in both Galois's and Riemann's thought. The connection had been established in the case of Riemann surfaces long before then and found in Weyl's book on a Riemann surface:

> The group of cover transformations, regarded as an abstract group, expresses purely and completely everything in the relation between the normal covering surface $\overline{\mathcal{F}}$ and the base surface $\mathcal{F}$ which has the character of analysis situs [topology]. This group is also called the *Galois group* of $\overline{\mathcal{F}}$. It is in fact the analog of the Galois group of a normal algebraic field (of finite degree) over a base field. (Weyl 2013, p. 58)

In a way, this passage defines the trajectory of this paper by connecting the mathematical practice as philosophy of Galois, Riemann, and Grothendieck.

For the moment, Grothendieck's concept of étale mappings gives a sufficient number of additional open sets to define adequate cohomolgy groups for some coefficients, for algebraic varieties over finite fields. For complex varieties, one recovers the standard cohomology groups with coefficients in any constructible sheaf. The category of étale mapping is a topos, a concept that is, for now, the most abstract form of spatiality, at least as spatial algebra. As became apparent later, étale cohomologies could be defined for most practical uses in simpler settings. The concept of topos remains crucial, however, especially in the present context, because it can be seen as the concept of a covering space over a Riemann surface converted into the (spatial-algebraic) concept of topology of the surface itself, and then generalized to any algebraic variety. The concept of topos also came to play a major role in mathematical logic. The subject cannot, however, be addressed here, except by noting that mathematical logic has radical epistemological implications

concerning the nature of mathematical reality, to which the concept of topos, as a *logical* concept, may be connected. On the other hand, Grothendieck's use of topoi in algebraic geometry is *ontological* rather than logical. His philosophical position concerning the nature of mathematical reality remains somewhat unclear, for example, as concerns whether it conforms or not to mathematical Platonism. My focus for the moment are the mathematical technologies that the concept of topos, whatever its ontological status, enables, such as étale cohomology. Such technologies may suggest a possible break with the possibility of the ultimate ontological description of mathematical reality, again, assuming that any *ultimate* reality independent of human thought, say, again, of the type considered in physics, is even possible in mathematics, as opposed to what can be specified by one's thought, possibly in one's unconscious (Plotnitsky 2019, 2020). I shall return to this subject in closing.

It would not be possible to present topos theory here in its proper abstractness and rigor, sometimes prohibitive even for those who are trained in the field of contemporary algebraic geometry. The essential philosophical nature of the concept, may, however, be sketched. First, very informally, consider the following way of endowing a space with a structure, generalizing the definition of topological space in terms of open subsets, as mentioned above. One begins with an arbitrarily chosen space, X, potentially any given space, which may initially be left unspecified in terms of its properties and structure. What would be specified are the relationships between spaces applicable to X, such as mapping or covering one or a portion of one, by another. This structure is the arrow structure $Y \rightarrow X$ of category theory, where X is the space under consideration and the arrow designates the relationship(s) in question. One can also generalize the notion of neighborhood or of an open subspace of (the topology of) a topological space in this way, by defining it as a relation between a given point and space (a generalized neighborhood or open subspace) associated with it. This procedure enables one to specify a given space not in terms of its intrinsic structure (e.g., a set of points with relations among them) but "sociologically," throughout its relationships with other spaces of the same category, say, that of algebraic varieties over a finite field of characteristic p (Manin 2002, p. 7). Some among such spaces may play a special role in defining the initial space, X, and algebraic structures, such as homotopy and cohomology, as Riemann in effect realized in the case of covering spaces over Riemann surfaces, again, one of the inspirations for Grothendieck's concepts of a topos and more specifically étale topos.

To make this a bit more rigorous (albeit still informal), I shall briefly sketch the key ideas of category theory. It was introduced as part of cohomology theory in algebraic topology in 1940 and, as I said, later extensively used by Grothendieck in his approach to cohomological algebra and algebraic geometry, eventually leading him to the concept of topos. Category theory considers multiplicities (which need not be sets) of mathematical objects conforming to a given concept, such as the category of differential manifolds or that of algebraic varieties, and the arrows or morphisms, the mappings between these objects that preserve this structure. Studying morphisms allows one to learn about the individual objects involved, often to learn more than we would by considering them only or primarily individually.

Categories themselves may be viewed as similar objects, in relating which one speaks of "functors" rather than "morphisms." Topology relates topological or geometrical objects, such as manifolds, to algebraic ones, as in the case of homotopy and cohomology groups. Thus, in contrast to geometry, topology, almost by its nature, deals with functors from categories of topological objects, such as manifolds, to those of algebraic objects, such as groups. Mathematical practice (as philosophy) becomes that of functoriality: it no longer primarily concerns objects but categories of objects.

A topos in Grothendieck's sense is a category of spaces and arrows over a given space, used for the purpose of allowing one to define richer algebraic structures associated with this space, as explained above. There are additional conditions such categories must satisfy, but this is not essential at the moment. To give one of the simplest examples, for any topological space S, the category of sheaves on S is a topos. The concept of topos is, however, very general and extends far beyond spatial mathematical objects (thus, the category of finite sets is a topos); it replaces the latter with a more algebraic structure of categorical and topos-theoretical relationships between objects. On the other hand, it derives from the properties of and categorical relationships between properly topological objects. The conditions, mentioned above, that categories that form topoi must satisfy have to do with these connections.

Beyond enabling the establishing of a new cohomology theory for algebraic varieties, as considered above, topos theory allows for such esoteric constructions as nontrivial or nonpunctual single-point "spaces" or, conversely, spaces (topoi) without points (first constructed by Deligne), sometimes slyly referred to by mathematicians as "pointless topology." Philosophically, this notion is far from pointless, especially if considered within the overall topos-theoretical framework. In particular, it amplifies a Riemannian idea that "space," defined by its relation to other spaces, is a more primary object than a "point" or, again, a "set of points." Space becomes a Leibnizean, "monadological" concept, insofar as points in such a space (when it has points) may be seen as a kind of monad, thus also giving a nontrivial structure to single-point spaces. These monads are elemental but structured entities, "spaces," rather than structureless entities (classical points), or at least are defined by structures associated to them (Cartier 2001). Naturally, this appeal to monads is metaphorical. Leibniz's monads are elemental *souls*, the "atoms" of soul-ness. One might, however, say, getting more mileage out of the metaphor, that the space thus associated with a given point is the soul of this point. In other words, not all points are alike insofar as the mathematical nature of a given point may depend on the nature or structure of the space or topos to which it belongs or with which it is associated. This approach gives a much richer architecture to spaces with multiple points, and one might see (with caution) such spaces as analogous to Leibniz's universe composed by monads. It also allows for different (mathematical) universes associated with a given space, possibly a single-point one, in which case a monad and a universe would coincide. Grothendieck's topoi are possible universes, possible worlds, or compossible worlds in Leibniz's sense, without assuming, like Leibniz (in dealing with the actual world), the existence of only one of them, the best possible.

One might also think of this ontology as an assembly, a "democracy," of ontologies (Grothendieck's concept of topos is, again, ontological, rather than logical), in the absence of any encompassing ontology. Topoi are multiple universes, defined ontologically, in the absence of a single encompassing domain, such as that found in Weil's "imperial" approach. According to David Reed's summary:

> Grothendieck's constructions [beginning with schemes replacing the varieties of Weil] are far from the 'palaces' which Weil suggest belong to algebraic geometers by birthright. Rather than refurbishing and renewing the old constructions he has instead created an entirely new type of architecture and in the process is forced to make extensive use of the 'makeshift algebraic constructions' Weil carefully thought to avoid placing reliance upon. Both Weil and Grothendieck seek generality in order to be able to analyze rigorously geometric situations to which the older algebraic geometers paid little attention. This generality permits consideration of 'degenerate' cases, objects with singularities, 'non-reduced' objects and the like. Whereas Weil sought this generality through a large 'domain' in which the result of his constructions could be contained, Grothendieck seeks generality by expanding the category of objects under considerations until a certain 'self-sufficiency' is acquired, i.e. until a natural reflexivity can be found so that the objects in the category can be related to each other and operations on them can be undertaken without going outside the original category. Furthermore, the objects do not require a predetermined 'fixed point' outside of the category for their specification. . . . This relativity also extends to the set theoretical foundations of the theory. Grothendieck introduces a set theoretic 'universe' which provides the sets used in his constructions and then studies those properties which remain invariant under a change of universe! (Reed 1995, pp. 131, 177n.36)

One, thus, encounters two very different (mathematical) ontological philosophies and, hence, forms of mathematical practice as philosophy – that of unification or encompassment and that of multiplicity, relational, but without a single encompassing domain. Grothendieck's generality or, better (Reed is right to shift to this term), *relativity* of construction is a kind rigorous mathematical instantiation or model of the capacity of mathematics as a human endeavor, for which Riemann's work on foundations of geometry is, arguably, the main modern precursor. This model is that of the creation of mathematical realities (possibly as mental ontologies, of which Grothendieck's topoi are examples), on the basis of, but creatively transforming, previously established realities, in the absence of an ultimate underlying or encompassing Platonist reality. Phenomenally, it is indeed a kind of creation of universes (now plural!) out of nothing, as Bolyai said, insofar as it is a creation of new mathematical worlds out of elements of the stuff of thinking. It is the same "stuff as dreams are made on," the stuff of all our thinking, conscious and unconscious, given a rigorous mathematical structure, rather than the worlds created by matter from its "dark materials," as John Milton aptly described them, admittedly, referring to God's power to create "worlds" out of his dark materials (*Paradise Lost*, Book 2, l. 916).

On the other hand, our thought and thus these mathematical worlds are ultimately created from and by these dark materials, and thus, by *everything*: they are the creations of our bodies and brains, which are the products, creations, of evolutionary history of life, in turn made possible by the whole history of the universe, composed

of its dark materials, from the Big Bang on. It is true that our thought is likely to reach little of how all this was made possible. We certainly do not and possibly cannot know or even conceive of how, by means of what material processes, our bodies and brains make thinking and, hence, mathematics, or science, philosophy, and art, possible. That, however, still leaves us plenty to think about and create in these fields, including in dealing with the nature and practices of thinking. There is still so much in them to be invented and so much that we need from them in our life beyond them.

6 Cross-References

- ▶ Philosophy Behind Mathematics
- ▶ Platonism, *De Re*, and (Philosophy of) Mathematical Practice
- ▶ The Ghost and the Spirit of Pythagoras

References

A'Campo N, Ji L, Papadopoulos A (2016) On Grothendieck's construction of the Teichmüller space. In: Papadopoulos A (ed) Handbook of Teichmüller theory, volume VI. European Mathematical Society, Zürich, pp 35–70

Cartier P (2001) A mad day's work: from Grothendieck to Connes and Kontsevich, the evolution of concepts of space and symmetry. Bull Am Math Soc 38(4):389–408

Connes A (2003) Symétries Galoisiennes et renormalization. In: Duplantier B, Rivasseau V (eds) Poincaré seminar 2002: vacuum energy, renormalization, Progress in mathematical physics 30. Birkhäuser, Boston/Cambridge

Connes A, Kreimer D (2001) Renormalization in quantum field theory and the Riemann-Hilbert problem. II. Comm Math Phys 216(1):215–241

Connes A, Marcolli M (2004) Renormalization and motivic Galois theory. Int Math Res Notices 76: 4073–4092

Connes A, Marcolli M (2007) Noncommutative geometry, quantum fields, and motives. American Mathematical Society, Hindustan Book Agency

Corfield D (2006) Towards a philosophy of real mathematics. Cambridge University Press, Cambridge

Corfield D (2020) Modal homotopy type theory: the prospect for a new logic of philosophy. Oxford University Press, Oxford

Corry L (2003) Modern algebra and the rise of mathematical structures. Birkhäuser, Boston

De Bianchi S, Wells JD (2015) Explanation and the dimensionality of space Kant's argument revisited. Synthese 192(1):287–303

Deleuze G (1995) Difference and repetition (trans: Patton P). Columbia University Press, New York

Deleuze G, Guattari F (1987) A thousand plateau (trans: Massumi B). University of Minnesota Press, Minneapolis

Deleuze G, Guattari F (1996) What is philosophy? (trans: Tomlinson H, Burchell G). Columbia University Press, New York

Doxiadis A, Mazur B (eds) (2012) Circles disturbed: the interplay of mathematics and narrative. Princeton University Press, Princeton

Ferreirós J (2006) Riemann's Habilitationsvortrag at the crossroads of mathematics, physics, and philosophy. In: Ferreirós J, Gray J (eds) The Architecture of modern mathematics: Essays in history and philosophy. Oxford University Press, Oxford, pp 67–96

Ferreirós J (2010) Labyrinth of thought: a history of set theory and its role in modern mathematics. Birkhäuser, Boston
Gray J, Ferreirós J (2021) Epistemology of Geometry, The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), EN Zalta (ed). URL = https://plato.stanford.edu/archives/fall2021/entries/epistemology-geometry/
Heisenberg W (1962) Physics and philosophy: the revolution in modern science. Harper & Row, New York
Hilbert D (1979) Foundations of geometry, 10th ed (trans: Unger L). Open Court, La Salle
Husserl E (1970) The origin of geometry. In: The crisis of European sciences and transcendental phenomenology: an introduction to phenomenological philosophy (trans: Carr D), Northwestern University Press, Evanston, pp 353–378
Ji L, Papadopoulos A, Yamada S (eds) (2017) From Riemann to differential geometry and relativity. Springer, Berlin
Kant I (1997) Critique of pure reason (trans: Guyer P, Wood AW). Cambridge University Press, Cambridge
Kant I (2012[1747]) Thoughts on the true estimation of living forces. In: Watkins E (ed) Immanuel Kant: Natural science. Cambridge University Press, Cambridge
Kontsevich M (1999) Operads and motives in deformation quantization. Lett Math Phys 48(1): 35–72
Kuhn T (2012) The structure of scientific revolutions. University of Chicago Press, Chicago
Lakatos I (1980) Mathematics, science and epistemology, philosophical papers volume 2 (Worral J, Currie G, eds). Cambridge University Press, Cambridge
Langlands R, Shelstad D (2007) Descent for transfer factors. In: Cartier P, Katz NM, Manin YI, Illusie L, Laumon G, Ribet KA (eds) The Grothendieck Festschrift. Birkhäuser, Boston, pp 485–563. https://doi.org/10.1007/978-0-8176-4575-5_12
Laugwitz D (1999) Bernhard Riemann: turnings points in the conception of mathematics (trans: Shenitzer A). Birkhäuser, Boston
Manin Y (2002) Georg cantor and his heritage, ArXiv.math.AG/0209244 v1. 19 Sept 2002
Manin Y (2019) Time and periodicity from Ptolemy to Schrödinger: paradigm shift vs. continuity in history of mathematics. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Berlin, pp 129–138
Negrepontis S (2019) Plato on geometry and the geometers. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Berlin, pp 1–88
Nietzsche F (1966) Beyond good and evil: a prelude to a philosophy of the future (trans: Kaufmann W). Vintage, New York
Pambuccian V, Struve H, Struve R (2017) Metric geometries in an axiomatic perspective. In: Ji L, Papadopoulos A, Yamada S (eds) From Riemann to differential geometry and relativity. Springer, Berlin, pp 414–455
Papadopoulos A (2010) Introduction. In: Lobachevsky N, Pangeometry (edit and trans: Papadopoulos A) heritage of European mathematics. European Mathematical Society, Zürich, pp 229–234
Papadopoulos A (2017) Riemann surfaces: reception by the French school. In: Ji L, Papadopoulos A, Yamada S (eds) From Riemann to differential geometry and relativity. Springer, Berlin, pp 237–294
Pesic P (2007) Introduction. In: Beyond geometry: classic papers from Riemann to Einstein. Dover, Mineola, pp 1–22
Picard É (1891–1896) Traité d'analyse, 3 volumes, Paris, Gauthier-Villars
Plotnitsky A (2017) 'Comprehending the connection of things:' Bernhard Riemann and the architecture of mathematical concepts. In: Ji L, Papadopoulos A, Yamada S (eds) From Riemann to differential geometry and relativity. Springer, Berlin, pp 329–363
Plotnitsky A (2019) On the concept of curve: geometry and algebra, from mathematical modernity to mathematical modernism. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Berlin, pp 153–212

Plotnitsky A (2020) The ghost and the spirit of Pythagoras: the 20th and 21st century mathematics between and beyond geometry and algebra. In: Sriraman B (ed) Handbook of the history & philosophy of mathematical practice. Springer/Nature, London

Plotnitsky A (2021a) Returns of geometry: from the Pythagoreans to mathematical modernism and beyond. In: Papadopoulos A (ed) Topology and geometry: a collection of essays dedicated to Vladimir G. Turaev. EMS Press, Berlin

Plotnitsky A (2021b) Reality without realism: matter, thought, and technology in quantum physics. Springer, Heidelberg

Polya G (1990) Mathematics and plausible reasoning, volume 1: induction and analogy in mathematics. Princeton University Press, Princeton

Reed D (1995) Figures of thought: mathematics and mathematical texts. Routledge, London

Riemann B (1854) On the hypotheses that lie at the foundations of geometry. In: Pesic P (ed) Beyond geometry: classic papers from Riemann to Einstein. Dover, Mineola, pp 23–40

Riemann B (1991) Gesammelte mathematische Werke und wissenschaftlicher Nachlass (Narasimhan R, ed). Springer, New York

Russell B (1897) Essays on foundations of geometry. Cambridge University Press, Cambridge

Silverman J, Tate J (2015) Rational points on elliptic curves. Springer, Heidelberg/New York

Stäckel P (1913) Wolfgang und Johann Bolyai, Geometrische Untersuchungen, vol I. Teubner, Leipzig/Berlin

Teller P (1995) An interpretive introduction to quantum field theory. Princeton University Press, Princeton

Voevodsky V et al (2013) Homotopy type theory: univalent foundations of mathematics. Univalent Foundations Program, Princeton

Vuillemin J (1962) La Philosophie de l'algèbre. Tome I : Recherches sur quelques concepts et méthodes de l'Algèbre moderne est un ouvrage d'une grande technicité, à la fois mathématique et philosophique, Presses universitaires de France, Paris

Weil A (1946) Foundations of algebraic geometry. American Mathematical Society, Providence

Weil A (1962) Foundations of algebraic geometry, 2nd edn. American Mathematical Society, Providence

Weyl H (2013) The concept of a Riemann surface (trans: MacLane GL). Dover, Mineola

Weyl H (2021) Philosophy of mathematics and natural science. Princeton University Press, Princeton

"That Which a Minority Construct": Abelian Mathematics in Abel, Galois, Noether, and Grothendieck

Arkady Plotnitsky

We might as well say that minor no longer designates specific [mathematics] but the revolutionary condition for every [mathematics] within the heart of what is called great (or established) [mathematics].
—*Gilles Deleuze and Félix Guattari,* Kafka: Toward a minor literature. *(Deleuze and Guattari 1996, p. 18; paraphrased)*

Contents

1 Introduction 750
2 Abel and Galois: Minority Mathematics and the Rise of Modern Algebra 766
3 The Three Mathematics of Emmy Noether 774
4 Topos Theory and Grothendieck's Parliamentary Mathematical Ontology 786
5 Conclusion 796
6 Cross-References 797
References 797

Abstract

This chapter considers the nature of radical transformations of mathematics, enabled by minority mathematics. It will be particularly concerned with modern mathematics, which emerged roughly around 1800, as abstract mathematics – abstracted from mathematics' relations to the natural world and physics, relations that previously dominated mathematics. As, however, defined here (transferring Gilles Deleuze and Félix Guattari's concept of a minor(ity) literature, as exemplified by F. Kafka's work), a minority mathematics is not something that exists entirely outside a *major* mathematics, to be distinguished here from a majority

A. Plotnitsky (✉)
Literature, Theory, and Cultural Studies Program, Philosophy and Literature Program, Purdue University, West Lafayette, IN, USA
e-mail: plotnits@purdue.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_140

mathematics. Instead, it is a mathematics that, while still exterior to the major mathematics to which it juxtaposes, constructs itself within and even at the very core of this major mathematics. This may be seen as merely a special form of revolutionary vis-à-vis normal mathematical practice in T. Kuhn's sense. I shall argue, however, by using the work of N. H. Abel, É. Galois, E. Noether, and A. Grothendieck, as my main cases, that this "special" type of revolutionary practice is the primary and even the only form of revolutionary practice possible in mathematics. I designate this mathematics Abelian mathematics, the term commonly associated with formal mathematical properties (such as commutative group or abelian categories), because Abel's work was, arguably, the first manifested case of a minority mathematics in this sense in modern mathematics.

Keywords

Functoriality · Major mathematics · Minority mathematics · Modern algebra · Modern mathematics · Concept-form

AMS Classification

00A30 · 01A65 · 01A60 · 81P05

1 Introduction

This chapter considers the nature of radical transformations of mathematics, enabled by minority mathematics. It will be particularly concerned with modern mathematics, which emerged roughly around 1800, as abstract mathematics – abstracted from mathematics' relations to the natural world and physics, relations that previously dominated mathematics. As, however, defined here (transferring Gilles Deleuze and Félix Guattari's concept of a minor(ity) literature, as exemplified by Franz Kafka work) a minority mathematics is not something that exists entirely outside a *major* mathematics, to be distinguished here from a majority mathematics. Instead, it is a mathematics that, while still exterior to the major mathematics to which it juxtaposes itself, is created, constructed within and even at the very core of the major mathematics it transforms. This may be seen as merely a special form of revolutionary vis-à-vis normal mathematical practice in Thomas Kuhn's sense. I shall argue, however, by using the work of Niels Henrik Abel, Évariste Galois, Emmy Noether, and Alexandre Grothendieck, as my main cases, that this "special" type of a revolutionary practice is the primary and even the only form of revolutionary practice in mathematics. I designate this mathematics Abelian mathematics, the term commonly associated with formal mathematical properties (such as commutative group or abelian categories), because Abel was, arguably, the first manifested case of a minority mathematics in this sense in modern mathematics. (To highlight this difference, Abelian will be capitalized when designating a minority mathematics.)

I would like to stress from the outset that in the present view all such transformations of a major mathematics by a minority mathematics, and more generally all mathematical developments, continuous or discontinuous, are assumed to be local, even though they always have their (more or less) extended history. They are the product of a *technical* or *technological* mathematical practice: Each is a local invention of new mathematical technology or a local development of an already existing one, without having their origins in Platonist-like archetypal preexisting mathematical forms or any others preexisting mental reality, or being governed by a unique preordained form (however multifaceted) of mathematics in its historical developments. By the same token, the history of mathematics and its transformations becomes multiple. It is comprised of a manifold of histories proceeding along different trajectories, with multiple forms of interplay between continuities and discontinuities even within each such history, if one can ever rigorously, rather than provisionally, speak of a single history of any mathematical entity – a concept, a theory, or a field. In this chapter, I am specifically concerned with transformations enacted by means of what I define as minority mathematical practices. As I shall argue, however, such minority practices are always involved in and shape transformations, especially radical transformations, of mathematics.

In speaking of a minority mathematics, I adopt Deleuze and Guattari's concept of minor or, as I shall term it (for the reasons explained below) minority literature, introduced by them in their analysis of Kafka's work, and expand this concept to other creative endeavors, with mathematics as my main subject. This expansion is virtually inherent in their concept and transpires in their treatments of many figures in different fields, including mathematics, for example, Abel and Galois, or Bernhard Riemann.[1] Of course, a minority practice is specific to the nature of its field, such as literature and art, philosophy, or mathematics and science. Kafka's work exemplifies both the generality of the concept and the specificity of its application to literature, and further specificity that accompanies the work of each particular figure, in this case Kafka vis-à-vis other modernists, such as James Joyce or Virginia Woolf. To cite Deleuze and Guattari's definition of the concept, as applied to mathematics (which I substitute for literature in their formulation):

[1] Deleuze and Guattari saw Riemann's concept of a manifold (a form of minority mathematics, but its status as such would require a separate analysis) as heralding a revolutionary *philosophical* change, manifesting a minority philosophy, versus dialectic as a (post-Hegelian) major philosophy. This view is helped by the French term "*multiplicité*" as a translation of Riemann's *Mannigfaltigkeit* (which the English translation of Deleuze and Guattari's works often renders, understandably but incorrectly, as "multiplicity"). They say: "It was a decisive event when the mathematician Riemann uprooted the manifold from its predicative state and made it a noun, manifold [*multiplicité*]. It marked the end of dialectics and the beginning of the typology and topology of manifolds" (Deleuze and Guattari 1987, p. 483; translation modified). It could be shown, however, that to leave dialectics behind requires a traversal through dialectical thinking, which, or in any event, thinking in terms of unities (rather than multiplicities, even if unities that contain multiplicities within them) remains the dominant, *major*, philosophical thinking. Philosophically, this chapter advocates the philosophy of the multiple, ultimately uncontainable by a unity. For a discussion of Riemann's conceptual thinking, including as parallel to that of Galois, see Plotnitsky (2022), which also contains further references.

> A minor [*minorité*] [mathematics] doesn't come from a minor [mathematical discourse]; it is rather that which a minority constructs within a major [mathematical discourse]. But the first characteristic of minor [mathematics] in any case is that in it [mathematical discourse] is affected with a high coefficient of deterritorialization. . . . We might as well say that minor no longer designates specific [mathematics] but the revolutionary condition for every [mathematics] within the heart of what is called great (or established) [mathematics]. (Deleuze and Guattari 1989, pp. 18–19)

A minority is an exteriority: A minority thinking comes from the outside of the majority or (as explained below, they are not the same) major thinking to which it relates, even if its representative is an "insider," as, say, Riemann was at Göttingen in his time, rather than is an "outsider," as Abel or Galois was. But a minority mathematics (or a minority practice in any field in this definition) is an *exteriority from within a major mathematics*: It works *through* the interiority of a major mathematics, by interacting with a major mathematics and using it to create new mathematics. I do, thus, introduce a terminological asymmetry, which is, however, suitable and may be necessary here. The term "minority mathematics" is arguably more fitting. It is in effect suggested by Deleuze and Guattari by speaking of a minor literature as "that which a *minority construct* within a major (or established literature)" [emphasis on minority added]. Thus defined, a minority mathematics qua mathematics, such as that of Abel or other figures considered here, is not minor or marginal (any more than a minor literature qua literature is), quite the contrary, although it may be and may remain marginal for a while, even a long while. ("Minoritarian" has been used as a translation of the French "*minorité*," but the term has other meanings in English, such as related to a political power of a minority in a society, which can make it misleading.) Hence, the term "major mathematics" (or "major literature") is a more accurate reflection of the situation, insofar as a minority mathematics (or a minority literature, such as that of Kafka) transforms the central or the most advanced part of the existing mathematics, which may not necessarily be a majority mathematics. Accordingly, I shall adopt this asymmetrical terminology as more precise in reflecting what a minority or Abelian mathematics does. "Minor mathematics" is a linguistically more graceful term, but "minority mathematics," or "minority mathematical practice," is more precise in reflecting the role of a minority, sometimes consisting of a single person. A few immediate qualifications are in order, beginning with the concept of a single-person minority.

A minority thinking is never entirely individual, even if the corresponding practice is individual or nearly so, as it was in the case of Abel or Galois. Either case is probably as close to a single-person minority in mathematics as one can find. One's mathematical thinking, however, is always, to one degree or another, affected by the preceding history of mathematics, because even the most innovative concepts or theories have their history (which is always collective, however individualized it may become), and by extramathematical, such as cultural, factors, which give any individual thinking collective dimensions. Such exterior elements are manifested more in literature, especially, because of the role of ordinary language in shaping the difference between a major (or in this case, a majority) and minority literature (as German vs. Czech, or Hebrew, in Kafka). The role of extramathematical factors

impacting mathematical thinking and practice has been extensively considered in historical studies of mathematics. Indeed, the analysis of "exterior" personal and cultural, including political, factors in the workings and development of mathematics and science has been dominant in the historical scholarship of mathematics and science during the last half a century or so, especially in the so-called constructivist school or rather schools. There have been several such schools, sometimes different in their approach, and the number of works in these areas is massive by now, although most of these works are devoted to science, rather than mathematics.[2] These works challenge the uncritical separation (previous common in the history and philosophy of mathematics and science) of the "interior" and the "exterior" in mathematics and science. On the other hand, the role of (conventionally) "exterior" factors is often difficult definitively to justify as concerns their weight relative to the role of (conventionally) "interior" considerations, the role of mathematics itself. The latter can still be sufficiently demarcated as such, even if one adopts a more critical view of the exterior and the interior of mathematics as relative, case-dependent, rather than absolute. Was Galois' concept of group or Grothendieck's concept of topos impacted by extramathematical factors, for example, given that in both cases one deals with figures with strong political views, and in Galois' case with a political revolutionary? Perhaps, and even likely! The degree of this impact is, however, usually difficult to establish, in contrast to the interaction of either figure with the institutional mathematical establishment, although even this aspect of either case still poses complexities and has been debated, especially, in the case of Galois. In any event, the present chapter is not this kind of study. Mathematics itself, specifically as the interaction between a minority and a major mathematics, is my main concern, and it takes precedence throughout. While exterior elements will be brought up, for example, in commenting on the minority situation of the figures considered in relation to disciplinary or institutional mathematics, my main emphasis will be on mathematical thinking and primarily individual mathematical thinking. I will suggest certain political implications of my argument. This part of my argument is, however, very different from those of the works with the constructivist orientation just mentioned. It is not my aim to challenge this orientation and projects that it led to, some of which have been convincing and important. My aim is to offer a different type of view on the nature of the political in connection with mathematics.

It is essential that, as defined here, a minority mathematics as the transformative, revolutionary, force for mathematics always arises *within* a given major mathematics, rather than (by definition) only in interaction with the latter from its intellectual exterior, as it does in the case of Abel or Galois. It would, however, be difficult to

[2] I mention here some key precursors or founding figures, such as Imre Lakatos (one major philosopher in this field who was primarily concerned with mathematics), Thomas Kuhn, Paul Feyerabend, and a few key recent figures, such as Ian Hacking (who has also written on probability theory), Bruno Latour, Peter Galison, and the Edinburgh school, founded by David Bloor. While addressing minority figures or groups (in the conventional sense), none of these authors or, to my knowledge, any other authors in this field consider concepts similar to those of a minority mathematics (or science) in the present sense.

think of Karl Friedrich Gauss, Riemann, and David Hilbert, whose thinking led to revolutionary mathematical transformations, as minority figures in the way Galois and Abel were (for both mathematical and extramathematical reasons). Nevertheless, even in cases like those of Gauss, Riemann, or Hilbert, there must be (in their case was) some minority mathematics for a revolutionary transformation of mathematics to be possible, although Riemann was more a minority figure overtly. Such a transformation is only possible when a minority thinking restructures or reterritorializes major mathematics, whether this minority thinking is that of a manifested outsider or that of an "official" insider, who must move to a minority-type outside to change mathematics. Cases of the first type, however, such as those of Abel, Galois, Noether, and Grothendieck, allow one to see better how, to return to Deleuze and Guattari's formulation, "minor [is] no longer a specific [form of mathematics] but [is] the revolutionary condition for every [mathematics] within the heart of what is called great (or established) [mathematics]," all major mathematics. Conversely, no minority mathematics is possible without the major mathematics which it affects. A minority mathematics is a revolutionary condition for every mathematics, but by the same token, it can only be this condition *within* a major mathematics. The exteriority of a minority mathematics is an exteriority from within the major mathematics with which this minority mathematics interacts.

The term "deterritorialization," uncommon outside the context of Deleuze and Guattari's work, needs to be explained. While Deleuze and Guattari's *concept* so designated, used by them in a broader set of contexts, would require an extended treatment, its meaning in the present context is reasonably straightforward. It means expanding a given territory, while also reshaping, *reterritorializing*, this territory by means of new conceptions and practices, such as minority ones, although it can also be done otherwise. A clear example is the deterritorialization/reterritorialization of the part of algebra dealing with solving polynomial equations by radicals by making it the mathematics of group theory by Galois. In this respect, Galois' work went beyond Abel's already radical deterritorializing and then reterritorializing (restructuring) the field of mathematics in the context of this problem itself, although, as discussed below, Abel's overall thinking was reaching much further as well. In any event, Galois' group theory is one of the greatest mathematical deterritorializations and reterritorializations ever. The deterritorializations and reterritorializations enacted by Noether and Grothendieck's restructurings of, respectively, algebra and algebraic geometry are far reaching as well, although it is hard to match the concept of group in this respect.

While, however, arguing for the interaction and in this respect continuity between minority and major mathematics, I do maintain that the practice of mathematics contains both continuity and discontinuity, in a complex and (from case to case) varied interplay, including between minority and major mathematics. This interplay always, in each local case (in the present view all such cases are, again, local), contains both a continuity and a discontinuity, even a radical, revolutionary discontinuity, just as it does in other fields, such as literature. Modernism, such as that of Kafka (or James Joyce, Samuel Beckett, and others in literature, Pablo Picasso and Wassily Kandinsky in art, or Arnold Schoenberg in music), was a revolution that

transformed established major literature and art, while retaining continuity with it. Although certain key modernist trends were shared by these figures, the interplay between the major and minority art was also specific in each case, given the individual thinking of each figure and the particular cultural situation in which their minority practice took place. The same is true in mathematics, for example, in Abel and Galois' cases, which had both shared and strictly individual aspects in their thinking and their cultural situations. It may not always be possible, and arguably is never ultimately possible, to *fully* track how new trajectories emerge from and against previous trajectories of thought, including all elements of minority thinking involved in any revolutionary practice. This limitation, however, leaves room for meaningful assessments of this emergence in terms of both continuities and breaks. Algebraic geometry, as transformed by Grothendieck, is a case in point. Grothendieck took advantage of and, first, transformed, by a minority thinking, established concepts of category theory, cohomological algebra, and sheaf theory, already in interaction with each other, even before he transferred them into algebraic geometry, where his work was most transformative.

Continuities do not prevent new theories from breaking with old ones and changing the nature of mathematics in the process but instead help these breaks. Continuities with a major mathematics are, by definition, necessary for breaks enacted by a minority mathematics, although never sufficient for these breaks. Galois' invention of the concept of group was helped by the major algebra of polynomial equations, but as a minority concept a group comes from elsewhere in a broader landscape of Galois' thinking. It was a very complex landscape, far exceeding mathematics and shaping the "elsewhere" that gave rise to the idea of a group. Also, previous major practices may continue alongside revolutionary changes or may change more continuously with previous practices. Not all major literature was transformed into modernism by a minority literature, such as of Kafka or Joyce. In both cases the role of minority (they were not minor!) languages, Czech and and Irish, were crucial. Joyce expressly played with this linguistic multiplicity, by adding other languages into that of his novel, in his final masterpiece, *Finnegans Wake*. Realist literature and "milder" modernist, or mixed, literature, including in good and important works, have continued to remain prominent or even dominant, and they still do.

It may be that, as Yu. Manin contends, there is more continuity in mathematical practice than scientific ones, specifically in physics (Manin 2002, 2010, 2019).[3] It is

[3] As an illustration of his thesis concerning the continuity of mathematical practice (keeping in mind that his article accomplishes a great deal more), Manin offers an analysis of the history of the mathematical theory of time and periodicity, from Ptolemy's epicycles to Schrödinger's quantum amplitude interference and Feynman path integrals, with a few key junctures in between, such as Fourier sums and integrals, which are also central to quantum theory (Manin 2019). One cannot deny this overarching continuity and local continuities with it, elegantly traced by Manin. It is also true that the role of such continuities deserves more attention than it is commonly given, versus revolutionary breaks within this history, which tend to have attracted most commentaries. I would argue, however, that it is difficult to see this history apart from several breaks, sometimes radical

certainly true that major accomplishments are possible by means of primarily, even if not entirely (which, I would argue, may not be possible), continuities with preceding developments, for example, with a major mathematics. This can, however, happen in physics as well: There are cases when the continuity succeeded, at least more so than expected, while attempts at revolutions failed, which also may (but need not) mean that the failure was in not finding a right revolutionary path.[4] In general, it is not always clear, as Kuhn realized and has shown in the case of Max Planck's discovery of quantum theory (Kuhn 1987), and may be a matter of interpretation and debate where, or whether, a normal practice ends and a revolutionary practice begins. I would, accordingly, argue that the balance between them is contingent, shaped by the circumstances of a given case. Thus, while set theory and its concepts are part of the history of category theory and its concepts, as Manin contends, there is, I would argue (contesting Manin's view) just as much, if not more, discontinuity between them. Grothendieck's use of category theory in his work,

breaks, even as concerns the emergence of revolutionary mathematical concepts involved, such as those of Fourier's analysis or the calculus of variations, used by Schrödinger, or the role of complex variables. (Each of these concepts has its own history of continuities and breaks with their predecessors.) These mathematical concepts also reflect and have shaped an even more radical change in physics during this history brought in by quantum physics versus classical physics. Manin acknowledges that such developments and their use of mathematics are "more sensitive to respective 'revolutions' and 'paradigm shifts,' whereas those parts that are closer to 'pure mathematics' show rather a kind of continuous development, as [he] argue[s] in [his] paper" (Manin 2019, p. 130). It is, as I said, possible that there is more such continuity in mathematics overall. I would, nevertheless, argue that either history, that of pure mathematics and that of physics, including in its use of mathematics, shows the interplay between continuities and breaks (again, more complex than that between normal and revolutionary practices), different at different junctures. Some of them in mathematics exhibit changes, such as those defined by new concepts, that are as revolutionary as in physics. This is, again, not to say that such changes do not have history and, hence, continuities with preceding concepts or ways of thinking, including those defined by the interactions between major and minority mathematics found throughout this history. In physics, too, one can see quantum mechanics as, mathematically, a reconstruction (it was called quantum-theoretical reinterpretation at the time) of Hamiltonian and Lagrangian mechanics, the dominant formulation of classical mechanics then. Mathematically, quantum mechanics or, in high-energy regimes, quantum field theory is a major and a majority quantum theory now. Attempts at alternatives, however, never stopped, primarily in view of the epistemological discontent with the theory, most famously expressed by Albert Einstein, but they have not thus far achieved the transformative minority role that quantum mechanics had. See also Note 1 above.

[4] The history of renormalization in quantum electrodynamics (QED) offers an instructive example of this situation. It was believed by many, including Paul Dirac, the founder of the theory, that its problems, having to do with the appearance of infinities (divergent integrals) in QED, could only be resolved by means of an alternative theory, and hence by a revolutionary break, akin to that from classical to quantum theory. As it happened, the renormalization program was able by 1950 to handle these difficulties within QED itself, as it existed. In fact, QED is now the best confirmed physics theory ever, although the discontent with its nature, as requiring renormalization, has never entirely disappeared (Dirac never came to terms with it) and is still around. See Schweber (1984) for a helpful historical account of QED and Plotnitsky (2021, pp. 273–305) on the epistemological aspects of quantum field theory. See also Plotnitsky (2022, pp. 271–274) on recent approaches to the mathematics of renormalization via motivic Galois theory, bringing both Galois and Grothendieck's mathematics into it.

vis-à-vis previous history of algebraic geometry, as in part based on set theory, as in case of André Weil, who, as will be seen, disparaged category theory (Weil 1962, p. 358). Grothendieck's use of category theory or (and in) sheaf cohomology, by then part of a major mathematics in algebraic topology, was a minority mathematics there. (Admittedly, Jean-Pierre Serre already started to bridge both fields in terms of sheaf theory.) The same case can be made for Galois' use of group theory or Noether's rethinking of Dedekind's ideal theory, an established part of the major algebra then, or, more radically, Noether's algebraically defined minority ventures into topology. Her insights, with the help of her students and followers, such as Heinz Hopf and Pavel Alexandrov, had a crucial impact on algebraic topology.

Abel and Galois' mathematics might be seen as more radically *exterior* minority mathematics (still, again, inevitably related to the major mathematics of their time), while that of Noether and Grothencieck as being closer to the major mathematics on their respective situations, except for Noether's minority intervention (from the exterior territory of abstract algebra) into algebraic topology. Still, topology needed topologists, some of which were Noether's students or followers, to build on her ideas to make them part of major topology. A minority mathematics and the major mathematics (or whatever field where both types of thinking are at work) need and depend on each other. Only through a minority mathematics, however, can mathematics transform itself. To return to Deleuze and Guattari's expression of this relationship, minority mathematics is the revolutionary condition for every mathematics in the heart of great (or established) mathematics, in the heart of a major mathematics.[5]

In the remainder of this introduction, I outline a general philosophical view of mathematics, especially creative mathematics, as conceptual mathematics, following (Plotnitsky 2022). Focused (in addition to the role of concepts) on epistemological questions, that study did not consider, except by implication, the role of a minority mathematics, which is, as I argue here, essential to all creative mathematics, including the invention of concepts. This chapter primarily concerns modern mathematics, as abstract mathematics. This mathematics began to emerge around 1800, consolidated, as the dominant form of major mathematics, around 1900, and has continued as such to our own time. The abstraction in question is double: Modern mathematics is abstracted from (A) the physical world, mathematically representing which defined modern physics; and (B) from our daily phenomenal intuition. (A) and (B) are connected because our intuition also contains the phenomenal representation of the external world. It was in fact this representation that was mathematically idealized by classical physics, from which modern mathematics aimed to abstract

[5] Manin might have been more in agreement with the present view of the role a minority mathematics in transforming a major mathematics, which retains some interactive continuity between them or thus a general continuity of mathematical practice. While, however, admitting this element of continuity within the interaction between a minority and a major mathematics, I do maintain that the practice of mathematics contains both continuity and discontinuity, in a complex and varied interplay, including between minority and major mathematics.

itself as well. A phenomenal reality could of course be mathematically represented apart from physical reality, beginning with the reality of mathematics itself.

The term "concept" is often used uncritically, without being defined or by merely assuming some common understanding of it, in mathematical (or scientific) or even philosophical literature. There is no single definition or concept of a concept. The concept of a concept assumed here is parallel to and in part follows that of Deleuze and Guattari's definition of a philosophical concept as grounding a creative practice in mathematics (Deleuze and Guattari 1996). This practice is that of the invention of new philosophical concepts, in juxtaposition to mathematical and scientific concepts. The latter, while casually mentioned sometimes, are in fact not defined by Deleuze and Guattari. They even deny that mathematics or science have concepts in their sense and see them defined by logical propositions and formal mathematical or scientific elements, such as functions, a surprisingly conventional view for these unconventional philosophers, including in other aspects of their view of mathematics and science. "The concept," they say, "belongs to philosophy and only to philosophy" (Deleuze and Guattari 1996, pp. 11–12, 33–34). By contrast, I shall define mathematical or scientific concepts on the model of their concept of a philosophical concept, following the argument given in (Plotnitsky 2022, pp. 52–76), which also offers a critique of Deleuze and Guattari's view of the role of concepts in philosophy versus mathematics and science.

This chapter, accordingly, gives preference to the role of conceptual thinking rather than in logical propositions or calculations, which, however, could also be shaped by a minority mathematics, as were some of Abel's methods of calculations. Both logic and calculations are essential to mathematics, including conceptual thinking there, because the structure of mathematical concepts is logical, and because calculations require concepts and can lead to new concepts. Nevertheless, while respecting these features, this chapter views creative mathematical thinking as most essentially defined by the invention of new concepts, which, I further argue, is only possible by means of a minority mathematics.[6]

In the present *general* understanding, following, as a general understanding, that of Deleuze and Guattari, a concept is not merely a generalization from particulars (which commonly defines concepts in linguistics, analytic philosophy, or cognitive psychology) or a general or abstract idea, although a concept may contain such generalizations or ideas, specifically abstract mathematical ideas, where they may also be concepts in their own right. A concept is a multicomponent structure, defined

[6] There are still other aspects of mathematical thinking, for example, a narrative one, the role of which, as constitutive rather than merely auxiliary (which has always been recognized), has received a considerable amount of attention during recent decades. For a representative collection of essays on the subject, see Doxiadis and Mazur (2012). For a full disclosure, the present author was among the contributors, although his article questioned, on epistemological grounds, the degree of the efficacy of narratives in mathematical invention or argumentation, advocated by other contributors (Plotnitsky 2012). In any event, my emphasis in this chapter is on concepts as the primary creative vehicles of mathematics, and the role, and one indeed might say, the narrative of minority mathematics, a narrative not discussed by the contributors in (Doxiadis and Mazur 2012).

by the organization, *composition*, of its components, and some of these components may be concepts in turn. Of course, every concept, no matter how innovative, has a history defined by other concepts (or other forms of thought) from which this concept emerges and without which it would not be possible. Just as any theory (which, while it may have other components, is always a conglomerate of concepts), a new concept is a product of an interplay of continuities with and breaks from preceding concepts. Some concepts considered here, such as those of a group by Galois or a topos by Grothendieck, were more radically innovative. Still, however, they had connections with earlier concepts, such as (putting aside the earlier history of which) those of Adrien-Marie Legendre in the case of a group and those of category theory in the case of a topos.

Riemann's thinking and practice, discussed by this author on several previous occasions (e.g., Plotnitsky 2022, pp. 99–135), provide arguably the most prominent example of the conceptual approach to mathematics, especially in geometry, as based on the concept of a manifold [*Mannigfaltigkeit*]. This approach led him beyond all previous approaches to geometry, including those, axiomatic in nature, that led to non-Euclidean geometry. Riemann shunned axiomatic thinking in favor of conceptual thinking, in accord with the present general definition of a mathematical concept. Riemann himself did not offer such a general definition, but his concepts, such as and in particular that of manifold, fully conform to the definition adopted here and may, as indicated above (Note 1), be seen as a model for Deleuze and Guattari's concept of a (philosophical) concept. Conceptual thinking plays, however, an essential role in the thinking of all the main figures considered here, although the relative balance between conceptual thinking and other forms of mathematical thinking, such as an axiomatic one, is different in each case. For example, axiomatic thinking plays a greater role in Grothendieck's work, although the minority part of his thinking was essentially conceptual, also shaping his view of axiomatization. It is, again, not my intention to diminish the role of the logical or calculational aspects of mathematics or science, or to take anything away from the contribution of philosophical investigations, such as those in the analytic philosophy of mathematics, that focus primarily on logical and propositional structures of mathematics or science.[7] My aim is to give a proper emphasis to the role of concepts, especially in *creative* thinking in mathematics or science, but not only in creative thinking, because working with already established concepts is indispensable in all mathematical or scientific practice. The unconditional opposition between the logical-propositional and conceptual structures, or between calculations and concepts, is not so easy to maintain, even in considering figures, such as Galois, Riemann, or Dedekind, who gave concepts a dominant role in their work. While their thinking was primarily governed by concepts vis-à-vis calculations, they were perfectly capable of and did perform difficult calculations. Besides, calculations involve

[7] There are new powerful approaches to logic in mathematics, such as homotopy type theory (Voevodsky et al. 2013; Corfield 2020). Concepts, however, still play an essential and unavoidable role in them.

concepts and lead to new concepts, as they did in the case of Abel, or Leopold Kronecker, who insisted on their primary role in mathematics. I would still argue, however, that the invention of new concepts is the primary feature of creative mathematical practice, accompanied and defined (the degree may vary) by the role of a minority mathematics, again, in the present definition, working within a major mathematics.

The definition of a concept given above was only a general, *abstract*, definition. What meaningfully establishes a concept is the specific character of its compositional structure, defined by both the nature of each component and the character of their relations within this composition, by how these components relate to each other in the structure of the concept. This composition is unique in a new concept, and it makes this concept unique. The creative essence of concepts is in their compositional individuality, which is only possible by virtue of their general compositional nature but is defined by the specificity of this composition. All musical works are compositions, but each gives us a different music. A group is composed of its elements, but in accordance to very specific rules, which give it its structure and role in mathematics in group theory and beyond, and then it multiplies further to still more specific concepts of a group, a finite group, a Lie group, an algebraic group, and so forth, and then into further concepts within each of these. A single-component concept is a product of a provisional suspension or cutoff of its multicomponent organization. In fact, there are always cutoffs in delineating a concept, which result from assuming some of the components of this concept to be primitive entities whose structure is not specified. These components could, however, be specified by alternative delineations, leading to a new concept, containing a new set of primitive (unspecified) components.

Consider the mathematical concept of space. Emerging historically at the intersection of mathematics, physics, philosophy, and general phenomenal intuition, and defined by such constitutive concept-components as point, line, plane, or distance, this concept or each of these components has a long history of definitions, modification, and transformations. As a general phenomenal concept, it is arguably as old as human consciousness. Kant did not see space (or time) as a concept, but rather as an a priori form of phenomenal intuition. He had a point, even if one rejects (as many have done) Kant's a priori view of space and time: In his view, these phenomenal forms precede any concept of space and time. As a geometrically representable physical concept, space is at least as old as ancient Greek geometry, well preceding Euclid, who did not define, although, arguably, implied, the concept of space in the *Elements*. The history of the concept of space, as a mathematizable physical and eventually (in the nineteenth century) purely mathematical concept, extends from the Pythagoreans to Euclid to Descartes (a coordinate space) to Riemann (a space defined as a manifold) to Felix Hausdorff (topological space) to Grothendieck (topos), with further versions within each. In the first four cases, the concept of space qua space was, while mathematizable, still physical. After each transformation, which was a product of a minor mathematics, the resulting new concept of space became part of a major mathematics, or physics.

A concept, then, is an assemblage of its components connected by its composition, which define its specific individual and, at the time of its invention, unique

nature. Some of its components may be unique but need not be. The components of a concept need not be new, but its composition must be new for a concept to be new. In this respect, a concept is *akin* to a work of art, always defined by its composition, in accord with the ancient Greek meaning of *poiein* (a making of an entity), as making something new, in which the term poetry originates. A theory, or a disciplinary field, mathematical, scientific, or philosophical, such as geometry, classical physics, or Kant's philosophy (philosophy may be more individualized), becomes an organized assemblage of concepts. It is a product of a broader *poiein*, which in mathematics or science may be commonly collective, although a theory may be initiated, but never contained, by a single new concept, such as Galois theory, initiated by the concept of group. In the present view, however, a *creative* conceptual *mimesis* (which may borrow, "imitate," elements of previous concepts but combines them in a new way) is a partial *technical* mimesis of previous concepts, rather than a Platonist mimesis of something that belongs to a preexisting independent reality.

Of course, theoretical concepts (or theories) are, in their compositional nature and functioning, different from works of art. For one thing, a work of art need not contain concepts, although so-called conceptual art always does. In functioning, a concept can be used directly or with minor modifications in, again, Kuhn's terms, the normal practice in mathematics and science. Philosophy, at least a certain form of philosophy, for example, as defined by Deleuze and Guattari (1996), may be associated strictly with the invention of concepts as a revolutionary and hence a minority practice, which thus becomes its normal or major practice. In this respect, this practice is closer to that of art, which, while it may use a tradition, is compelled to be innovative and in this case a minority art, but only in this respect, because the nature of the conceptual and artistic composition remains different.

What is shared in the creative nature of art, philosophy, and mathematics and science is *poetically* (again, in the ancient Greek sense of *poiein*) the role of composition, and *politically* the minority nature of creative thought, the political essence of which is freedom. All these fields and all creative human endeavors first assume, as Deleuze and Guattari argue, the necessity of a war they, by their different means, wage against the dogmatism of opinion or any dogmatism, the greatest enemy of creative and, hence, free thought, which is always a minority thought (Deleuze and Guattari 1996, p. 202). Although Deleuze and Guattari did not expressly make this point, their argument concerning a minor (minority) literature, certainly in Kafka's case, clearly implied it. A minority literature is, by definition, a creative practice of literature, and as such is always at war with, and is a war against opinion, and it poses a great danger to opinion and majority thinking, or even major thinking, when the latter becomes governed by opinion. The dogmatism of opinion is, in the first place, dogmatism, to which Hegel juxtaposes true thinking, as defined by a free creative mediation. "*Dogmatism*," he says in the Preface to *The Phenomenology of Spirit*, "the way of thinking, in both knowing or in the study of philosophy, is nothing but the opinion that truth consists in a proposition which is a fixed result or else in a proposition which is immediately known [*der unmittelbar gewußt*]" (Hegel 2019, p. 31). In other words, dogmatism forgets or disregards the mediating and especially creative (all creation is mediation, but a special one) role of

thinking and makes one surrender one's capacity for true thinking. The dogmatism of opinion was an enemy of thought for all figures considered in this chapter. They had to confront and fight opinion, not the least by means of new concepts, as Riemann argued in closing his Habilitation lecture. New concepts, such as those his lecture introduced, are, he said, needed "to insure that [our] work is not hindered by unduly restricted concepts and that progress in comprehending the connection of things is not obstructed by traditional prejudices" (Riemann 1854, p. 33). Such restricted, and *restricting*, concepts are often part of a majority mathematics and sometimes even of a major mathematics. On the other hand, new concepts or all creative thought always require one or another degree of a minority thinking and hence a freedom in deciding to pursue a minority thinking rather than follow a majority thinking, even when it is a major thinking. A decision is, I would argue, a better category than choice, at least a free choice, because there may always be factors, psychological, social, or cultural, that shape our decisions, restricting and sometimes precluding their freedom. It does not follow, however, that there is no freedom in our decisions, although this type of argument is made sometimes (usually in the form of denying the existence of free will). The category of decision allows for the *possibility* of freedom, which, while always relative, is crucial. Also, in saying that the *essence* of minority thinking is freedom, the possibility of freedom, I *do* refer to its political *essence*. I am not saying that mathematics (or science) always involves politics, either within mathematics itself or in relation between mathematics and culture. This may be true, but it means little, without giving such claims due specificity. My point here is specific and different in nature: The *essence* of minority thinking in mathematics or elsewhere is the possibility of freedom, because a minority thinking and hence creative mathematics would not be possible otherwise. I shall return to this point in closing this chapter to argue that mathematics allows for an extraordinary and even unique degree of freedom, and that a minority mathematics, defining all creative mathematics, is *governed* by this freedom rather than by a majority mathematics, even when it is a major mathematics. Deleuze and Guattari, in this connection, also speak of a majority mathematics or science, governed by a major mathematics, as "royal" or "state" mathematics (or science), and a minority mathematics or science, as nomadic, specifically juxtaposing Gaspar Monge's (state) geometry to Girard Desargues' (nomadic) projective geometry (Deleuze and Guattari 1987, pp. 363–365). Creative mathematics is grounded in and is an affirmation of a minority thinking and of freedom in mathematics. There is no creative mathematics otherwise. Although creative mathematical decision-making may, as Georges Polya argued, be based on plausible reasoning (Polya 1990), we still have the possibility of freedom to decide on the line of thought to pursue. There is no possibility of new mathematics without a minority mathematics and without the possibility of freedom to pursue it.

Now, according to Deleuze and Guattari, a philosophical concept, as defined by them, is also a problem, a positing of a problem. This view has a mathematical genealogy, although the concept of a problem preceded mathematics and was adopted by it and given a mathematical specificity by ancient Greek mathematics. A problem in this sense is not something that, like a theorem, is derived from

assumed axioms by means of logical rules but is something that is posed, created, along with and as a concept. According to Deleuze and Guattari's definition of these connections, however, as applied to philosophical concepts, a concept as a problem persists and thus remains a problem, rather than disappears, or is *solved*, in its solutions (Deleuze and Guattari 1996, p. 5). This is a complex conception, which is not easy to interpret, especially given the paradoxical nature of this formulation, and I shall not attempt to do so. Nor shall I adopt this concept as such, in part, because this conception is applied by them to philosophical concepts as problems, rather than to mathematical problems, which, in their view, are always, in principle, soluble, rather than continue to remain problems in their solutions as philosophical concepts do. Instead, I shall use a different mathematical concept proposed in (Plotnitsky 2022), that of "a concept-form," without claiming that Deleuze and Guattari would have subscribed to this concept, either in relation to their concept of a philosophical concept or otherwise.

A (new) concept-form, such as that of a group by Galois or a topos by Grothendieck, would be *introduced*, just as any concept, in order to solve an existing problem, which defied a solution by means of already existing concepts or had a solution that was unsatisfactory, conceptually, technically, or even aesthetically. Finding the algebraic solution to a polynomial equation (or proving the impossibility of doing so) in Galois' case was such a problem, leading him to the concept of group. There are thus two approaches (with gray areas between them and sometimes combining both) to solving a problem, especially a difficult one: The first is essentially technical, which aims to solve it by means of logical and technical manipulations of previously existing concepts; and the second, adopted by Galois, is essentially conceptual, which uses a new concept or set of concepts, although technical and logical manipulations remain necessary, as they are in all mathematics. Such a concept may be relatively simple, such as that of a group, although this simplicity is retrospective. But, even if it is, such a concept usually arises in response to a complex problem. This was certainly so in the case of the concept of group, which was invented by Galois to find an algebraic solution to a polynomial equation (and, as it happened, proving the impossibility of doing so for the equations of degree five and higher). This was a very complex problem at the time. By means of this concept, Galois redefined this problem by establishing that whether such a solution exists or not is related to the structure of a group of permutations associated with the roots of the polynomial, known as the Galois group of this polynomial. Abel's approach had elements of this thinking by posing the question whether a solution exists or not, given the structure of the equation, and finding that in general it does not. This finding is now known as the Abel-Ruffini theorem or Abel's impossibility theorem. Paolo Ruffini gave an incomplete proof in 1799, and Abel was the first to prove it in 1824. Galois accomplished much more, by making a group a concept-form.

As a concept-form, however, a new concept does more than merely solve a given (initial) problem. It defines a new problem-posing field in which new types of actual problems emerge and are being solved, while the concept-problem that initiated and governs this field is not something that is "solved" in the sense of no longer being

necessary for this field to continue to develop, which it does by defining new concepts and concept-forms, some of which are new forms of the original concept-form. In this way, one can maintain a certain affinity with Deleuze and Guattari's understanding of a concept as a problem persisting in its solutions, rather than disappearing in them. I am, again, not claiming that the concept of a concept-form offers an interpretation of Deleuze and Guattari's concept of a concept-problem. In fact, it cannot be because the latter concept is philosophical, while the concept of a concept-form is mathematical. All that matters here is that the concept of a concept-form works, as I think it does, in mathematics and science. Galois' concept of group not only gave the Abel-Ruffini theorem a more general grounding but also gave algebra a new field, Galois theory, in which new problems were subsequently posed, as they still are, which give Galois theory its continuing life. The same is true about Grothendieck's topos theory. This type of invention of a new concept in response to a concrete problem not only in solving this problem but also making this concept a concept-form can happen as an outcome of a deliberate strategy or it might happen contingently, in a complex interplay between chance and necessity.

The concept of a concept-form is both a tribute to Plato's thinking and a critique and ultimately abandonment of all Platonism, including Plato's own, which is different from many other Platonisms, including mathematical Platonism, a twentieth-century conception, quite different from the way of Plato's thought, including about mathematics. It is a tribute insofar as it affirms a creation of new conceptual forms which then shape the subsequent development of thought in a given field and beyond it. Thus, the concept of group has shaped and continues to shape both mathematics and other fields, such as physics or other natural sciences, and even social sciences. On the other hand, against Platonism (now including Plato's own) a concept-form is not something that preexists the history of thinking related to it and is then discovered by human thought, which may only conceive of this something itself partially or approximated. It is still human thought that assumes the existence of this something as a real preexisting human thought. This assumption is Platonism in its ultimate form. By contrast, as explained, each concept-form emerges in and shapes this history in the complex interplay of chance and necessity in the work of individual figures and communities. It comes from, *emerges in*, a contingent history, even though it can shape the subsequent history in, one might say, quasi-Platonist fashion. I qualify by "quasi" because new concepts or concept-forms shaped by this concept-form are not mimetic shadow-images of preexisting (eternal) forms, as in Plato (Plotnitsky 2022).

By the same token, a concept-form need not be seen as containing or fully determining a problem-posing field it creates, even if only as a potentiality. This field emerges in a contingent development, as a complex interplay of causal chains and chance occurrences. In general, in adopting a problem-posing versus axiomatic-theorematic view of mathematics, one is more likely to reject a Euclidean, essentially Platonist, view in which axioms are assumed as pregiven truths rather than created as concepts first, as axioms are from the present perspective, a distinction that was brought into the foreground by the discovery of non-Euclidean geometry. The present view is close to Imre Lakatos' conception of a quasi-empirical theory versus

a Euclidean theory, modeled on Euclidean geometry. ("Quasi" reflects the fact that mathematics is "empirical" in dealing with mental rather than material phenomena.) Lakatos sees creative mathematics as quasi-empirical, consistently with or even implying the key role of a minority mathematics in it.

Reciprocally, the present view is quasi-empirical, even though some and even most of the figures considered here, such as Abel, Noether, or Grothendieck, might have adopted a more Euclidean philosophy of mathematics. Lakatos also argues that no Euclidean ideology can avoid quasi-empirical practice, especially as creative practice, any more than a creative major mathematics can avoid becoming a minority mathematics. According to Lakatos:

> The development of a quasi-empirical theory is very different. It starts with problems followed by daring solutions, then by severe tests, refutations. The vehicle of progress is bold speculations, criticism, controversy between rival theories, problem shifts. Attention is always focused on the obscure border. The slogans are growth and permanent revolution, not foundations and accumulations of eternal truths.
>
> The main pattern of Euclidean criticism is suspicion: Do the proofs really prove? Are the methods used too strong and therefore fallible? The main pattern of quasi-empirical criticism is proliferation of theories and refutations. (Lakatos 1980, pp. 29–30)

This proliferation invites and even gives a central role to a minority mathematics. As, correlatively, both quasi-empirical and grounded in the role of a minority mathematics, the present view of the history of mathematics is that of multiple trajectories, heterogeneous yet interactive, that is, trajectories with multiple connections between them, without assuming or even precluding unifying them in a single "grand" narrative, as Manin appears to see as at least possible to do (Manin 2010, p. 242).[8]

Every new mathematics, which, as new, is always a creation of a minority mathematics, takes a chance on the future of mathematics, just as every new minority science, literature, such as that of Kafka, or philosophy takes a chance on becoming a mathematics or science, art, or philosophy of the future. "A philosophy of the future" is Friedrich Nietzsche's phrase describing his own philosophy, manifestly a minority philosophy, and all true philosophy in his subtitle to *Beyond good and evil: A prelude to a philosophy of the future* (Nietzsche 1966). To become a philosophy of the future, a philosophy, Nietzsche in effect argues, must start as a minority philosophy, which is quite true about that of most major figures in the history of philosophy – Plato, Aristotle, Descartes, Leibniz, Kant, and Hegel. One can take the same view of creative mathematics, such as that of Abel, Galois, Noether, and Grothendieck. Robert Langlands (of the Langlands program fame) compared Grothendieck to "Nietzsche's *Philosoph der Zukunft*" (Langlands and Shelstad 2007, p. 486). This is no less true about Abel, Galois, or Noether, and other key figures mentioned here, or Langlands himself, as the Langlands program has proven.

[8] The concept of a grand narrative is borrowed by Manin from Jean-François Lyotard, who saw it as challenged by the present-day, "postmodern," world, in part in view of twentieth-century mathematics and science, which has shaped this world (Lyotard 1984).

2 Abel and Galois: Minority Mathematics and the Rise of Modern Algebra

Over half a century ago, Jules Vuillemin offered an influential analysis of Abel and Galois, and of the rise of (modern) algebra in *La philosophie de l'algèbre* (Vuillemin 1962), on which, along with other sources, such as Charles Lautman, Deleuze built his commentary on Abel and Galois in his 1968 *Difference and repetition* (Deleuze 1995). Vuillemin's approach had a strong structuralist orientation shaped by the Bourbaki program (the book is dedicated to, among others, Pierre Samuel, a member of Bourbaki), an orientation different from the one of this chapter, on four counts. First, a minority mathematics was never considered, even by implication, by Vuillemin, while it is, as I argue, crucial for understanding both Abel's and Galois' roles in algebra and modern mathematics in general. Deleuze does not consider a minority mathematics in this earlier book either. The concept of minor literature was developed two decades later in the book on Kafka, cowritten with Guattari, building on the ideas of their earlier joint works. Second, this chapter adopts a conceptual approach versus a structuralist one. Third, Vuillemin's book and the corresponding part of Deleuze's book (a broader philosophical project overall) aim to transpose the methods of modern algebra, in particular those of Abel and Galois, into philosophy, which is not my aim here. Fourth, "ideas," in *Difference and repetitions* and then concepts in Deleuze and Guattari's *What is Philosophy?* (Deleuze and Guattari 1996), are defined strictly as philosophical ideas and concepts, even when they are operative in mathematics, specifically as problems that persist rather than disappear in "solutions" to them. By contrast, as explained above, mathematics is, for Deleuze and Guattari, always about solutions of mathematical problems (Deleuze 1995, pp. 163–164, 179–182). One might readily agree that mathematics is commonly in pursuit of solutions to its problems and that there may be a philosophical dimension to a new problem-posing field defined by a new mathematical concept-form, such as that of a group by Galois. In the present view, however, a concept-form is a *mathematical* concept rather than a *philosophical* concept. Concept-forms still enable one to solve, or to advance the possibility of solving, concrete problems, such as, in Galois' case, the problem of finding the algebraic solution to a polynomial equation, which gave rise to the concept of group. This is made possible by embedding this original problem in a new mathematical concept-form, which defines a new problem-posing mathematical field, or even a new field of mathematics, such as Galois theory. An exact definition possible in mathematics is difficult to achieve in a philosophical concept, the rigor of which needs to be, accordingly, defined otherwise, which is a separate issue that will be put aside here.

These differences notwithstanding, it may be helpful to summarize Vuillemin and Deleuze's arguments concerning Abel and Galois' work. According to Deleuze:

> Abel elaborated a whole method according to which solvability must follow from the form of the problem. Instead of seeking to find out by trial and error whether a given equation is solvable in general, we must determine the conditions of the problem which progressively specify the fields of solvability in such a way that 'the statement contains the seeds of the

> solution'. This is a radical reversal in the problem-solution relation, a more considerable revolution than the Copernican. It has been said [by Vuillemin] that Abel thereby inaugurated a new Critique of Pure Reason, in particular going beyond Kantian 'extrinsicism'. This same judgement is confirmed in relation to the work of Galois: starting from a basic 'field' (R), successive adjunctions to this field (R′, R″, R‴ . . .) allow a progressively more precise distinction of the roots of an equation, by the progressive limitation of possible substitutions. There is thus a succession of 'partial resolvents' or an embedding of 'groups' which make the solution follow from the very conditions of the problem: the fact that an equation cannot be solved algebraically, for example, is no longer discovered as a result of empirical research or by trial and error, but as a result of the characteristics of the groups and partial resolvents which constitute the synthesis of the problem and its conditions (an equation is solvable only by algebraic means—in other words, by radicals, when the partial resolvents are binomial equations and the indices of the groups are prime numbers). The theory of problems is completely transformed and at last grounded, since we are no longer in the classic master-pupil situation where the pupil understands and follows a problem only to the extent that the master already knows the solution and provides the necessary adjunctions. For, as Georges Verriest remarks, the group of an equation does not characterize at a given moment what we know about its roots, but the objectivity of what we do not know about them. Conversely, this non-knowledge is no longer a negative or an insufficiency but a rule or something to be learnt which corresponds to a fundamental dimension of the object. The whole pedagogical relation is transformed—a new *Meno* [Plato's dialogue dealing with teaching mathematics]—but many other things along with it, including knowledge and sufficient reason. Galois's progressive discernibility unites in the same continuous movement the processes of reciprocal determination and complete determination (pairs of roots and the distinction between roots within a pair). It constitutes the total figure of sufficient reason. (Deleuze 1995, pp. 179–180)

Manifestly the product of a minority mathematics, this is a radical transformation of mathematical, and, on this pattern, for Deleuze, of philosophical practice, including as the pedagogy of ideas or, in his later works, concepts. This view also implies that the pedagogy of mathematics should be transformed, certainly as it functions in mathematical practice (which, just as that of philosophy, always involves a teaching of concepts) and possibly in the actual teaching of mathematics. The latter, primarily defined by teaching even more a majority mathematics than a major mathematics, would benefit by giving more attention to minority thinking in mathematics. In the absence of a contemporaneous alternative education (possibly informal), such a teaching is not likely to lead to the emergence of mathematical thought such as that of Abel, Galois, and Noether. When B. L. van der Waerden's *Modern Algebra* (van der Waerden 1930) was published in 1930, it was also an introduction to a new way of thinking in algebra, developed by Noether and Emil Artin, and at the time (and for a while) it helped new thinking in algebra and beyond, including as pursued by Bourbaki. But subsequent textbooks on algebra rarely perform this function, and *Modern Algebra* no longer in general does so either. I shall return to this subject below.

Galois' thinking had an even more revolutionary significance for both Vuillemin and Deleuze, given the role of the concept of group, which was (in present terms) a new concept-form and, as such, an even more consequential invention. All concept-forms require a minority practice to emerge, but Galois' concept of group, or his thinking in general, more dramatically exhibited a minority mathematics. As noted

from the outset, the concept of group has radically deterritorialized and reterritorialized mathematics, well beyond algebra. This deterritorialization is, again, not addressed by Deleuze in *Difference and repetition* but is suggested by his later appeal to Galois' "abstract machine," alongside Riemann's "abstract machine," the term designating a vehicle (also in its literal sense of enabling a movement) of deterritorialization (Deleuze and Guattari 1987, p. 142). The term also designates a nomadic, and hence, minority, versus state or royal (major and often majority), practice of mathematics.

I shall consider first Abel's work, arguably the first manifested case of a minority mathematics in modern mathematics, prompting me to designate all minority mathematics as Abelian mathematics. According to Vuillemin, Abel transforms algebra by the invention of "a general method":

> In the first place, one analyzes, in its most general form, a mathematical relation or a defined set of such relations, which make it possible to determine a property, of which we do not yet know whether or not we can attribute it to a class of beings: by example the character of being algebraically solvable, of being convergent, of being expressible by a defined number of functions of a certain class. In the second place, one considers the class of beings to which it is a question of attributing this property . . .; we analyze these beings from a general point of view, we define the relationships to which their nature allows them to be subjected. Finally, this double examination reveals the cases of incompatibility (demonstrations of impossibility) and, possibly, indicates the way to find the new required relationships in cases of possibility. (Vuillemin 1962, p. 214; translation mine)

Thus, rather than looking for the solution to equations of the fifth degree, one asks whether such a solution is possible, in the first place, for any polynomial equation. This shift is the key (minority) move by Abel. Abel's new "idea of a general method consists in giving a problem a form such that it is always possible to solve it" (Vuillemin 1962, p. 209; translation mine). Abel, thus, reversed the method of his predecessors and his contemporary major mathematicians, who would proceed from the particular to the general to proceeding from the general to the particular. As a result, Abel becomes one of the creators of the modern algebraic method. Vuillemin sees this method as, in effect, structuralist in Bourbaki's sense. Against Kant, who limited his construction of the proof of impossibility by positing the possibility of an experience as an exclusive criterion of knowledge, "general demonstrations, in the sense of Abel, change the modality of the proof. Particular demonstrations are actual [*réelles*]: they assume as their grounding principle the possibility of the experience given in the affect of sensation. The general proofs deal with the possible and start from the concept alone, disregarding the restrictive conditions of sensitivity" (Vuillemin 1962, p. 216; translation mine). This initiating concept is, again, defined not by a generalization from particulars, but as a new structured entity, in accord with the present definition of concepts. In this way, Abel introduced a new *critique* of pure reason (in Kant's sense of a critique a rigorous examination of the legitimacy and limits of philosophical claims, such as those concerning causality, and hence concepts), at least in mathematics, but, as Vuillemin and Deleuze argue, extendable beyond mathematics. Mathematics needed Galois in algebra (a group) and then

Riemann in geometry (a manifold) to be expressly rethought in this way, or perhaps *more* expressly. For, one might see Abel's thinking as reaching further in this direction and in general as a more radically transformative than Vuillemin and Deleuze suggest. His minority mathematics was even more momentous in its transformative and, hence, deterritorializing potential. Mikhail Gromov considers Abel to be "a major figure, if not the major figure, in changing the course of mathematics from what could be visualized and immediately experienced to the next level, a level of deeper and more fundamental structure" and hence toward modern mathematics (Raussen and Skau 2010, p. 401).[9] Abel's thinking also becomes even closer to that of Galois in general, beyond the problem of solving polynomial equations. Gromov continues:

> [Abel] changed the perspective on how we ask questions. I do not know enough about the history of mathematics but it is obvious that the work of Abel and his way of thinking about spaces and functions has changed mathematics. I do not know enough history to say exactly when this happened, but the concept of underlying symmetries of structures comes very much from his work. We still follow that development. It is not exhausted yet. This continued with Galois theory and in the development of Lie group theory, due to Lie, and, in modern times, it was done at a higher level, in particular by Grothendieck. This will continue, and we have to go through all that to see where it brings us before we go on to the next stage. It is the basis of all we do now in mathematics. (Raussen and Skau 2010, p. 401)

As such, Abel's minority mathematics becomes a crucial case of a deterritorialization of modern mathematics. Gromov's statement is a reflection on this aspect of Abel's work as well, even if without considering Abel's works strictly from this perspective. My designation of all minority mathematics as Abelian mathematics fits well with Gromov's assessment and with his comments, on the same occasion, on mathematical education. In responding to the question concerning mathematical education or education in general now, Gromov gives a somewhat unexpected answer, by using Abel as an example:

> ***Raussen and Skau****: Education is apparently a key factor. You have earlier expressed your distress about realizing that the minds of gifted youths are not developed effectively enough. Any ideas about how education should change to get better adapted to very different minds?*
>
> **Gromov**: Again I think you have to study it. There are no absolutes. Look at the number of people like Abel who were born two hundred years ago. Now there are no more Abels. On the other hand, the number of educated people has grown tremendously. It means that they have not been educated properly because where are those people like Abel? It means that they have been destroyed. The education destroys these potential geniuses – we do not have them! This means that education does not serve this particular function. The crucial point is that you have to treat everybody in a different way. That is not happening today. We don't have more great people now than we had one hundred, two hundred, or five hundred years ago, starting from the Renaissance, in spite of a much larger population. This is probably due to education. ...

[9] This assessment was given in the wake of Gromov's reception of the Abel prize, which might have inflected Gromov's remarks, without, in my view, undermining Gromov's essentially correct assessment itself.

> ***Raussen and Skau****: You point out that we don't have anybody of Abel's stature today, or at least very few of them. Is that because we, in our educational system, are not clever enough to take care of those who are exceptionally gifted because they may have strange ideas, remote from mainstream?*
>
> **Gromov**: The question of education is not obvious. There are some experiments on animals that indicate that the way you teach an animal is not the way you think it happens. The learning mechanism of the brain is very different from how we think it works: like in physics, there are hidden mechanisms. We superimpose our view from everyday experience, which may be completely distorted. Because of that, we can distort the potentially exceptional abilities of some children. There are two opposite goals education is supposed to achieve: firstly, to teach people to conform to the society they live in; on the other hand, to give them freedom to develop in the best possible way. These are opposite purposes, and they are always in collision with each other. This creates the result that some people get suppressed in the process of adapting them to society. You cannot avoid this kind of collision of goals, but we have to find a balance between the two, and that is not easy, on all levels of education. (Raussen and Skau, p. 402)

Gromov is right to qualify the question if education is not obvious and to note this conflict at the core of the Enlightenment ideal of education. In effect (whether or not Gromov had in mind or even was aware of this connection to the Enlightenment), Gromov describes this ideal by referring to giving people "freedom to [individually] develop in the best possible way." (Conforming to the society one lives in is common to most conceptions of education.) This concept and, with it, the conflict in question emerge with the Enlightenment, especially following Kant, who introduced the term "Enlightenment [*Aufklärung*]," in part by grounding this ideal in, to use the title of his second critique, a "*critique of practical reason*" (Kant 2015). Two influential examples may be noted here. The first, before (and influencing) Kant, Jean-Jacques Rousseau's *Émile*, *ou de l'éducation* [1762] emphasized the importance of this freedom for an individual to develop in the best possible way and the difficulty of achieving the aim of doing so in a society (Rousseau 1997). (Because of one section of the book, it was banned in Paris and Geneva, and publicly burned when published in 1762.) The second, following (and yet displacing) Kant, Friedrich Schiller's *Über die ästhetische Erziehung des Menschen in einer Reihe von Briefen* [Letters on the Aesthetic Education of Man] (1795) (Schiller 2004), aimed at reconciling both aims within a framework, which, however, poses major questions of its own.[10]

The question of education is indeed far from obvious. Abel and others to whom Gromov is referring under an apt collective name "Abels," certainly Galois, or any number of Renaissance figures who could be mentioned, such as Galileo, were not uneducated, quite the contrary. Nor were Abel and Galois merely self-educated in mathematics either, even if to some degree they, especially Galois, were more so than others. But if one considers the likes of Gauss and Riemann, one encounters a reasonably conventional mathematical, as well as broader, education. (Kafka was remarkably educated and not only when it comes to knowledge of literature, and

[10] For a critique of Schiller's "aesthetic ideology," including his misreading of Kant, see (De Man 1996) and for the connections between the general problematic of "aesthetic ideology" and the idea of algebra, in connection with Paul de Man's analysis (Plotnitsky 2001).

Gromov, perhaps not coincidentally, brings up Russian literature in his interview, and some of its greatest figures, comparable to Kafka, including as, in some cases, representing a minority literature.) Gromov must have been aware of these qualifications. His point was, I would surmise, different, and it may be expressed in terms of a minority thinking in mathematics or elsewhere. As indicated above, even more than individuality, our education resists and discourages orientations toward a *minority* thinking, not the least because while transforming major thinking it challenges even more majority thinking. As I argue, however, without a minority thinking in mathematics and a chance, and a freedom, to put it into practice the likes of Abel and Galois are not possible or in any event not likely to emerge. The problem is equally acute elsewhere, including in literature and art, which has managed to still produce their "Abels," such as Kafka, throughout much of the twentieth century, although there are now no more Kafkas or Joyces around either. Still, minority literature or art *might* have a greater chance, a chance diminished by the contemporary nature of education in mathematics and science, although there are (hence, I qualify by "might") societal and specifically educational forces operative against pursuing minority ways in literature and art.[11] Mathematics or other fields could also be transformed *more* within major practices – more, but never entirely so. Instead, such cases show (and hide) more incremental multistep punctuated processes, with a lesser impact of minority thinking entering from a manifested exterior, such as that exhibited in Abel and Galois' thinking.

Galois is an equally exemplary case in both his uniqueness as a minority mathematician and in the role of "the concept of underlying symmetries of structures" in mathematics invoked by Gromov, who mentions Galois theory as an example following Abel's "revolution." It would, however, be difficult to argue that Galois theory merely follows Abel in this regard, and Gromov would probably admit its greater overall significance. It was the first expressly formulated actual theory of the type Gromov invokes. For one thing, it is as a product of a group theory,

[11] More generally, as I argue here, if there is no opening toward minority thinking, there is no creative thinking in any domain, but only a proliferation of opinion and forms of programing according to preestablished rules, which are essentially equivalent to or are generative of opinions. Although I mean the term "programming" in a broader sense here, I do imply that the recent development of and claims for the so-called artificial *intelligence* (a misleading term), such as, now ubiquitous, ChatGPT and its avatars, have nothing to do with thought, but are merely forms of programing. In this respect, the term artificial intelligence (AI) is misleading, although it is difficult to replace it at this point, for example, by a much more accurate "digital robotics." As such, such programs merely propagate knowledge defined by or defining opinions, with no space for a minority thinking and hence creativity. Even if such programs use things derived from minority thinking gathered from Internet archives, they will convert them to preestablished information or rigid articulations of ideas. While recognizing many extraordinary accomplishments of digital technology, one should not confuse a clever programing with thinking. A remarkable and in many respects distressing aspect of ChatGPT and related software is that they do what they do *without* thinking. Their use in mathematics, much discussed recently as well, is a separate issue, which I put aside. I merely note that, in my view, it does not seriously challenge my main point concerning the difference between programing and creative thinking, which is always a minority thinking. ChatGPT can only help to propagate opinions.

which, as I have stressed throughout, enacts a broad deterritorialization of modern mathematics, as well as physics and other sciences, natural and social. It is hard to think of a broader one. Some elements of group-theoretical thinking are found in Abel, or Legendre, earlier, and Abel's thinking in general has deeper affinities with Galois in rethinking the fundamentals of mathematics as such. But the concept of group carries Galois' signature. It was also accompanied by another important concept, that of an action of a group on another structured entity, such as an algebraic field or vector space. Besides, Galois theory is much more than only part of group theory.

For Vuillemin, deeply steeped in Bourbaki's ideology, the emphasis is on the idea of structure, and he sees the concept of group as the first formal algebraic structure in the history of mathematics. Galois' thinking, then, is both akin to that of Abel as concerns Abel's philosophy of solving problems and reaches beyond Abel, even if it is, as Gromov argues, not as far beyond as is sometimes claimed. It does so, again, by bringing in the role of a structure and the concept of group as an entity (we would now say a set) with a structure, a relationally organized multiplicity. Galois, for Vuillemin, becomes the first *modern* algebraist, who associates with solving a problem the analysis of the structure from which this problem arises and which makes it possible or not to solve it, such as the problem of finding algebraic solutions by radicals of polynomial equations. Vuillemin's structuralist perspective is not out of place. One cannot avoid structures or merely dismiss structuralism, an important movement, conceptually and historically (Corry 2006). For one thing, concepts are also structures. On the other hand, one might, as I do, see structures as emerging from and as concepts, which thus precede structures in creative thinking in mathematics.

Indeed, as in the case of Abel's thinking, one of Vuillemin's main points in considering Galois is more about concepts than structures, keeping in mind that a concept is a structure, too, and conversely any given structure is a concept. Sometimes, the structure of a mathematical concept, such as that of a group, may coincide with the mathematical structure defined by this concept. Still, a structure qua structure must be defined and thus comes from a concept, possibly a concept-form. Vuillemin distinguishes two types of abstraction: The first is conventional, defined by an increasing generalization from the particular to the general, and the second, found in group theory, is conceptual in the present sense. It defines the differences between specific individual cases because it constructs these cases from the initial concept. Vuillemin sees the second abstraction, which he terms "formalization," as a double abstraction: The first is that of the formal presentation of the group, as a concept, by algebraic symbols and their relations; the second is that defined by providing a method that makes it possible "to construct individuals, no longer in intuition according to imperfect [arbitrary] schemes, but [determinately] in the concepts themselves" (Vuillemin 1962, pp. 288–289; translation mine).

Thus, one deals with an algebraically formalized conceptual practice and redefines algebra accordingly, as a deterritorialization of algebra by means of a minority algebra. In Galois' case and, more extensively, in Abel's case, one also deals with a parallel shift to the conceptual algebraization of analysis, as well as geometry, where

this shift was initiated with the analytic geometry of Fermat and Descartes. In noting that, in the wake of Galois, "modern mathematics is therefore regarded as based upon the theory of groups or set theory rather than upon differential calculus," Deleuze adds: "Nevertheless, it is no accident that Abel's method consisted in giving a problem a form such that it is always possible to solve it, concerned above all the integration of differential formulae" (Deleuze 1995, pp. 180–181).

Galois' concept of group as a multiplicity with a structure was an entirely new type of concept, created by a minority mathematics from within and transforming, reconstructing, the major mathematics of his time and of the future. As a concept-form, it defined a new problem-posing field through the new mathematical formulations of existing problems (such as that of the solution of polynomial equations by radicals) and posing new problems, and the simultaneous establishing of their conditions of solubility. This new thinking changes the future of mathematics as a field of solving problems, because these problems themselves and the conditions of their solubility change, by asking under what conditions a given problem can be solved. While Abel and already Ruffini did something similar, there was no specific concept of the type a group is in its power and generality. Galois, again, also creates the concept of an action of a group on another structural entity, such as a field in algebra or a manifold in geometry.

One has now a new abstract concept, an invention of thought, that of a group. This new concept governs a previously established object, an equation, or eventually governs a field (another new concept-form), or a class of such objects, *similarly* to the way a Platonic form precomprehends a given concrete reality. As indicated earlier, however, there are also major differences (hence my emphasis) between these two conceptions. The objects considered, such as equations or fields, are not "shadows" of a group as this new abstract form, which does not preexist them, as Plato's form or idea would. Each such object or instantiation of a concept, such as an equation or a field, has its own definition. Just as that of a group, a more general concept of an equation (which has a long and complex history) or a field, like algebra, only preceded a particular equation, or area of algebra, locally, and not as a Platonic form, preexisting them in some "eternal" domain. (Where, one wonders, would a mathematical Platonist place such a domain?) In the present view, any new concept is a local invention of human thought and, to one degree or another, is always a product of a minority thinking. As such, it relates and handles the reality already in place and the problem this reality posed and, by doing so also, establishes a new problematic field, such as Galois theory in the case of a group. In fact, in the process it creates a new mathematical reality (Plotnitsky 2022, pp. 83–95). It always does so, however, locally and contingently, in the interaction between a minority and a major mathematics, rather being governed by one or another preexisting Platonist reality or a single transforming Ur-reality governing all possible local transformations. Such Ur-reality would imply a possibility, at least in principle, of a single major mathematics to which a minority mathematics could contribute but that it could not change by an independent act of creation. A minority mathematics could, in this case, only participate in transformations of a single major mathematics, Ur-mathematics, defined by this Ur-reality. In the present view, the possibility of

either this type of Ur-reality or this type of Ur-mathematics is precluded in principle. So is, it follows, a grand narrative of mathematics, replaced by multiple narratives of its continuous and discontinuous developments, with multiple mathematical practices (minority, major, and majority ones, among them) in multiple interactions.

3 The Three Mathematics of Emmy Noether

Emmy Noether, a *major* figure in twentieth-century mathematics, is also an important case of a minority mathematics. Her case has additional complexities, first, because the relationships between a major and a minority mathematics in her work in different areas of mathematics were different, and second, because of the "geography" of mathematics at the time, with Göttingen, where she worked, as one of its centers or even the center of major mathematics in most key areas of modern mathematics then. Although her achievements and the importance of her work have always been acknowledged by mathematicians and historians of mathematics, her role as a shaping figure in twentieth-century mathematics has remained an underappreciated subject for a long time, although the situation has changed in recent years. Described by some of the greatest twentieth-century mathematicians and physicists, Hilbert, Einstein, and Weyl, among them, as the most important woman mathematician in the history of mathematics, she was once referred to, in *The New York Times*, as "the most significant mathematician you've never heard of" (Angier 2012). A decade old now and no longer as applicable, this assessment still reflects both the status of mathematics in our culture and the status of women in mathematics. While the second subject has been extensively considered in literature, including in connection with Noether's work, the first has not received much attention, although it has been given more attention in recent years as well. This chapter does not intend to offer an analysis of either subject as such. It might, however, suggest a new perspective on both by considering Noether's work as a minority mathematics.

In geometry, the primary field of her earlier work, she followed the tradition, defined by the role of symmetry and group theory, established by Riemann, Felix Klein, Sophus Lie, and others, and by that time also shaping physics, specifically general relativity, in which Noether worked as well. In algebra, the primary field in her later work, she followed the tradition initiated by Galois and developed especially by Dedekind. She also made important and, in some respects, revolutionary contributions, to algebraic topology. All three fields were by then parts of major mathematics, even though continuous innovations within each involved (in the present view, unavoidably) minority mathematical thinking and practices. Noether's thinking and practice were minority ones both in their purely mathematical aspects and given her social and specifically institutional situation, because she was a female mathematician in a (vastly) male dominated field, as mathematics was then, including in all areas of her work. Here, I shall be primarily concerned with her mathematics, without offering an analysis of the second aspect of her minority practice, addressed in several recent studies, including her biographies. I shall argue that,

while all three facets of her work still represented, in view of their innovative nature, minority mathematics, each case was different. Her work on her celebrated theorems was more firmly within what was major mathematics in geometry at the time, her work in algebra had much more significant minority dimensions, and her contribution to algebraic topology was the most pronounced manifestation of her minority thinking in mathematics. All three of *Noether's* mathematics, however, also confirm that a minority mathematics depends on and is not possible without working with a major mathematics. One might read in this way Noether's refrain "*Es steht alles schon bei Dedekind* [this is all already in Dedekind]," even if she herself meant it more directly, a claim that is difficult literally to apply to Noether's ideas in algebra, many of which were too original to be thus attributed to Dedekind (McLarty 2006, p. 212). Instead, these connections to Dedekind, whose work was by then a well-established part of major algebra, reflect the dependence of Noether's minority mathematics on its interaction with her contemporary major mathematics in making possible her contributions, profoundly innovative as they were. This, I argue here, is always the case. In sum, this section presents a new angle, that of a minority or Abelian mathematics, on Noether's work.

During the last half a century or so, Noether, has become especially known, including in her recent public recognition, for her theorems, published in 1918. (Most of the studies devoted to her work focus on these theorems.) By contrast, she has always been world-famous for her contributions to modern algebra and (to a somewhat less degree) algebraic topology among mathematicians and historians of mathematics. The significance of her contribution in topology was due to Noether's realization that algebraic operations, related to homotopies and homologies, in topology should be considered in terms of abstract algebraic concepts, such as groups. This view, quickly adopted by her contemporary topologists and, at least in the case of homotopies, by earlier figures, such as Henri Poincaré, had proven exceptionally fruitful for the development of algebraic topology and has come to define both cohomology and homotopy theories there. (Both, especially cohomology theory, were also developed in algebra and algebraic geometry.) As discussed below, however, Noether's thinking in this field and in algebra itself was more innovative and more radical, giving her use of what Wolfgang Krull, a major algebraist in his own right, referred to as "Noether's principle" (Krull 1935, p. 5; McLarty 2006, p. 218). The principle is grounded in the view that abstract mathematical concepts, such as groups, rings, or modules, should be primarily defined, in terms of their structure, a view that came to define modern algebra and much of modern mathematics, in part following Bourbaki's structuralism. This understanding as such was not especially new, although Noether was a key figure in establishing it. Noether's principle, however, was new. It says that the study of these objects should focus on the relationships, *morphisms*, between algebraically structured objects (sets), such as groups, rings, or modules, rather than merely on these objects themselves.

Innovative as it was in algebra, Noether's work was even more so in topology at the time, although it has become routine since, including as translated into categorical terms. Noether's principle and, in the first place, Noether's ideas concerning the

role of groups in considering homotopies and homologies of topological spaces were developed in algebraic topology by her disciples and followers, such as Heinz Hopf and Pavel S. Alexandrov, as well as independently by others, such as Leopold Vietoris. Noether's principle is a *precursor* of later categorial and functorial principles, as considered earlier in this study, especially in Grothendieck's work. It would, in my view, be difficult to claim that Noether herself did go that far. There is, in her work, no suggestion of the relations, functors, between categories or protocategories that she considered. In retrospect, algebraic topology is inherently categorical. What made it an exact mathematical discipline was the fact that one can associate algebraic structures (initially numbers and algebraic operations, eventually groups and other algebraic structures) to the architecture of spatial objects that are invariant under continuous transformations, independently of their geometrical properties, associated with measurements. It thus relates, functorially, objects of topological and algebraic categories. This nature of algebraic topology had, however, *to be made* categorical in its rigorous sense, which took a while and happened about a decade after Noether's death (in 1935). Her thinking in algebraic topology was, nevertheless, a contribution to this history. Alexandrov, one the topologists influenced by Noether, commented on this new, *minority*, thinking (addressing a Russian audience), in commemorating Noether, first by comparing her work on algebra and differential geometry with that of Sophia Kovalevskaya:

> It is often forgotten that in this period Emmy Noether obtained excellent results concerning the concrete algebraic problems of Hilbert. These results and her work on differential invariants would have been enough by themselves to earn her the reputation of a first-class mathematician and are hardly less of a contribution to mathematics than the famous research of S. V. Kovalevskaya. (Alexandrov 1936, cited in Dick 1981, p. 156)

He added, giving Noether a greater credit for her contribution to algebra and then algebraic topology:

> But when we think of Emmy Noether as a mathematician, we have in mind not these early works, important though they were in their concrete results, but rather the main period of her research, beginning in about 1920, when she became the creator of a new direction in algebra and then leading, the most consistent and prominent representative of a certain general mathematical doctrine–all that which is characterized by the words 'begriffliche *Mathematik*' [abstract mathematics]. (Alexandrov 1936, cited in Dick 1981, p. 186)

It is worth noting that "*begriffliche*" means conceptual, which does not exclude abstract, but is not quite the same (regardless of what Alexandrov thought).

I shall, however, consider first Noether's celebrated theorems by way of contrast as an example of a primarily major mathematics, while, again, keeping in mind that, as any innovation, they contained ingredients of a minor mathematics, even in purely mathematical terms, aside from the gender aspects of Noether's minor mathematics, which would require a separate analysis. Noether's theorems, especially Noether's first theorem, were mathematical generalizations of previously known mathematical expressions of physical conservation laws (those of momentum, energy, and angular

momentum) in classical mechanics and relativity. General relativity introduced mathematical complexities into the formulation of the energy-conservation law, which complexities were handled by Noether in her article introducing her theorems. Her theorems were, however, formulated abstractly (a minority move), and they can be considered apart from any reference to physics, in the style of symplectic geometry. (The latter originates in the Hamiltonian formulation of classical mechanics where the phase spaces of certain systems assumed the structure of a symplectic manifold, defined by a closed, nondegenerate 2-form.) When applied in physics, Noether's theorems state that every differentiable symmetry of the action of a physical system has a corresponding conservation law. The action of a physical system is an integral over time of a Lagrangian, which allows one to determine the system behavior by using the principle of least action (the quantity defined as the integral of the momentum over a given distance traveled by a body). The theorems are an application of the principle, arguably, stated first by Pierre Louis Maupertuis in 1774 and mathematically developed by Leonhard Euler around the same time, and then by Joseph Louis Lagrange and Sir William Rowan Hamilton. Leibniz had related ideas earlier. Philosophically, the idea that nature follows the most efficient path possible can be traced to much earlier, even to ancient Greek thinking. According to Yvette Kosmann-Schwarzbach's account, closely following Noether's article:

> Noether's first theorem, a generalization of several conservation theorems that were already known in mechanics ... She considers a multiple integral,
>
> $$I = \int \dots \int f\left(x, u, \frac{\partial y}{\partial x}, \frac{\partial^2 y}{\partial x^2}, \dots\right) dx,$$
>
> of a higher-order Lagrangian f that is a function of n independent variables, $x_1, \dots, x_\lambda, \dots, x_n$, and of μ dependent variables, $u_1, \dots, u_i, \dots, u_\mu$, as well as of their derivatives up to a fixed but arbitrary order, κ. She then considers a variation of u, $\delta u = (\delta u_i)$, and derives identity (3),
>
> $$\sum_{i=1}^{\mu} \psi_i \delta u_i = \delta f + \mathrm{Div}\, A,$$
>
> where the ψ_i are the Lagrangian expressions, which is to say the components of the variational derivative (Euler–Lagrange derivative) of f, and where the components A_λ of A are linear in the variation δu and in its derivatives. The opposite of the quantity A is now called the Legendre transform of the Lagrangian f. Here Div is the ordinary divergence, $Div\, A = \sum_{\lambda=1}^{n} \frac{\partial A_\lambda}{\partial x_\lambda}$, of $A = (A_1, \dots, A_n)$ considered as a vector in n-dimensional space, and δf is the variation of f corresponding to the variation δu of u, while the variation of x is assumed to vanish. Identity (3) is obtained by an integration by parts. In the case where $n = 1$, the case of a simple integral, Noether gives an expression for A for an arbitrary μ, first for $\kappa = 1$, which yields what she calls Heun's "central Lagrangian equation," then for an arbitrary κ, and then she states her theorem:

I. *If the integral I is invariant under a [group] $\mathfrak{G}_\rho$, then there are ρ linearly independent combinations among the Lagrangian expressions which become divergences—and conversely, this implies the invariance of I under a [group] $\mathfrak{G}_\rho$. The theorem remains valid in the limiting case of an infinite number of parameters.*

Noether explains that "in the one-dimensional case," that is, when $n = 1$, one obtains first integrals, while, "in higher dimensions," i.e., when $n > 1$, "one obtains the divergence equations which, recently, have often been referred to as conservation laws." By the "limiting case" included in the statement of Theorem I is meant the case in which the elements of the group depend on an infinite but denumerable set of parameters, as opposed to the case dealt with in her Theorem II (Kosmann-Schwarzbach 2011, p. 57).

Noether's second theorem, in which this limitation no longer applies, is:

II. *If the integral I is invariant under a [group] $\mathfrak{G}_{\infty\rho}$ depending upon arbitrary functions and their derivatives up to order σ, then there are ρ identities among the Lagrangian expressions and their derivatives up to order σ. Here as well the converse is valid.* (Kosmann-Schwarzbach 2011, p. 4)

The physical meaning of the theorems easily follows given that the mathematics used (ideally) *represents* the key physical features of the systems considered, in an essentially geometrical way, notwithstanding the algebra of the theorems, that of symmetry groups. For systems with suitable Lagrangians (which allow for this representation), as most classical and relativistic systems are, symmetry under continuous translations in time implies the conservation of energy, symmetry under continuous translations in space implies the conservation of linear momentum, and symmetry under continuous rotations implies the conservation of angular momentum. These formulations apply straightforwardly in classical physics, with the mathematics of the proof automatic by Noether's first theorem. As noted, general relativity involves certain mathematical complexities, in part leading Noether to her second theorem. Noether did important work on general relativity earlier, and her work on her theorems was in part prompted by Hilbert's questions concerning energy conservation there, on which he consulted Noether and that Noether resolved by a subtle argument in her article (Noether 1918; Kosmann-Schwarzbach 2011, pp. 63–64). Noether closes her article, by citing Klein's famous pronouncement that relativity is an invariance under group transformation, a view that of course applies in mathematics itself (Noether 1918; Kosmann-Schwarzbach 2011, p. 22, note 27, pp. 63–64). These complexities do not affect the geometrically representational (realist) character of her theorems. Just as is classical physics, relativity, special and general, remains a realist theory, which means that Noether's theorems or symmetry principles apply in essentially the same way as in classical physics. In principle, her theorems are, again, independent of physics, with which, however, they have been primarily associated, including by Noether herself. The situation becomes different when Noether's theorems appear in quantum mechanics or quantum field theory, where one can still use them. More accurately one can use certain *versions* of these theorems, because they have to be reformulated even for the same dynamical variables – momentum, angular momentum, and energy – as those used in classical

physics in view of the nature of the formalism of quantum theory versus that of classical physics or relativity. In quantum theory, these variables are no longer functions of real quantities but are operators in Hilbert spaces over $\mathbb{C}$, in the case of these variables, infinite-dimensional Hilbert spaces. The theorems can also be generalized to variables, especially discrete variables, such as spin, and symmetries that are strictly quantum and have no classical counterparts. The subject will be put aside, because these versions of Noether's theorems have only remote connections to Noether's work; they are no longer *Noether's* Noether's theorems. I have considered this subject in detail in (Plotnitsky 2022, pp. 205–210). My point here is that her theorems as such, *Noether's* Noether's theorems, were by and large the products of what was then a major mathematics, important and innovative as these theorems were and, as such, unavoidably dependent on a minority mathematical thinking in their discovery. Thus, as noted, formulating them *abstractly* apart from any reference to physics was a minority move, characteristically Noetherian in nature, arguably, a link to her work in abstract algebra and then algebraic topology, where, however, the minority nature of her thinking becomes much more pronounced.

Noether's work in abstract algebra is one of the greatest representations of the idea, and the ideal, of modern algebra as abstract algebra. Algebra was part of and shaped modern mathematics in the present definition, as abstract mathematics in general (vs. representing natural objects, specifically in physics), from around 1800 on. On the other hand, the term "modern algebra" has immediate connections to and even originates in Noether's work, because it appears to have been used for the first time in the famous book, under this title, "Modern Algebra [*Moderne Algebra*]," by her close associate and collaborator, van der Waerden (1930). The book was based on lectures given by Noether and Emil Artin. It was one of the first books to use an abstract (axiomatic) approach to algebraic structures, such as groups, rings, and fields, and as such became, transitioning from a minority to a major approach, one of the most influential texts on algebra in the twentieth century. It reflected Noether's mathematical practice, again, initially as a minority practice, quickly transformed by Noether and her followers, into a major practice, with Galois and especially Dedekind as the main precursors.

The grounding principle of modern algebra is what Vuillemin, as discussed above, defined, in considering Galois' concept of group (but clearly following post-Noetherian algebra, as shaped by Bourbaki structuralism), as a double abstraction, or "formalization." The first abstraction is that of the formal presentation of the group, as a concept, in terms of algebraic symbols and their relations; the second is that defined by a method that makes it possible "to construct individuals, no longer in intuition according to imperfect [more arbitrary] schemes, but [determinately] in the concepts themselves" (Vuillemin 1962, pp. 288–289; translation mine). In other words, this is an abstraction as a creation of new general (abstract), but in their construction concrete, *intrinsically* concrete, mathematical concepts. Such concepts can also borrow elements from elsewhere and mathematize them, thus abstracting them from their source areas. Symmetry is an example of such a concept: It was borrowed from general phenomenal spatial considerations, but made into invariance under transformations, which may form a group, an algebraic concept that can be

used, as it was by Galois, outside geometrical considerations, and, as a mathematical concept, defined by that of invariance, it came to geometry from algebra.

In Noether's work, then, and, in part under the influence of her work, in general, modern algebra is practiced as abstract algebra – algebra defined by means of its abstract concepts and structures. Commonly, as in van der Waerden's book, and in most commentaries on Noether, the emphasis is on structures and axioms. I would argue, however, that concepts are equally important to Noether, thus making Noether an heir of Riemann as well as Dedekind, who was of course himself an heir of Riemann. For one thing, as noted, a structure is a concept or a set of concepts first, even though concepts are in turn defined by their structure as an organization, *composition*, of their elements. As such a concept and a structure often coincide in the case of purely mathematical concepts, but the concept of concept and the concept of structure are still different. According to van der Waerden's (reported) recollection: "The maxim by which Emmy Noether was guided throughout her work might be formulated as follows: 'Any relationships between numbers, functions, and operations become transparent, generally applicable, and fully productive only after they have been isolated [abstracted] from their particular objects and been formulated as universally valid concepts'" (Dick 1981, p. 101). Structures are, however, essential for defining morphisms between (in present-day terms) categorical entities, and morphisms are central for Noether's work in algebra and algebraic topology.

Thus, integers (which do not form a field) become studied as a commutative ring – a concept with a formally defined abstract structure. One of the outcomes of this approach is Noether's proof of the unique factorization theorem, the Lasker-Noether theorem. The theorem extends the fundamental theorem of arithmetic, which states that every positive integer can be factored uniquely into primes, which is not so for the ideals of commutative rings in general, but is true for Noetherian rings, a very beautiful mathematical discovery. A Noetherian ring satisfies the ascending chain condition on left and right ideals. The ring is left Noetherian or right-Noetherian if the chain condition is satisfied only for the left or the right ideals, which means that every increasing sequence $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ of left or right ideals has the largest elements, that is, there exists n such that $I_n = I_{n+1} = \ldots$. Equivalently, a ring is left or right Noetherian if every left ideal (or right ideal) is finitely generated, or Noetherian if both conditions are satisfied. The Lasker-Noether theorem states that every Noetherian ring is a Lasker ring, a ring in which every ideal can be decomposed as an intersection, called primary decompositions, of a finite number of primary ideals (which are not the same as prime ideals, which share many key properties of prime numbers).[12] The theorem was proven by Emanuel Lasker (famous as a world chess champion) in 1905 for polynomial rings and convergent power series rings, and in full generality by

[12] An ideal Q of a commutative ring A is primary if whenever xy is an element of Q then x or y^n is also an element of Q, for some $n > 0$. In the ring of the integers $\mathbb{Z}$, (p^n) is a primary ideal if p is a prime number. An ideal P of a commutative ring R is prime if: (a) If a and b are two elements of R such that their product ab is an element of P, then a is in P or b is in P; and (b) P is not the whole ring R.

Noether in 1921. Lasker's proof, fully in accord with an established major algebra then, was no small achievement, but it was technical (calculational) and cumbersome, and of course limited. Noether's proof is conceptual and elegant. It was also an example of how conceptual abstraction worked in commutative algebra as, at the time, a minority mathematics.

Noether's later remarkable and highly influential article *Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern* [*Abstract Structure of the Theory of Ideals in Algebraic Number and Function Fields*] (Noether 1927) characterized the rings in which the ideals have unique factorization into prime ideals as the Dedekind domains: integral domains that are Noetherian, 0- or 1-dimensional, and integrally closed in their quotient fields. This article also contained what is now known as the isomorphism theorems, which concern certain fundamental natural isomorphisms, and other key results on Noetherian and Artinian modules. Noether was "the first to develop a general representation theory of groups and algebras, valid for arbitrary ground fields," presented in the important article "*Hyperkomplexe Grössen und Darstellungstheorie*" (Noether 1929). According to van der Waerden:

> [The article] has had a profound influence on the development of modern algebra. . . . In the introduction Emmy Noether states that in recent publications the structure theory of algebras and the representation theory of finite groups have been separated completely. She, on the other hand, aims at a purely arithmetical foundation, in which the structure theory and the representation theory of groups and algebras appear as a unified whole, namely as a theory of modules and ideals in rings satisfying finiteness conditions. (Van der Waerden 1985, p. 244)

In her later work, Noether also made important contributions to the study of noncommutative algebras, including division algebras.

The discussion just offered only considered a very limited set of examples of Noether's work in abstract algebra, sufficient to argue that the high degree of abstraction in her approach could be seen as, at the time, a minority mathematics. I should add, however, that even these examples show (while the methods of modern algebra sometimes hide) the history that shaped these concepts, as a historical part of their intrinsic mathematical concreteness, and as such, also the history of the interaction between established major and minority mathematical practices. Only a few prominent names were mentioned here as part of this history, such as Galois, Dedekind, Hilbert, and Artin. Dozens of other figures, however, even if one thinks only of major ones, are part of this history, even if some of them, such as Kronecker, fought against abstractness, in part, I would argue, by failing to appreciate the intrinsic concreteness of this abstractness. I now move to Noether's use of the *concepts* of abstract algebra, especially that of a group, in algebraic topology, a more dramatically minority intervention, if localized to the relationships between a "minor" role of abstract algebra and the major mathematics in topology.

The history of Noether's engagement with topology is well known, although there is more to Noether's thinking than most accounts convey, and none of them considered this engagement from the viewpoint of a minority mathematics. The situation is complicated by the fact that Noether did not publish on the subject. She

only mentioned it in passing, as "an application of a group theory," in a single article, devoted to another subject: "*in den Anwendungen des Gruppensatzes – z.B. Bettische und Torsionszahlen in der Topologie* [in an application of a group theory—for example, Betti and torsion numbers in topology]" (Noether 1925). This mention is, however, revealing in conveying her contribution, that of bringing group theory into topology and thus helping to make it more fundamentally algebraic, although others, such as Vietoris, did this independently around the same time. Yet again, however, Noether's contribution was more innovative and far-reaching than only that of rethinking homology theory by means of group theory. Equally significant was the way in which she did so, nearly, even if not quite, functorially. Her ideas had a great immediate influence, quickly reshaping the major practice of topology, through the work of such figures as Alexandrov and Hopf. According to a widely circulated statement by Alexandrov:

> When [Noether] first became acquainted with a systematic construction of combinatorial topology, she immediately observed that it would be worthwhile to study directly the groups of algebraic complexes and cycles of a given polyhedron and the subgroup of the cycle group consisting of cycles homologous to zero; instead of the usual definition of Betti numbers, she suggested immediately defining the Betti *group* as the complementary (quotient) groups of the group of all cycles by the subgroup of cycles homologous to zero. This observation now seems self-evident. But in those years (1925–1928) this was a completely new point of view. (Alexandrov 1936, cited in Dick 1981, p. 174)

It was also a minority *algebraic* thinking in topology, or rather a new minority algebraic thinking, because there is no algebraic topology without algebra. How can topology be algebraic otherwise? On the other hand, what algebra brings into topology may be and in this case was a matter of a minority thinking, and the very emergence of algebraic topology was of course a minority thinking. Noether reportedly also noted that the idea of a Betti group makes the Euler-Poincaré formula easier to understand (Dick 1981, pp. 174–175). There is a clear consensus, which included Alexandrov and Hopf themselves, that their work greatly benefited from Noether's insights, unpublished but extensively discussed in her circle at Göttingen at the time. On the other hand, there is less consensus and some debate concerning whether homology groups were used before, specifically by Poincaré, who clearly used homotopies as groups and spoke of them as such. A helpful historical and conceptual discussion and references are found in several articles by Colin McLarty (2006, 2011, 2017). In assessing Noether's contribution, McLarty goes beyond previous arguments concerning Noether's work. He credits Noether with a more radical and significant innovation than only using groups. He sees her main achievement by connecting this work to Noether's developing of, in her words, "set-theoretical foundations of algebra" (McLarty 2006, p. 212).

This would, in present terms, be seen as a more radical example of a minority mathematics by virtue of introducing homomorphisms (rather than merely groups) to bear on topological maps. The "set-theoretical topology" of his title "Emmy Noether's set-theoretical topology" of (McLarty 2006) is a pun: Part of his argument concerns how *algebraic* topology (rather than set-theoretical topology, also known

as general topology) was affected by Noether's concept of set. Luitzen E. J. Brouwer, who worked on these topological maps, becomes part of this history.

McLarty's opening is of some interest in the present context. He uses, as an epigraph, a famous statement by Weyl: "In these days the angel of topology and the devil of abstract algebra fight for the soul of each individual mathematical domain" (Weyl 1939, p. 500). The devil, as a fallen angel in the first place, is a quintessentially minority figure. McLarty then comments as follows:

> If Hermann Weyl ever put faces on these spirits they were his good friends, the angel Luitzen Brouwer, and the devil Emmy Noether. Weyl well describes the scope of their ambition. But topology and algebra were not fighting each other. They would come to share the soul of most of mathematics. And the angel and devil had worked together. In 1926 and 1927 Emmy Noether induced young topologists gathered around Brouwer to use her algebra to organize the kind of work on topological maps that Brouwer taught. This created the still current basis for 'algebraic topology.' It was a huge advance for the structuralist conception of mathematics. And it turned Noether from a great algebraist to a decisive figure for twentieth-century mathematics. (McLarty 2006, p. 211)

This history is an example of the interplay of geometrical and algebraic thinking (in the broad sense, so as to include topology), which sometimes haunted and sometimes inspired mathematics from the Pythagoreans on, as discussed in detail in Plotnitsky (2022). It is true that the angel (of topology) and the devil (of algebra) sometimes worked together. But they also fought against each other, and changed their role, with algebra becoming a major (but not necessarily majority) angel and topology a minority (but not necessarily minor) devil, keeping here the asymmetry of terms that is part of my argument.[13]

Formally, my emphasis is on concepts and the *conceptual* conception of mathematics, rather than on structures and the structuralist conception of mathematics (à la

[13] This suspicion concerning algebra is often found in geometrically oriented mathematicians, such as Michael Atiyah, who goes as far as stating, quite beyond Weyl's more genial ironic assessment, that "algebra is the offer made by the devil to mathematicians," which even makes one "stop thinking" (Atiyah 2002, p. 7). This is not something that Grothendieck, let alone Noether (even at her early geometrical phase), would have been likely to agree with. It is not easy to so unconditionally distinguish algebra and geometry either, as the present chapter shows and as Atiyah indeed acknowledged, still, however, holding on to his negative sentiment toward algebra. It may well be, however, that, even for Atiyah and certainly for Weyl, the devil of algebra was only the devil of algebra as the "machine" (Atiyah's word) of calculation, algebra deprived of geometry, or topology. In the present view, algebra may also embrace and create geometry and topology, just as geometry and topology embrace and sometimes create algebra. It would in fact be difficult to find greater manifestations of this interplay than Weyl or Atiyah's own work, including that in Atiyah's case on the mathematics of the Yang-Mills theory, which grounds the standard model. The Yang-Mills theory then played a key role in the topology of four-dimensional manifolds, courtesy of the work of Atiyah's student Simon Donaldson, on which Atiyah reflects at some length in his article (Atiyah 2002, pp. 12–13) and which was considered in this context by this author in (Plotnitsky 2022, pp. 275–278). See also Athanase Papadopoulos' article on René Thom, who was equally suspicious about algebra (Papadopoulos 2022, p. 15). Papadopoulos also commented on Atiyah's view under discussion (Papadopoulos 2022, p. 15).

Bourbaki), which governs McLarty's analysis, in part via (Corry 2006), referred to throughout his article. As noted above, the present emphasis does not invalidate this perspective on "modern algebra," for which Noether was one of the founding figures, or the role of structures elsewhere. They are important. Morphisms, central to Noether's (re)algebraization of modern algebra, require structures, as do functors, which are maps between categories as (higher-level) structures. Nevertheless, the present view shifts a perspective, including in Noether's case. Morphisms and functors are concepts first. The same considerations would apply to Grothendieck, who appears to be seen by McLarty along structuralist lines in this article and on several other occasions (McLarty 2006, p. 218, n. 14; 2018). McLarty's analysis of Noether could be easily accommodated by the present view. He summarized his argument as follows:

> The time is important because Noether was just then developing what she called her set-theoretic foundations for algebra. This was not what we now call set theory. It was not the idea of using sets in basic definitions and reasoning. She took that more or less for granted, as did other Göttingers by the 1920s. Rather, her project was to get abstract algebra away from thinking about operations on elements, such as addition or multiplication of elements in groups or rings. Her algebra would describe structures in terms of selected subsets (such as normal subgroups of groups) and homomorphisms. Alexandro[v] applied her tools to Brouwer's use of continuous maps—though Vietoris was the first to make it work.
>
> Noether brought something much deeper and more comprehensive to topology than just the use of homology groups. The next section (and a more technical appendix) will show that groups were familiar in homology before her. She brought an entire programme of looking at groups, and other structures in algebra, and other structures outside of algebra like topological spaces, in terms of the homomorphisms between them. She called this 'set-theoretic foundations.' (McLarty 2006, p. 212)

Noether's "chief tools for making this work," McLarty adds, are "her homomorphism and isomorphism theorems" (McLarty 2006, p. 212). In other words, abstract algebra is now the algebra of morphisms, which is "Noether's principle," as Krull called it. As McLarty says, citing Krull:

> Noether used the homomorphism theorems to prove isomorphism theorems, which show that certain relations among subsets imply that certain morphisms are isomorphisms. These and other ideas served Noether's well-known goal, in Krull's terms:
>
> > Noether's principle: base all of algebra so far as possible on consideration of isomorphisms. (Krull 1935, p. 4)
>
> Van der Waerden captured the method exactly: To understand ring ideals is to understand their analogy with normal subgroups, for which 'we proceed from the concept of homomorphism!' (van der Waerden 1930, p. 55). (McLarty 2006, p. 218)

It is worth noting that van der Waerden, who, as noted, invoked concepts in reflecting on Noether's work, speaks here of the *concept* of a homomorphism, which, or that of a morphism, is a concept-form in the present definition, and is seen as such by Noether, or along with the concept-form of a functor, functorially by

Grothendieck. This point does not affect McLarty's argument as such, because Noether's use of homomorphism implied that abstract algebraic concepts such as a group, field, ring, or module, in fact, function as concepts-forms, without, in the present view, implying the Platonist view of these concepts (even if Noether had such a view, which is unclear). McLarty concludes:

> Noether emphasized homomorphisms, and her influence on homology forced topologists and algebraists to bring their methods together. So algebra shifted to an ever-greater focus on homomorphisms. All this algebra looked too abstract to a tough-minded geometer like Solomon Lefschetz. In the 1920s, he made tremendous but obscure progress applying homology in algebraic geometry. It was a major source of his reputation for never stating a false theorem or giving a correct proof. Realizing that his arguments needed serious improvement, he went to work on the homology of topological spaces. His first book on it says:
>
> > The connection with the theory of abstract groups is clear. . .. Indeed everything that follows in this section can be, and frequently is, translated in terms of the theory of groups. It is of course a mere question of a different terminology. (Lefschetz 1930, p. 29)
>
> But Poincaré had said, in promoting analysis situs, that we must not "fail to recognize the importance of well-constructed language" (Poincaré 1908, p. 180). Whether or not Lefschetz noticed this passage, he shared the thought.
>
> Soon his topology was heavily algebraic (Lefschetz 1942). He kicked off the spate of articles on what became category theory when he asked for an appendix by Eilenberg and Mac Lane (1942). This became the standard foundation for algebraic topology and for the huge proliferation of cohomology theories in the 1950s and since. The unprecedentedly vast machinery gave unprecedented power in solving concrete problems from topology to abstract algebra to number theory. Lefschetz wrote:
>
> > As first pointed out by Emmy Noether, the proper and only adequate formulation of the relation between chains, cycles, . . . requires group theory. (Lefschetz 1949, p. 11, Lefschetz's ellipsis)
>
> In fact, it required Noether's formulation of group theory [in terms of morphisms], and that soon required functors. (McLarty 2006, p. 230).

It is difficult to say how close Noether came to functoriality, and McLarty does not appear to claim that she ever reached this concept. The concept, concept-form, of a functor is the next step, which is, however, far from obvious and was only, and could have only been, created as part of category theory. It then became one of Grothendieck's key concepts, which, while not invented by him, was made by him his own, as I shall discuss in the next section. One might say that in Noether's remarkable approach to algebra, or algebraic topology, there may already have been, if implicitly, categories, those of groups, rings, or modules, even if there were no functors between categories. In Grothendieck, by contrast, functoriality becomes dominant, as one of the most important concept-forms of his mathematics, as both (initially) a minority mathematics and (eventually) a major mathematics, a mathematics that became central to the advancement of mathematical thought in algebraic geometry and beyond.

4 Topos Theory and Grothendieck's Parliamentary Mathematical Ontology

Grothendieck's work on algebraic varieties over a finite field, ultimately allowing one to study them by the tools of algebraic topology, is among the greatest achievements of the twentieth-century *major mathematics* and a remarkable case of a *minority mathematics*, thus also a manifested example of the essential connections between major and minority mathematics, emphasized throughout this chapter. The project of extending these tools to finite fields was initiated by André Weil, a key figure in bringing together algebraic geometry and number theory, in which he was an heir of Fermat (and probably saw himself as one), as well as of Kronecker (in this case, Weil did see himself as one). Arguably, Weil was less of a practitioner of a minority mathematics, except to the degree the latter is always involved in a creation of new mathematics, and Weil has created plenty of it. The project has a complex history, including the development, beyond algebraic geometry itself, of algebraic and differential topology, group theory, elliptic functions, and other fields, in which Weil did important work, eventually leading him to his work on algebraic geometry, culminating in his 1946 *Foundations of Algebraic Geometry*, with the second edition published in 1962 (Weil 1962). The book is distinguished by many important technical contributions, which would be difficult and, for the present purposes, not necessary to discuss here. Eventually (in 1949), Weil suggested that a cohomology theory for algebraic varieties over finite fields, now known as Weil cohomologies, could be developed by analogy with the corresponding theories for complex algebraic varieties or topological manifolds (considered in algebraic topology) in general. Weil's motivation was a set of conjectures (going back to Gauss, they were stated for algebraic curves by Emil Artin in 1924, and hence are part of the history of the major mathematics in this field), known since as the Weil conjectures. They concern the so-called local ζ-functions, which are the generating functions derived from counting numbers of points on algebraic varieties over finite fields. These conjectures, Weil thought, could be attacked by means of a proper cohomology theory, although he did not propose such a theory itself. The idea, stated somewhat vaguely in the first place, was initially met with much skepticism. Obstacles were formidable, and the way of overcoming them advanced by Grothendieck was remarkable, almost miraculous. This way is, however, less of a miracle if viewed from within Grothendieck's mathematical thinking and practice, as minority ones, in several respects unique to him. In this regard, his thinking was similar to that of Abel and Galois, both among earlier predecessors of Grothendieck's work.

As far as Weil's work is concerned, my focus, defined by the differences between Weil and Grothendieck's philosophy of algebraic geometry, is Weil's concept of a "universal domain," assumed by him to be necessary, versus Grothendieck's multiple ontologies of topos theory. As against previous approaches to algebraic geometry, that of Weil required, in addition to the base field of definition of the objects (algebraic varieties) considered, an assumption of an (algebraically closed) "universal domain" encompassing all fields that may appear in any constructions made over the base field. The Platonist implications of this claim are tempered by his view that

such a domain need not be unique. Weil argued that the theorems of algebraic geometry will remain the same regardless of the choice of this domain if it has an infinite transcendence degree and is closed. Grothendieck abolished the idea of a universal domain on his way to solving a problem of Weil's cohomologies posed by Weil, by creating a remarkable array of concept-forms, eventually leading to several concept-forms, such as a topos, that allow one to define Weil's cohomologies. These concept-forms gave Grothendieck's work a unique place in algebraic geometry and beyond. Grothendieck's definition of "Weil's cohomologies," or more accurately Grothendieck's cohomologies that accomplished the task in question, was a result of his, in its key aspects, minority mathematical practice, even when adopting major concepts of major mathematics of his time. This practice and the ways of thinking grounding it were initiated in his earlier work in functional analysis, his original field of research, although developed, as more directly connected to algebraic geometry, in his work on cohomological algebra and Teichmüller theory.[14]

A remarkable manifestation of this thinking in algebraic geometry was his invention of the concept of scheme, a radically innovative concept, a concept-form, even though it was *almost* a natural extension of by then the standard view of algebraic varieties. Nobody, however, thought about it at the time, which fact made it a minority concept. This type of event (which is not uncommon in creative mathematics and science) also requires a rethinking of the idea of "natural," the naturalness of the natural. Grothendieck's key idea was replacing maximal ideals with prime ideals of a ring, a *natural*, Noetherian, move in commutative algebra, by then an established major field. It is, however, a different matter, when one connects it to algebraic varieties, replacing the latter by schemes, thus introducing a new form of multiplicity insofar as the same algebraic variety can arise from different schemes. It also allows one, which may be seen as another post-Noetherian move, to define algebraic varieties over any commutative rings, such as integers. In its ultimate formality (favored by Grothendieck), a scheme is a topological space with commutative rings associated to all of its open sets, by gluing together spectra (defined as spaces of prime ideals) of commutative rings along their open subsets. Thus, a scheme is locally a spectrum of a commutative ring. This also leads to a relational view, according to which algebraic geometry should be developed for a morphism $X \to Y$ of schemes (designated as a scheme *X over Y*), rather than for an individual scheme. In some cases, the family of all varieties of a given type can itself be viewed as a variety or scheme, a moduli space. The approach led to an enormous amount of new mathematics, much of which quickly became major mathematics, although not majority mathematics. As is, again, always the case, these concepts, even apart from their interconnections, have their history (as do some of these interconnections, while others are new), for example, in Noether's work or in more immediately

[14] The genealogy of his work on Teichmüller space originates in Riemann's moduli problem, recast by Grothendieck in his framework. The theory provided an important example of Grothendieck's use of functoriality, a major concept by then, but given a new life by his minority use of it. For the discussion of this work, see A'Campo et al. (2016).

preceding works in algebraic geometry by Jean-Pierre Serre. A subtler earlier link is the Stone representation theorem, which states that every Boolean algebra is isomorphic to a certain field of sets. The theorem emerged from the *spectral* theory of operators in a Hilbert space, a subject close to Grothendieck's early work in functional analysis.

The concept of topos was introduced nearly simultaneously with that of scheme, by relating the history of algebraic geometry just sketched to the history of algebraic topology and adopting the method of the latter in algebraic geometry over the finite field. This (unlike their automatic application to algebraic manifolds over $\mathbb{C}$) proved to be impossible to do merely by shifting from varieties to schemes. To be able to do so, one needed a proper topology, which was nontrivial because the objects in question are topologically discrete. A "native" topology that could be defined by them, known as Zariski's topology, did not work, because it had too few open sets, even if one uses schemes instead of varieties. The decisive insights came from Grothendieck, helped by sheaf-cohomology theory and category theory, "cohomological algebra," by then the standard tool of algebraic topology, and as such part of major mathematics, already transformed by Grothendieck previously, on which I shall comment below. Using these tools, a hallmark of Grothendieck's thinking throughout his career, and his previous concepts, such as that of "scheme," led him to topos theory and étale cohomology as a viable candidate for Weil's cohomology, which it had quickly proven to be. As Grothendieck said: "The two crucial driving ideas [*idées forces*] in launching and developing the new geometry were that of scheme, and that of topos. Appearing almost at the same time and in close symbiosis with each other, they were as if a single sinew in the spectacular flight of the new geometry, and this from the very year they appeared" (Grothendieck 2022, p. 50, cited in McLarty 2018, p. 109; McLarty's translation). Soon thereafter, he introduced a more general concept of motivic cohomology, a kind of ultimate cohomological concept-form. By using étale cohomology, Grothendieck (with Michael Artin and Jean-Louis Verdier) and Pierre Deligne (Grothendieck's student) were able to prove Weil's conjectures, a bit later generalized by Deligne.

Grothendieck's key insight was to generalize, in terms of category theory, the concept of "open set," beyond a subset of an algebraic variety. This was possible because the concept of sheaf and of cohomology of sheaves could be defined for any category, rather than only that of open sets of a given space. Étale cohomology is this type of replacement, specifically by using the category of étale mappings of an algebraic variety or a scheme, which become "open subsets" of the finite unbranched covering spaces of the variety, a vast generalization of Riemann's concept of a covering space. Grothendieck was in part building on ideas of Serre. The origin of this generalization was in the fact that the fundamental group of a topological space, for example, and (in the present context) in particular, a Riemann surface, could be defined in two ways. First, it can be defined, more geometrically but still by means of algebra, as a group of equivalence classes of sets of loops at a given point, with the equivalence relation given by homotopy. Alternatively, a fundamental group can be defined even more algebraically, as a group of transformations of covering spaces over the topological space considered. In this second definition, it is analogous to the

Galois group of the algebraic closure of a field, as Serre was first to consider for finite fields, which prepared grounds for Grothendieck's work on étale cohomology. The idea has, thus, both Galois' and Riemann's concepts as important parts of its genealogy. The connection had been established in the case of Riemann surfaces long before Serre's work, and found, for example, in Weyl's 1913 book on a Riemann surface, undoubtedly known to Serre:

> The group of cover transformations, regarded as an abstract group, expresses purely and completely everything in the relation between the normal covering surface $\overline{\mathcal{F}}$ and the base surface $\mathcal{F}$ which has the character of analysis situs [topology]. This group is also called the *Galois group* of $\overline{\mathcal{F}}$. It is in fact the analog of the Galois group of a normal algebraic field (of finite degree) over a base field. (Weyl 2013, p. 58)

Étale mappings give a sufficient number of open sets to define adequate cohomolgy groups for some coefficients for algebraic varieties over finite fields. In the case of complex varieties, one recovers the standard cohomology groups with coefficients in any constructible sheaf. The category of étale mappings is a topos, a concept that is, for now, the most abstract algebraic form of spatiality in mathematics. It was seen as such by Grothendieck himself, who compared his contribution to the idea of space in mathematics to that of Einstein in general relativity in physics and that of Erwin Schrödinger in quantum mechanics (Grothendieck 2022, p. 68).[15] Étale cohomologies could be defined for most practical uses in simpler settings, without using the concept of topos. The concept of topos has, however, never lost its importance. It also came to play a major role in mathematical logic. The subject cannot be addressed here, apart from noting that mathematical logic is the culminating development of the trajectory, initiated by Georg Cantor's work, with radical epistemological implications, concerning the nature of mathematical reality, to which the concept of topos, as a *logical* concept, may be connected (e.g., Plotnitsky 2022, pp. 82–95). By contrast, Grothendieck's use of his topoi in algebraic geometry is

[15] A reference to quantum mechanics may appear strange, but it is astute. In quantum mechanics, one deals with Hilbert spaces over $\mathbb{C}$, in the case considered by Schrödinger in his discovery of the theory, infinite-dimensional ones, using which enables one to predict the probabilities of quantum events. (Adding the so-called Born's rule, which is, essentially, a form of complex conjugation, to this formalism over $\mathbb{C}$ allows one to move to probabilities, which are real numbers.) The astuteness of Grothendieck's reference comes from the fact that quantum mechanics establishes an entirely new relationship between a purely mathematical space, here Hilbert spaces over $\mathbb{C}$, and the actual (three-dimensional) physical space, mathematically represented as a real manifold, as considered in detail in (Plotnitsky 2021). Technically, it was Werner Heisenberg (not mentioned by Grothendieck) who defined this situation as such in his discovery of quantum mechanics a few months before Schrödinger (who discovered it independently in a different but mathematically equivalent form). Ironically, Schrödinger became discontented with and repudiated quantum mechanics on the account of its epistemological features defined by the situation just described, in particular the role of complex numbers and the probabilistic nature of the theory. He came to see quantum mechanics, as did Einstein, as merely a convenient method of calculations rather than as a proper fundamental theory of the ultimate constitution of nature (Plotnitsky 2021, pp. 145–166). I am not sure whether Grothendieck was aware of this attitude on the part of Schrödinger and Einstein.

essentially ontological rather than logical. His overall position concerning the nature of mathematical reality remains somewhat unclear, for example, as concerns whether it conforms or not to mathematical Platonism. On the other hand, as discussed below, the mathematical architecture of Grothendieck's algebraic geometry, as different from that of Weil's algebraic geometry, as defined by a universal domain, may, even if against Grothendieck's own grain, lead to questioning all known forms of Platonism.

To better see why this is the case, I shall offer brief informal remarks on category and topos theory, in connection (direct in the case of topos theory) with Grothendieck's thinking. First, consider the following way of endowing a space with a structure, generalizing the definition of topological space in terms of open subsets. One begins with an arbitrarily chosen space, X, which may initially be left unspecified in terms of its properties and structure. What would be specified instead are the relationships between spaces, which are applicable to X, such as mapping or covering one or a portion of one by another. This structure is the arrow structure $Y \rightarrow X$ of category theory, where X is the space under consideration and the arrow designates the relationship(s) in question. One can also generalize the notion of neighborhood or of an open subspace of (the topology of) a topological space in this way, by defining it as a relation between a given point and space (a generalized neighborhood or open subspace) associated with it. This enables one to specify a given space not in terms of its intrinsic structure (e.g., a set of points with relations among them) but "sociologically," throughout its relationships with other spaces of the same category, say, that of algebraic varieties or schemes over a finite field of characteristic *p* (Manin 2002, p. 7). Some among such spaces may play a special role in defining the initial space, X, and algebraic structures, such as homotopy and cohomology, as Riemann realized in the case of covering spaces over Riemann surfaces. (Proper homotopy and cohomology theories were developed later.) Riemann's idea of a covering space was, as noted, one of the inspirations for Grothendieck's concept of étale topos.

To make this description more grounded, I shall briefly, and again, informally, sketch the key ideas of category theory. It was introduced as part of cohomology theory in algebraic topology in 1940 and extensively used by Grothendieck in his approach to cohomological algebra and algebraic geometry. Category theory considers multiplicities (which need not be sets) of mathematical objects conforming to a given concept, such as the category of differentiable manifolds or that of algebraic varieties, and the arrows or morphisms, the mappings between these objects that preserve this structure. Studying morphisms allows one to learn about the individual objects involved, often to learn more than we would by considering them only or primarily individually. As discussed earlier, a shift toward studying morphisms, even if not categories, was one of the key new features of Noether's work in both algebra and topology. In a certain sense, in his Habilitation lecture (Riemann 1854), Riemann already thinks in terms of morphisms and even categories (although he did not of course have either concept) because he does not start with a Euclidean space. Instead, the latter is just one specifiable object of a large protocategorical

multiplicity, that of differentiable or, more narrowly, Riemannian manifolds, defined by a particularly simple way we can measure the distance between points.

Categories themselves may be viewed as objects defined by the relationships between them, and in this case one speaks of "functors" between categories rather than "morphisms" between objects in a given category. Functoriality is a crucial aspect of category theory and, as emphasized here, of Grothendieck's work. Grothendieck did not invent the concept of functor, but he made it his own and used it with great effectiveness in his work, including in his invention of new concepts, often, as in the case of his cohomological and homotopical concepts, defined by functoriality, as are, however, cohomology theories in general. Topology relates topological or geometrical objects, such as manifolds, to algebraic ones, such as homotopy and cohomology groups. Thus, in contrast to geometry (which relates its spaces to algebraic aspects of measurement), topology, by its very nature, in effect deals with functors between categories of topological objects, such as manifolds, and categories of algebraic objects, such as groups, rather than only with morphisms, which of course remain essential. There are no functors without morphisms, while morphisms do not require functors. Functors, however, give morphisms their categorical meaning and, as a result, a deeper and more fundamental mathematical meaning and significance.

With Grothendieck, mathematical practice becomes that of functoriality: It no longer primarily deals with objects but categories of objects. Although, arguably, less of a minority thinking insofar as it had a tradition by then, it was still highly transformative vis-à-vis previous uses of functoriality, and much of it was new and hence a minority mathematics. Grothendieck's use of it, as a concept-form, led him to numerous conceptual innovations, beginning with those in *Sur quelques points d'algèbre homologique*, a long article, now known as the "Tohôku paper" (Grothendieck 1957). The article, as part of its minority and, thus, Abelian mathematics, also introduced abelian categories, an important new concept, and used this concept to show that sheaf cohomology can be defined as certain derivative functors in this setting. The name "abelian category" merits a brief comment in the context of this chapter. The concept was independently introduced earlier (in 1955), by David Buchsbaum, under the name of "exact categories." Grothendieck's well-known penchant for selecting his mathematical names with both flair and precision (schemes, étale mapping, and motives, among them) also served him here, even apart from the present context, in which the name carries additional resonances linked to the idea of a minority mathematics. It is unlikely, although not inconceivable, that Grothendieck also had in mind Abel's thinking in defining modern mathematics, as an example of Abelian mathematics in the present sense of a minority mathematics. The origin of the name had most likely to do with the fact that a prototypical example was the category of abelian (commutative) groups. Grothendieck gave an axiomatic definition, which is secondary for the moment. An abelian category is a category in which objects and morphisms can be added and in which kernels and cokernels exist (and possess certain useful properties). Their significance in Grothendieck's works and elsewhere in algebraic geometry, cohomology theories, or category theory, as such, has to do with their strong stability

aspects, such as the fact that the category of chain complexes or the category of functors from a (small) category to an abelian category are abelian. Among the examples of abelian categories is that of finitely generated left modules over a left-noetherian ring, which implies that the category of finitely generated modules over a noetherian commutative ring is abelian. This is one of several junctures in which Noether's work in commutative algebra enters the history leading to Grothendieck's work, as part of the sequence considered here from Abel to Galois to Noether to Grothendieck. There are, again, many other proper names in between and beyond (or before) this history, some of which, but only a small sample, were mentioned here in this chapter. As already in his earlier work in functional analysis (where he introduced several new concepts, such as topological vector spaces), but more significantly in the present context, *Sur quelques points* confirmed the essential role of inventing new concepts, in particular new concept-forms, in creative mathematics. So did Grothendieck's subsequent work, defined by a massive proliferation of new concepts and concept-forms. Even those (a large number as well) of such concepts and concept-forms that are contained in thousands of pages of his unpublished works, kind of reinstating his work as minor mathematics, had major impacts during the last half of a century and continue to do so. Those of scheme and topos, however, remain arguably the greatest examples, again, seen as such by Grothendieck himself.

A topos in Grothendieck's sense is a category of spaces and arrows over a given space, used for the purpose of allowing one to define richer algebraic structures associated with this space, as explained above. There are additional conditions such categories must satisfy, but this is not essential at the moment. As a simple example, for any topological space S, the category of sheaves on S is a topos. The concept of topos extends far beyond spatial mathematical objects (thus, the category of finite sets is a topos); it replaces the latter with a more algebraic structure of categorical and topos-theoretical relationships between objects. On the other hand, it derives from the properties of and (arrow-like) categorical relationships between properly topological objects. The conditions, just mentioned, that categories forming topoi must satisfy have to do with these connections.

One might think of this ontology as an assembly, a "parliament," of ontologies (Grothendieck's concept of topos is, again, ontological, rather than logical, as is the concept of topos in mathematical logic), in the absence of any ultimate ontology, or even, as against physics, any ultimate mathematical or otherwise mental reality, as explained in Plotnitsky (2022, pp. 82–95). As against Weil's universal domain of algebraic geometry, topos theory enables a new mathematical ontology by allowing for different (mathematical) universes associated with a given space, possibly a single-point one. Grothendieck's topoi are possible worlds, or compossible worlds in Leibniz's sense, without assuming, as Leibniz did (in dealing with the actual world), the existence of only one of them, the best possible one. Hence, I refer to this ontology as a parliamentary assembly of possible ontologies, in the absence of any encompassing ontology, implied by the idea of a universal domain, assumed in Weil's "imperial" vision of algebraic geometry. According to David Reed's summary:

> Grothendieck's constructions [beginning with schemes replacing the varieties of Weil] are far from the 'palaces' which Weil suggests belong to algebraic geometers by birthright. Rather than refurbishing and renewing the old constructions he has instead created an entirely new type of architecture and in the process is forced to make extensive use of the 'makeshift algebraic constructions' Weil carefully thought to avoid placing reliance upon. Both Weil and Grothendieck seek generality in order to be able to analyze rigorously geometric situations to which the older algebraic geometers paid little attention. This generality permits consideration of 'degenerate' cases, objects with singularities, 'non-reduced' objects and the like. Whereas Weil sought this generality through a large 'domain' in which the result of his constructions could be contained, Grothendieck seeks generality by expanding the category of objects under considerations until a certain 'self-sufficiency' is acquired, i.e. until a natural reflexivity can be found so that the objects in the category can be related to each other and operations on them can be undertaken without going outside the original category. Furthermore, the objects do not require a predetermined 'fixed point' outside of the category for their specification. ... This relativity also extends to the set theoretical foundations of the theory. Grothendieck introduces a set theoretic 'universe' which provides the sets used in his constructions and then studies those properties which remain invariant under a change of universe! (Reed 1995, pp. 131, 177n.36)

In question, then, are two very different (mathematical) ontological philosophies – that of unification or encompassment and that of multiplicity, which is *relational* but has no single encompassing domain. In other words, this multiplicity is a multiplicity without unity, "the multiple without One," as Alain Badiou would have it (Badiou 2006). This multiplicity was made possible in part by functoriality as a key part of Grothendieck's mathematical practice, initially a minority mathematical practice. It is true that Grothendieck often speaks of "unity" (also associated with "simplicity") as more profound than "generality" (e.g., Grothendieck 2022, p. 47; McLarty 2018, p. 124).[16] This unity, however, appears to me more as a unity of his vision, although I cannot exclude that Grothendieck might have envisioned an ontology (different from that of Weil) unifying his multiple universes. I am, however, more concerned here with a philosophical view that Grothendieck's mathematics may support, even if against his own grain.

From this perspective, Grothendieck's generality or, better (Reed is right to switch to this term), *relativity* of construction, a crucial aspect of his functorial thinking, is a mathematical instantiation or model of the capacity of mathematics as a human endeavor. This model is that of the creation of mathematical realities (as mental ontologies, of which Grothendieck's topoi are examples), on the basis of, while creatively transforming, previously established realities, in the absence of an ultimate underlying or encompassing Platonist reality. It is a creation of mathematical worlds (plural!) from the materials of human thought, by endowing them with logically rigorous conceptual structures, the novelty of which is especially striking and prolific in Grothendieck.

[16] McLarty's article as well as his (McLarty 2016) offers an informative account of these concepts in Grothendieck, although neither article addresses the question of the multiple, or the implications of Grothendieck's conceptions of mathematical multiplicities for the question of reality and realism in mathematics, considered, also in connection with Grothendieck's thinking, in Plotnitsky (2022).

Topoi are conceptual multiplicities by virtue of bringing together the concept of algebraic geometry and algebraic topology, a strategy that, while already in place in general, was, ironically, motivated by Weil's cohomological ideas in the case of étale cohomology. One needed to get rid of the concept of a universal domain, but one also needed, in order to do so in the first place but also to be able to work in multiple domains, new concept-forms. The creation of such concept-forms was, again, prolific in Grothendieck's mathematical practice, the grounding principle of which was functoriality, just as the grounding principle of Noether's mathematical practice, Noether's principle, was defined by the mathematics of morphisms between sets with structures (rather than only the structures of sets). It might be added that functoriality was not found in Weil either. *Foundations of algebraic geometry* (still following Noether's approach focusing on morphism vs. functors) mentions categories and functors once, somewhat condescendingly and that in 1962 (obviously aware of Grothendieck's breakthroughs by using categories and functors): "Any reader interested in expressing these facts [concerning morphisms of algebraic varieties over topological fields] in the language of categories and functors will presumably be able to do so" (Weil 1962, p. 358). There is, however, no Weil cohomologies without functoriality, implicit in any cohomology theory, as Weil must have been aware. As noted, Weil cohomologies can be defined in simpler settings. However, they were *invented* through functoriality, a key tool of Grothendieck's mathematical practice helping many of his conceptual inventions. The concept of functor itself was not one of them, but, as I said, Grothendieck made it his own by his innovative use of it throughout his work.

Grothendieck's philosophy of *mathematical practice*, and perhaps his *philosophy* of mathematics, especially as manifested in the relativity of topos theory, may be seen as mirroring the ultimate task of mathematics, that of creating new mathematical worlds. The multiple ontologies of Grothendieck's topos theory may be seen as all having their possible voices in the parliament of topos theory, and a similar view can be taken concerning mathematics in general, making it a parliament of various mathematics. These voices, again, can only be given to them by us, by our mathematical thought, individual or collective (Plotnitsky 2022, pp. 82–95).[17]

[17] As I argue on the same occasion, an analogous multiplicity is also a consequence of Gödel's undecidability, as exemplified by Cantor's continuum hypothesis. This hypothesis was crucial not only for the question of continuity but also for the question of Cantor's hierarchical order of infinities (the infinity of which was one of his discoveries, in, it might be added, his own minority mathematics at the time) and thus for the whole edifice of Cantor's set theory. The hypothesis was proven undecidable by Cohen in 1961, with Gödel contributing a crucial part earlier. It follows, however, that one can extend classical arithmetic in two ways by considering Cantor's hypothesis as either true or false, that is, by assuming either that there is no such intermediate infinity or that there is. This allows one, by decisions of thought, to extend arithmetic into mutually incompatible systems, ultimately infinitely many such systems, because each of them will contain at least one undecidable proposition. Gödel's incompleteness theorems may not be a necessary condition of the multiplicity of possible mathematics, as Grothendieck's ontological relativity suggests. The existence of Cantorian and non-Cantorian set theories, or the existence of Euclidean and the various

Grothendieck's relativistic ontological architecture "formally" captures or provides a model for the architecture of the mathematical practice as allowing the space for alternatives to a dominant major mathematics. It is not only a *metaphorical* model because this type of mathematical practice can be enacted within an architecture of the type defined by Grothendieck. In this chapter, this practice is expressly considered not merely as some alternative mathematics but as a minority mathematics defined as operative inside and transforming what is the established major mathematics at the time. I further argue that a major mathematics cannot be transformed otherwise, however hidden or implicit a minority practice may be in these transformations. This transformational dynamic is not contained by a model strictly based on Grothendieck's relativistic architecture but instead underlies this architecture as a more general conception. His own mathematics, including that defining the ontological multiplicity in question, was a minority mathematics aimed to affect and transform an important and very rich part of the major mathematics of his time. As he realized, he was not always successful in this aim at the time when some of his new concepts were introduced (especially after he left *Institut des Haute Études Scientifiques*, where his role was akin to that of Hilbert in Göttingen in the early twentieth century) although some of his, initially marginal, ideas became more widespread subsequently. The view of a minority mathematics as something inside a major mathematics, rather than outside it, also helps to explain why Grothendieck continued to reflect on the nature of the major mathematics of his time, especially in the fields related to his work but, at least by implication, in general. Our major mathematics, he thought, did not pay enough attention to foundational issues, did not ask the right foundational questions, as Abel would have it.[18]

non-Euclidean geometries established earlier, does not depend on these theorems. Gödel's theorems, however, imply the inevitability of such a multiplicity in principle, on mathematical grounds, whether one pursues such alternatives or not, in the way say mathematics has pursued both Euclidean and non-Euclidean geometry. In this case, however, one could still unify them with the same broader framework, say, as defined, following Riemann, by the concept of manifold. In the case of undecidable propositions, the only thing that these alternative forms of mathematics share is the fact they are forms of mathematics, which are, however, incompatible with each other. The role of a decision of thought, in general important for my argument here, especially in the case of Gödel's theorems, is a key aspect of this situation. Allowing one, by such a decision and a freedom, even if relative, of making it, to extend arithmetic or any system containing it into two mutually incompatible systems, and (by iteration) to infinitely many such systems, has been seen as a troubling and even intolerable situation for some. One might, however, also find an appeal in it, as the present author or Badiou does (Badiou 2006, p. 99).

[18] His long later manuscripts, such as, famously, *Pursuing Stacks,* a long letter to Daniel Quillen (Grothendieck 1983) may be seen along these lines. (Quillen's work, inspired by that of Grothendieck, is another example of a creation of new category- and functoriality-based concept-forms). On the other hand, Grothendieck's return to Montpellier university, his alma matter, after he left Paris, might be seen as (or at least conjectured to be) aimed to reinstate his position as a minority mathematician and to affirm a minority mathematics, as a revolutionary condition of all creative mathematics and hence mathematics itself, as I argue here. An enormous amount of new mathematics, including new concept-forms, was created by him in his post-Paris years.

5 Conclusion

"Mathematics is *a thought*," Badiou says (Badiou 2006, p. 43). I give this statement a meaning beyond Badiou's own, correctly juxtaposing mathematics to merely a rigorous and refined form of logic, which is necessary but not sufficient for mathematics. While I fully agree with this view, I use this statement more in accord with Deleuze and Guattari's view of *thought* in *What is Philosophy?* as creative thinking (rather than merely referring to our mental states), specifically in art, philosophy, and mathematics and science (Deleuze and Guattari 1996, pp. 202–281). As I argue here, however, in accord with their view of creative thought as a minority thought, as exemplified by Kafka's literary thought, all creative thought depends on the possibility, and freedom, of pursuing a minority thinking, sometimes against odds. It was not easy for Kafka or Abel and Galois to have this freedom, but they had the capacity and boldness to be free and create new literature and new mathematics. Any instance of creative thinking is, in the present view, an instance of a minority thinking, and only a minority thinking is creative. If there is no minority thinking (however relative in its scope), there is no creative thought and, hence, no creative literature or mathematics, which still needs to be created, *made* by us.

I have argued here that the creative, or *poetical*, essence of creative thought and hence a minority thought in mathematics is composition, in the sense of the ancient Greek *poeien* as making, just as it is for poetry or all creative endeavors, respecting their specificity as such; I have also argued here that the *political* essence of creative or minority thought in mathematics or elsewhere is freedom. In this sense, a minority mathematics and hence all mathematics are political: It requires the politics of freedom, manifested in the possibility of a minority mathematics. Hence, the essence of mathematics is ultimately the power of compositions, conceptual or other and the freedom to create them, not rules, logical or other, necessary as such rules are for mathematics as a human endeavor and a disciplinary field. In spite, however, and sometimes because of its logical and other constraints, mathematics gives human thought an extraordinary freedom. Mathematics is a combination of an extraordinary precision or exactitude and an extraordinary freedom. I would be tempted to say the ultimate precision or exactitude, if not the ultimate freedom, except that as creative thinking, science, philosophy, and literature and art have their own ways of precision and exactitude, and all creative thought is a manifestation of freedom, again, always the freedom of minority thought.

This freedom is, as I argued, not merely a free choice in our decision in pursuing our projects, including as a minority mathematics, because, as in all human decisions, such decisions may involve factors that shape and complicate our choices, including as concerns their freedom. A choice, at least a free choice, is difficult, if ever entirely possible as free, even in a democracy, mathematical or other. This fact, as I explained, makes decision a better category. As I also argued, however, this does not mean there is no freedom in our decisions. At stake, politically, is creating the *possibility* for freedom in our endeavors, for pursuing a minority thinking, the necessary condition of creative thinking. No creative thought or justice anywhere is possible otherwise. It is this *possibility* of freedom that is the political essence of a

minority mathematics and hence of mathematics itself, which cannot advance by creating new mathematics and thus remain alive without a minority thinking and a freedom to pursue it.

6 Cross-References

- ▶ Grothendieck: A Short Guide to His Mathematical and Philosophical Work (1949–1991)
- ▶ Mathematical Practice as Philosophy, with Galois, Riemann, and Grothendieck
- ▶ Mathematics and the Method of Abstraction
- ▶ One Mathematic(s) or Many? Foundations of Mathematics in Twentieth-Century Mathematical Practice
- ▶ Ontology in the History and Philosophy of Mathematical Practice: An Introduction
- ▶ The Notion of Space in Grothendieck
- ▶ What Mathematicians Do: Mathematics as Process and Creative Rationality

References

A'Campo N, Ji L, Papadopoulos A (2016) On Grothendieck's tame topology. In: Papadopoulos A (ed) Handbook of Teichmüller theory, vol VI. European Mathematical Society, Zürich, pp 35–70

Alexandrov P (1936) In memory of Emmy Noether. In: Dick A (ed) Emmy Noether 1982–1925. Birkhäuser, Boston, 1981, pp 152–179

Angier N (2012, 26 March) The mighty mathematician you've never heard of. New York Times

Atiyah M (2002) Special article: Mathematics in the 20th century. Bull Lond Math Soc 34:1–15

Badiou A (2006) Briefings on existence (trans: Madarasz N). SUNY Press, Albany

Corfield D (2020) Modal homotopy type theory: the prospect for a new logic of philosophy. Oxford University Press, Oxford

Corry L (2006) Modern algebra and the rise of mathematical structures. Birkhäuser, Boston

De Man P (1996) Kant and Schiller. In: De Man P Aesthetic ideology. University of Minnesota Press, Minneapolis, pp 129–162

Deleuze G (1995) Difference and repetition (trans: Patton P). Columbia University Press, New York

Deleuze G, Guattari F (1987) A thousand plateaus (trans: Massumi B). University of Minnesota Press, Minneapolis

Deleuze G, Guattari F (1989) Kafka: toward a minor literature (trans: Polan D). University of Minnesota Press, Minneapolis

Deleuze G, Guattari F (1996) What is philosophy? (trans: Tomlinson H, Burchell G). Columbia University Press, New York

Dick A (1981) Emmy Noether 1982–1925. Birkhäuser, Boston

Doxiadis A, Mazur B (eds) (2012) Circles disturbed: the interplay of mathematics and narrative. Princeton University Press, Princeton

Eilenberg S, Mac Lane S (1942) Appendix A: on homology groups of infinite complexes and compacta. In: Lefschetz S Algebraic topology. American Mathematical Society, Providence, pp 344–349

Grothendieck A (1957) Sur quelques points d'algèbre homologique. Tôhoku Math J 9(2):119–221. https://doi.org/10.2748/tmj/1178244839

Grothendieck A (1983) Pursuing stacks (a letter to D. Quellen) *thescrivener.github.io*. Retrieved 07/01/2023
Grothendieck A (2022) Récoltes et semailles, I, II. Réflexions et témoignage sur un passé de mathématicien. Galimard, Paris
Hegel GWF (2019) Hegel's phenomenology of spirit (trans: Pinkard T). Cambridge University Press, Cambridge
Kant I (2015) Critique of practical reason (trans: Gregor M). Cambridge University Press, Cambridge, MA
Kosmann-Schwarzbach Y (2011) The Noether theorems: invariance and conservation laws in the twentieth century (trans: Schwarzbach B). Springer, Berlin
Krull W (1935) Idealtheorie. Julius Springer, Berlin
Kuhn T (1987) Black body theory and the quantum discontinuity, 1894–2012. University of Chicago Press, Chicago
Lakatos I (1980) Mathematics, science and epistemology. Philosophical papers, vol 2 (ed.: Worral J, Currie G). Cambridge University Press, Cambridge
Langlands R, Shelstad D (2007) Descent for transfer factors. In: Cartier P, Katz NM, Manin YI, Illusie L, Laumon G, Ribet KA (eds) The Grothendieck festschrift. Birkhäuser, Boston, pp 485–563. https://doi.org/10.1007/978-0-8176-4575-5_12
Lefschetz S (1930) Algebraic topology. American Mathematical Society, Providence
Lefschetz S (1942) Algebraic topology, 2nd edn. American Mathematical Society, Providence
Lefschetz S (1949) Introduction to topology. Princeton University Press, Princeton
Lyotard J-F (1984) The postmodern condition: a report on knowledge (trans: Bennington G, Massumi B). University of Minnesota Press, Minneapolis
Manin Yu (2002) Georg Cantor and his heritage. arXiv.math.AG/0209244 v1
Manin Yu (2010) What then? Plato's ghost: the modernist transformation of mathematics. Notices Am Math Soc 57(2):239–243
Manin Yu (2019) Time and periodicity from Ptolemy to Schrödinger: paradigm shift vs. continuity in history of mathematics. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer/Nature, Berlin, pp 129–138
McLarty C (2006) Emmy Noether's set-theoretical topology: from Dedekind to the rise of functors. In: Ferreirós J, Gray J (eds) The architecture of modern mathematics: essays in history and philosophy. Oxford University Press, Oxford, pp 211–236
McLarty C (2011) Emmy Noether's first great mathematics and the culmination of first-phase logicism, formalism, and intuitionism. Arch Hist Exact Sci 65:99–117. https://doi.org/10.1007/s00407-010-0073
McLarty C (2016) How Grothendieck simplified algebraic geometry. Notices of American Mathematical Society 63(3):256–65
McLarty C (2017) The two mathematical careers of Emmy Noether. In: Beery J, Greenwald S, Jensen-Vallin J, Mast M (eds) Women in mathematics. Association for Women in Mathematics series, vol 10. Springer, Cham. https://doi.org/10.1007/978-3-319-66694-5_13
McLarty C (2018) The large structures of Grothendieck founded on finite-order arithmetic. Rev Symb Log 13(2):296–325
Nietzsche F (1966) Beyond good and evil: a prelude to a philosophy of the future (trans: Kaufmann W). Vintage, New York
Noether E (1918) Invariante Variationsprobleme, Göttinger Nachrichten (1918), pp 235–257 (presented by F. Klein at the meeting of 26 July 1918); Abhandlungen, pp 248–270. Abstract by the author in Jahrbuch über die Fortschritte der Mathematik, 46 (1916–1918), vol 1, IV.15 (Variationsrechnung), p 770
Noether E (1925) Ableitung der Elementarteilertheorie aus der Gruppentheorie, 27. Januar 1925, Jahresbericht derDeutschen Mathematiker-Vereinigung 34 (Abt. 2)
Noether E (1927) Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern [Abstract structure of the theory of ideals in algebraic number fields]. Math Ann 96(1):26–61. https://doi.org/10.1007/BF01209152, S2CID 121288299

Noether E (1929) Hyperkomplexe Größen und Darstellungstheorie [Hypercomplex quantities and the theory of representations]. Math Ann 30:641–692. https://doi.org/10.1007/BF01187794, S2CID 120464373
Papadopoulos A (2022) René Thom: from mathematics to philosophy. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer/Nature, Switzerland AG, Cham
Plotnitsky A (2001) Algebra and allegory: nonclassical epistemology, quantum theory, and the work of Paul de Man. In: Cohen T, Cohen B, Miller JH, Warminski A (eds) Material events: Paul de Man and the afterlife of theory. University of Minnesota Press, Minneapolis, pp 49–92
Plotnitsky A (2012) The disaster of the diagonal. In: Doxiadis A, Mazur B (eds) Circles disturbed: the interplay of mathematics and narrative. Princeton University Press, Princeton
Plotnitsky A (2021) Reality without realism: matter, thought, and technology in quantum physics. Springer/Nature, Heidelberg
Plotnitsky A (2022) Logos and Alogon: thinkable and the unthinkable in mathematics, from the Pythagoreans to the Moderns. Spinger/Nature, Heidelberg
Poincaré H (1908) Science et méthode. Flammarion, Paris
Polya G (1990) Mathematics and plausible reasoning, volume 1: induction and analogy in mathematics. Princeton University Press, Princeton
Raussen M, Skau C (2010) Interview with Mikhail Gromov. Notices Am Math Soc 57:391–403
Reed D (1995) Figures of thought: mathematics and mathematical texts. Routledge, London
Riemann B (1854) On the hypotheses that lie at the foundations of geometry. In: Pesic P (ed) Beyond geometry: classic papers from Riemann to Einstein. Dover, Mineola, 2007, pp 23–40
Rousseau J-J (1997) Emile, or on education (trans: Bloom A). Basic Books, New York
Schiller F (2004) On the aesthetic education of man (trans: Snell R). Dover, Mineola
Schweber SS (1984) QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga. Princeton University Press, Princeton
Van der Waerden BL (1930) Moderne algebra. Springer, Berlin
Van der Waerden BL (1985) A history of algebra: from al-Khwārizmī to Emmy Noether. Springer, Berlin
Voevodsky V et al (2013) Homotopy type theory: univalent foundations of mathematics. Univalent Foundations Program, Princeton
Vuillemin J (1962) La Philosophie de l'algèbre. Tome I: Recherches sur quelques concepts et méthodes de l'Algèbre moderne. Presses universitaires de France, Paris
Weil A (1962) Foundations of algebraic geometry, 2nd edn. American Mathematical Society, Providence
Weyl H (1939) Invariants. Duke Math J 5:489–502
Weyl H (2013) The concept of a Riemann surface (trans: MacLane GL). Dover, Mineola

Pavel Florensky and His World

Athanase Papadopoulos

They call it "Silver" Age, but it is comparable to the Renaissance in Europe
(Valentin Poénaru told me).

Contents

1 Introduction 802
2 Vita 803
3 Imaginaries 820
4 Dante 823
5 Time, Relativity, and Literature 828
6 Bugaev and Florensky: Mathematics and Philosophy of Discontinuity 832
7 Luzin: Spiritual Crisis, Function Theory, Luzitaniya, and Disgrace 836
8 Silver Age: Mathematics, Symbolism, Art, and Poetry 845
9 Florensky and Akhmatova 849
10 Contradiction and Antinomy in Mathematics, Philosophy, and Theology 853
11 The Philosophers 858
12 Reverse Perspective Again 859
13 In Guise of a Conclusion 861
References 862

Abstract

This is an overview of the life and works of Pavel Florensky, an important and singular figure of the period rightly described as the *Silver Age of Russian mathematics*, with a substantial overlap with the *Silver Age of Russian literature, poetry and philosophy*. Florensky is certainly among the great philosophers-theologians of the twentieth century, with a very atypical trajectory of life. His

A. Papadopoulos (✉)
Institut de Recherche Mathématique Avancée (IRMA) and Centre de Recherche et d'Expérimentation sur l'Acte Artistique (ITI CREAA), Université de Strasbourg and CNRS, Strasbourg Cedex, France
e-mail: athanase.papadopoulos@math.unistra.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_147

work in philosophy is imbued with mathematical ideas. Talking about his life and works is also an opportunity to reflect upon the Russian mathematical school of the first third of the twentieth century, its philosophical foundations, and the conflicts it underwent. It is also an occasion for discussing poetry, literature, and art during the Russian Silver Age.

Keywords

Pavel Florensky · Mathematics and philosophy · Mathematics and theology · Mathematics and poetry · Contradiction · Antinomy · Mathematics and poetry · Mathematics in Dante · Mathematics of the Russian Silver Age · Nikolai Vasilievich Bugaev · Anna Akhmatova · Nikolai Luzin · Philosophy of mathematics · Philosophy of discontinuity · Mathematics in the twentieth century

AMS Classification

01A60 · 01A20 · 00A30 · 03A10 · 03A05

1 Introduction

Pavel Aleksandrovich Florensky (1882–1937) is a complex and singular figure in the landscape of twentieth century culture. His complete works contain more than 1000 items (see Florensky 2012, Note p. 7). In the last three decades, his books and manuscripts have been continuously edited and published in several languages and the literature on him has been growing at an exponential rate. Symposia on his case are regularly organized in Europe and elsewhere in the world, and a number of doctoral dissertations on his work have been written. Biographers and people who knew him dubbed him the Russian Leonardo,[1] and others called him the Russian Pascal,[2] sometimes the Gulag Pascal. Cédric Villani, in the preface to his book (Florensky 2016), describes him as "one of the most original thinkers of his times." Florensky is invariably presented as a mathematician, physicist, chemist, botanist, theologian, engineer, philosopher, art historian, specialist in electrodynamics, marine botany, linguistics, languages, symbol, esoterism, gnoseology, logic, esthetics, semiotics, and other fields of knowledge. Strictly speaking, I am not sure he may be called a mathematician, since I do not know of any mathematical discovery he made, and I cannot really judge about most of the other fields I mentioned. But without any doubt, his writings and thought encompass all these fields, and mathematics is always present there as a background. His highly unusual versatility led Nikolai

[1] In (Lossky 1952, p. 178), the philosopher Nikolai Onufriyevich Lossky (1870–1965) quotes an article by B. Filistinsky titled *A Russian Leonardo da Vinci in a concentration camp*, in which the latter says about Florensky: "A new Leonardo da Vinci was standing before us and we all were conscious of it."

[2] See Vassili Rozanov (Rozanov 1984, p. 413).

Lossky to describe him as possessing "inhuman erudition" (Lossky 1926). Perhaps this is where his strength and originality lie. When I started writing this chapter, my intention was to concentrate on Florensky's mathematical inclination, since the chapter is addressed primarily to mathematicians. But I soon realized that it is impossible to isolate the mathematical dimension of his thought from the rest of his works and vision of life. Florensky certainly had a mathematician's liking in his view of the world; it is this view, and some reflections that go with it, that I would like to outline.

Speaking of Florensky, I cannot avoid to begin by saying a few words about his life, and that is what I shall do in the next section.

2 Vita

Pavel Aleksandrovich Florensky was born on January 22 (Old Style: January 9), 1882, deep in the steppe, near the town of Yevlakh, today in Azerbaijan, to an Armenian mother, Salomé Saparian, from an aristocratic family, and a Russian father, Aleksander Ivanovich Florensky, who was working as an engineer in the construction of the Caucasus railroads. Azerbaijan was part of the Russian Empire, and Aleksander Florensky was in charge of the railway section that passed through Yevlakh. Pavel was born before the city's railroad station was built, and the family lived in freight cars "dressed in oriental carpets," he recalls.[3] Thus, it is possible that it is in a train carriage that Pavel was born. His childhood was spent in Tbilissi, then in Batumi,[4] another Georgian town, under the guidance of a father who had a "distaste for convention." Family life, Pavel says, "was like being on an island cut off from the world, that is, a desert island, because we didn't particularly like people and we always tried to keep ourselves out of the way. People would have taken hold of the purity, the unchanging calm and the rigorism of this insular paradise, which is why they were barely tolerated, and only occasionally." During these childhood years, Pavel was fascinated by the abundant nature of the Caucasus, of which he would later speak on several occasions. Especially, he recalled this landscape with particular emotion many years later, while he was a deportee in the Solovki Archipelago of the White Sea, the Gulag where he spent the last 4 years of his earthly existence. It is conceivable that it was because of the powerful influence of nature, especially that of the Caucasus, which fascinated him as a child, that Florensky remained deeply attracted, throughout his life, by the study of natural sciences and all the phenomena of the living.

[3] In this section, unless otherwise specified, the sentences in quotation marks are taken from Florensky's autobiographical notes addressed to his children, which he wrote between 1916 and 1926 (Florensky 2007). Here and elsewhere, I have translated from the French (I cannot do it from the Russian).

[4] Batumi, situated on the Black Sea, experienced its expansion with the construction of the Batumi-Tiflis-Baku Transcaucasus Railway, completed soon after Pavel Florensky's birth. Incidentally, Batumi is a town where Joseph Stalin lived, organizing strikes since 1901.

At school, Florensky was a gifted pupil, interested in all subjects, even if attending school every day bored him. Those who knew him at that age described him later on as a solitary child, extremely shy and very much an insomniac. In his recollections, we read: "I would fight long and hard not to go to bed, and once in bed, I would go on for hours without falling asleep, turning from side to side and studying the pattern of the wallpaper and blanket for a million times. These were hours of almost torture, when I was in bed unable to sleep." Pavel reads a lot, sets up a chemistry laboratory in his room, and works alone at night. He is interested in ancient languages. Later, he was able to speak most European languages, besides being proficient in Latin and ancient Greek, of course, but also in Hebrew and Syriac, which he learned at the Theological School. Several years later, from his exile in the Gulag, he wrote to his 15-year-old son Mikhal (Florensky 2021, Letter no. 77, p. 497, dated 23.X.1936):

> I wish I could have been with you to teach you how to work properly and accumulate knowledge. [. . .] When I was your age, every minute lost seemed to me like a misfortune or a crime, and I tried to fill my time as much as possible. I had diaries where I tried to record the essentials of my reading and reactions to books, notebooks of interesting quotations, albums of sketches from nature, copybooks where I methodically recorded the results of experiments, and others for onsite observations. Every evening, I would write myself a (motivated) note for the day's work. In this way, I acquired a wealth of knowledge, practical skills and, above all, the ability of making my own judgements based on what I had observed. I tried to compare and synthesize the data I collected in the form of tables, diagrams and curves, as this condensed aspect made them easier to understand; they become alive and acquire meaning. Immediately and automatically, you get an 'empirical generalization.'

In 1899, Florensky, who was 17, arrived to Moscow, where he was to enter university. At his arrival, he underwent a profound existential crisis. Having been brought up away from religion, in a positivist climate and an atmosphere of secular humanism, he realized the limitation of the natural sciences and became interested in social, religious, and moral issues, in the life of the ordinary and poor people, and in the human soul, striving for a meaning to life. The crisis is recounted in the last part, titled *Avalanche*, of the autobiographical text (Florensky 2007). We read there: "On top of my own magical idealistic understanding of the world, a film had formed that prevented me from breathing freely; the idea of continuity, fundamentally excluding the miraculous, was trying to take possession of my intelligence." He was already seeing the world with the eye of a mature mathematician. The ideas of continuity and discontinuity, which for him had become philosophical concepts, were to follow him for the rest of his life. Florensky's existential and religious crisis was often compared to that of Pascal, who, at age 29, experienced a similar experience which led him to a life centered on philosophical and theological reflection, without leaving science. Pascal had, like Florensky, a strange personality. T. S. Eliot described him as "a man of the world among ascetics, and an ascetic among men of the world" (Eliot 1951, p. 411).

Under pressure from his parents, Florensky enrolled as a mathematics student at the Faculty of Physics and Mathematics of the Imperial University of Moscow. His

mentor there was Nikolai Vasilievich Bugaev (1837–1903), a well-known mathematician who was the president of the Moscow Mathematical Society, of which he had been a cofounder. Bugaev was engaged in the philosophy of mathematics and its repercussions in the social, historical, and religious spheres. For him, questions like continuity, discontinuity, the finite, and the infinite were intimately linked to philosophical issues such as free will, scientific and social revolutions, evolution in history, etc. I shall talk later in this chapter about Bugaev, whose ideas were to have a definite influence on Florensky. The other teachers of Florensky included Boleslav Kornelievich Mlodzievsky (1858–1923), a representative of the Moscow school of metric geometry who, like Bugaev and several other Russian mathematicians of that period, spent some time in Western Europe on several occasions, in cities like Göttingen, Paris, and Zürich, where he came into direct contact with Western mathematicians and philosophers. According to I. H. Anellis in (Anellis 1994), Mlodzievsky gave the first course on set theory and the theory of functions at Moscow University, two topics which became important parts of Florensky's background. Anellis adds that it was Mlodzievsky who established the Russian terminology in these subjects. According to Demidov (Demidov 2014), Mlodzievsky's synopsis of his lectures was preserved thanks to notes compiled by Florensky (Medvedev 1986).

Florensky's teachers also included Leonid Kuz'mich Lakhtin (1863–1927), a student of Bugaev, who studied nonorientable surfaces, in particular the question of representing them by explicit equations, see (Lakhtin 1904). We shall see below how nonorientable surfaces became part of Florensky's later interests, mathematically and symbolically. A nonorientable surface represents the fusion of two contraries. In a note in his booklet titled *The imaginaries in Geometry*, which we mentioned in the introduction, which we shall discuss at length below and in which the notion of nonorientability plays a key role, Florensky says that he wrote that booklet during the month of August 1902 while he was a student, and that he showed it to Lakhtin (see Florensky 2016, p. 79). Among the Russian mathematicians of the same epoch who studied nonorientable surfaces, one should mention Nikolai Borisovich Delaunay (1856–1931) who addressed the problem of finding a locally Euclidean parametrization of the Möbius band, in his memoir (Delaunay 1898) published in 1898 in the Bulletin of the French Mathematical Society, and which Florensky also mentions in a note to his work *Imaginaries in geometry* (see Florensky 2016, p. 82). In the same note, Florensky declares that in the *Encyclopedia of mathematics* of which he was in charge, he wrote an article on the division of surfaces in orientable and nonorientable.

Below, I shall discuss the booklet *Imaginaries in Geometry*, which was essentially written in 1902, at the time Florensky was a mathematics student.

The young Florensky became quickly interested in formal logic. Questions of contradiction, antinomy, and the nature of the relationship between mathematics and the physical world found an echo in him, both philosophically and theologically. He writes (Florensky 2007, p. 42):

> In mathematics, Fourier series and other topics which represent any complex rhythm as an infinite set of simple objects resonate with me, speak to me, correspond to my innermost

> feelings. Continuous functions without derivatives and discontinuous functions where everything unravels, where all elements are *standing*, hold a dear language for me. If I listen to the depths of my being, I discover in the rhythm of my inner life, in the sounds that fill my consciousness, the rhythm of the waves that have been retained forever, inscribed in my memory for all time, and I know that they seek their conscious expression in me through the schema of these mathematical concepts.

Leonid Sabaneeff,[5] who knew Florensky since they were both mathematics undergraduates at the same university, writes in a set of recollections (Sabaneeff 1961) that Florensky rarely attended lectures, preferring to read books at home. He also recalls that Florensky seemed oddly old for his age, to the point that one of their fellow students said about him: "he looks as if he had already lived a thousand lives." Sabaneeff continues: "One had an impression of genius, of unusual depth and power of thought, and at the same time there was something of black magic, dark, devoid of divine grace." He had the feeling that Florensky possessed an immense spiritual experience and was endowed with hypnotic power. He recalls Florensky speaking of "many facets of any true thought and of the compatibility of contradictions at the deepest levels," asserting that "every perceived law generates its own negation" and that "logic is valid for the earthly life, for the lower levels; but the true world is one where contradictions are compatible—a realm of antinomies." Sabaneeff adds that Florensky "obviously regarded antinomy as the basic law of the universe, encompassing all others." In Florensky's mind, the notions of contradiction and antinomy have no negative connotation. Contradiction, nonprovability, and paradoxes in set theory fascinated him, and he often referred to them in his mathematical and nonmathematical writings.

At university, Florensky, besides studying mathematics, attended lectures at the philosophy department, by Lev Mikhailovich Lopatin (1855–1920), a psychologist and philosopher inspired by Leibniz' monadology, and by Sergei Nikolaevich Trubetskoy (1862–1905), who belonged to the Russian tradition of universal theism, trying to unite Christianity with Platonism. Both philosophers were imbued with spiritual ideas, in search of universal wisdom,[6] and their ideas had a definite impact on Florensky.

Bugaev died in 1903. Dmitri Fedorovich Egorov (1869–1931), who had been a student of Bugaev, took over the direction of the mathematical school founded by his

[5] Leonid Sabaneeff (1881–1968) graduated in mathematics from the Imperial University of Moscow and studied music under Rimsky-Korsakov and Taneyev. He became a composer and musicologist. He wrote an authoritative and influential book on Skriabin. He taught at the State University of Moscow and he founded the Moscow Institute of Musicology. He remained close to Florensky and to Nikolai Nikolaevich Luzin (1883–1950), whom I discuss in the sequel, until the year 1926, when he left the Soviet Union. He lived in Paris, London, the United States and Nice, where he is buried.

[6] Florensky, in the notes of his major theological work, *The pillar and ground of the truth (*Florensky 1975*)*, quotes several works by Lopatin, among which are *The positive tasks of philosophy*, *The philosophical worldview of N. V. Bugaev*, and *The present and future of philosophy*. He also quotes books by Trubetskoy, among which are *The metaphysics of ancient Greece*, *Foundations of idealism,* and *The doctrine of Logos*.

mentor. At that time, Florensky had already become friends with Bugaev's son, Boris Nikolaevich Bugaev (1880–1934), 2 years his senior, poet and writer, known by the pseudonym Andrei Belyi, who was a leader and theorist of the Russian Symbolist movement. Belyi's novel *Petersburg* (first edition completed in 1913) made him later one of the greatest writers of the twentieth century. Like Florensky, Belyi was interested in the impact of mathematics in life sciences and philosophy.

During his study years at the Mathematics and Physics Department of the University of Moscow, Florensky initiated a "Circle of mathematics students" within the Russian Mathematical Society, with the collaboration of his friend and fellow student Nikolai Luzin, who was to become later one of the best-known Russian mathematicians of his generation, respected in Russia and France. On this occasion, he wrote a *Sketch for a discourse for the inauguration of the Circle of mathematics students of the Moscow Mathematical Society* which was published in 1990 (Florensky 1989), discussing there, among other topics, Cantor's set theory and Bugeav's arithmology (the theory of discontinuous functions). At about the same time, Florensky founded a "Religious-Philosophical Society of Writers and Symbolists" and a "Union of Christian Struggle," a movement dedicated to social Christianity, opposing the then current subjugation of the Church to the State, which had started under Peter the Great.

I will say more about Luzin in Sect. 7.

In 1904, Florensky completed his mathematics studies with a dissertation titled *On singularities of plane curves as loci violating their continuity.* Part of this dissertation is published in (Demidov and Parshin 1986). According to Oppo (Oppo 2022, p. 72), this dissertation was part of a wider doctoral project titled *The Idea of discontinuity as part of a postulated worldview.* Florensky wrote an introduction to this work and published it under the title *On a certain premise of a worldview* (Florensky 1904b) in the symbolic journal *Vesy.* Oppo gives further details on this publication in a note (Oppo 2022, p. 87).

Florensky's dissertation was well received by his teachers. Egorov considered him as one of the most brilliant and original mathematics students. One of Florensky's teachers, Nikolai Yegorovich Zhukovsky,[7] offered him the possibility of pursuing a specialization curriculum which would lead him to a doctorate. But to the surprise of his teachers, Florensky turned down this opportunity and enrolled as a student at the Moscow Theological Academy. He delegated the function of secretary of the Students' Mathematical Circle to Luzin (Demidov 2014). The Theological

[7] Nikolai Yegorovich Zhukovsky (1847–1921) was a mathematician, physicist, and engineer. He is considered to be the founder of the Russian schools of hydromechanics and aeromechanics. In 1904, he established, near Moscow, the first aerodynamical research institute in Europe. He is often called the "father of Russian aviation"; see (Grigorian 1970). In the theory of conformal mappings, the transformation $z \to \frac{1}{z}$ is called the *Zhukovsky transformation.* Zhukovsky's work had a definite impact on Mikhail Alekseyevich Lavrentieff (1900–1980), the Russian founder of the theory of quasiconformal mappings whom we shall mention later in this chapter. Lavrentieff was a specialist of partial differential equations and hydrodynamics and the founder of the Institute of Hydrodynamics of the Siberian Division of the Russian Academy of Sciences.

Academy is situated in Sergiyev Posad, a city about 70 km northeast of Moscow which also hosts the Trinity-Saint Sergius Lavra.

Even though Florensky distanced himself from the profession of mathematician, mathematics remained for him, for the rest of his life, a way of thinking and a basis for his philosophical and religious work.

Florensky's strained relations with the state police started in 1906. That year, he was sentenced to 3 months' jail for having called for the abolition of the death penalty[8] and for having protested, along with some fellow students, against the execution of Lieutenant Piotr Petrovich Schmidt, a naval officer who had been one of the leaders of the Black Sea Fleet revolt in Sebastopol during the 1905 revolution.[9] Florensky spent a week in the Tagansky prison in Moscow (March 23–30, 1906), before being released after the intercession of several people, including the rector of the Theological Academy. During the week he spent in prison, he wrote a paper titled *On the elements of the α number system* (see Demidov and Ford 2005, p. 601).

A few words on the 1905 social movements are in order, because these movements had a definite impact on several of the figures we will be talking about.

The origin of the 1905 social movements goes back several years. In 1881, the year before Pavel Florensky was born, Tsar Aleksander II, who was ready to give his country a constitution, was assassinated by members of the *Narodnaya Volya* ("People's Will"), a late nineteenth-century revolutionary socialist political organization. With this assassination, Russia started being in the grip of a political effervescence and a state of siege that culminated in the bloody events of 1905 that left their mark on Florensky, as they did on a large number of Russians. The 1905 revolt started with a peaceful demonstration by workers who had a request for the Tsar, and it ended in a terrible shoot-out. What happened during these events was later called the First Russian Revolution and was regarded as the beginning of Imperial Russia's agony.

In a sense, the 1905 revolution was a culmination of the idealistic traditions of the Russian intelligentsia. Chapter 2 of Pasternak's *Doctor Zhivago*, titled *A Girl from a Different World*, starts with the words "The war with Japan was not yet over when it was unexpectedly overshadowed by other events. Waves of revolution swept across Russia, each greater and more extraordinary than the previous."[10] In relation with the events we are recalling, it is also interesting to know that, in 1925–1927, Pasternak wrote two long poems dedicated to the 1905 revolution, one of them titled *Lieutenant Schmidt*.

[8] Florensky made his statements against the death penalty in a homily, first delivered at the Lavra and then published clandestinely. Demidov and Ford note in (Demidov and Ford 2005) that "the Moscow Theological Academy had the practice of allowing seminarians to preach in the Cathedral."

[9] Lieutenant Piotr Petrovich Schmidt, an idealist sympathizer of the 1905 revolution, was a typical representative of the Russian intellectuals during this episode. He founded a *Union of People's Friends Officers* and addressed several revolutionary meetings that culminated in a general insurrection of the Black Sea fleet crews on November 11, 1905.

[10] Modified English translation from (Pasternak 1958).

Florensky obtained his theology diploma in the spring of 1908, with a dissertation titled *On spiritual truth*, a work that was expanded later to a treatise titled *The pillar and ground of the truth: An essay in orthodox theodicy in twelve letters* (Florensky 1975), which became his first doctoral dissertation, defended in 1914. This work had a definite impact on the Russian theological circles.

Upon his graduation in 1908, Florensky was entrusted with the teaching of history of philosophy at the Moscow Theological Academy, a position he held until the Academy's closure in September 1918. Figure 1 is extracted from the album of illustrations of his course on the history of classical Greek philosophy.

Florensky was ordained priest in 1911, after his marriage. Just as philosophy accompanied him during his mathematical studies, mathematics constituted an important element of his background during his theological studies. The discussions around the mathematical notion of discontinuity that were taking place at that time, the recent developments of set theory, the language of mathematical logic, and other mathematical subjects were at the basis of his treatise *The pillar and ground of the truth*. Since the book was addressed to theologians rather than mathematicians, Florensky included there several sections containing technical mathematical explanations. For instance, Section XV is titled *Certain concepts from the theory of infinite*, Section XVII is titled *Irrationalities in mathematics and in dogma*, Section XIX is titled *The concept of identity in mathematical logic*, Section XXX is titled *The basic symbols and elementary formulas of symbolic logic*, etc. I shall say more about this work below.

In 1917, the teaching of theology in Russia entered a period of upheaval as a result of the October revolution, and the Academy closed for good in September 1918.[11] Florensky could no longer teach philosophy there. He became a teacher of history of Byzantine art at the MIKhIM.[12] In October 1919, he was appointed scientific secretary of the committee for the Preservation of Art Monuments and Antiquities of the Sergiyev Posad Lavra, which belonged then to the Commissariat of People Instruction. He prepared a treatise on religious art, titled the *Reverse perspective* (Florensky 1992), which was the occasion for him to expand his ideas on space in drawing and painting. I will say more about this work in Sect. 3 below. This was also the epoch where Florensky became close to the writer Vassili Rozanov. The latter wrote about Florensky in a letter to the art critic Mikhail Spasovsky,

[11] The Moscow Theological Academy reopened in 1944. In 1930, the name of Sergiyev Posad, in reference to Saint Sergius of Radonezh, considered as the heart of Russia – it was from there that Russian culture expanded – was changed to Zagorsk, named after Vladimir Mikhailovich Zagorsky, a revolutionary who had been assassinated and whose real name was Wolf Mikhelevich Lubotsky. Zagorsky was a party activist who had taken part in the 1905 revolution. He was Secretary of the Moscow Committee of the Russian Communist Party (Bolsheviks) when he was killed, on September 25, 1919, by a bomb thrown by an anarchist. It was only in 1991 that the town of Zagorsk recovered its original name of Sergiyev Posad.

[12] This is the acronym of Moskovskii Institut Istoriko-Khudozhestvennykh Izyskanii i Muzeevedeniia, that is, the Moscow Institute of Historical and Artistic Research and Museum Studies.

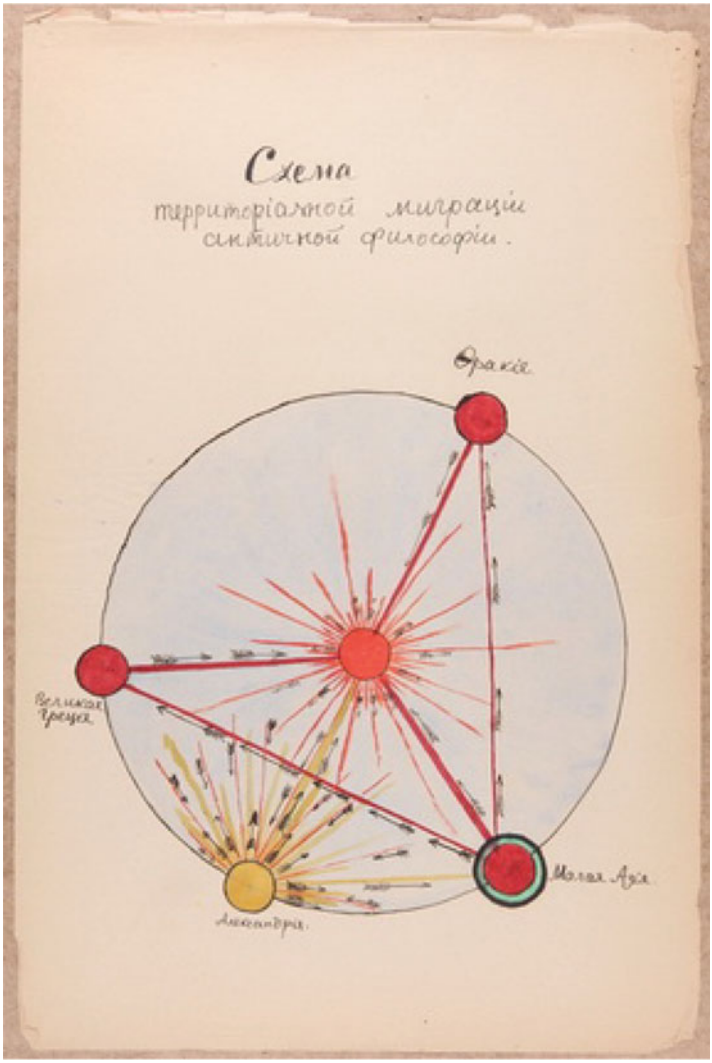


Fig. 1 From the album of illustrations of P.A. Florensky's course on history of classical Greek philosophy 1908–1909. The translation of the page title is "The scheme of territorial migration of ancient philosophy." Around the circle, starting from the top, clockwise, one reads: Thrace, Asia Minor, Alexandria, and Great Greece. Paper, watercolor, pencil, and ink. Multimedia Art Museum, Moscow

quoted in part in (Rozanov 1984, p. 413) and in the introduction of the French version of the *Imaginaries* (Florensky 1992, p. 15):

> He is the Pascal of our time, the Pascal of our Russia. You know, sometimes I even get the impression that he is a *saint*, so exceptional is his mind, so out of the ordinary. . . . I think and I am firmly convinced that he is even infinitely greater than Pascal, both in his essence—on

> the level of a Plato with all that is so *exceptional* about him—and in his intellectual discoveries, intellectual visions.

By that time, Florensky had already produced a large amount of writings. He signed a contract with the Pomor'ye publishing house to have his works appear in 19 volumes. Eventually, only his book *Imaginaries in geometry* (Florensky 2016), which I shall review in Sects. 3 and 4 below, was published by Pomor'ye before its closing.

Pavel Florensky, with his family, lived in a small wooden house near the Sergiyev Posad monastery. For the rest of his life, including the years he spent in prison or in a Gulag, he used his scientific skills to work as a specialized technician in a plastics factory, as a teacher and encyclopedia editor, and as an engineer in several research institutes. He also taught mathematics and physics at the Zagorsk Pedagogical Institute. With all the troubles he had to deal with, work was what allowed him to survive psychologically, besides supporting his family. At some point, he was called as an engineer on an important project to electrify Russia, which was known under the name GOELRO.[13] There, he had the occasion to meet Stalin, who considered him, like many intellectuals and clerics, an enemy of the people, but who tolerated him because the state needed his skills. Trotsky, at some point, took an interest in Florensky and to some degree protected him, before he was driven out himself. The following episode is reported in the book by V. Chentalinsky who collected KGB documents concerning literary figures when the KGB archives were made available, in the 1990s (Chentalinsky 1993, p. 174):

> One day, long before his disgrace, Leon Trotsky[14] was astonished to see Florensky's white cassock.
>
> "Who's that?" he asked.
>
> The leader of the world revolution, who thought he was as wise as King Solomon, held a multitude of positions. He was also head of the *Glavelectro*,[15] and as such visited the institute where the priest worked.
>
> "This is Professor Florensky", came the reply.
>
> – "Ah yes, I know..." Trotsky approached him and invited him to an engineering conference.
>
> "Naturally, he said, you wouldn't come dressed like that...
>
> – I haven't renounced my vows and can't wear any other clothes," Florensky replied.
>
> "– Oh, you can't, . . . so come in this outfit"
>
> When Florensky stepped up to the podium to address the conference, a murmur of astonishment ran through the hall: a priest at the pulpit! And, more than by his brilliant communication, the audience was struck by the enigma he represented: a cleric, and therefore an obscurantist, who possessed such knowledge of the exact sciences!

[13] GOELRO is the transliteration of the Russian abbreviation for "State Commission for the Electrification of Russia."

[14] Trotsky was banished in 1929.

[15] This is the Central Directorate of Electricity.

Trotsky's indulgence for Florensky would later prove fatal, as we shall recall below.

Florensky, besides carrying out several assignments for the state, was teaching drawing and history of art, and most of all, he was writing and writing. At some point, he was offered to leave the country on one of the "philosophers' boats"[16] that had already taken a large part of the Russian Intelligentsia.[17] He refused, declaring that he wanted to stay in his country. Of this refusal, Sergei Bulgakov, who was his disciple and friend and who had been expelled, in 1922, said in an address he gave in Paris in 1943, when the news of Florensky's death reached him (Bulgakov 1989): "He had not the shadow of naivety and idyllic primitiveness, [. . .] but he loved this native land, like the Mother of all men, the Demeter of the Ancients." In staying in his country, Florensky knew what he was expecting. He was not put straight into prison, and he was always given new tasks, even when he was in the Gulag.

At dawn on May 21, 1928, Florensky was arrested and taken to the Lubyanka.[18] The episode is recounted in the book (Chentalinsky 1993). He was required there to fill in a questionnaire. To the question of whether he had ever been prosecuted, he replied: "Yes, in 1906, for a sermon against Lieutenant Schmidt's death sentence." This could have played in his favor, but at the same time he added: "I have practically nothing to say on political questions. Because of my character, my occupations and my conviction, based on the lessons of history, that events still develop according to as yet unclarified laws of social dynamics and not according to the will of the participants, I have always avoided the political struggle." In those times, it was dangerous to declare oneself politically neutral. Thus, Florensky was arrested along with 79 other people, accused of a plot in connection with the shooting of a revolutionary. The minutes of his arrest at that date and of later arrests are published in the book (Chentalinsky 1993).

Florensky was placed under house arrest in Nizhny Novgorod. The photo reproduced in Fig. 2 was taken there. A few months later, he was granted early

[16] The "philosophers' boats" is the name that was given to the steamships that transported intellectuals expelled by the Soviet government in 1922, soon before the creation of the Union of Soviet Socialist Republics. Leon Trotsky had decided that "there was no pretext for shooting these people, but it was no longer possible to tolerate them," cf. (Nikolski 2012).

[17] Among those who were expelled and with whom Florensky was closely associated, we mention the symbolist poet Vyacheslav Ivanov (1866–1949), the religious philosopher Nikolai Berdyaev (1874–1948), and Sergei Bulgakov (1871–1944), who was a theologian, philosopher, and economist. We also mention the philosopher Simon Frank (1877–1950), the theologian Vladimir Nikolaievich Lossky (1903–1958), the writer and journalist Mikhail Andreyevich Osorgin (1878–1942), the sociologist Pitirim Alexandrovich Sorokin (1889–1968), and Valentin Fyodorovich Bulgakov (1886–1966), a Christian anarchist and antimilitarist who had been the last secretary and biographer of Leo Tolstoy. A large number of the persons expelled participated in the 1905 revolution and belonged to the Socialist Revolutionary Party.

[18] This is the building in Moscow that hosts the headquarters of the Political Police, called State Political Directorate (known by the acronym GPU), with its affiliated prison.

Fig. 2 Pavel Florensky in exile in Nizhny Novgorod, 1928

release, after the intervention of influential people, among them Ekaterina Peshkova, Maxim Gorky's ex-wife. Peshkova was at that time director of the political Red Cross and was on good terms with Stalin. Florensky returned to Moscow and resumed his work for the state. In 1930, he was made deputy director of the Electrical Engineering Institute K. A. Krug in Moscow and in 1931 member of the central directorate for the study of electro-insulating material. Meanwhile, he continued to teach philosophy clandestinely at the Danilovski and Petrovski monasteries, where the Theological Academy continued to exist informally, after it was officially closed in 1918. Florensky continued to publish in scientific journals until his deportation. In 1932, he published several articles in the journal *Socialist Reconstruction and Science*, among them an article in which he describes an analogical calculator for finding approximate solutions of algebraic equations of high degree [see (Demidov and Ford 2005), p. 608 and the reference there].

Between 1927 and 1933 (the year of his deportation), Florensky filed almost 50 patent applications, but most of them were rejected because he was a priest, automatically accused of obscurantism.

The existence of a cleric, always dressed in cassock, who was a scientist and useful to the nation, became more and more unbearable for the Soviet authorities, and at some point they ran out of patience. At dawn on February 26, 1933, Florensky was arrested at his home and beaten up at the Lubyanka, again accused of plotting and antirevolutionary propaganda. V. Chentalinsky, based on the KGB documents, writes about his arrest (Chentalinsky 1993, p. 164): "Today we can see Florensky's true face, but his case file presents him as a dangerous criminal for society, as an obscurantist placed under permanent and strict surveillance until his death."

This time, Florensky spent 6 months at the Lubyanka. During his captivity, he worked around the clock. On July 26 of the same year, after a sham trial based on false accusations, he was sentenced to 10 years in the Gulag, with the obligation to continue his scientific work.[19] He was again offered the possibility to leave Russia. In fact, Tom Masaryk, the president of the Czechoslovak Republic, offered to take him in, so that he could continue his scientific work there. But once again, Florensky refused to leave his country. In August of the same year, he was transported to Eastern Siberia on a train full of common criminals. He arrived at his destination, a camp whose name translates as "free," exhausted, starving, and penniless: Everything he possessed was stolen during the journey. In February 1934, that is, a few months after his arrival to the camp, he learned from a courier that his library had been confiscated during a search at his family home in Zagorsk. Shocked, he wrote to the OGPU:[20]

> My whole life has been devoted to science and philosophy, and I have never known rest, leisure or pleasure. Not only did I devote all my time and energy to the service of humanity, but I also spent a large part of my modest salary on buying books, etc. My library was not simply a collection of books, but selections for specific, established subjects. One might say that many of the works I intended to write were already half-finished in the form of notes in books whose arrangement on the shelves, the intellectual key, was known to me alone [. . .] My whole life's work is now destroyed [. . .] This is for me far worse than physical death. (Misler 2006)

In November 1934, after seeing for the last time his family who was visiting him, Florensky was forbidden visitation rights and sent to the Solovki Islands, an austere place in the White Sea, 160 km from the Arctic Circle. There lies an ancient monastery, known under the name *Solovetsky*, surrounded by an enclosure of walls and towers and encircled by immense frozen expanses. In the days of the Tsars, it was the place where recalcitrant monks were sent. In 1920, the monastery was transformed into a Gulag, and it remained so until 1937, the year Florensky was shot. Prisoners there were crammed into the former monks' cells, several to a cell. Later, the monastery came to be regarded as the most terrible of the Soviet camps,

[19] It is probably time to recall that the word Gulag is an acronym for "Glavnoye Upravleniye Lagerey," which means Main Administration of Camps. In theory, the Gulags were "camps for re-education through work."

[20] The OGPU is an acronym for "Joint State Political Directorate," which was the Soviet Union's state police.

where half a million people perished. Aleksandr Solzhenitsyn called this place "the mother of Gulags."

About Florensky, Solzhenitsyn wrote that he was "perhaps the most remarkable person devoured by the Gulag." Chapter 2 of the second volume of his *Gulag Archipelago* opens with a historical account of the Solovki Islands, where the story of the Gulag starts (Solzhenitsyn 1975):

> On the White Sea, where the nights are white for half a year at a time, Bolshoi Solovetsky Island lifts its white churches from the water within the ring of its bouldered kremlin walls,[21] rusty-red from the lichens which have struck root there—and the grayish-white Solovetsky seagulls hover continually over the kremlin and screech.

On October 10, 1934, Florensky sent his first letter to his family from his new exile. He was still in Kem, 200 km as the crow flies from the Solovki Islands. He wrote (Florensky 2021, Letter No. 1, p. 13):

> On arrival at the camp, I was robbed by an armed attack and threatened with three axes, but as you can see, I am safe and sound, albeit without luggage or money; some of my belongings have been found. All this time I have been cold and hungry. It was much harder than I had imagined when I left the Skovorodino train station.[22] I was supposed to go to Solovki, which in itself would not be so bad, but I am held up in Kem where I am busy making and filling out control forms.

In the Gulag of the Solovki Islands, a new life for Florensky started, although he spent his time on exactly the same activity as before: working. I have reproduced here a recent photograph of the Solovetsky monastery; see Fig. 3.

In a letter dated 8–9.VII.1936, addressed to his wife (Florensky 2021, Letter No. 67, p. 435), he writes: "I never see the sun or nature, even through the window: I am constantly busy. All my time is spent developing and deepening technological processes, equipping the algae workshop as much as possible, given the increasing demands for quality and quantity." Besides, Florensky also set up a theater for the detainee and taught classes there, keen to passing on his knowledge to the other prisoners. His work force sustained and guided him through all his hardship, as the enormous correspondence with his family shows. The prison authorities, aware of Florensky's talents and keen to exploit them to the full, arranged workshops for him and let him organize his courses and lectures. In a letter addressed to his son Vassili, dated 27–28.IV.1935, he writes (Florensky 2021, Letter no. 17, p. 111): "I just gave a talk on Maxwell's theory and related questions. I got home at 11:30 p.m. and it was perfectly clear outside. [. . .] The light here is transparent: everything seems to be seen in a dream, and in fact non-existent. This is particularly true at night, but also during the day, when the sun is shining." He would have liked to do mathematics, but

[21] The Russian word "kreml" (kremlin) refers to a fortified complex or citadel, and not just the one in Moscow.

[22] Skovorodino is a city situated on the Trans-Siberian Railway, 7306 kilometers from Moscow.

Fig. 3 The Solovetsky Monastery, where the prisoners were crammed together. This recent and romantic photograph does not give any idea of the atmosphere in the monastery during the time it was used as a Gulag, between 1926 and 1939, but the sky above it and the sea surrounding it have not changed. (Photo: Cultural heritage object in Russia, No. 2910179018, GitaOggg)

this was not easy. In a letter to his mother, dated 24.VI.1936, Florensky writes (Florensky 2021, Letter no. 65, p. 422):

> I deal a bit with mathematical questions; for more serious work in this field, I have neither time, nor the necessary books, and finding everything in one's head is too unproductive a task. With the most advanced workers at Idoprom, where I work, i.e., with those who have completed eight years of school or technical institute, or a few years of higher education, I have organized a physics circle with the aim of learning how to reactivate one's knowledge of physics, mathematics and a little chemistry, really assimilate them, deepen them and complete them. So, the sum of the various tasks and activities occupies all the time, not only during the day, but also most of the night. I try, after dinner, from 7–7.30, to get about two hours of sleep, so that I can work with a fresh mind at night.

He accepted his new situation philosophically. Perhaps with a touch of humor, but also with great sadness, he wrote to his wife on 7.XI.1934 (Florensky 2021, p. 17):

> I have long since come to the conclusion that desires in our lives are realized, but with great delay, in a caricatured way that makes them unrecognizable. In recent years, I wanted to live in the immediate vicinity of a laboratory, and that came true; in Skovorodino. I wanted to take care of the soil, and that came true here too. Before that, I wanted to live in a monastery,

> and that came true, but in Solovki. As a child, I used to fantasize about living on an island, seeing the tides, handling seaweed, and now here I am on an island, there are tides and I may soon be handling seaweed. But the realization of desires is such that one no longer recognizes them, and it happens when the desire has already passed. I send you a big kiss, my dear Annoulia, don't lose courage.

In fact, early in 1920, he had written in his autobiography about this call of the sea (Florensky 2007, p. 41):

> The call of the sea, the fractional sound of the surf, the infinite, glittering surface in which I could distinguish ever-smaller flakes, down to the tiniest particles, but which never merged, all this resonated endlessly in my soul. My body was craving the saltiness of the sea, the salty, iodine-soaked air, but also the fractionated air carrying the tiniest crystals: sometimes it feels good to hold at least one vial of tincture of iodine close to you. I crave the taste of the sea. . . It seems to me that if I could find a pile of seaweed, I would devour them all.

Florensky set up a laboratory to process seaweed and extract iodine and agar-agar, a substance used as a glue, and also in cooking. He conducted research in biology and organic chemistry. During his stay in Solovetsky, he registered seven patents. Among the prisoners were engineers, scientists, and academics, educated people as well as workers; he gave courses to all of them, working at an inhuman pace. Since he worked more than double the required time, he was allowed to send three letters a month to his family. He sent long letters, sometimes a dozen of pages long, to his wife, mother and, children, and also, from time to time, to friends such as the pianist Maria Yudina, who was Stalin's favorite pianist and who had become Florensky's spiritual daughter. He also corresponded with Vladimir Ivanovich Vernadsky, a member of the Russian and Soviet Academies of Sciences.[23] In his letters to his family, he used to describe his everyday life in the Gulag, giving advice to each family member.

In some of his letters, Florensky gave mathematics and physics lessons to his children. For instance, in the letter dated 5.XII.1935, to his son Vassili (Florensky 2021, Letter no. 40, p. 245), he formulated a problem in mathematical physics and he developed its solution. In the letter dated 24.III.1936 (Florensky 2021, Letter no. 54, p. 341), he explained a mathematical problem to his daughter Olga. Several times, in his letters, he wanted to know about his confiscated books. On September 16, 1935, he wrote to his wife: "On occasion, enquire about the fate of my work on permagel and my mathematical articles, and in particular the articles on four-color maps." The works were never found. On February 25, 1936, he wrote, again to his wife (Florensky 2021, Letter No. 50, p. 311):

[23] Vladimir Ivanovich Vernadsky (1863–1945) was a famous mineralogist and geochemist. He is the author of a book titled *The Biosphere*, published in 1936, in which he defends the ideas that life and human cognition are the forces that are at the origin of the Earth's development and the transformation of the biosphere. Vernadsky was one of the fathers of the Soviet atomic bomb. He never sought to hide his friendship and respect for Pavel Florensky.

> Now that I live in the laboratory, I can also work at night, especially as the work is done 24 hours a day and needs to be supervised all the time. At the moment, it is already nearly 3 o'clock. Next to me, a worker is cutting the agar into pieces for freezing, and on the other side, pumped water is gurgling in a large tub. [. . .] Try not to get bored and to find interest in work in addition to that which is immediately essential. Life must be embellished.

He tried to keep abreast of developments in all the sciences, and every small information he could get was important for him. In the same letter, he wrote:

> In the Pravda of February 9, 1936, No. 39 (6645) p. 5, I came across the composition of a liquid for toothache, that of Professor Hartman, an American from Columbia University. As this liquid is very easy to obtain and, according to the newspapers, gives 80% positive results, you need to have some on hand. Here is the formula: one part etyl alcohol, 2 parts ether, 1.25 parts timol.

He is worried again about his books: "Have you received my books and materials from P. N.? Ask him to send them to you, because I have spent a whole year putting them together, and I want the boys to use them: I am already giving them nothing, so do not neglect the little they can get." He is thinking not only about his past, but also about the past of the universe. In one of his last letters, dated 20.22.1937, (Florensky 2021, letter No. 71, p. 455), Florensky writes, to his son Kyrill: "A distant past absorbs my thoughts: the formation of the island's rocks some two billion years ago, the history of the formation of the Earth's crust in these very special places, the appearance of new rocks based on electronic waste, the movements of the glacier, its advance and retreat." With his meager salary, he is also sustaining his family. In the same letter, he writes to his son: "Did you receive the small sum I sent you?"

V. Chentalinsky describes the last days of Father Florensky. We are in June or July 1937. The camp had been recently reorganized under a new management which had recently arrived. The iodine production plant, where Florensky was the leading scientist, was closed. The detention conditions were tightened. Prisoners were herded off to an unknown destination. Florensky was arrested after a report, made by an informer. Chentalinsky records this denunciation in the following terms (Chentalinsky 1993, p. 191):

> Florensky and another inmate walked to the library, talking loudly and gesturing forcefully. The informer "Comrade" follows them closely and listens attentively. They are talking about the upcoming war:
>
> "The predictions of Trotsky, the eminent Party strategist and ideologist, according to which war will soon break out seem justified to me, says Florensky. That is the rule: war breaks out periodically every fifteen or twenty years . . . "
>
> This short, seemingly innocent report had unfortunate consequences. It was used to draw up a special attestation on Florensky's conduct:
>
> "Engaged in counter-revolutionary propaganda in the camp, praising the enemy of the people Trotsky." [. . .]

Very soon after this, Florensky witnessed the destruction of his research facilities and learned that his discoveries on algae (in Solovki) and permagel (in Siberia) have been attributed to others. References to him in the *Technical Encyclopedia* of which he was an editor were canceled in the second edition [see (Demidov and Ford 2005), p. 608]. Florensky was again sentenced, this time to death, by a three-judge emergency court. He was immediately transferred to the mainland, in the Leningrad region, with a large group of prisoners from Solovki. The execution act was signed, and on December 8 of that year, he died riddled with bullets, along with 508 other people, on the outskirts of Leningrad. The execution act is reproduced in Chentalinsky's book (Chentalinsky 1993, p. 194): "The ruling of the troika of the Leningrad region NKVD leadership concerning the death row inmate, Pavel Aleksandrovich Florensky was executed on December 8, 1937, as attested by the present deed."

Florensky's body, along with those of the other shot, was dumped in a mass grave, which has since become a place of memorial. We were in the middle of the period known as the "Great Stalin Terror" (1937–1938), during which 750,000 people were executed in the USSR.

Meanwhile, life went on at the mathematics departments of Leningrad university, a few kilometers from that place. At the beginning of his historical review (Zalgaller 2012), V. A. Zalgaller writes:

> In 1937, enrolment in the Mathematics and Mechanics Faculty of Leningrad University (MatMekh) was unusually high: 210 students. The "big Chemistry lecture hall" in the main building on the Neva embankment hardly contained us. At the beginning, analytic geometry was taught by Professor A. R. Kulisher (1878–1937). But in September, he was already arrested and replaced by a 25 year old Professor Aleksandr Danilovich Aleksandrov. [. . .]

On May 5, 1958, the Moscow City Court overturned Florensky's conviction of July 16, 1933.

The rehabilitation court order reads: (Chentalinsky 1993, p. 184):

> Florensky's arrest (like that of the others judged for the same case) is not based on any document in the case file. Witnesses were not questioned, and those who took part in the investigation of the case were subsequently convicted of forgery. Florensky (and the others) were unjustly convicted, in the absence of proof of their guilt.

On March 5, 1959, the Arkhangelsk Oblast Court overturned his conviction of Nov. 15, 1937. It was only in 1989 that Florensky's family and descendants learned, through a letter from the KGB, of the date and place of his execution.

In this section, I tried to draw, with words, some features of Florensky's face. For an extended biography, the reader can refer to (Pyman 2010). A complete biography should take several volumes and is still to be written. There are a number of portraits of him. I have reproduced three of them here, in Figs. 2 and 4.

Fig. 4 Pavel Florensky, shortly after his arrest in 1933

3 Imaginaries

The book, for which I shall refer to the French version (Florensky 2016), is titled *Imaginaries in geometry*, with two subtitles: *Extension of the domain of geometric images to two dimensions* and *An essay of a new materialization of the imaginaries*. The work is complex, as are all of Florensky's writings, because of the continuous intermingling of philosophy, science, and symbolism. In the following pages, I shall extract a few ideas from this book, which deserves a very thorough analysis that would go beyond the scope of the present chapter. Villani, who wrote the preface to the French edition, declares there: "We find the exposition of a brilliant dreamy dissertation on the notion of duality, illustrated with advanced mathematical concepts, all designed to lead the reader's mind to the point where it can absorb a daring metaphysical thesis evoking a duality between matter and ideas." The physicist Pierre Vanhove, in the introduction to the same edition, writes: "This singular book should not be read as a work of mathematics or physics, but rather as an attempt to use the geometry of the imaginary to analyse the fundamental notion of duality between the real and spiritual worlds." At the same time, the *Imaginaries* is an illustration of the fact that Florensky's mathematical works do not consist of theorems and proofs, as one would usually expect from a mathematician; the mathematics that it is about is nonmainstream, close to art and to symbolism; and it is the language of a new imaginary world. To mention the opinion of a contemporary to Pavel Florensky, let me quote a fragment from a letter that Luzin wrote to

his wife about Florensky, after a visit he made to him in 1908; the fragment is reproduced in an article by C.E. Ford in (Ford 1998): "[. . .] But as soon as he showed me his works in mathematics, my old opinion came back to me: that all his works are of no value in the area of mathematics. Suggestive hints, beautiful analogies—very attractive and promising, provocative, beckoning, but without any results." On the connection between Luzin and Florensky, see the article (Demidov et al. 1989) by Demidov, Parshin, Polovinkin, and Florensky Jr., the discussion in (Sinkevich 2019) by Galina Sinkevich, and Sect. 7 below. Establishing a relation, which, in Florensky's *Imaginaries*, takes the form of a mathematical duality, between matter and ideas, between the rational and spiritual worlds, etc., are objectives that Florensky sought to achieve throughout his life.

Let me start with a clarification concerning the word "Imaginaries" in the title of the work. This word, in the first few pages of the book, has the meaning that mathematicians give it, when they talk about an "imaginary number," that is, the "imaginary" refers to one of the two coordinates of a complex number. Later in the book, this word designates the world of Platonic ideas, which Florensky considers as living in a space which is "inverted" with respect to our own. The fact that a complex number is the sum of a real and an imaginary number becomes a symbol of the fact that the world consists of a combination of concrete and abstract elements.[24]

In the same way as Florensky declares that the real and imaginary numbers constitute the two sides of the same surface, he sees the two worlds, the physical and the spiritual, as the two sides of the same reality. Gradually, in the book, the word "imaginary" takes on a literal meaning, as something that does not exist in reality, but rather as a tale that belongs to the realm of our imagination. As a result, the discussion quickly turns philosophical and symbolic, and the book becomes a poetic, historical, and cosmic metaphor on space and time, and on man and mankind. The book culminates in a fabulous reference to Dante's *Divine Comedy*, discussing the Florentine poet's geometrical vision. Talking about geometry in Dante's work has caused considerable ink flow in recent years, by authors who have taken up Dante's allegory, without any reference to Florensky and missing an essential point of what he actually discovered; I will discuss this in Sect. 4.

The *Imaginaries* has also a cosmological side, and it may come as a surprise to the reader that Florensky, in the last section of the work, invoking Einstein's theory of relativity, defends Ptolemy's geocentric system against that of Copernicus.[25] This section of the book, which is its culmination, should be taken (as several parts of the work) as a metaphor, continuing there the exposition of Dante's poetic and spiritual worldview. In his commentary, Vanhove, restoring Florensky's considerations in

[24] Let me add here the following remark made to me by Stelios Negrepontis: "The sensible entities according to Plato are the imperfect approximations of the true Being to which they participate. The eventual construction of an intelligible true being employs a quadratic conjugation between binomial and apotome lines, according to Theaetetus' Book X of Euclid's *Elements*, a conjugation that inspired Rafael Bombelli around 1558 in his discovery of complex numbers, which Bombelli called 'binomio', exactly because of this similarity."

[25] Florensky had to explain and defend himself before the state authorities on this issue; see Footnote 36.

context, writes that "far from having misunderstood the relativity revolution, Florensky wants to make Einstein's formalism compatible with his spiritual vision of the world."

Let me review the book in some more detail.

I start with the part that concerns complex numbers.

In this part, Florensky, showing his great erudition, refers to a host of mathematicians who worked on the foundations of the theory of complex numbers and their various uses. He mentions, in Sect. 1 and especially in the first and second notes, works of Gregory, Marie, Truel, Gauss, Argand, Kühn, Wessel, l'Abbé Buée, Français, Warren, Lambert, Mourey, Bellavitis, Cauchy, Matzka, Scheffler, Möbius, Riecke, Hoüel, and Gomes Teixeira. I know of no other work on this topic that contains such a large number of references.[26] "The various *interpretations* of the complex variable are," he says, "interpretations through which this variable manifests itself symbolically." This leads him to a literary and poetic digression (Florensky 2016, p. 36):

> We know that several translations of the same poetic work in one or more languages, far from hindering each other, actually complement each other, although none of them can totally replace the original. Scientific representations of the same reality can and should also be increased without detracting from the truth. Knowing this, we have learned not to blame this or that interpretation for what it does not give, but to be grateful to it whenever we have the opportunity to use it.
>
> Yet we are forced to point out the limitations of an interpretation as soon as we notice the hypertrophy of this or that translation which attempts to identify itself with the original and substitute itself for it, i.e., which monopolizes a certain entity jealously cutting off all other interpretations.

I will address later in this chapter the question of interpreting poetry. The problem of the "limitations of an interpretation" will also appear in the discussion on theology.

Talking about functions of a complex variable, Florensky discusses the works of Lejeune Dirichlet, Riemann, Weiertrass, and others. He notes in passing that the notion of function as a "correspondence" is "insufficient, for if we take into consideration only the **content** (the 'material cause') of the function, we miss the essential point, the function itself as a whole, as a **form** linking this content into a whole (the 'formal cause')"[27] (Florensky 2016, pp. 36–37).

Then comes a part on projective geometry, a geometry dear to the Renaissance mathematicians and artists,[28] and it is not surprising that it fascinated Florensky, especially later on; he taught perspective drawing, and he wrote articles and a book on this topic. The importance of projective geometry for the rest of this work will

[26] Rafael Bombelli, whom we mentioned in Footnote 24 before, is however missing in this list.

[27] The bold face and the quotation marks, here and in other quotes, are Florensky's.

[28] Leonardo da Vinci and Albrecht Dürer come to mind first, but there are many others, and I would like to mention just a few of those discussed by Florensky in his important book on religious art, *The Reverse Perspective* (Florensky 1992): Leon Battista Alberti, Filippo Brunelleschi, Paolo Uccello, Piero della Francesca, Donatello, Masaccio, Fillippo Lippi, Raphael (Raffaello Sanzio), Michelangelo Buonarotti, and there are others. The history of perspective is developed by Florensky in this book and in other articles referred to there.

also become clearer when Florensky will explain, in the last section of the *Imaginaries*, with Dante's poem in the background, that the projective plane is an important part of our physical universe. I will comment on this below. Florensky proposes to develop analysis (in the mathematical sense of the term) on two sides of the same surface, to which he refers as a real side and an imaginary side. He attempts to extend the duality of projective geometry to a philosophical context. Villani, in his Preface to the book, insists on the fact that Florensky's interest lies not only in the mathematical side of things: "We must seek the reason for this duality in something else, something deeper." Florensky expounds a metaphor concerning the real world and the imaginary one, and the visible world and the spiritual one, which finds its apogee in the evocation of the *Divine Comedy* that I shall review below.

In a relatively long section of the book, titled *Explanation of the cover*, Florensky comments on the book cover,[29] the work of the artist Vladimir Favorsky, a well-known graphist, woodcut illustrator, and painter who was his friend and colleague at the VKhUTEMAS.[30] In this section, Florensky, reviewing the phenomena of perception and the psychology of vision, talks about a duality in concrete visual perception which is an expression of the dualistic nature of the geometric plane, where the visual images correspond to the real side of this plane, and the abstractly visual images to its imaginary side.

Since we are talking here of projective geometry, let me also note that Florensky's *Reverse Perspective* (Florensky 1992) contains several mathematical and historical notes on this topic and its complex and long development, with mention of works of Ptolemy who, in his *Geography*, described the stereographic projection from the sphere onto the plane, together with other projections that were used in the art of drawing geographical maps, as well as works of Lambert, Loria, Aschieri, Enriques, Chasles, Poncelet, von Staudt, Fiedler, Wiener, Kupfer, Burmeister, Wilson, and the writings of the historian of mathematics Moritz Cantor on this subject.

Florensky completed his book in July 1922. At that time, he was teaching at VKhUTEMAS. It was the occasion for him to explain his ideas on space and time in art. Vassily Kandinsky was teaching at the same institute. The scientist and the artist, who were both art theorists as well, were friends. Kandinsky was influenced by Florensky; there are several articles on this topic (see Taroutina 2016; Sokolova 2015). The book (Misler 2006) contains a collection of writings of Florensky on art.

4 Dante

Section 9 of the *Imaginaries*, which is the last section of the booklet, is dedicated to the 600th anniversary of the death of the founder of Italian poetry. This event was celebrated in Russia on September 14, 1921. I will comment now on this section of Florensky's work.

[29] This cover won a Silver Medal at the 1925 *Exposition des Arts Décoratifs* held in Paris.

[30] VKhUTEMAS is an abbreviation for the Russian expression "Higher Art and Technical Studios." This is the state art and technical school where Florensky lectured from 1921 to 1924.

A certain number of modern mathematicians and other authors have quoted passages from the *Divine Comedy*, in order to point out a mathematical idea that they consider Dante expressed there, though in a hidden way, namely, that the universe is a three-dimensional sphere. Let me mention some of these authors, approximately in a chronological order: The physicists I. Ozsvàth, and E. Schücking, in the article *An anti-Mach metric*, published in the book *Recent Developments in General Relativity* (Ozsvàth and Schücking 1962), and the same authors 34 years later, in the article *The world viewed from outside* (Ozsvàth and Schücking 1998); M. Peterson, in his paper *Dante and the 3-sphere* (Peterson 1979), and the same author some 30 years later, in his paper *The geometry of paradise* (Peterson 2008) published in the Mathematical Intelligencer; R. Osserman, in his book *Poetry of the Universe: A Mathematical Exploration of the Cosmos* (Osserman 1995); W. Egginton, in his paper *On Dante, Hyperspheres, and the Curvature of the Medieval Cosmos* (Egginton 1999); M. Wertheim, in her book *The pearly gates of cyberspace: A history of space from Dante to the internet* (Wertheim 1999); G. Mazzotta, in his article *Cosmology and the Kiss of Creation (Paradiso 27–29)* (Mazzotta 2005); S. L. Lipscomb in his book *Art Meets Mathematics in the Fourth Dimension* (Lipscomb 2011); Curt McMullen in his survey, *The evolution of geometric structures on 3-manifolds* (McMullen 2010), published in the Bulletin of the American Mathematical Society;[31] M. Bersanelli, in his article *From Dante's universe to contemporary cosmology* (Bersanelli 2016); C. Rovelli, in his article *Michelangelo's stone: an argument against Platonism in mathematics*, published in a philosophical journal (Rovelli 2017); E. Ghys, in his book *A singular mathematical promenade* (Ghys 2017); and K. Sunada, in the article *From Euclid to Riemann and beyond: how to describe the shape of the universe*, published in a book I edited (Sunada 2019).

Let me also mention the interesting and beautifully illustrated article (Dreyer 1921) by J. L. E. Dreyer, *The cosmology of Dante*, in *Nature*, in which the author surveys Dante's cosmological ideas expressed in his poem (with no mention of the topology of the space described there). The article ends with the words: "To the student of the history of science it is a never-failing source of pleasure to find medieval cosmology so beautifully illuminated in the writings of the great Florentine poet" (Dreyer 1921).

I think that all these authors failed in two ways: First of all, none of them refers to the work of Florensky which was written between 40 years and almost a century earlier, although for several among them, this is understandable because Florensky's booklet, until its first translation in French in 2016, existed only in Russian.[32] Let me

[31] McMullen's survey starts with the sentence: "In 1300, Dante described a universe in which the concentric terraces of hell—nesting down to the center of the Earth—are mirrored by concentric celestial spheres, rising and converging to a single luminous point. Topologically, this finite yet unbounded space would today be described as a three-dimensional sphere."

[32] First edition, Moscow, Pomor'ye 1922 (1000 printed copies), new edition, Moscow, Lazur, 1991.

mention incidentally that one rarely finds authors on this list who quote others on the same list who made equivalent statements. Among the few authors who cited others are Rovelli, Sunada, and Ghys, who quote Peterson.

The second way I think these authors failed is that all of them declared, omitting an important part of Dante's poem which was highlighted by Florensky, that the space which is described by the Florentine poet is the three-dimensional sphere. In fact, they missed the part on the nonorientability and projectivity. The only paper I am aware of where the authors mention nonorientability is a commentary on Florensky's *imaginaries* by Bayuk and Ford (2003), in which the mathematical phrasing is very shaky but where the authors reach the conclusion that "[. . .] All these considerations brought Florensky to the conclusion that the geometry of Dante's cosmos is very similar to that of the single-sided elliptical surface described by Riemann and Klein."

Before developing this, I would like to point out that Florensky, with his acute intellectual honesty, noted that three authors before him had observed that the geometry of the *Divine Comedy* is non-Euclidean, namely, the American mathematician George Bruce Halstead (1853–1922), the German mathematician Heinrich Martin Weber (1842–1913), and the German historian of mathematics Maximilian Simon (1844–1918).[33] For the fact that Dante's world is non-Euclidean (which presumably means, in this setting, that it carries a geometry of positive curvature), Florensky refers to a passage in the *Divine Comedy* where the author speaks of "triangles whose angle sum is not equal to two right angles." Indeed, at some point of the exposition, Dante quotes King Solomon who asks the Lord whether there exists a triangle inscribed in a semicircle, with one side being the diameter, and no right angle (*Paradiso*, Canto XIII, v. 101–102):[34]

non *si est dare primum motum esse*,	nor *si est dare primum motum esse*,[35]
o se del mezzo cerchio far si puote	nor if in a semicircle a triangle can be formed
triangol s'ı ch'un retto non avesse.	without its having one right angle.

The question of nonorientability needs a longer quote and is based on Florensky's analysis, which I will review now.

Florensky starts by recalling Dante's description of the Ptolemaic model of the universe, composed of spheres: the Earth, the celestial bodies and the spheres on which they stay, together with the Empyrean stratum, surrounding the whole universe and occupied by the element of fire. Dante recounts a long imaginary journey he makes on his way to visit God, who dwells in the last circle after the Empyrean. The path he takes leads successively through Hell, Purgatory, and Heaven. In this journey, Dante is accompanied by Virgil, the Roman poet from the first century BC, author of the *Aeneid*, who guides him through Hell and most of the Purgatory. On

[33] Personally, I looked for written texts where these authors made such a statement, but I could not find any.

[34] Transl. Henry Wadsworth Longfellow (1807–1882).

[35] "nor if we have to admit a prime motion" (Dante wrote this line in Latin).

their way, the two poets start by going downward, traversing one after the other the spirals of Hell. It is important to note that during this descent, they always maintain a standing position, their heads pointing toward the point from which they started, that is, the city of Florence, and their feet directed toward the center of the Earth. When the two poets find themselves at the bottom of Hell, they see Lucifer with legs up and head down. It is conceivable (and mathematically equivalent) that at this point they are themselves turned upside down, with their feet directed toward the surface of the Earth and their heads on the opposite side (see Fig. 5). Here is the passage of the *Divine Comedy* quoted by Florensky (*Inferno*, Canto XXXIV, v. 76–96):

Quando noi fummo l dove la coscia	When we were come to where the thigh revolves
si volge, a punto in sul grosso de l'anche,	Exactly on the thickness of the haunch,
lo duca, con fatica e con angoscia,	The Guide, with labour and with hard-drawn breath,
volse la testa ov'elli avea le zanche,	Turned round his head where he had had his legs,
e aggrappossi al pel com'om che sale,	And grappled to the hair, as one who mounts,
sì che 'n inferno i' credea tornar anche.	So that to Hell I thought we were returning.
"Attienti ben, ch per cotali scale",	"Keep fast thy hold, for by such stairs as these,"
disse 'l maestro, ansando com'uom lasso,	The Master said, panting as one fatigued,
"conviensi dipartir da tanto male"	"Must we perforce depart from so much evil."
Poi uscì fuor per lo fóro d'un sasso,	Then through the opening of a rock he issued,
e puose me in su l'orlo a sedere;	And down upon the margin seated me;
appresso porse a me l'accorto passo.	Then tow'rds me he outstretched his wary step.
Io levai li occhi e credetti vedere	I lifted up mine eyes and thought to see
Lucifero com'io l'avea lasciato,	Lucifer in the same way I had left him;
e vidili le gambe in s tenere;	And I beheld him upward hold his legs.
e s'io divenni allora travagliato,	And if I then became disquieted,
la gente grossa il pensi, che non vede	Let stolid people think who do not see
qual quel punto ch'io avea passato.	What the point is beyond which I had passed.
"Lvati s", disse 'l maestro, "in piede:	"Rise up," the Master said, "upon thy feet;
la via lunga e 'l cammino malvagio,	The way is long, and difficult the road,
e gi il sole a mezza terza riede".	And now the sun to middle-tierce returns."

The two poets, after having avoided the center of the world, continue, following the same way, but now they find themselves moving upward, until they traverse a crater track, to find themselves in the clear world. This is illustrated here in Fig. 5, from a manuscript in the Vatican Library (Inferno, Canto XXXIV, v. 130–139):

Lo duca e io per quel cammino ascoso	The Guide and I into that hidden road
intrammo a ritornar nel chiaro mondo;	Now entered, to return to the bright world;
e sanza cura aver d'alcun riposo,	And without care of having any rest
salimmo s, el primo e io secondo,	We mounted up, he first and I the second,
tanto ch'i' vidi de le cose belle	Till I beheld through a round aperture
che porta 'l ciel, per un pertugio tondo.	Some of the beauteous things that Heaven doth bear;
E quindi uscimmo a riveder le stelle.	Thence we came forth to rebehold the stars.

Fig. 5 Illustration of canto XXXIV of Dante's *Inferno*, with Lucifer upside down, from a manuscript in Rome, Vatican Library, Ms. Urb. Lat. 365, fol. 95v. Ferrara, 1474–1482

Beyond this limit, Dante begins the ascent toward the celestial spheres, always following the same line, passing through the Empyrean where he is able to contemplate the Glory of God, and eventually ending up again in Florence, where he began his journey. In other words, the poet, having traveled along a (non-Euclidean) straight line in the three-dimensional world, finds himself at the same place and in the same position, but after having undergone a complete reversal. The surface on which he has traveled is nonorientable. Obviously, Dante did not have the technical means to express this fact in the way a mathematician would today and, likewise, he did not have the technical means to say that the universe is a three-dimensional sphere. Florensky commented on Dante in his own way, using the tools he had, and letting his imagination do the rest.

Florensky calls Dante "the greatest master of world design" (Florensky 2016, p. 70), and the last sentence of the *Imaginaries* reads: "Thus, beyond time, the *Divine Comedy* is not behind, but in front of contemporary science" (Florensky 2016, p. 78).

I said earlier that all the modern authors I mentioned failed to understand Dante's cosmology, and I explained why. In fact, Florensky himself considered that his

predecessors failed to understand Dante's cosmology, for another reason. Let me quote him on this. Before he gives his own description to which I have just referred, he writes, on p. 71 of (Florensky 2016):

> The following image is usually given in commentaries on the *Divine Comedy*. The terrestrial sphere is surrounded by the sphere of celestial bodies, by the sky of unchanging, immovable stars, by the crystal sky and finally by the Empyrean, and Dante's path from the bosom of the Earth follows a broken line spiraling through the concentric spheres and curving 180° towards the zenith of Sion. But this drawing corresponds neither to Dante's narrative, nor to the foundations of cosmology.

I noticed long ago that Dante occupies a special place in Russian culture. Talking with some Russian colleagues and friends, I realized that some of them can recite passages of the *Divine Comedy*; they learned it at school. I cannot resist mentioning the first three lines of this poem, which my Russian colleague, Mitia Millionschikov, to whom I recently told about the article I was writing, spontaneously recited:

Nel mezzo del cammin di nostra vita,	Midway upon the journey of our life,
mi ritrovai per una selva oscura	I found myself within a forest dark,
ch la diritta via era smarrita.	For the straightforward pathway had been lost.

I will mention Andredi Tarkovsky later in this chapter. The opening of the *Divine comedy* is recited (twice) in his film *The mirror*.

Let me end this section by mentioning that in his major theological work, *The pillar and ground of the truth*, which I will quote several times, Florensky refers to another grand epic poem in which a long journey is recounted, the *Odyssey*.

5 Time, Relativity, and Literature

In the *Imaginaries* (Florensky 2016), after the passage on Dante, comes a discussion of Einstein's theories of special and general relativity, an important component of the book. Florensky's goal is to show that, in the same way as "space must be represented as an elliptic space and recognized as limited, time is limited and is closed in on itself." (Florensky 2016, p. 73].

Sabaneeff, who kept notes of his conversations with Florensky, recalls that the latter had the following thoughts (Sabaneeff 1961, p. 322–323):

> Time, like space, can be not only straight but also curved (Einstein's time is curved). It can be infinite, but it can also be closed (analogous to a closed line). The concept of a single future is nothing but a theory of predestination. Furthermore, at every moment, the future divides itself into a multitude (infinity) of actual new futures. The "closed" time, after having run its cycle, "returns" and is inevitably drawn into the "sphere of eternal repetition," into a vicious circle, which he was inclined to identify with the idea of "hell" in religion.

I think that every sentence in this quotation could be analyzed in the context of scientific and philosophical culture of those times, and of today. In fact, one could

make relations with the idea of eternal return which is associated with Pythagoreanism, Platonism and Stoicism in Ancient Greece, then Proclus, Maximus the Confessor, and many others. Closer to us, the idea of eternal return was revived by Nietzsche, late in his life, in his search for a way to combat nihilism.

Florensky understood early on that the theory of relativity has changed the worldview of scientists, and that in this context, Copernicus' heliocentric system was no more valid than the geocentric theory of the Ancients. More than that, he maintained the following (Florensky 2016, p. 75):

> It would be a serious error to declare that the two systems, Copernicus' and Ptolemy's, are equally legitimate: they are so only in terms of abstract mechanics, but in terms of all the data, it is the latter that is true and the former that is false. This is a direct confirmation of the great poem, albeit more than six hundred years later.

Florensky is talking here about "data." Ptolemy collected a huge amount of data for the drawing of geographical maps, in his *Almagest*, for celestial maps, and his *Geography*, for terrestrial maps. It is true that Ptolemy's data, and the maps that were drawn using them, are amazingly precise, despite the rudimentary means used for collecting the needed astronomical observations, and despite the fact that the theory is based on a geocentric system. Florensky's defense of Ptolemy's system, a defense which was not being accepted by some of his contemporaries, was the cause of subsequent troubles with the authorities who accused him of obscurantism. These problems are reported and commented on in an interesting addendum to the French edition of the *Imaginaries* (Florensky 2016, p. 95–118).[36] Although at some point Florensky gives a reasonable exposition of some elements of the theory of relativity, he introduces other thoughts that are part of the expression of his atypical view of science, which irritated the authorities, accusing him of antimodernism, and which were termed irrational by readers of that work. At the end of his book, after his excursus on the *Divine Comedy*, Florensky, returning to the *imaginaries*, writes:

> The domain of the imaginary is real, conceivable, and in Dante's language it is called the **Empyrean**.[37] We can represent all of space as **double**, being formed of surfaces with

[36] The Postface to the French edition of the book, written by Pavel Vassilievich Florensky, the author's grandson, contains the translation of a letter sent by Pavel Aleksandrovich Florensky to the *Political Department*, dated September 13, 1922. The letter starts by the words: "This is not to claim everyone's 'right' to print whatever comes to his mind, and while recognizing that, in the name of goals concerning the whole state, the authorities cannot yield to the interests of literature, but I assume that in the case at hand, of my little brochure on *Imaginaries*, there is a *quid pro quo*, which can be explained by the novelty of the object, which is why I will take the liberty of disturbing the censors in order to set out the following considerations. . . ." The passage that was censored is the one about Dante, and Florensky recalls in his petition the context in which he wrote this passage (the 600th anniversary of the death of the Florentine poet) while defending the use in geometry of poetic images as the expression of a certain *psychological factor*. He ends his letter with a reminder of the work he is conscientiously doing to serve the State. His request not to have the book censored was eventually accepted.

[37] The bold face letters are Florensky's.

> Gaussian coordinates, real surfaces and imaginary ones corresponding to them, but the transition from the real to the imaginary surface is only possible through the **fracture** of space and the **turning** of the body inside itself. In the meantime, we only represent this process through the acceleration of motion, perhaps through the motion of certain small parts of the body, beyond the speed limit c, but we have no proof that other means are impossible.

In this passage, the expressions *fracture* and *turning* that Florensky emphasizes refer to the discontinuous operation of orientation-change that occurs after one describes a complete geodesic in the projective plane. In Fig. 6, I have reproduced an illustration of Florensky's *Imaginaries in geometry*, of a man walking in a nonorientable world. As to the mathematical notion of discontinuity, I already mentioned that it was already at the service of the philosophy of Bugaev, Florensky's first mentor who exerted a strong influence on him. I shall elaborate on this in Sect. 6. Let me quote here another passage of the *Imaginaries* in which Florensky talks about discontinuity (Florensky 2016, p. 76):

> What does the speed limit $c = 3 \times 10^{10}$ cm/second actually mean? It does not mean that velocities equal to or greater than c are impossible, but it does mean that with them would appear entirely new conditions of life that we cannot yet visually represent, and, perhaps, forms transcendent to our Kantian terrestrial experience. But this does not at all mean that such conditions are impossible, and perhaps as the realm of experience expands, they will become representable. In other words, the life of a world with velocity equal to c is qualitatively different from that observed with velocities below c, and *the transition between the domains of this qualitative difference can only be thought of as discontinuous*.

In a manuscript on religious art, written in 1922 and published in Russia in 1999 and in French translation a little bit later (see the historical note in Florensky 1992, p. 121), Florensky expresses a similar idea:

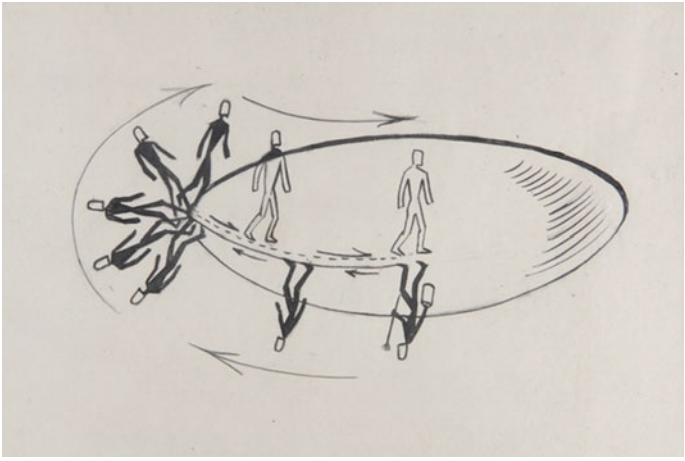


Fig. 6 Illustration from Florensky's *Imaginaries in geometry.* 1922. Paper, ink

> [. . .] While everyone agrees, even without knowing the principle of relativity, that every system (at least in relation with the case under consideration) has its own time, its own rhythm and its own measure, very few people have thought about the possibility of time running at an **infinite** speed, and even turning around completely, once this infinite speed has been exceeded, and **reversing** its course. But time can really be instantaneous, flowing from the future to the past, from consequences to causes, teleologically speaking, and this is what happens when our lives move from the visible to the invisible, from the real to the imaginary.[38]

The last three quotes express similar ideas, in different contexts, poetry, science, and the imaginary. The reversal of time is a counterpart to the idea of reversal of space, expressed in the *Reverse perspective*.[39]

The *Imaginaries* had an impact on the symbolist and Neoplatonic trend in Russian philosophy. In the Postface, written by Florensky's grandson (p. 104–105 of the French edition), there is a passage on the influence of this work on the philosopher Aleksei Fedorovich Losev (1893–1988), who referred to the last section of the *Imaginaries* for the possibility of a motion at a velocity greater than that of light, where "the body collapses into itself; its surface folds up within itself, turning inside out, through itself, and time begins to flow backwards, i.e., the consequence precedes the cause . . . The cause is transformed into a goal, an ideal."

I would like to end this section by a few words on the influence of the *Imaginaries* on literature, and for this I will quote another passage from the Postface of the French edition (Florensky 2016, p. 105–106). It concerns Mikhal Bulgavov, one of the best Russian writers of the twentieth century:[40]

> Mikhal Bulgakov's widow, Elena Sergeyevna Bulgakova, testifies that the writer liked Florensky, and that his library contained a copy of the *Imaginaries*. This book was one of the writer's first purchases in Moscow, and he would hide it from onlookers. One of the most inspired and lyrical passages of the novel *The White Guard*, written in 1923–1924, appears as a paraphrase of the concluding paragraph of *The Imaginaries*. [. . .] While working on his novel *The Master and Margarita*, Bulgakov reread the *Imaginaries* more than once. His copy of the book contains underlined words and exclamation marks in the margins of the pages of paragraph 9. Elena Bulgakova recounts that when the first listeners of *The Master and Margarita* heard, dumbfounded, the novel's finale, Bulgakov pointed out to them the

[38] Here, Florensky refers to the work of Carl Du Prel (1839, 1899), who wrote about instantaneous time (without mentioning inverted time).

[39] It may be useful to recall here that "reverse perspective," or "inverted perspective," is a technique of drawing, used in Byzantine and Russian icon painting, which is in opposition with the rules of perspective that are followed in Western classical art and theoreticized by the Renaissance artists. In reverse perspective, objects depicted that are supposed to be further away from the viewer of the painting are drawn as larger, whereas closer objects are drawn as smaller. In reverse perspective, parallel lines that move away of the viewer of the picture are drawn as diverging instead of converging lines, as they are drawn in the usual linear perspective drawing.

[40] Mikhail Afanasyevich Bulgakov (1891–1940) is best known for his novel *The master and margarita*, one of the important pieces of literature of the twentieth century. It was written under the Stalin regime. Bulgakov was a medical doctor. During his lifetime, the novel belonged to the underground literature in the Soviet Union. It was published posthumously, for the first time as a book, in Paris, in 1967.

> pages of Florensky's book that scientifically confirmed his ideas. In Florensky's mathematical and philosophical interpretation of Dante's journey with Virgil, Bulgakov saw the basis for the final chapters of his novel.

We have been naturally led to literature, and I will say more about this in Sect. 8. I would like to end this section by quoting again Sabaneeff, from his article (Sabaneeff 1961, p. 321), on the mathematical tendency of Florensky:

> [Florensky was] an excellent mathematician who concentrated his attention primarily on the latest trends in mathematics, he expressed time and again the idea that mathematics was broader and more comprehensive than the human mind. Mathematics, according to him, "is more intelligent than the human brain; it leads man farther ahead into regions inaccessible to the brain and not corroborated by sensual images." It was clear to me that mathematics was his guide even in the area of mystic speculations: it helped him not only through the elementary language of numbers (as in the case of many earlier mystics) but by means of the whole panoply of the modern mathematical apparatus: analysis, the theory of sets, and all the latest theories on the boundary between physics and mathematics.

6 Bugaev and Florensky: Mathematics and Philosophy of Discontinuity

If one wishes to mention a single mathematician who exerted a definite influence on Pavel Florensky during the years he was a student at the Imperial University of Moscow, Bugaev's name comes naturally. Nikolai Vasilievich Bugaev taught mathematics at Moscow University from 1867 till his death. Besides being a founder and president of the Moscow Mathematical Society, he was also an active member of the Moscow Psychological Society. He is considered as the main founder of the Moscow school of function theory, about which I will say a few words in Sect. 7. Some of his students became first-order mathematicians; these include Nikolay Yakovlevich Sonin (1849–1915), Konstantin Alekseevich Andreev (1848–1921), and Dmitri Egorov, whom we already mentioned. He also exerted a strong influence on Boleslav Kornelievich Mlodzievsky, who was one of Florensky's teachers and whom we already mentioned as well.

Besides being a mathematician, Bugaev was a philosopher steeped in humanism and Christian philosophy. For him, mathematics was a means of expressing philosophical ideas. Notions like continuity and discontinuity, which formed the basis of the theory of functions that was at the heart of his mathematical investigations, were for him, at the same time, part of the language of metaphysics, used in theories concerning free will, continuity in history, social revolutions, etc. It was through his influence that Florensky made some of these mathematical notions the basis of his philosophical thought. Sabaneef writes about Florensky, in the article we mentioned (Sabaneeff 1961, p. 321): "It seems to me that the dominant motive of his life was the idea of the alliance and fusion of science and revelation, the termination of the antagonism between these two spheres that had developed in the course of the historical process."

To give an idea of Bugaev's world and his influence on Florensky, I will make a short overview of the talk he delivered at the 1897 International Congress of Mathematicians held at Zurich.[41] The talk was naturally in French and is titled *Les mathématiques et la conception du monde au point de vue philosophie scientifique* (Mathematics and the conception of the world from the point of view of scientific philosophy) (Bugaev 1898a). The aim of this talk is to develop the idea that some fundamental principles of mathematics are at the basis of important philosophical conceptions of the world. Thus, the subject of the talk involves science, philosophy, and the description of the world.[42] The mathematical concepts that are referred to are those that accompanied Florensky in his later philosophical writings: discontinuity, contradiction, the infinite, antinomy, etc.

Bugaev declares in his talk:[43]

> It is very important to examine presently the influence of these parts of mathematics on the scientific and philosophical conception of the world. [. . .] At present, this need for number and measurement seems to be the vital issue of the day, not only in science, but also in art and human relations. Finding the measure in the realm of thought, will and feeling is the problem of the contemporary philosopher, politician and artist.

Bugaev, in his speech, wants to show how man, with the help of "number and measure," can get out of the realm of instinct. He considers that what is needed for that purpose is the theory of functions. He makes a distinction between *analysis*, the theory of continuous functions, and *arithmology*, the theory of discontinuous ones.[44] He states that mathematical analysis is distinguished by its generality and universality while arithmology bears the stamp of an original individuality, and is attractive because of its mysterious character and striking beauty. He believes that this explains the fact that some thinkers have linked various questions of mystical philosophy to questions related to the integers. He adds that the elegant truths of arithmology awaken in the scientist a sense of scientific beauty that fully satisfies him. He then talks about the applications of mathematical analysis in philosophy. Continuous functions explain mathematically some natural phenomena. It is interesting that he quotes a poem by the poet Sichler

[41] Initially, the first International Congress of Mathematicians was scheduled for the year 1900, and the Zurich congress of 1897 was planned as a trial congress. But eventually, the latter was so successful that it was considered later as the first ICM.

[42] This kind of philosophical topic, chosen as the subject of a lecture delivered at an ICM, was not exceptional. Poincaré, in his lecture to the same congress, titled *On the relationship between pure analysis and mathematical physics*, (Poincaré 1898) makes a comparison between mathematics cultivated for its own sake and that which has applications and used as a language for physics. He discusses these analogies in mathematics, the reason why mathematics needs other sciences to develop, questions of rigor, and, like Bugaev, issues of continuity and discontinuity.

[43] My translation from the French.

[44] In his article at the same congress, Poincaré also talks about the ideas of continuity and discontinuity. In particular, he says: "The only natural object of mathematical thought is the [natural] number. It was the external world that imposed the continuous on us, which we undoubtedly invented, but which it forced us to invent. Without it, there would be no infinitesimal analysis; all of mathematical science would be reduced to arithmetic or the theory of substitutions."

(translated from the Russian), expressing the analytic conception of the world. Having a closer look at natural phenomena, he concludes that the point of view of discontinuity is necessary, in particular in atomic physics, chemistry, biology, and sociology. "Discontinuity always appears, he says, where independent and autonomous individuality manifests itself," and further: "The highest human interests are often linked to problems of arithmology. The arithmological worldview indicates that final causes also play a role in the world's manifestations." Bugaev refers to Leibniz, as a source of inspiration for such questions. He explains why until present times only the analytical point of view was used in the scientific and philosophical considerations on the world and nature. The problem, he says, is that until now it was assumed that any scientific question assumes only one answer. But in arithmology, one finds functions that are multiple-valued, and sometimes these values are infinite. Such functions, he says, arise in nature, and he gives examples from sensation theory. He is led then to talk about antinomy. From the mathematical point of view, it is impossible to reconcile the concepts of continuity and discontinuity. But to have a complete understanding of facts of nature, we have to take equally into account the two notions. He says that in logic, biology, psychology, history of philosophy, and sociology, the universal and the individual, the concrete and the abstract, and person and society mutually complete each other. These thoughts are very much within the cosmic Russian stream of that epoch, according to which man and the cosmos are in harmony; this is why they called man a microcosmos, a terminology that was used in Ancient Greece by Anaximander, Plato, the Stoics, and others.

The use of the idea of discontinuity and its necessity in the description of nature, in Bugaev's teaching, did not escape Florensky. The title of the dissertation he presented at the Faculty of Mathematics and Physics in 1904, first under the supervision of Bugaev and, after the latter's death, under that of Egorov, is informative: *The idea of discontinuity as an element of a worldview*. He quotes this work several times in his later writings; see in particular Note 183 of the *Pillar* (Florensky 1975, p. 485), where he mentions the numerous existing works "in the domain of the formal investigation of the idea of discontinuity, i.e., in mathematics and in logic," referring to the bibliography given in the above manuscript. He then refers to the "works of the Moscow mathematical school," and in particular to Bugaev's *Introduction to the theory of numbers* and *Mathematics and the scientific-philosophical worldview* (Bugaev 1898b). In the same note, Florensky talks about the idea of discontinuity in thermodynamics, leading to the construction of the discontinuous thermodynamic surface, and in the works of Gibbs on phases, of Korzhinsky on heterogenesis, and of neo-Lamarckian and neovitalist researchers in biology, "opening broad horizons of the discontinuous evolution of organisms and their discontinuous adaptation." He also mentions researches in psychology and on the subconscious and superconscious psychic life, where one analyzes discontinuous changes in consciousness, in creativity, in inspiration, etc. In a note to Letter X of his *The pillar and ground of the truth*, titled *Sophia*, Florensky refers again to two papers by Bugaev, *Geometry of arbitrary quantities* and *Discontinuous geometry*.

In Sect. 2, I mentioned a passage from Florensky's autobiographical notes (Florensky 2007) where, talking about his 1899 crisis, he declares that the notion of continuity "suffocated him." This shows the importance of this topic for him, and at the same time, it shows how he experienced mathematics physically, in the same way he experienced sound – we shall talk about that in Sect. 8.

With Florensky, the notion of discontinuity also entered the realm of theology, which is the main subject of *The pillar and ground of the truth*. In Letter V, titled *The Comforter*, Florensky talks about "the discontinuity of the metahistorical revelation of the Spirit" (Florensky 1975, p. 101). In Letter VI, titled *Contradiction*, talking about the Apostle Paul, Florensky writes: "His brilliant religious dialectic consists of a series of discontinuities; it jumps from one affirmation to another, where each successive affirmation is antinomic with respect to the preceding one. Sometimes an antinomy is even embodied in a stylistic discontinuity of exposition, in an external asyndeton.[45] Rationally contradictory and mutually exclusive judgments have their sharp edges directed against each other." Antinomy and discontinuity are present in all of Florensky's theological works; I will say more about this in Sect. 10. Regarding the philosophical implications of the concept of discontinuity in Florensky's works, and, more generally, in Russian philosophical thought, let me mention the article (Oppo 2022) by Andrea Oppo, titled *Conceptualising discontinuity: Pavel Florenskii's* Preryvnost' *as a universal paradigm of knowledge*, in which the author discusses, among other things, Florensky's interpretation of Shakespeare's *Hamlet* in his essay (Florensky 2006a), and Florensky's passage on Dante in the *Imaginaries* (Florensky 2016).

To conclude this section on Bugaev, and talking again about the ideas of continuity and discontinuity beyond mathematics, let me mention a paper by Yuri Manin titled *Time and periodicity from Ptolemy to Schrödinger: paradigm shifts vs. continuity in history of mathematics* Manin (2019), in which the author discusses the idea of "scientific revolution," central in Thomas Kuhn's philosophy, opting rather to a theory of continuity for what regards the history of pure mathematics. In the same article, Manin mentions Ptolemy's system, also dear to Florensky.[46]

[45] By the word *asyndeton*, Florensky refers to a form of speech in which coupling conjunctions are dropped or replaced by commas, in order to speed up speech. An example, in the writings of Saint Paul to whom Florensky refers, given by F. Godet in (Godet 2002), is the following passage (*2nd Epistle to the Romans*, v. 9 and 10): "Tribulation and anguish, upon every soul of man that doeth evil, of the Jew first, and also of the Gentile; but glory, honour, and peace, to every man that worketh good, to the Jew first, and also to the Gentile." Godet comments: "As always, the asyndeton indicates a more energetic reaffirmation of the preceding idea." The notion of *asyndeton* is present at several places in the explanations that Godet gives in his book (Godet 2002).

[46] For a related discussion, the interested reader is also referred to the article on Manin in the present Handbook (Papadopoulos 2023). For a recent point of view on (the philosophy of) continuity and discontinuity in mathematics and in the relation between mathematics and physics, let me mention Plotnitsky's article *Continuity and discontinuity in, and between, mathematics and physics* (Plotnitsky 2023). Let me also take this opportunity to mention another paper by Plotnitsky, titled *The Image of time and the specter of the clock, from modernist literature to cinema* (Plotnitsky n.d.), in which he discusses continuity and discontinuity in Belyi's famous symbolist novel *Petersburg* (1913), and in cinema, in particular in the work of Andrei Tarkovsky, who was very close spiritually speaking, and also in his films, to Florensky, as we shall recall in Sect. 13. Plotnitsky tells me that Belyi's novel *Petersburg* is not only a founding novel of symbolism and of modernism in general, but it is also a novel about anarchism. In fact, the action takes place during the 1905 revolution. Plotnitsky points out the following passage from this novel which makes a relation with cinema, which in those times was one of the rising arts, seen by some (Lenin was among them) as the most important communist art: "We [our brains] are the unrolling of a cinematographic film subject to the minute action of occult forces; should the film stop, we will be fixed forever in an artificial pose of terror."

We have mentioned Bugaev's influence on Florensky. If, in another direction, we want to name a mathematician on whom Florensky had a definite influence, the first name that comes to mind is Luzin, of whom I will talk in the next section.

7 Luzin: Spiritual Crisis, Function Theory, Luzitaniya, and Disgrace

Nikolai Nikolaevich Luzin and his mentor, Dmitri Fedorovich Egorov, are usually considered as the founders of the Moscow school of functions of a real variable. Demidov, in his paper *On an early history of the Moscow school of theory of functions* (Demidov 1999), writes that this school was born with Egorov's *Comptes Rendus* note *Sur les suites des fonctions mesurables* (On the sequences of measurable functions) (Egorov 1911). In this paper, Egorov proves that any convergent sequence of measurable functions is uniformly convergent up to neglecting a set of measure ε, where ε may be chosen arbitrarily. The theorem was stated without proof by Lebesgue in his paper (Lebesgue 1903) and in his book (Lebesgue 1904).[47]

The development of the theory of functions of a real variable relies in an essential way on set theory, and more precisely, on the field known as "descriptive set theory," whose goal was to set up firm foundations for the topological and set-theoretical methods used in the theory of functions. Descriptive set theory may be considered as a subarea of mathematical logic, to a large extent based on the works of Georg Cantor (1845–1918), whose name is closely associated with the use of infinity in mathematics and who, incidentally, was Russian born.[48] Luzin was one of the architects of this emerging area, and an extensive development of this theory is contained in his 1930 book *Leçons sur les ensembles analytiques et leurs applications* (Lectures on analytic sets and their applications) (Luzin 1930). The French mathematicians that were thoroughly involved in this field include Emile Borel (1871–1956), René Baire (1874–1932), Jacques Hadamard, (1865–1963), Henri Lebesgue (1875–1941), Arnaud Denjoy (1884–1974), and Paul Painlevé (1863–1933). Florensky was a mathematics student at the time of the development of the foundations of function theory, with set theory in the background, the theory which remained for him a constant subject of reflection.

Luzin became familiar with the main problems of set theory and measure theory during a stay he made in Paris after the first Russian revolution and the bloody events in its aftermath (1905–1906). In Paris, he attended lectures on function theory by the leading mathematicians there. Functions were classified in terms of set-theoretic properties and limit operations. New sets emerged and were given names. Since open and closed sets are not stable under simple operations involving infinity, G_δ and

[47] Egorov attended Lebesgue's course at the Collège de France in 1903; see (Kuznetsov 1971).

[48] See the article (Sinkevich 2015) for a concise biography of Cantor, together with notes on Florensky, Mlodzievsky, and other actors of the Silver Age.

F_δ-sets were introduced as countable unions (respectively, intersections) of open (respectively, closed) sets. *Borel sets* are obtained from open or closed sets under the operations of countable union, countable intersection, and taking complement. Among the other classes of sets that appeared, we mention *A-sets* (A stands for "analytic"), later called *Suslin sets*, whose existence arose with a remark that Suslin[49] made, saying that a continuous image of a Borel set is not necessarily Borel, correcting a mistake made by Lebesgue on this metter.[50] Suslin introduced the class of *analytic sets* in his note (Suslin 1917). Luzin, who gave Suslin the impetus for the introduction of these sets, continued the latter's work on this topic after his death. *Baire functions* arose, as functions obtained from continuous functions by a transfinite repetition of simple limits. Baire first and second category sets[51] and *Luzin sets*[52] were introduced. Luzin introduced the notion of *projective set*, that is, a set obtained from a Borel set in *n*-dimensional Euclidean space by operations of projection and taking complement. The notion of *effective sets* (a set whose definition does not use the axiom of choice) was born. Lebesgue introduced the notion *of nameable set*, a set that can be "named." Luzin presents this notion, in his *Mémoire sur les ensembles analytiques et projectifs* (Luzin 1926). He writes:

> One of the fundamental ideas that we owe to Mr. Lebesgue is the precise distinction between the two notions of a nameable set and unnameable set. We already know what, according to Mr. Lebesgue, is meant by a nameable set: it is the set of points that can be named, i.e., characterized without possible confusion, by means of a suitable definition. "Nameable" sets are precisely those which, according to Mr. Lebesgue, are the necessary object of Mathematical Analysis. As we know, in the current state of Science, unnameable sets are those that can be obtained by means of Mr. Zermelo's Axiom or similar reasoning. It is these unnameable sets that Mr. Lebesgue considers doubtful: they can only be introduced into Science under the necessary condition that they show their interest, applicability, usefulness and efficiency in solving some classical problem. But, in general, it is the domain of nameable beings that is, par excellence, the content of Mathematical Analysis. This is the conception of Mathematical Analysis proposed by Mr. Henri Lebesgue.

Luzin was aware of the difficulties which affected the theory of functions that arose from set theory, including the use of transfinite constructions and the continuum hypothesis, the discovery of hierarchies between infinities, etc. Lebesgue refused the use of the axiom of choice, considering it as counterintuitive. He considered that in order for a mathematical object to exist, one has to be able to

[49] Mikhail Yakovlevich Suslin (1894–1919) was Luzin's student. He died at age 25 from typhus, an epidemic that followed the Russian Civil War.

[50] The story is recounted by Lebesgue in his preface to Luzin's monograph *Leçons sur les ensembles analytiques et leurs applications* (Luzin 1930).

[51] A *Baire first category set* is a subset of a Baire topological space contained in a countable union of closed subsets which all have empty interior. A *Baire second category set* is a subset of a Baire topological space which is not a Baire first category set.

[52] A *Luzin set* is a subset of the real numbers which is uncountable but whose intersection with every set of the first Baire category is countable.

explicitly "name" a property that defines it in a unique way. Bourbaki in (Bourbaki 2007, p. 53ff) has an interesting passage on this matter:

> For Lebesgue, who broadens the debate, it all boils down to knowing what we mean when we say that a mathematical being "exists"; for him, we have to "name" explicitly a property that defines it in a unique way (a "functional" property, we would say), and already the question of choosing a single element in a set seems to him to raise difficulties: we need to be sure that such an element "exists", i.e., that we can "name" at least one of the elements in the set. Can we then speak of the "existence" of a set in which we do not know how to "name" each element? Already, Baire has no hesitation in denying the "existence" of the set of parts of a given infinite set; in vain does Hadamard observe that these requirements result in abandoning even the idea of talking about the set of real numbers: it is indeed this conclusion that E. Borel eventually rallies to. With the exception of the fact that the countable seems to have acquired the right to exist, we have pretty much returned to the classical position of the opponents of "actual infinity."[53]

On the questions of existence and nonexistence of certain mathematical objects, Lebesgue, in the preface he wrote to Luzin's book, *Leçons sur les ensembles analytiques et leurs applications* (Luzin 1930), recounts the following story: "When I was a student, we used to have coffee and talk about general ideas; the discussion was getting hot and seemed bound to go on forever, when one of us shouted: 'First of all, you, do you exist? I say *you* for convenience, but even, do I exist?'. . . ."

In France, like in Russia, a group of religiously inclined mathematicians[54] raised questions such as how, and to what extent, infinity – which is an attribute of God – can be used in human and mathematical contexts. It was a period where mathematicians realized that the theory of transfinite numbers and the use of the axiom of choice may lead to contradictions. Lebesgue did not exclude the existence of sets that are neither finite nor infinite.[55] This indicates again the fact that the notion of infinity, in the mathematical context, must be used with care. It was natural that Florensky, even after he quitted the Mathematics Department of Moscow University, given his interest in all areas of science, and especially mathematics, kept abreast of all these developments which involve at the same time science, philosophy, and religion.

Luzin was also tormented by philosophical and religious thoughts. At the time he was a mathematics student, he underwent, like Florensky a few years before him, a spiritual crisis that shook what he called his "materialistic worldview," an expression which, at that time, meant atheism. Eventually, his relation with Florensky became an important factor in the resolution of this crisis and his return to mathematics. C. E. Ford, in his article *The influence of P. A. Florensky on N. N. Luzin*, writes (Ford 1998, p. 332–333): "The early history of the Moscow school has been greatly illuminated

[53] The clear distinction between actual and potential infinity was made by Aristotle. According to Aristotle, actual infinity cannot exist because it is paradoxical. Only potential infinity exists.

[54] In Russia, these mathematicians include Bugaev, Luzin, and Florensky. In France, the religious tendency among mathematicians is represented by Cauchy, Hermite, and Appell.

[55] Lebesgue writes in a letter to Borel: "Although I very much doubt that we will ever name a set that is neither finite nor infinite, the impossibility of such a set does not seem to me to have been proved"; see (Hadamard 1905). The same sentence is quoted in Luzin's ICM talk (Luzin 1929).

by the recent discovery that a close relationship had developed between Luzin and Florensky.. . This enabled Luzin to launch his career as a mathematician." The article (Ford 1998) is based on the correspondence between Luzin and Florensky which has been preserved, and I would like to quote some passages from it. This will confirm the fact that Florensky had a definite influence on one of the main founders, if not *the* main founder, of the twentieth-century Moscow mathematical school.

In a letter to Florensky sent from Paris, dated May 1, 1906, quoted in (Ford 1998, p. 335), and translated from (Demidov et al. 1989), Luzin writes:

> Life is too depressing for me, sometimes agonizingly depressing. I am left with nothing, no solid worldview; I am unable to find a solution to the "problem of life." My self-image is so frequently changing that life has become pure torment....You would not believe how precious your letter and your concern for my inner life were to me. It is horrible, horrible, infinitely horrible to feel yourself surrounded by egoism, nothing but unrelieved egoism.. .. By now you will have understood how precious to me, against the background of the vast and empty night that oppresses the soul, the heart, and the brain to madness, how precious to me is your concern for my inner life....I cannot be satisfied any more with analytic functions and Taylor series. To see the misery of people, to see the torment of life, to wend my way home from a mathematical meeting, to wend my way through the Aleksandrovsky Garden, where, shivering in the cold, some women stand waiting in vain for dinner purchased with horror – this is an unbearable sight. It is unbearable, having seen this, to calmly study (in fact to enjoy) science... . Then came the terrible days for us all [the 1905 revolution], for me in particular...I could not work in science, and I seem to have begun to lose my mind as a result of the impossibility of living quietly and understanding where, where the truth is. Yes, I now understand that "science," in essence, is metaphysical and based on nothing. The scientific quest for "differential resolvents for equations of degree six" is absolutely incomprehensible...At the moment my scholarly interests are in principles, symbolic logic, and set theory. But I cannot live by science alone. . . I have nothing, no worldview, and no education. I am absolutely ignorant of the philological sciences, history, philosophy...It is painfully clear to me that it makes no sense to settle for a "scientific" education.

Luzin's crisis lasted 3 years and was eventually resolved with the decisive help of Florensky, who persuaded Luzin to return to mathematics (Ford 1998). Upon his return to Russia, Luzin frequently visited the Sergiyev Posad monastery and valued his meetings with Florensky. In a letter to his wife, dated June 29, 1908, talking about Florensky's *Pillar*, of which he read a preliminary version, he writes (Ford 1998, p. 337):

> I read it all at once in a single day—skipping a lot, but the impression was overwhelming. As I read it, I was STUNNED the entire time by blows from a battering ram against a stronghold...Our life, the life of the XIX and XX centuries, is full of confusion, full of contradictions. The work of Florensky is one solution to this impossible situation.

Luzin was interested in all of Florensky's writings. In a letter dated June 19, 1919, he asks him: "I would like to have, if possible, all your published works. I do have many, but do not have your big works, the latest lectures on philosophy (it is said that they are already published) and above all, I do not have your thesis *The Pillar and Foundation of Truth*."

On his return to Moscow, Luzin joined the Faculty of Mathematics of Moscow University. He was appointed professor there in 1917, just before the revolution.

A few words on Luzin's mathematical interests are in order. Luzin was strongly influenced by French mathematics. In 1911 and 1912, Egorov and Luzin published a series of papers in the *Comptes Rendus* which confirmed the French school's support for the Russian (Egorov 1911; Luzin 1912; Luzin 1914). In particular, Egorov's result in (Egorov 1911) on sequences of measurable functions was a substantial improvement of previous works by Borel and Lebesgue. This result says that any sequence of measurable functions is uniformly convergent up to a set of arbitrarily small measure. Luzin's result in his paper (Luzin 1914), titled *Sur un problème de M. Baire* (On a problem of Mr. Baire), answers a question raised by Baire. Precisely, Luzin proved that under the assumption that the cardinality of the continuous is aleph one, there exists in the interval (0,1) a set E having the cardinality of the continuous such that any perfect set which is non-dense in (0,1) contains at most a countable set of points of E.

Luzin was an atypical teacher. Mikhail Alekseyevich Lavrentieff, who was his student, in a memorial article on his teacher (Lavrentieff 1974), writes that around the years 1922–1926, Luzin introduced to the mathematical division of Moscow university a new style of unprepared lectures, far from the standard "good lecture," and that the result was a brilliant, deep, and fascinating way of teaching, encouraging the development of originality and independent thinking. Lavrentieff also writes that the essential characteristic of the school founded by Luzin was "the fostering of independent thought, the capacity to crack problems, to find new methods and to pose new problems." He reports that Luzin prepared his lectures only in outline, that he was often late to his class, and that he believed that having a strict schedule was not compatible with doing mathematics.

Luzin's 1926 *Mémoire sur les ensembles analytiques et projectifs* (Memoir on analytic and projective sets) (Luzin 1926) starts with the words:

> The current state of mathematical analysis provides ample proof of the desirability of establishing a precise demarcation between mathematical beings regarded as "existing" and others whose reality is only apparent. . . .This distinction was the subject of the famous *Five letters on set theory*[56] by Messrs. Hadamard, Baire, Lebesgue and Borel, on whose necessity Mr. Emile Borel insisted with extreme exactitude in his later writings. This distinction between mathematical beings that are truly real and others that appear to be real, but which have no substratum and to which no intuition corresponds, is currently very necessary, but its imminent realization seems unlikely in the present state of Science.

In 1927, Luzin was elected corresponding member of the USSR Academy of Sciences. The following year, he was elected full member, but, surprisingly, at the Department of Philosophy. In 1928, he was Vice-President of the ICM (Bologna). He presented there a communication titled *Sur les voies de la théorie des ensembles* (On the paths of set theory) (Luzin 1929) in which he discussed the problems that arose in the theory of functions as a consequence of the questions that were raised in set theory regarding infinity and the definition of real number. He used in this communication the philosophical tone that Bugaev had, 31 years before him, in

[56] Luzin refers to the letters reproduced in (Hadamard 1905).

his ICM communication. The problems he addressed are the same as those that have preoccupied Florensky. In the introduction, Luzin writes:

> In order not to get bogged down in metaphysical discussions on the meaning of the word infinite, I shall confine myself here to discussing a particular point that no doubt attracts the attention of all mathematicians: this is an attempt to satisfactorily resolve the so-called continuum problem, posed by Georg Cantor some sixty years ago.

In the conclusion of the talk, after a tour of several problems concerning the foundations, Luzin ends with the sentence: "Over the centuries, the great debate between d'Alembert and Euler on the notion of arbitrary function now becomes the debate on arbitrary real numbers and even arbitrary positive integers."[57]

In 1930, Luzin published his monograph *Leçons sur les ensembles analytiques et leurs applications* (Luzin 1930), with an appendix by Sierpiński and a preface by Lebesgue. The book was published in Paris.

Lebesgue is full of praise for Luzin in the preface. At the end of this preface, he writes:

> [Luzin] has philosophical worries, and he admits this. For him, mathematical and philosophical imperatives are constantly associated, one might even say merged. Although his book is a mathematical work, written by a mathematician for mathematicians, this intimate association of philosophical and mathematical thought is clearly evident on almost every page, giving the volume exceptional importance and appeal.

At Moscow University, Luzin was a charismatic figure. A mathematical community was formed around him, affectionately called Luzitaniya, whose members were linked not only by mathematics but also by culture. S. S. Demidov and V. D. Esakov write in (Demidov and Lëvshin 2016, p. 3) that Luzin was "the acknowledged leader of the Moscow School of function theory, the heart and soul of the famous Luzitaniya, [. . .] a true brotherhood united around their adored teacher." They add (Demidov and Lëvshin 2016, p. 10) that "[Luzin] had a charismatic personality that was attractive to young people" and that "[he] felt most in his element when working in a tight circle organized like a family gathered around an adored father-teacher." Alexey Sossinsky tells me that no such mathematical community has ever existed since, except perhaps for the early Bourbaki members. Let me add that from what we know, Pythagorean communities were the same.

Luzin's last years were saddened as he became persecuted by the state and rejected by most of his former students. His fate was not comparable in severity to that of his old friend Pavel Florensky, but still, he was betrayed, he fell into depression, and he never recovered. I will say a few words on this now.

The violent campaign against Luzin started in the summer of 1936. The purpose was to expel him from the Academy of Sciences. The period was that of a relative political quiteness, just preceding Stalin's Great Purges, which started a year before

[57] Luzin refers to a controversy which involved some of the brightest mathematicians of the eighteenth century, among whom are Euler, d'Alembert, Lagrange, and Daniel Bernoulli, on the question "what is a function?" I have surveyed this question in §10 of the article (Papadopoulos 2017).

Florensky was shot. A document reproduced in the book (Demidov and Lëvshin 2016) shows that, as early as in 1931, a confidential letter signed by Ernst Yaromirovich Kol'man (1892–1979),[58] who became president of the Moscow Mathematical Society after Egorov was arrested in 1930, was sent to the divisions of the Central Committee, in which he accuses Luzin of several offences, among them, that of being closely connected with Emile Borel, "an active member of the military [French] establishment."[59] In the same letter, Luzin is accused to be an idealist militant, and as a proof of that, Kol'man quotes a passage from a report in which Luzin writes: "The series of natural numbers does not seem to be an absolutely objective structure. It seems to be an artifact of the brain of the mathematician who happens to be speaking about the natural numbers. There seem to me, among the problems of number theory, some that are absolutely unsolvable" (Demidov and Lëvshin 2016, p. 12).

The official public campaign against Luzin started with two anonymous articles in the *Pravda*, with the titles *Answer to Academician N. Luzin* (July 2, 1936) and *On enemies in Soviet masks* (July 3, 1936). The same year, a "Committee concerning N. Luzin" was created at the Presidium of the Academy of Sciences of the USSR. The committee was headed by the Vice President of the Academy, G. Krzhizhanovsky, and its members included several well-known mathematicians, among them I. Vinogradov, L. Shnirelman, S. Sobolev, P. Aleksandrov, and A. Khinchin. Kolmogorov and Lyusternik, without being members of the Committee, were invited to attend the meetings. Most of the committee members were Luzin's former students or students' students.

The academic accusation claimed that Luzin, by the end of the 1920s, showed a lack of interest in the mathematical schools that were emerging in the Soviet Union, and was only interested in his own field. The prosecution also claimed that he was putting pressure on his students to include his name in their publications, or to make them acknowledge his help in their work. We already mentioned that Luzin continued the work of his student Suslin, after the latter's death. In fact, he was accused of stealing this work. No one will ever know if there is any truth in this, because no one will ever know to what extent Suslin shared unpublished ideas with Luzin. I learned from Sossinsky that some of Luzin's students tried to appropriate Suslin's results after the latter's death. In any case, the claims that Luzin stole others' ideas contradict what Lebesgue wrote in the preface to Luzin's book (Luzin 1930): "Mr. Luzin is perfectly happy only when he has managed to attribute his own discoveries to someone else," adding that this is a "strange mania, forgivable, it seems, since it is hardly to be feared

[58] Kol'man was a virulent Marxist who, in his own way, brought order to Moscow's mathematical community in the 1930s. He is among those who attacked Florensky's *Imaginaries*, ridiculing "mathematics in the service of religion," "mathematics in the service of priestcraft," and "mathematics in the service of obscurantism" it contains (Demidov and Ford 2005, p. 608). He published an article in the Moscow journal *Bolchevik* titled *Against the recent revelations of bourgeois obscurantism*. See pages 114 ff. of (Florensky 2016).

[59] Borel had been the Minister of Naval affaires, between April 17 and November 28, 1925, in the government of Paul Painlevé, the French mathematician who became prime minister (in those days this function was called Président du conseil).

that Mr. Luzin will be a model in this respect." Luzin was also blamed, by the Academic court, for the fact that he published his papers in French rather than Russian journals, which was claimed to be a disloyalty to his country.

In the end, and surprisingly, Luzin was not condemned as an "enemy of the people," as was planned by his detractors. The common explanation is that the physicist Piotr Leonidovich Kapitsa[60] (1894–1984), who was the head of the Physics section of the Russian Academy of Sciences, wrote a letter to Stalin, saying that Luzin might yet be useful to the government; see (Hoffman 2009) where the author comments: "It is not clear why Stalin listened, but his whim ensured the future of a discipline." Still, after this affair, Luzin remained rejected by an important part of the mathematical establishment and the case left a deep wound in the Russian mathematical community.[61] He continued working, but in isolation.[62] In 1946, Luzin published a 900-page textbook on calculus for technological universities. In 1953 (the year Stalin died), 300,000 copies of the book were already sold.

This so-called "Luzin case," a combination of a generation dispute and personal interests, is considered to be one of the darkest episodes of the history of Soviet mathematics, if not the darkest one; see the details in (Demidov and Lëvshin 2016). Discussion of these events remained a taboo in the Soviet Union, until their reexamination several decades later. In 2012, the Russian Academy of Sciences reversed the decision that was taken against Luzin in 1936. A reconstruction of the minutes of the various trials that took place is carried out in book (Demidov and Lëvshin 2016) by Demidov and Lëvshin. The reader may also consult Demidov's article *The Moscow school of the theory of functions in the 1930s* (Demidov 1993) which constitutes an excellent concise report on the school founded by Luzin. The Lavrentieff memorial book (Dobretsov et al. 2000) contains a chapter dedicated to the Luzin case.

Today, it is clear that the reason behind Luzin's persecution was his religious convictions. The same thing had happened to Egorov. The latter, who had not been politically active, fell into disfavor with the Soviet authorities 6 years before Luzin. Demidov and Esakov write in (Demidov and Lëvshin 2016, p. 10) that "[Egorov] disliked the new regime and made no attempt to hide his negative attitude." The

[60] Kapitsa was awarded the Nobel prize in physics in 1978. In 1946, he refused to work on the Soviet atomic weapons, and this led to his dismissal from the Academy, a position he took up again in 1955, 2 years after Stalin's death. At his death in 1984, Kapitsa was the only member of the Russian Academy of Sciences who was not a member of the Communist Party.

[61] In an email I received on April 29, 2019, Alexei Sossinsky writes: "When the Committee concluded its work, it was expected that it would declare that Luzin is 'an enemy of the people', which meant that he would be sent to the camps or even be given the death penalty. But it didn't! Until a few years ago, we did not know why this happened. The publication of certain documents now gives us the correct story, which involves Kapitsa's letter, Kol'man's falling into disfavor, Krzhizhanovsky, Molotov, and Stalin, the final decision apparently being taken in a tête-à-tête between Krzhizhanovsky and Stalin."

[62] I learned from Yuri Neretin that in 1937–1939, Luzin published an ingenious work on some types of bendings of two-dimensional surfaces, solving a problem that had been discussed during 50 years; see (Luzin 1939). Unfortunately for him, the solution was negative and his work closed the topic and had no direct continuation. Sabitov wrote a review of this problem in (Sabitov 2016), in which he says: "a truly dramatic end to this direction of geometry came with Luzin's paper (Luzin 1939)."

campaign against him started in 1929 when Egorov, at the first All-Union Congress of Mathematicians in Kharkov, opposed a proposal to send greetings on behalf of the Mathematical Congress to the Sixteenth Party Congress; see (Demidov and Lëvshin 2016, p. 11). The same year, after having been President of the Moscow Mathematical Society for 7 years, he was dismissed from his position at the University. The next year, he was jailed for being a "religious sectarian." About Egorov's religiousness, see (Kolyagin and Savvina 2010). In fact, Egorov was accused of being part the so-called Catacomb Church (Demidov and Lëvshin 2016, p. 12).[63] Little is known about this sad story, and the information about Egorov's last years contains contradictions. Most probably, he died after a hunger strike. According to Demidov, he died in the hospital of the Institute for the Continuous Education of Doctors (a branch of Kazan University) (Demidov 1999). I learned from V. N. Berestovskiĭ that Egorov was buried in the same cemetery as N.I. Lobachevsky, in Kazan.

On the Luzin case, I have benefited a few years ago from correspondence with Yuri Neretin and Alexey Sossinsky. There are many comments on this in Neretin's article (Neretin 1914). Let me end this section by a last information about Luzin, which I learned from Sossinsky. The year before his death, Luzin thought he had a counterexample to the Goldbach conjecture. He sent his result to the *Doklady Akademii Nauk*. The example was numerical, with huge numbers involved. It took a large amount of time for the persons in charge of checking the result to find an error in the calculations.

The book *Naming infinity: A true story of religious mysticism and mathematical creativity* by Graham and Kantor (Graham and Kantor 2009), despite many historical, mathematical, and theological inaccuracy,[64] contributed in the rising of interest among Western readers in the history of Soviet mathematics.

[63] The "Catacomb Church" is a generic name for underground Christian communities in the Soviet Union who no longer felt part of the official Orthodox Church after the state took over its activities. The name makes reference to the Christians worshipping in the catacombs of Rome during the period of the Roman persecutions against Christians.

[64] Regarding the inaccuracy, see the corrective reviews by Hoffman (Hoffman 2009), Trimble (Trimble 2021), and Katasonov (Katasonov 2009), to mention just a few. See also the interesting bachelor thesis by Eriksson (Eriksson 2009). Katasonov (Katasonov 2009) argues against the claim which is repeated all over the book that Name Worshippers consider that naming things create them, whereas these worshippers never associated naming with creating. Trimble (Trimble 2021) notes that Graham and Kantor have no knowledge and a false reading of Orthodox theology. For instance, they conflate the hesychastic tradition as a whole with Name Worshippers. Eriksson writes that the book by Graham and Kantor with, "lamentably, very narrow view of the cultural context, attempts to analyse the connection between culture and mathematics." She adds that in their treatment, "Bugaev is left out, and in this way the mathematical roots of the tradition are cut. Graham and Kantor even seem to be unaware of Bugaev's contributions, as Bugaev is not mentioned at all in the discussion about continuity and discontinuity in Florensky's work." Personally, I do not see any reason for Graham and Kantor's insistence on homosexuality in the Russian mathematical community they describe, and for me their suggestion that Florensky was homosexual is bewildering. I think that it is based on a misinterpretation of his religious style of writing and a misunderstanding of his personality. In any case, homosexuality was severely condemned in Russia during the Soviet era, and Florensky was being very closely monitored. If there had been the slightest hint of his homosexuality, this would have been known and recorded by the KGB, but there seems to be nothing to suggest this in the documents found and published in (Chentalinsky 1993) and elsewhere.

8 Silver Age: Mathematics, Symbolism, Art, and Poetry

They call it "Silver" Age, but it is comparable to the Renaissance in Europe
(Valentin Poénaru told me).

In the chapter *The Practice of Mathematics in a Totalitarian Society* (Sossinsky 2023a) published in the present Handbook, Alexei Sossinsky talks about his past as a student at Mekhmat, the Faculty of Mechanics and Mathematics of Moscow State University, and then as a mathematician at this faculty, during the period called the "Golden Age of Russian mathematics." This is the epoch of an unprecedented development of mathematics in the Soviet Union, and it involves fields such as geometry, topology, number theory, function theory, probability theory, mathematical logic, and others. With this period are associated the names of Pavel Sergeyevich Aleksandrov (1896–1982), Mikhail Alekseyevich Lavrentieff (1900–1980), Andre Nikolaevich Kolmogorov (1903–1987), Israel Moiseyevich Gelfand (1913–2009), Yuri Ivanovich Manin (1937–2023), and several others, forming at least two generations of first-rank Russian mathematicians. The impact of their work on the development of mathematics worldwide was enormous. The reader interested by the history of this period of Russian mathematics may refer to the book (Zdravkovska and Duren 2007).

The Golden Age of Russian mathematics was preceded by another, sometimes referred to as the "Silver Age of Russian mathematics," which has been the subject of much attention in recent years among historians of mathematics; see the articles (Demidov 2021; Demidov 1993; Demidov 1999), and (Ford 1998). This is the period dominated by the figure of Luzin, which we presented in Sect. 7. Luzin mentored and had a significant impact on several of the key players of the Golden Age. This Silver Age of mathematics overlaps, by a few years, with a period also known as the "Silver Age," for poetry, literature, and art. It is a period which saw the birth of the Russian Symbolist movement, one of whose main protagonists was Andrei Belyi (Bugaev's son), whom we mentioned in Sect. 6. In 1910, Belyi published an important collection of articles bearing the characteristic name *Symbolism*, to which Florensky refers in several writings and in particular in his major theological work, *The Pillar and ground of the truth* (Florensky 1975).

Let me add that Andrey Andreyevich Markov (1856–1922) the outstanding Russian mathematician who worked in Saint Petersburg, best known for his work on stochastic processes, and Andrei Kolmogorov, the student of Luzin whom we already mentioned and who became the world leader of probability theory, algorithmics, complexity, and information theory, were both interested in the statistical approach to verse analysis in Russian poetry, a topic which was started by Andrei Belyi. Both Markov and Kolmogorov applied this approach to Pushkin's poetry; see (Markov 1913). More recently, Sossinsky wrote a paper on this topic, titled *Belyi–Kolmogorov theory* (Sossinsky 2023b).

Florensky, who belonged to the group of mathematicians of the Silver Age, was connected at the same time with the Russian symbolist movement, with its renewed interest for classical literature and for the significance of language. Since the times he was a student at Moscow University, he had close relations with Moscow's Symbolist literary circle. He published several articles in symbolist periodicals such as

Novyi Put' (New Way) and *Vesy* (Scales). In 1903, he published in *Novyi Put'* an article titled *Superstition* (Florensky 1903), which is concerned with supernatural phenomena and the psychology of their perception. In 1904, he published in the same journal an article titled *On symbols of infinity: Sketch of the ideas of G. Cantor* (Florensky 1904a), to which he referred later in the *Pillar*. This article is the first essay ever published in Russian on Cantor's ideas on infinity. In the same year, Florensky published in *Vesy* a series of poems, and an essay titled *On a certain premise of a worldview* (Florensky 1904b). In 1905, he wrote the essay we mentioned, with symbolist inclinations, *Hamlet*, on Shakespeare's play which many have seen as a work in which the author anticipates in several ways the modern world. This essay was meant to be published in the symbolist journal *Vesy* but was never published before it was included in Volume I of Florensky's Collected Works published in Moscow in 1994. It was recently translated into French, in the form of a booklet; see (Florensky 2006b).

Talking about symbolism in the Russian Silver Age, one should mention the names of some leading symbolist poets of that period: Vyacheslav Ivanovich Ivanov (1866–1949), a very complicated figure with a long and serpentine biography, Valery Bryusov (1873–1924), and Aleksander Blok (1880–1921). I like the following description of symbolism by Ivanov who, according to Sabaneeff (Sabaneeff 1961, p. 314), knew Florensky and had a high esteem for him: Symbolism is "a cult of eternity and all-embracing unity of the multicolored reflections of the moments" (Strakhovsky 1959).[65] Finally, in referring to the Russian Silver Age in art, one may mention the names of Chagall, Kandinsky, Stravinsky, Prokofiev, and Skriabin, to name but a few of the best known figures.

Florensky described symbolism as "the necessity of seeing the concrete incarnation of the soul." In his autobiography dedicated to his children, he writes: "All my life I have pondered a single problem, that of the symbol. [...] Positivism repelled me, but abstract metaphysics repelled me just as much. I wanted to see the soul, but I wanted to see it incarnate. Someone will say that this is materialism, but it is not materialism that is at stake, but the necessity of the concrete, of symbolism." In the same autobiography, we can read: "I have always been a symbolist. Symbols are created by the mind according to precise laws, and with an inner necessity, and this happens whenever certain aspects of the mind begin to function in a particularly lively way. What symbolizes and what is symbolized are not fortuitously interrelated." With all this, it is not surprising that Florensky's major theological and philosophical work, *The Pillar and ground of the truth*, is written in the language of symbols. He writes there, in the Xth letter (Florensky 1975, p. 236): "I am compelled to use a metaphysical terminology, but in my speech these terms have not a strictly technical sense but a conventional or rather a symbolic one. They have the significance of colors by means of which inward experience is painted." The language he uses is also the one of the *Imaginaries*.

[65] In this sentence, Strakhovsky in (Strakhovsky 1959) is translating Ivanov, from his article *Of the Merry Craftsmanship and Smart Revelry*, published in the ephemeral periodical *Zolotoe Runo* (1906–1909) (Ivanov 1907, p. 54).

Speaking of art and poetry, I would like to mention Florensky's inclination toward poetry and music. Sabaneeff writes, in an article I have already mentioned several times, that Florensky "was too far removed from music and from art in general." I think Sabaneeff was wrong.[66] In his autobiography and correspondence, Florensky regularly talks about music and poetry. As a child, he received a musical education, and everybody in his family listened to and played music. Let me quote a small passage from his recollections (Florensky 2007, p. 80): "As a child, my musical inclinations also led me to poetry. I was less interested in the meaning of verses than in their sound and rhythm. Endowed with an almost absolute memory, I was able to memorize everything that interested me, especially poems, instantly and with total precision." Florensky's writings are so full of musical references, thoughts on music and poetry, that one could write a book on the subject: *Music and poetry in the works of Pavel Florensky.*

Florensky also wrote poetry. Many of his poems were confiscated and are still lost, probably forever, but the volume of correspondence (Florensky 2021) contains an unfinished 60-page poem, ORO, which he wrote during his exile. It is translated into French by Françoise Lhoest.

Regarding sound and music, let me quote a passage from Florensky's correspondence (letter addressed to his mother dated 28–29.VII.1936) (Florensky 2021, Letter no. 70, p. 447):

> An acquaintance of mine asks me why I never write about the sounds of Solovki, but only about colors and shapes. This is because everything here is soundless, like a dream. This is the realm of silence. Not in the literal sense, of course; there is more than enough unwanted noise, and one would want to get away from it. But you cannot hear the inner sound of nature, you cannot perceive the inner words of beings. Everything glides along, in a shadow theater, and sounds are added from outside, like an appendix or a disturbing noise. It is hard to explain why nothing produces sound, why there is no music in things and in life, and I do not really understand, but all the same, there is no music.

Finally, talking about sound and music, I would like to quote another passage, this time from Florensky's autobiography, in which he describes how he experiences, in both a physiological and symbolic way, Chladni's figures, the acoustic lines formed on a plate rubbed by a bow[67] (Fig. 7). He writes (Florensky 2007, p. 66):

> I was feeling almost physically like a string, or rather like a Chladni plate, over which nature passes like a bow: in my soul, or rather in my whole organism vibrated, almost audible to the ear, a pure, high, elastic sound, and in my thoughts were ordered schematic figures, like Chladni figures, like symbols of worldly phenomena. I write and I am almost certain to be misunderstood. People will want to hear in these words comparisons and poetic essays, whereas I want to express, to draw from myself, the soberest and most literal description

[66] Sabaneeff based his opinion on the fact that when the poet Vyacheslav Ivanov introduced Florensky to the world of Aleksander Skriabin, Florensky did not fit in there; see (Sabaneeff 1961, p. 315).

[67] Ernst Florens Friedrich Chladni (1756–1827) was a German physicist, sometimes considered as the founder of modern acoustics. He conducted experiments consisting in making a plate vibrate by scraping it on the side with a bow, after covering it with a layer of fine sand, and observing the patterns the sand produces: geometric figures, always regular, but very different, depending on where and how the plate is excited. These are the *Chladni figures* mentioned by Florensky in the passage.

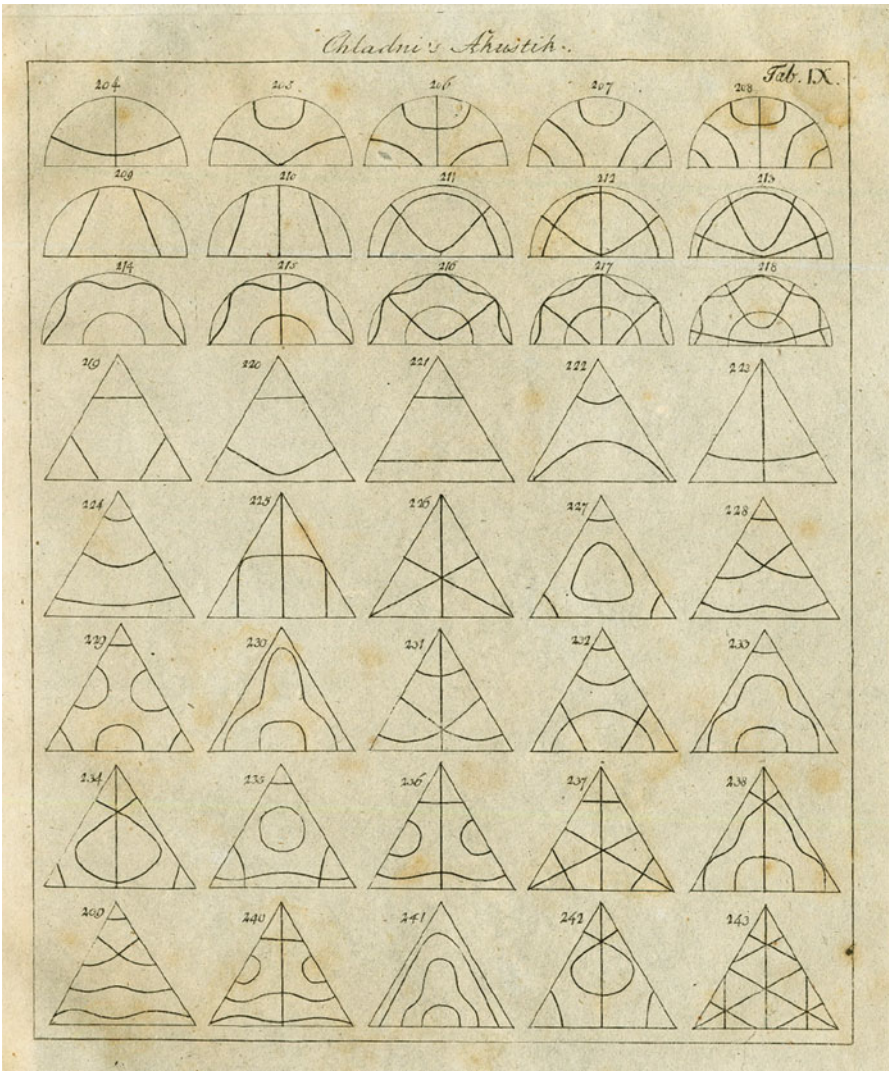


Fig. 7 Chladni acoustic figures, from Chladni's *Akustik* (Chladni 2015)

> possible, a kind of physiological picture. It consisted of this: everything in me, every vein, was full of an ecstatic sound that concentrated within itself *my* knowledge of the world, generating patterns of a rather mathematical order: my categories of knowledge.

Talking about poetry, let me mention Pasternak again. Chentalinsky, in (Chentalinsky 1993, p. 164), quotes him describing Florensky as a "hostage of eternity; captive of time." Arkady Plotnitsky tells me that the line comes from a well-known and much quoted poem from the 1950s by Pasternak, and that this poem is addressed to an artist in general, and not particularly to Florensky. But the poem fits Florensky very well. Let me quote here the corresponding stanza and the one before:

Don't sleep, don't sleep, work
Don't interrupt your labor,
Don't sleep, struggle with a slumber,
Like a plain pilot, like a star.

Don't sleep, don't sleep, artist
Do not submit to sleep,
You are a hostage of eternity,
In time's captivity.

In this section, I have talked about poetry in the Silver Age. In the next section, I will continue with, arguably, the greatest poet of the Silver Age.

9 Florensky and Akhmatova

While the expression "Silver Age of Russian mathematics" is used in reference to the Golden Age that followed it, the expression "Silver Age of Russian poetry" is used in comparison with the Golden Age that *preceded* it, that is, the age of Pushkin and Lermontov. Some consider that the expression "Silver Age of poetry" was coined after certain verses in Akhmatova's poem *Poem without a hero*, namely,

And the silver month is bright
Frozen over the Silver Age.

Akhmatova, who was dubbed "the Soul of the Silver Age," was close to Florensky in many respects, and I will talk about this in the present section. Let me start by quoting a poem she wrote on Dante (Akhmatova 1997):[68]

From death there will be no return
To the old Florence he adores.
And leaving, back forever turned,
For him I sing this song, this lore.
Dark night, fire, torches, last embrace,
Beyond the Gate, the howl of Fate.
From Hell he sent his final curse,
In Heaven Florence he will not forget.
And barefoot, wearing shirt of shame,
With flaming candles lightly carried,
He walks in Florence all the same,
So trech'rous, base, but so desired.

In the middle section of her *Poem without a hero*, Akhmatova has the following epigraph:

[68] The following translation is by Alexei Sossinsky.

. . . a jasmine bush that bloomed
Where once Dante walked and the air is mute.

These two lines by Akhmatova are a variation on (and a truncation of) the following lines in a poem about Akhmatova's silence during several years, by Nikolai Klyuev:[69]

Akhmatova—a jasmine bush,
Scorched by asphalt and gray pitch—
Is she lost on the path through the dark pits
Where Dante passed and the air was parched,
Or is she a nymph spinning flax into crystals?
Among Russian women, remote and subtle
Anna is a cloud at sunset, she reflects
Streaks of gray light like rockets!

Dante was for Akhmatova, like for Florensky, a source of inspiration and reflection. They both regarded the fourteenth century Florence *Podestà*, which condemned Dante to an exile with no return, as a prelude for the twentieth century brutal Soviet regime. The long journey that Dante recounts in the *Divine Comedy* symbolized for Akhmatova all the exiles she had known, people she loved who emigrated to Europe, and those who were sent to the Gulag. Among the latter, there were poets, writers, and her son.

The resemblance with Florensky is striking. Akhmatova had Eurasian origins: Her father, like Florensky's, was an engineer, of Greek descent, and her mother was of Tatar descent. Those who knew her personally described her as a strange woman: shy, gentle, sad, indifferent to worldly things, and with an acute taste for contemplation. All this coincides with what we know of Pavel Florensky. Contradictions were noted in her social behavior. Her first husband, the prominent poet Nikolai Gumilev, called her an enchantress. It was said that she was both a saint and a harlot, and that she had diabolical charisma.[70] Apart from the harlot part of it, Florensky's social image was similar. I already quoted Sabaneeff, who wrote that there was in him "something of black magic, dark, devoid of divine grace." In the same article, we read (Sabaneeff 1961): "Beyond any doubt, there was a demoniac element in him, and as indubitably he was an extraordinary man, an outstanding personality,

[69] The symbolist poet Nikolai Alekseevich Klyuev (1884–1937) was arrested and sent to the camps in 1933 (the same year as Florensky). He was shot to death in 1937 in Tomsk (2 months before Florensky) and rehabilitated posthumously in 1957 (1 year before Florensky). The translation here is from the article (Mager 2005). Although the details would take us far beyond the scope of this chapter, let me mention that I learned from Plotnitsky that Akhmatova's truncation and the change of meaning of Klyuev's passage was likely to be mediated by Mandelstam (with Klyuev's poem still as the original source), and that there is a controversy about this because Akhmatova changed her poem several times; see the article (Leving 2009).

[70] On August 14, 1946, the decree signed by Andrei Zhdanov, who was Stalin's minister of culture, which banned Akhmatova from publishing, accused her (in the present case, with an outright malice) of mixing debauchery with prayer, describing her as a "bourgeois lady gone mad, tearing between her boudoir and her altar . . . half nun, half harlot, or rather both nun and harlot."

quite beyond comparison with any other prominent man of his time... . Lucifer was closer to him than Christ." Sabaneeff adds: "This increased rather than diminished his fascination for me."

Detached from everything, Akhmatova was trying to understand life through poetry. Like Florensky, she refused emigration, even though she had friends in Italy and France. For this, let me quote the opening of her *Requiem*[71]:

> No, not far beneath some foreign sky then,
> Nor with foreign wings to shelter me, —
> I was with my people then, close by them,
> Where my luckless people chanced to be.
> 1961

Akhmatova was neither exiled nor shot like Florensky, but she experienced exile and execution within herself: Her first husband, Nikolai Gumilev, accused of conspiration, was shot in 1921. Her son, Lev, endured the Siberian camps. Her lover, the art historian and critic Nikolai Punin, died in the Stalin camps, and one could continue this gloomy list.[72] We already talked about Florensky's very insecure position regarding publication: He was deprived of the opportunity to publish in scientific journals, and his notes and manuscripts were confiscated. Akhmatova's writings were judged by the authorities, but also by the Writers' Union, as "socially too irrelevant," and her poetry was banned from publication, starting in 1922, for over 30 years. She was led to recite her poems secretly to close friends, asking them to memorize them. Thus, it was her works that were deported instead of her.

Akhmatova had a religious soul. Korney Ivanovich Chukovsky dubbed her "the unique and last Orthodox poet" (Chukovsky 1999).[73] Like Florensky, she was

[71] Transl. Robin Kemball.

[72] An exception to this "gloomy list" is her last lover, Isaiah Berlin, an English diplomat and academician, born in Riga, which belonged to the Russian Empire, who was 20 years her junior and who came to see her in Leningrad while he was completing an internship at the British Embassy in Moscow. Reportedly, when Stalin learned of this affair, he said: "Now the nun is receiving foreign spies at night!" Akhmatova was deeply moved by Berlin, whom she dubbed the "Guest from the future," and to whom she dedicated four cycles of poems. In one of her conversations with him, she declared, "I have only darkness and Rus' by my side" (Victoroff 2016). Berlin was also deeply moved by Anna. On his return to England, he arranged for her to be made Doctor Honoris Causa of Oxford University.

[73] Korney Ivanovich Chukovsky (1882–1969) was a journalist, poet, translator, literary critic, and professor of Russian literature. With Akhmatova, he belonged to the same artistic and intellectual circles, before the revolution. Chukovsky was the illegitimate son of Emmanuel Solomonovich Levinson, whose (legitimate) daughter, Henrietta Emmanuilovna Levenson, was the mother of the outstanding Soviet mathematician Vladimir Abramovich Rochlin (1919–1984). Rochlin was the doctoral advisor of M. L. Gromov, A. M. Vershik, V. G. Turaev, O. Y. Viro, Y. M. Eliashberg, and several other bright mathematicians. Chukovsky's daughter, Lydia Korneyevna Chukovskaya (1907–1996), was a Russian and Soviet author and literary critic. Her second husband, Matvei Petrovich Bronshtein (1906–1938), was arrested as part of Stalin's Great Purge and was shot on February 18, 1938. Bronshtein was an outstanding Soviet theoretical physicist, working on quantum theory, astrophysics, semiconductor theory, cosmology, and quantum gravity theory. His doctoral dissertation, *Quantization of gravitational waves*, defended in 1935, was the first serious work on quantum gravity. He indicated there the main difficulties in this theory. In 1965, Lydia Chukovskaya

steeped in Orthodox culture, even though this culture was forbidden. Her poem *Requiem*, dedicated to "seventeen months waiting in prison queues in Leningrad,"[74] includes whole lines that are variations on Orthodox liturgical texts.

> Do not weep for me, Mother, when I am in my grave.
> A choir of angels glorified the hour,
> the vault of heaven was dissolved in fire.
> Father, why hast Thou forsaken me?
> Mother, I beg you, do not weep for me.

Each line here is taken from a liturgical text; Florensky has a poem on the same theme, titled *Lament of the Mother of God*, which he wrote in 1907; see (Oppo 2022, p. 89]. Akhmatova's poem continues:

> Mary Magdalene beat her breasts and sobbed,
> His dear disciple, stone-faced, stared.
> His mother stood apart. No other looked
> into her secret eyes. No one dared.

This is a faithful description of the Orthodox icon of the crucifixion.

Akhmatova was a twin sister of Florensky. The comparison and the close relationship between them could be the subject of another book.

There would still be a third book to write, about the relationship between Florensky and another twin sister: Simone Weil. In this respect, I would like to quote a passage from her book *La pesanteur et la Grâce* (1947), the chapter titled *Mystique du travail* (Mysticism of work) (Weil 1948): "The spirituality of work. Work makes us experience in a harassing way the phenomenon of finality bouncing back like a ball; working to eat, eating to work. . . If we look at one of the two as an end, or at both taken separately, we are lost. The cycle contains the truth."[75] This passage also contains another incarnation of the idea of eternal return, dear to Florensky, which I discussed in Sect. 5 above. Among the strong ties between Simone Weil and Pavel Florensky are a love of mathematics, a strong sense of self-observation, a passion for history and especially for Greek antiquity, extreme mysticism, and an intense religious feeling.[76]

published a novel, *Sophia Petrovna*, in which the son of the main character, Sophia Petrovna, was imprisoned in Leningrad's jail and later shot. In fact, Sophia Petrovna repeated in many respects the fate of Lydia Chukovskaya. (I learned most of the content of this note from V. N. Berestovskiĭ.)

[74] The dedication is to the people waiting for hours, often with no result, in this mile-long line to get any information about their relatives in prison or in a Gulag. Akhmatova was among these people waiting.

[75] Spiritualité du travail. Le travail fait prouver d'une manière harassante le phénomène de la finalité renvoyé comme une balle; travailler pour manger, manger pour travailler. Si l'on regarde l'un des deux comme une fin, ou l'un et l'autre pris séparément, on est perdu. Le cycle contient la vérité.

[76] I would like to thank S. Tanabé who suggested to me several relations between Pavel Florensky and Simone Weil.

10 Contradiction and Antinomy in Mathematics, Philosophy, and Theology

So far, I have tried to allude to Florensky's philosophy, and I shall try to say more about it in this section.

The Russian Silver Age, besides referring to mathematics, poetry, and art, was a period of fusion between philosophy and theology. The representatives of this alliance include Sergei Nikolaevich Bulgakov (1871–1944), Lev Isaakovich Shestov (1866–1938), Aleksei Fedorovich Losev (1893–1988), Nikolai Aleksandrovich Berdyaev[77] (1874–1948), and Dmitry Sergeyevich Merezhkovsky (1866–1941). The latter was one of the founders of the symbolist movement, to whom Florensky belonged. The others are also connected to this movement in various ways.

Florensky's philosophical worldview is based on antinomy, and I will mostly talk about this now.

Florensky considered that the world of ideas is based on antinomy rather than on logic as one intends it usually; in fact, truth, for him, is grounded on contradiction, and this holds in particular for the truths of theology. He considered that contradiction and paradoxes are unavoidable in theology. He believed that religious truth is beyond any formal logic, and more generally, beyond any formal categorization. He considered that the Holy Scriptures, including the Old and New Testament, contain contradictory statements that are nevertheless all true, and that these contradictions do not need to be resolved but on the contrary they are an integral part of Christian teaching. In Letter VI of the *Pillar*, he reviews some sources of antinomy, in Greek Antiquity, passing by Heraclitus, the Eleatics, and Plato. To understand Florensky's background, it is good to recall first a few facts about contradiction in mathematics, the first field that he studied in depth.

At the beginning of the twentieth century, at the time when Florensky was suffering his existential crisis, mathematics was traversing a *foundational crisis*. A strong indication of this crisis is Hilbert's "second problem," formulated in 1902 at the second congress of mathematicians in Paris, asking whether one can prove the coherence of arithmetic, i.e., whether one can demonstrate that the axioms of arithmetic are not contradictory. In reality, it was a question not only about arithmetic but also of geometry, and, by extension, of all mathematics, because it was shown before (notably, by Hilbert) that the coherence of the axioms of geometry depended on the coherence of the axioms of arithmetic. For the answer to Hilbert's Problem II, one had to wait for the work of Kurt Gödel, who demonstrated that the answer is negative. Gödel published his famous *Incompleteness Theorem* in (Gödel 1931), while Florensky was already in the glaciers, far away from it all. In any case, mathematicians who had thought about these questions had always felt that there might be problems of coherence in formal logic, and this had been so since Greek antiquity. (Even though formal logic did not exist yet, it was formally the same logic.) The paradoxes of Zeno

[77] It has been suggested that the term "Silver Age," in relation with Russian philosophy, derives from Berdyaev's work *The Brightest Lights of the Silver Age* (Berdyaev 2015).

of Elea are generally quoted in this respect. Let me recall here that Zeno was considered by Aristotle as the father of dialectic, that is, the art of asking questions and defending successively opposite theses with the aim of finding the truth. Florensky mentions Zeno in the *Pillar* (see p. 345 and p. 574), in relation with the continuous and the discontinuous, with references to his paradoxes of motion. He quotes classical paper on Zeno's aporias, by the French mathematician and historian of science Paul Tannery (Tannery 1885). As a matter of fact, Zeno's arguments are not contradictory; they are only paradoxical, and based on the paradoxes that are possible in mathematics. But these paradoxes are presented in Florensky's writings in a philosophical language that makes them appear practically contradictory. Regarding antinomy in theology, Florensky refers to the philosophical system of the monk Serapion Mashkin, from the Optina monastery.[78] In Note 837 (p. 574) of the *Pillar*, talking about Zeno's antinomies, Florensky writes: "It is precisely on these antinomies, and their overcoming that the entire philosophical system of Father Serapion is built." Antinomies, but also discontinuity: In the same page (note 836), Florensky writes that "one cannot fail to mention Father Serapion Mashkin's attempt to conceive space and time as composed of finite, further-indivisible elements."[79]

Mathematics since at least Euclid is based on definitions and postulates/axioms. The axioms are assumed to be internally consistent, not containing a contradiction. In case an axiomatic system contains a contradiction, the simple rules of logic (*Modus Ponens*) imply that every statement formed in the language of the system is a formal consequence, thus rendering the system most powerful but useless and mathematically unacceptable. Consequently, if a contradiction is detected within a system, then the system is rejected. Frege's foundation of set theory was famously found to be contradictory because it contained the rule of comprehension, from which Bertrand Russell constructed his paradox. However, it is a consequence of Gödel's Incompleteness Theorem (Gödel 1931) that internal consistency cannot be proved within the system if the system is reasonably strong. Let me refer here to a paper by Farmaki and Negrepontis, titled *The paradoxical nature of Mathematics*

[78] Father Serapion was a former naval officer. The name *Optina* refers at the same time to a small village and a monastery, in the Kaluga region, 250 km to the South-West of Moscow. The monastery is linked to episodes in the lives of several Russian writers and philosophers. In the summer of 1878, Vladimir Solovyov brought Fedor Dostoyevsky to the monastery following the tragedy of the latter's son's death. Dostoyevsky stayed at the monastery for 3 days. Several details of his novel *The Brothers Karamazov* appeared following this visit to Optina. The list of the other visitors of the Optina monastery includes Vasily Zhukovsky, Nikolai Gogol, Ivan Turgenev, Vasily Rozanov, and Leo Tolstoy. After the October Revolution, Soviets closed the Optina monastery and the monks of the monastery were asked to choose between deportation in the Solovki Archipelago or working in a secular agricultural cooperative. The monastery itself was declared a Gulag, and various kinds of prisoners were sent there. After these events, Florensky launched a campaign to "Save Optina"; see (Demidov and Ford 2005, p. 607).

[79] For a modern treatment of Zeno's aporias and of contradiction in mathematics from the historical point of view, let me mention the paper *Zeno's arguments and paradoxes are not against motion and multiplicity but for the separation of true Beings from sensibles* (Negrepontis 2023) by S. Negrepontis.

(Farmaki and Negrepontis 2021), in which these authors stand that in a sense all of Mathematics can be considered as a (hopefully) consistent approximation of the truly contradictory. It is not clear whether Florensky has in mind, even in some very intuitive sense, such a view of Mathematics; it would indeed be fascinating if he did.

In any case, the fact that *truth* is antinomic is one of the main themes of *The pillar and ground of the truth*. Florensky explained quite simply in his work that reason alone is sufficient to accept or not accept what is not antinomic, and that if faith is needed, it is precisely because truth is antinomic. Letter VI of the *Pillar* is titled *Contradiction*. In this letter, Florensky analyzes different occurrences of antinomy, in philosophy, theology, logic, mathematics, ethics, linguistics, and other fields. In short, he develops the thesis that contradiction and antinomy are the basis of knowledge. As a matter of fact, each of Florensky, Bulgakov, and Losev developed elaborate philosophical and theological systems based on antinomy.

At the time he was writing his book, Florensky was witnessing a renaissance in Russian theology, which found its sources in the Jewish Kabbalah and in the writings of the Desert Fathers, a theology which came to be known as *apophatic* theology, or *negative* theology. According to this theology, every time one tries to say something positive about God and the transcendent is inevitably led to contradictions or absurdities. Here, let me quote again Sabaneeff on Florensky:

> His mind was complex, many-storied, and to some extent even hostile to simplicity. One might even call it a cabalistic mind. . . . His extravagant and excessively luxuriant thoughts often contradicted one another, but this did not embarrass him in the least. . . . He even asserted that every perceived law "generates" its own negation, inasmuch as every law discerned by the logical apparatus of man is but a part of a real, comprehensive synthetic law embracing "all that exists" and "all that is possible and thinkable."

Let me talk now about the notion of infinity.

Florensky was fascinated by the idea of actual infinity which was highlighted by Cantor, and he saw in it a symbolic vision of God. He proposed a philosophical and symbolic interpretation of this concept based on Bugaev's mathematical theory of discontinuity. In 1904, he published the pamphlet titled *Symbols of infinity: Sketch of the ideas of G. Cantor*, which we already mentioned, dedicated to the analysis of Cantor's ideas on set theory and infinity. According to I. H. Anellis, Florensky's essay was the first article in Russia written on Cantor's set theory (Anellis 1994). In some sense, Florensky was a disciple of Cantor, who had already declared in his 1883 work, *Grudlagen einer allgemeinen Mannigfaltigkeitslehre*, that everything that has to do with the infinite has theological, mathematical, and philosophical connotations that cannot be separated. N. Misler writes in (Misler 2006, p. 35) that Florensky's proposed philosophy of symbol was in accordance with the wish of the editors of *Novy put'* to promote the relation between the intelligentsia and the Church. The logic of symbols is used by Florensky in his theological arguments. In his writings, the issues about contradictions in formal logic become issues about contradictions in theology. He declares in Letter VI of the *Pillar* (Florensky 1975, p. 106): "Truth is an antinomy. . . . . What is needed is a formal logical theory of

antinomy," and he goes on explaining a way to obtain such a theory. In Letter X, titled *Sophia*, he declares: "faith clothes its knowledge of dogmatic truth in a symbolic garment, in figurative language, which covers the higher truth and depth of contemplation in consistent contradictions" (Florensky 1975, p. 244). In Letter II, titled *doubt*, he writes that "[the rational mind] desires self-contradiction" (Florensky 1975, p. 17) and that "contradiction is everywhere and always" (Florensky 1975, p. 23). We can read in the same letter: "At the same time we get: A is A; A is not A" (Florensky 1975, p. 29). In Letter VI, titled *Contradiction*, he writes: "For rationality, truth is contradiction, and this contradiction becomes explicit as soon as truth acquires a verbal formulation. . . . truth is an antinomy, and it cannot fail to be so." (Florensky 1975, p. 109).

Florensky's references include Heraclitus, Zeno, Plato, and Origen. Regarding Plato, Florensky writes: "The majority of his dialogues are nothing but a gigantic antinomy, developed with all care and artistically dramatized. Plato's very predilection for the dialogic form of exposition, i.e., the form of the contraposition of convictions, already hints at the antinomic nature of his thinking" (Florensky 1975, p. 116).

Let me make a remark here. It is difficult to be certain of what Florensky had in mind when he referred to Plato. According to Stelios Negrepontis,[80] one might reasonably assume that Florensky had in mind, at least in part, Plato's most contradictory-looking statements of the dialogue *Parmenides*, especially in the second hypothesis (*Parmenides* 142–155). These statements are indeed scandalously antinomic/contradictory in their appearance. For instance, Plato's intelligible Being (i) is infinite and finite, (ii) is One and Many, (iii) is in motion and at rest, etc., and even worse: is older and younger than itself and the others, and is neither older nor younger than itself or the others! However, it would be a rush conclusion that these statements are in fact contradictory. They have a fully rational and mathematically consistent interpretation. According to an interpretation proposed by Negrepontis, Plato exploits in these passages the paradoxical, but fully consistent, nature of the infinite but periodic anthyphairesis and describes it employing what might be considered an abuse of language, to cause surprise and amazement. But Plato is always safe within consistent noncontradictory mathematics. The same remarks apply to Zeno's scandalous and fascinating paradoxes since they are of the same nature as Plato's paradoxes, and in fact form the primary source for Plato's paradoxical arguments. I refer the reader to the interpretation of Zeno's paradoxes by Negrepontis in (Negrepontis 2023). In the same page (Florensky 1975, p. 116), Florensky declares: "Knowledge of contradiction and love of contradiction, along with ancient skepticism, appear to be the highest achievement of antiquity." In the same chapter, he gives a detailed list of dogmatic antinomies. "The Holy Scripture, he says, is full of antinomies." The Apostle Paul's epistle to the Romans, he writes, is an "antinomy-charged bursting bomb against the rational mind." The paradoxical method of reasoning is justified by symbolic logic (Florensky 1975, p. 112). In

[80] Personal communication.

Chapter XIV of the *Pillar*, he writes: "There are an uncountable number of antinomies. There are as many of them as there can be acts of the rational mind. But, as we have already noted, antinomies are essentially reducible to the dilemma: 'finiteness or infiniteness'" (Florensky 1975, p. 344). We are back to mathematics. We already quoted Lebesgue saying that there might exist a set which is nameable and which is neither finite nor infinite (Footnote 55.)

In Letter VI (*Contradiction*), Florensky makes a list of 11 theological antinomies.[81] In this list, we find for example:

- *Retribution* – Thesis: Retribution applied to all according to their works (Rom. 2: 610; 2 Cor 5:10). Antithesis: Free forgiveness of the redeemed (Rom. 4:4, 9:11, 11:6).
- *Grace* – Thesis: "Where sin abounded, grace did much more abound" (Rom. 5: 20). Antithesis: "Shall we continue in sin, that grace may abound? God forbid" (Rom. 6:12). "If we say that we have no sin, we deceive ourselves, and the truth is not in us" (1 John 1:8).
- *Faith* – Thesis: Faith is free and depends on the free will of man (John 3: 1628). Antithesis: Faith is God's gift and is not found in human will but in the will of the Father Who draws us to Christ (John 6:44).

There are eight other such theological themes in his list, with antinomic approaches. I refer the interested reader to the detailed list together with Florensky's commentary in p. 121–123 of the *Pillar*.

Antinomy is found in Yahweh, the supreme order of the world. The God of the Old Testament, who punishes severely and looks very cruel, has two entirely antinomic faces. Florensky writes in Chapter VI of the *Pillar*: "The Book of Job wholly consists of such a concentrated experience of contradiction. This book is wholly constructed on the idea of antinomialness."[82] C. F. Jung, in his book *Answer to Job* (Jung 1952), reached conclusions similar to those of Florensky.[83]

What can we conclude?

Mathematics is the most exact science, and with the actual axioms, there is no proof that there is no contradiction in mathematics. If ever we find contradictions in mathematics, then there would be contradictions in all the other sciences. In theology, we already live with contradictions. This is expressed in a most dramatic style in *The Pillar and ground of the truth*, which is Florensky's major theological written work. The *Pillar* is a treatise on contradiction and antinomies.

From the more recent period, let me mention that the French philosopher and epistemologist Jules Vuillemin (192–2001), who was a professor at Collège de France, dedicated his course there, during 7 years (in the 1960s–1970s), to the

[81] "I deliberately exclude certain antinomies that I intend to examine in a future book," he writes.

[82] I am using the wording of the English translation (Florensky 1975).

[83] I owe this remark to S. Tanabé.

notion of mathematical antinomy,[84] after which he wrote a book on antinomy in the concept of God (Vuillemin 1971).

11 The Philosophers

There is a famous canvas, owned by the Moscow Tretyakov Gallery, representing Pavel Florensky with the philosopher Sergei Bulgakov, painted in 1917 by Mikhail Nesterov (1862–1942), a brilliant representative of the Russian symbolist movement. The painting is known under the name *The philosophers*. It is represented here in Fig. 8. The quiet landscape in the background is the countryside near Sergiev Posad. The contrast between the two philosophers is manifest in their face expressions and in the colors of their clothes. Bulgakov became priest in 1918 and was expelled from Russia, together with 160 other intellectuals, in 1922, on one of the *Philosophers' boats*. Florensky decided to stay in Russia. Nicoletta Misler wrote a remarkable article on this painting in the Tretyakov Gallery Magazine to which I refer the reader (Misler 2007). In this article, she quotes a passage by Bulgakov, from a Russian version of (Bulgakov 1989). I reproduce here an extended passage which I translated from the French version:

> A recollection, and with it the foreshadowing of events and achievements to come, never leaves me. It is the portrait of the two of us, painted by our common friend Nesterov on an evening of May 1917 in Father Paul's garden. For the artist, it was not just a portrait of two friends by a third, but a spiritual vision of the times. For the painter, the two faces expressed the same understanding, but for one as a vision of horror, for the other as peace, the joy of overcoming hardship. The painter himself had doubts as to whether he should portray the *first* character, and tried to remake the portrait by replacing the horror with idyll and the tragedy with placidity. But he immediately felt the unbearable falseness of this substitution, and returned to his first vision. On the other hand, he immediately found Father Paul's face, its artistic and spiritual evidence, which he had no need to modify. This was the clear artistic vision of two figures of the Russian apocalypse, of one and the other face of earthly existence, the first in struggle and disarray (and in my heart it was about the fate of my friend), the second in the victorious fulfillment we are now contemplating.

[84] The titles of Vuillemin's courses at the Collège de France include the following: 1962–1963: La notion d'antinomie: analyse des doctrines de Kant et Hegel; 1963–1964: La notion d'antinomie: l'origine des antinomies dans les mathématiques modernes; 1964–1965: La notion d'antinomie: les solutions de Russell en 1905; 1965–1966: Sur la notion d'antinomie: Découverte des antinomies "épistémologiques"; 1966–1967: Sur la notion d'antinomie (1890–1910): Conséquences des antinomies pistmologiques pour la conception du principe d'induction complète, logicisme, formalisme, intuitionnisme; 1967–1968: Sur la notion d'antinomie: Analyse de l'article de Russell "Mathematical Logic as based on the Theory of Types"; and 1968–1969: Sur la notion d'antinomie: Analyse de Russell: Préface de la première édition des "Principia mathematica" (1910), "On the Relation of Universals and Particulars" (1911–1912).

Fig. 8 Pavel Florensky and Sergei Bulgakov, by Mikhail Nesterov, painted in 1917, Tretyakov Gallery, Moscow

12 Reverse Perspective Again

Andrei Tarkovsky was close to Florensky, philosophically, spiritually, and also in his way of making films. Beatrice Barbalato writes about Tarkovsky: "Inspired by Florensky's thinking (*Reverse Perspective*, 1916), his filmic visual focuses radiate from the minute detail, both as seen in a dream and as occurs in experiential reality" (Barbalato 2021). Jaromír Blažejovský writes, on the same subject (Blaejovsk 2008):

> Tarkovsky used in what we have called—following Pavel Florensky—"reverse perspective," typical for Russian Orthodox icons: narration from various points of view and completed from relatively independent parts of time and space. The effect of this pattern is a feeling of infinity of the narrative world.

The Trinity icon – Abraham's hospitality – which is shown in the film's finale, is an epitome of inverted perspective. I have reproduced in Fig. 9 two paintings, in which the drawing is done in reverse perspective. In such a portrait drawing, the back of the face appears larger than usual, as does the hair.

Fig. 9 Two paintings drawn in reverse perspective. Left: The Archangel Michael, by Andrei Rublev, ca. 1414, egg tempera. Tretyakov Gallery, Moscow. Right: Portrait of Pavel Florensky, by Vladimir Alexeyevich Komarovsky. Oil on canvas, 1924. The Pavel Florensky Museum, Moscow

Tarkovsky quotes Florensky at several occasions. In chapter IV of his book (Tarkovsky 1987), titled *Cinema's destined role*, he makes some observations on reverse perspective in Florensky's writings, pointing out relations with his own works:

> I am reminded of a curious observation of Father Pavel Florensky's in his book, *The iconostasis*. He says that the inverted perspective in the works of that period was not the result of Russian icon-painters being unaware of the optical laws which had been assimilated by the Italian Renaissance, after being developed in Italy by Leon Battista Alberti. Florensky argues, convincingly, that it was not possible to observe nature without discovering perspective, it was bound to be noticed. For the time being, however, it might not be needed—it could be ignored. So the inverted perspective in ancient Russian painting, the denial of Renaissance perspective, expresses the need to throw light on certain spiritual problems which Russian painters, unlike their Italian counterparts of the Quattrocento, had taken upon themselves. (One account has it, incidentally, that Andrei Rublev had actually visited Venice, in which case he must have been aware of what Italian painters had been doing with perspective.)

Florensky writes in his book on reverse perspective, regarding the children's drawings (Florensky 1992, p. 219 of the English translation):

> The children's drawings, in their lack of perspective and especially their use of reverse perspective, vividly recall mediaeval drawings, despite the efforts of educators to instil in children the laws of linear perspective. It is only when they lose their spontaneous relationship to the world that children lose reverse perspective and submit to the schema with which they have been indoctrinated

In my article on Manin which appears in the present Handbook (Papadopoulos 2023), I mention Tarkovsky several times, especially from his diary (Tarkovsky 1993). Reverse perspective appears on several occasions in Manin's writings. It describes a world where there is no predetermined logic. Manin, like Florensky, notes that children draw, before they are formatted by the adults' reason, in reverse perspective. I refer to chapter III.4 of the French version of his book *Mathematics as a metaphor* (Manin 2007), titled *Spontaneous artistic activity, the origin of logograms and mathematical intuition*.[85] Let me add that the Swiss biologist, psychologist, logician, and epistemologist Jean Piaget (1896–1980) has an article titled *Le développement des perceptions chez l'enfant* (The development of perception in children) (Piaget 1960) whose first section is titled *Structures projectives (perspective)*. Piaget discusses there the evolution of the notion of perspective in child development. He states in particular that, from the study of children's drawings, it appears that they are unable to reproduce perspective until around the age of 9 or 10, which is the stage when children draw "what they know about an object, not what they see of it."[86]

Finally, let me mention a theory developed by R. K. Luneburg (Luneburg 1947) according to which objects in the region of binocular vision (which differs substantially from monocular vision only for nearby objects) are perceived in terms of Lobachevsky geometry. This is developed by B. V. Rauschenbach in his book (Rauschenbach 1980) and in his paper (Rauschenbach 1983) titled *On my concept of perceptual perspective that accounts for parallel and inverted perspective in pictorial art*.[87]

13 In Guise of a Conclusion

I wrote at the beginning of this chapter that Florensky has been compared to Leonardo and Pascal. The comparison is amply justified. I would like to add to this list two mathematicians of recent times, René Thom and Yuri Manin, for whom, like Florensky, mathematics, besides being the language of nature, is that of philosophy, linguistics, poetry, and magic. To conclude, let me quote some verses of a poem by Manin titled *Péniches sur le Rhin (élégie)* in which he mentions Florensky's name (Manin 2007, p. 523):

[85] The French version of this book is much more complete than the English.

[86] I learned about Piaget's article by S. Tanabé.

[87] I learned this from V. Berestovskiĭ. Rauschenbach's theory is commented on and used in Berestovskiĭ's paper on Lobachevsky geometry (Berestovskiĭ 2022).

Dieu est inconnaissable mais nommable,
Le nom de Dieu, c'est Dieu,
croyaient les adorateurs du Nom.
Le Starets Hilarion, Pavel Florensky ...
L'infini non plus n'est pas connaissable, mais il est nommable.
Qui connaît les péniches ? Qui leur donne un nom ?
On ne sait pas si c'est Lui — si c'est Elle ...

Acknowledgments I would like to thank V. N. Berestovskiĭ, Stelios Negrepontis, Ken'ichi Ohshika, Victor Pambuccian, Arkady Plotnitsky, Valentin Poénaru, Galina Sinkevich, Alexei Sossinsky and Susumu Tanabe for useful comments and suggestions on previous versions of this chapter. This work is supported by the lnterdisciplinary Thematic lnstitute CREAA, as part of the ITl 2021–2028 program of the University of Strasbourg, CNRS and Inserm (ldEx Unistra ANR-10-IDEX-0002), and the French lnvestments for the Future Program.

References

Akhmatova A (1997) Poems of Akhmatova, transl. S. Kunitz and M. Hayward. Houghton Mifflin Company, Boston

Anellis IH (1994) Translator's introduction: Father Pavel Florensky. Modern logic 4(2):157–161

Barbalato B (2021) Autour de Le Miroir de Tarkovsky: des rêves ou de la perspective inversée. In: Chehab M, Barbalato B (dir.), Auto/biographie: prémonitions, rêves, cauchemars, in Mnemosyne o la costruzione del senso, n. 14, PUL-Presses universitaires de Louvain

Bayuk DA, Ford CE (2003) Dante's cosmology revisited, Archives Internationales d'Histoire des Sciences, Mémorial Adolf Youschkevitch 58(160–161):69–88

Berdyaev N (2015) The brightest lights of the silver age, essays on Russian religious thinkers, transl. B. Jakim, New edition. Angelico Press/Semantron

Berestovskiĭ VN (2022) Lobachevsky geometry and stellar parallaxes. Siberian Mathematical J 63(5):834848

Bersanelli M (2016) From Dante's universe to contemporary cosmology. Rend Scienze Ins Lombardo 150:147–165

Blažejovský J (2008) Spirituality in cinema. PhD dissertation, Masaryk University, Faculty of Arts, Brno

Bourbaki N (2007) Eléments d'histoire des mathématiques, 1st edition, 1984, new edition. Springer, Berlin/Heidelberg/New York

Bugaev NV (1898a) Les mathématiques et la conception du monde au point de vue philosophie scientifique. Verhandlungen des ersten internationalen Mathematike-Kongresses, Leipzig, pp 206–223

Bugaev S (1898b) Mathematics and the Scientific-Philosophical Worldview, Vop. Fil. i Psikh., yr. IX vol. V (45) (Nov.-Dec.), p. 697–717

Bulgakov S (1989) Le père Pavel Florensky, Allocution l'Institut St Serge, Paris, April 11, 1943. Published in: La perspective inversée [39]. Russian text in: Sviashchennik Pavel Florensky, Slovo, no. 12, p. 29

Chentalinsky V (1993) La parole ressuscitée: Dans les archives littéraires du K.G.B., translated from the Russian by G. Ackerman and P. Lorrain. Robert Laffont, Paris

Chladni E (2015) Die Akustik, Breitkopf und Härtel, Leipzig, 1802. English translation by Robert T. Beyer: Treatise on acoustics. Springer

Chukovsky KI (1999) Anna Akhmatova. Kniga, Moscow

Delaunay N (1898) Sur les surfaces n'ayant qu'un côté sur les points singuliers des courbes planes. Bull Soc Math France 26:43–52

Demidov SS (1988) On an early history of the Moscow school of theory of functions. Philos Math 1:29–35
Demidov S (1993) The Moscow School of the theory of functions. In: Zdravkovska S, Duren P (eds) The golden years of Moscow mathematics. American Mathematical Society
Demidov SS (1999) Dmitrii Fedorovich Egorov, professor of Moscow State University, and the doctrine of the veneration of the name of God in Russia during the first third of the 20th century (Russian), Istor.-Mat. Issled. (2) No. 4 (39), pp 123–156
Demidov SS (2014) The origins of the Moscow school of the theory of functions, Fundamental Science, Technical transactions, pp 73–84
Demidov S (2021) Nikolai Nikolaevich Luzin at the crossroads of the dramatic events of the European history of the first half of the 20th century. Studia Historiae Scientiarum 20:317–335
Demidov SS, Ford CE (2005) On the road to a unified world view: Priest Pavel Florensky—Theologian, Philosopher and Scientist. In: Koestsier T, Bergmans L (eds) Mathematics and the divine: a historical study. Elsevier, Amsterdam, pp 595–612
Demidov S, Lëvshin BV (eds) (2016) The case of Academician Nikolai Nikolaevich Luzin. Russkii Khristianskii Gumanitarnyi Institut, St Petersburg, 1999. Trans. R. Cooke, American Mathematical Society, History of Mathematics series, 43
Demidov SS, Parshin AN (eds) (1986) Publikatsiia i primechaniia: Florenskii P. A. Vvedenie k dissertatsii "Ideia preryvnosti kak element mirosozertsaniia". Istoriko-matematicheskie issledovaniia 30:159–176
Demidov SS, Parshin AN, Polovinkin SM, Florensky PV (eds) (1989) Correspondence between N. N. Luzin and P. A. Florensky (1904–1922), in Russian. Publication and notes by S. S. Demidov, A. N. Parshin, S. M. Polovinkin and P. V. Florensky. Istoriko-Matematicheskiye Issledovanyia 31:116–191
Dobretsov NL, Marchuk GI, Privitz NA, Ermikov VD, Ibragimova ZM (eds) (2000) The age of Lavrentieff, Novosibirsk. Siberian division of the Russian Academy of Science Publishing House, Collection Geo. 456 p
Dreyer JLE (1921) The cosmology of Dante. Nature 107:428–430
Egginton W (1999) On Dante, hyperspheres, and the curvature of the medieval Cosmos. J Hist Ideas 60(2):195–216
Egorov DF (1911) Sur les suites de fonctions mesurables. Comptes Rendus Acad Sci Paris 152: 244–246
Eliot TS (1951) Selected essays, 1917–1932. Faber & Faber, London
Eriksson H (2009) Science cannot be grasped by the intellect. Bachelor thesis, Uppsala University, Department of Modern Languages
Farmaki V, Negrepontis S (2021) The paradoxical nature of mathematics. In: Papadopoulos A (ed) Topology and geometry. A collection of essays dedicated to Vladimir G. Turaev, IRMA Lectures in Mathematics and Theoretical Physics 33. European Mathematical Society, Berlin, pp 599–642
Florensky PA (1903) Superstition. Novyi Put 8:91–121
Florensky PA (1904a) Symbols of infinity: sketch of the ideas of G. Cantor, Novyi Put', pp 191–192
Florensky PA (1904b) On a certain premise of a worldview, Vesy, No. 9
Florensky PA. The pillar and ground of the truth. An essay in orthodox theodicy in twelve letters, Moscow, Put, 1914 (Russian), French translation: La colonne et le fondement de la vérité, translated by C. Andronikov, L'Âge d'homme, Lausanne, 1975. English edition, translated and annotated by B. Jakim, with an introduction by R. F. Gustafson. Princeton University Press, Princeton and Oxford, 1997
Florensky PA (1989) Sketch for a discourse for the inauguration of the Circle of mathematics students of the Moscow Mathematical Society (Russian). Edited by S M Polovinkin and P V Florensky Istoriko-Matematicheskiye Issledovanyia 31:116–125
Florensky PA (1992) La perspective inversée, suivi de L'Iconostase et autres écrits sur l'art, transl. Françoise Lhœst, L'Âge d'homme, Lausanne. English translation in [78]
Florensky PA (2006a) Hamlet, translated from the Russian by Evdokiya Sichov, Ed. Allia, Paris

Florensky P (2006b) Hamlet (extract), Preceded by: Out of Joint, E. Ghezzi (Preface) E. Sichov (Transl.) Allia, Petite Collection, Paris
Florensky P (2007) Souvenirs d'une enfance au Caucase, Original title in Russian: Detjam moim, vospominanija proshlyh dnej, (To my children, remembrances of the past days) (published Moscow, 1992), translated from the Russian by Françoise Lhoest. L'Âge d'homme, Paris/Lausanne
Florensky P (2012) Stupeur et dialectique, N. Valentini (Preface), E. Ponomareva (Translation), Payot et Rivages, Paris
Florensky PA (2016) Les imaginaires en géométrie. Trad. du russe par Françoise Lhoest et Pierre Vanhove. Préface de Cédric Villani. Zones Sensibles. English translation : Imaginaries in geometry, edited by Andrea Oppo and Massimiliano Spano, translated by Anna Maiorava and Karen Turnbull, Mimesis International, 2021
Florensky P (2021) Lettres de Solovki, 1934–1937, transl. and presentation by F. Lhoest, Lausanne, L'Âge d'homme, 2012. coll. Classiques slaves. English version: Letters from Russia's Far East Pavel Florensky, transl. Adela Lawless, Amazon Digital Services LLC - KDP Print US
Florensky P (2022) La nature magique de la parole, R. Nicolas (Preface and translation), Rivages, Paris
Ford C (1998) The influence of P. A. Florensky on N. N. Luzin, Historia Mathematica 25, 332–339. Reprined in Russian mathematicians in the 20th century, ed. Y. Sinai, World Scientific, 2003
Ghys E (2017) A singular mathematical Promenade. ENS ditions, Lyon
Gödel K (1931) Über formal unentscheidbare Satze der Principia Mathematica und verwandter Systeme I, Monatshefte fur Mathematik und Physik 38, p. 173–198
Godet F (2002) Commentaire sur l'épître aux Romains, first ed. 1883, New printing, Impact
Graham L, Kantor J-M (2009) Naming infinity: A true story of religious mysticism and mathematical creativity. Belknap Press of Harvard University Press
Grigorian AT. Zhukovsky, Nikolay Egorovich. In: Dictionary of scientific biography, New York, 1970–1990
Hadamard J (1905) Cinq lettres sur la théorie des ensembles. Bulletin de la S M F, t 33:261–273
Hoffman J (2009) Pursuing the infinite (a review of Graham and Kantor's book *Naming infinity*). Nature 458:971–972
Ivanov V (1907) O veselom remesle i umnom veselii, Zolote Runo, no. 5
Jung CG, Antwort auf Hiob, 1952, English transl. by R. F. C. Hull, Answer to Job, In: The Collected Works of C. G. Jung 11. Princeton University Press, Princeton, 1973
Katasonov VN (2009) Review of Graham and Kantor's Naming Infinity. Vestnik Pravoslavnogo Svyato-Tikhonovskogo Gumanitarnogo Universiteta. Seriia 1: Bogoslovie. Filosofiia Religiovedenie 3(27):133–140
Kolyagin YM, Savvina OA (2010) Dmitri Fedorovich Egorov. The path of a scientist and a Christian (Russian). PSTGU, Moscow
Kuznetsov PI (1971) Dmitri Fedorovich Egorov. Russ Math Surv 26(5):125–164
Lakhtin LK (1904) A note on one-sided surfaces (Russian). Mat Sb 24(2):178–193
Lavrentieff M (1974) Nikolai Nikolaevich Luzin (on the 90th Anniversary of his birth). Russ Math Surv 29(5):173–178. Translation of an article that appeared in Russian, Usp. Mat. Nauk 29, No. 5(179), p. 177–182 (1974). The content is part of Lavrentieff's memoirs published in Dobretsov NL et al. (2000)
Lebesgue H (1903) Sur une propriétié des fonctions. Comptes Rendus Acad. Sc. Paris 137:1228–1230
Lebesgue H (1904) Leçons sur l'intégration et la recherche des fonctions primitives, 2nd edn. GauthierVillars, Paris
Leving Y (2009) Whose is the seal-ring? Kliuev's subtexts in Mandelstam's poem "Give Tiutchev a dragonfly", The Slavic and East. Eur J 53(1):41–64
Lipscomb SL (2011) Art meets mathematics in the fourth dimension, 2nd edn. Springer
Lossky NO (1926) The successors of Vladimir Soloviev (Russian), Put' 2, No. 3, p 17
Lossky NO (1952) History of Russian philosophy. George Allen and Unwin Ltd, London
Luneburg RK (1947) Mathematical analysis of binocular vision. Princeton University Press, Princeton

Luzin NN (1912) Sur les propriétés des fonctions mesurables. Comptes Rendus Acad. Sc. Paris 154:1688–1690
Luzin NN (1914) Sur un problème de M. Baire. Comptes Rendus Acad Sc Paris 158:1258–1261
Luzin N (1926) Mémoire sur les ensembles analytiques et projectifs. Matem Sbornik 33(3): 237–290
Luzin N (1929) Sur les voies de la théorie des ensembles. In: Proceedings of the ICM Bologna, 3–10 Sept. 1928, Bologna, Nicola Zanichelli, t. 1, pp 295–299
Luzin NN (1930) Leçons sur les ensembles analytiques et leurs applications, Collection de Monographies sur la théorie des fonctions. Gauthier-Villars, Paris
Luzin NN (1939) Proof of one theorem in the theory of bendings (Russian). Izv Akad Nauk SSSR, Div Techn Sci (2):81–106; no. 7, p. 115–132; no. 10, p. 65–84
Mager DN (2005) Textual revision, Stalinist revisionism, and the obligations of memory: Situating Anna Akhmatova's Poem without a Hero, Issues in Integrative Studies, No. 23, pp 71–98
Manin YI (2007) Mathematics as metaphor. Selected essays of Yuri I. Manin. With foreword by Freeman J. Dyson. American Mathematical Society, Providence. French version: Les Mathématiques comme Métaphore (Essais choisis), Translated by Claire Vajou, Les Belles lettres, Paris, 2021
Manin YI (2019) Time and periodicity from Ptolemy to Schrödinger: paradigm shifts vs continuity in history of mathematics. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Cham, pp 129–138
Markov AA (1913) An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains, Lecture at the physical-mathematical faculty. Royal Academy of Sciences, St. Petersburg, 23 January 1913. English translation by Gloria Custance and David Link, published in: Science in Context, Volume 19, Issue 04, December 2006, p. 591–600
Mazzotta G (2005) Cosmology and the Kiss of Creation (Paradiso 27–29). Dante Studies, with the annual report of the Dante Society, No. 123, pp 1–21, The Johns Hopkins University Press, 2005
McMullen CT (2010) The evolution of geometric structures on 3-manifolds. In: Carlson J (ed) The Poincaré conjecture. Clay research conference, resolution of the Poincaré conjecture. Institut Henri Poincaré, Paris, June 89, 2010. Providence, RI: American Mathematical Society (AMS); Cambridge, MA: Clay Mathematics Institute, Clay Mathematics Proceedings 19, p. 31–46 (2014)
Medvedev FA (1986) O kurse lekcij B.K. Mlodzeevskogo po teorii funkcij dejstvitel'nogo peremennogo, prochitannyh osen'ju 1902 g. v Moskovskom universitete, IMI, vyp. 30, p. 130–147
Misler N (ed) (2006) P. Florensky, Beyond vision: essays on the perception of art, Transl. W. Salmond. Reaktion Books, London
Misler N (2007) Mikhail Nesterov and the Russian religious philosophers, Tretyakov Gallery Magazine No. 2
Negrepontis S (2023) Zeno's arguments and paradoxes are not against motion and multiplicity but for the separation of true Beings from sensibles. In: Papadopoulos A (ed) Essays in Geometry, Dedicated to Norbert A'Campo, European Mathematical Society. IRMA Lectures in Mathematics and Theoretical Physics 34. European Mathematical Society Press, Berlin, pp 947–996
Neretin Yu. Nikolay Luzin, his students, adversaries, and defenders (notes on the history of Moscow mathematics, 1914–1936). https://arxiv.org/pdf/1710.10688.pdf
Nikolski A. A boatload of thinkers, Voice of Russia, 9 Novembre 2012
Oppo A (2022) Conceptualising discontinuity: Pavel Florenskii's *Preryvnost'* as a universal paradigm of knowledge. Russ Lit 130:69–93
Osserman R (1995) Poetry of the universe: a mathematical exploration of the Cosmos. Anchor Books, Doubleday/New York
Ozsvàth I, Schücking E (1962) An anti-Mach metric, recent developments in general relativity. Polish Scientific Publishers, Warsaw, p 339–350
Ozsvàth I, Schücking E (1998) The world viewed from outside. J Geom Phys 24(4):303–333

Papadopoulos A (2017) Looking backward: from Euler to Riemann. In: Ji L, Papadopoulos A (eds) From Riemann to differential geometry and relativity. Springer, Cham, pp 1–94
Papadopoulos A (2023) Yuri Ivanovich Manin (1937–2023): Mathematics, philosophy, poetry, This Handbook
Pasternak B (1958) Doctor Zhivago, transl. M. Hayward and M. Harari, Wm. Collins Sons & Co. Ltd., London
Peterson M (1979) Dante and the 3-sphere. Am J Phys 47:1031–1035
Peterson MA (2008) The geometry of paradise. Math Intell 30(4):1419
Piaget J (1960) Le développement des perceptions chez l'enfant. Bulletin de psychologie, t 13(176): 529–535
Plotnitsky A (2023) Continuity and discontinuity in, and between, mathematics and physics. In: Papadopoulos A (ed) Essays in Geometry, Dedicated to Norbert A'Campo, European Mathematical Society. IRMA Lectures in Mathematics and Theoretical Physics 34. European Mathematical Society Press, Berlin, pp 881–946
Plotnitsky A. The Image of time and the specter of the clock, from modernist literature to cinema, to appear
Poincaré H (1898) Sur les rapports de l'analyse pure et de la physique mathématique, Verhandlungen der ersten internationalen Mathematiker-Kongresses in Zürich vom 9. bis 11 August 1897. Leipzig, pp 81–90
Pyman A (2010) Pavel Florensky: a Quiet Genius: the tragic and extraordinary life of Russia's unknown da Vinci, Forwarded by G. Hosking. Continuum International Publishing Group, New York
Rauschenbach BV (1980) Prostranstvennye postroyeniya v zhivopisi (Spatial composition in painting). Nauka, Moscow
Rauschenbach B (1983) On my concept of perceptual perspective that accounts for parallel and inverted perspective in pictorial art. Leonardo 16(1):28–30, MIT Press
Rovelli C (2017) Michelangelo's stone: an argument against platonism in mathematics. Eur J Philos Sci 7:285–297
Rozanov V (1984) Feuilles tombées, traduction, introduction et notes de J. Michaut, préface de G. Nivat. Editions l'Âge d'Homme, Lausanne
Sabaneeff LL (1961) Pavel Florensky – Priest, scientist and mystic. Russ Rev 20(4):312–325
Sabitov IKh (2016) The Moscow Mathematical Society and metric geometry: from Peterson to contemporary research, Trudy Moskovskogo Matematicheskogo Obshchestva, tom 77 (2016), vypusk 2. (Russian), English translation: Trans Moscow Math Soc 2016, pp 149–175
Sinkevich G (2015) The fate of Russian translations of Cantor. Notices ICCM 3(2):74–83
Sinkevich G (2019) Some clarifications on the history of Russian mathematics of the 20th century. In: Papadopoulos A (ed) Handbook of Teichmüller theory, IRMA Lectures in Mathematics and Theoretical Physics. European Mathematical Society, pp 340–347
Sokolova L (2015) Sacred Image in a New Form: Eastern orthodoxy at the core of Wassily Kandinsky's art and theory. MA thesis, Savannah College of Art and Design
Solzhenitsyn AI (1975) The Gulag archipelago, 1918–1956 an experiment in literary investigation. Translated from the Russian by T. P. Whitney Harper and Row, New York/Evanston/San Francisco/London
Strakhovsky LI (1959) The silver age of Russian Poetry: Symbolism and Acmeism. Canadian Slavonic Papers/Revue Canadienne des Slavistes 4:61–87
Sossinsky A (2023a) The practice of Mathematics in a Totalitarian Society, This Handbook
Sossinsky A (2023b) Belyi-Kolmogorov theory. Unpublished manuscript
T. Sunada, From Euclid to Riemann and beyond: how to describe the shape of the universe. Dani, S. G. (ed.) and Papadopoulos, A., Geometry in history. Cham: Springer. p. 213–304 (2019)
Suslin MY (1917) Sur une définition des ensembles mesurables B sans nombres transfinis. Comptes Rendus Acad Sc Paris 164:88–91
Tannery P (1885) Le concept scientifique du continu. Zénon d'Elée et Cantor, Revue philosophique de la France et de l'Etranger, vol 20, pp 385–410

Tarkovsky A (1987) Sculpting in time: reflections on the cinema, translated from the Russian by K. Hunter-Blair. Alfred A. Knopf, New York
Tarkovsky A. Journal, 1970–1986, French translation from the Russian by Anne Chilikov, Cahiers du Cinéma, Atelier 3, 1993. Revised and augmented edition, December 2004, definitive edition, 2017. English edition: Time Within Time—The Diaries 1970–1986, translated by Kitty Hunter-Blair, Faber and Faber, London, Boston, 1994
Taroutina M (2016) Iconic encounters: Vasily Kandinsky's and Pavel Florensky's "Mystic Productivism". Postmedieval: J Medieval Cultural Stud 7:55–65
Trimble W (2021) Claiming infinity: tokens and spells in the foundations of the Moscow mathematical school. Technol Language 2(1):37–53
Victoroff T (ed) (2016) Anna Akhmatova et la poésie européenne, Etudes romanes, Série: Nouvelle potique comparatiste. vol 36
Vuillemin J (1971) Le Dieu d'Anselme et les apparences de la raison. Aubier, Paris
Weil S (1948) La pesanteur et la grâce. Librairie Plon, Paris
Wertheim M (1999) The pearly gates of cyberspace: a history of space from Dante to the internet. W. W. Norton Company, New York/London
Zalgaller VA (2012) Memoirs on A. D. Alexandrov and his Leningrad geometry seminar, Siberian Electronic Mathematical Reports, tome 9, p. A.62A.80
Zdravkovska S, Duren PL (eds) (2007) Golden years of Moscow Mathematics, History of Mathematics series, 2nd edn. AMS and London Mathematical Society

The Practice of Mathematics in a Totalitarian Society

A. B. Sossinsky

Contents

1 The Unusual Case of Yuri Manin 872
2 The Opposite Case: Anatoly Vershik 872
3 The Ambiguous Case of Sergey Novikov 872
4 The Most Typical Attitude 873
5 A Paradoxical Principle 874
6 Fighting the Regime: My Resignation from Mekhmat 874
7 After Solzhenitsyn's Exile 875
8 Bella Muchnik's People's University 876
9 Aleksandr Esenin-Volpin and the Letter of the 99 876
10 Yuri Shikhanovich and the Chronicle of Current Events 877
11 My Late Friend Sasha Vinogradov (Before Italy) 877
12 My Friend Mitya Fuchs (Before California) 878
13 The Older Generation 878
14 Conclusion 879

Abstract

The author, who came to the Soviet Union in 1957 from the West, describes his own experience in the totalitarian regime in the USSR and the attitudes of the leading Russian mathematicians of his generation to the regime in the period from 1957 to 1987.

Keywords

Mathematics · Mathematicians · USSR · Totalitarian regime

A. B. Sossinsky (✉)
Strasbourg, France
e-mail: asossinsky@yandex.ru

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_69

AMS Subject Classification Codes

01.11 · 01.A60

In this essay, I will try to describe my own experience as a student, a researcher, and a teacher of mathematics in the Soviet Union in the 30-year interval 1957–1987, and analyze the different behavior patterns of several outstanding mathematicians of the times.

I came to Moscow in July 1957 from the United States, where I had received my BS from New York University with all kinds of honors and had been accepted to a PhD-track program at the Courant Institute, but did not take that program. Thus, my personal background and experience differ from that of Russian students, teachers, and researchers practicing mathematics in the USSR during that period, so that my account is perhaps more objective and certainly more understandable to the Western reader than the accounts that were, or could have been, written by my Russian colleagues stuck behind the Iron Curtain in 1957–1987.

From my very first years in the USSR, I was often asked by Western visitors what I felt as a mathematician in the repressive regime of Soviet Russia, and quickly developed a joking answer that usually ended conversations on that topic: "Well, any Russian boy who is smart and honest, cannot perform on the violin or play ice hockey, has only one career choice – to do theoretical physics or mathematics. I can't play the violin or ice hockey, yet I feel fine doing mathematics here."

This joking answer should be taken seriously – it is in fact the secret of success of mathematical creativity in that repressive society: only the smartest and most honest persons did mathematics. This key point will be developed in more detail below, but let me begin at the beginning.

My undergraduate years in Moscow. In September 1957 I became a second-year undergraduate at *mekhmat* (the Mechanics and Mathematics Faculty (or Department) of Moscow State University) as a transfer student from NYU. At that time, the following people were also mekhmat undergraduates: Yura Manin, Sasha Kirillov, Dima Arnold, Mitya Fuchs, Sasha Vinogradov, Galya Turina, Borya Mityagin, Serezha Novikov, Dima Anosov, Igor Girsanov, Gena Henkin, Tolya Karatsuba, Vitya Palamodov, Yasha Sinai. I did not realize it at the time, but one can safely say that no university before or after can claim a similar amount of undergraduates who eventually became world famous, or simply great, mathematicians.

At my arrival, I already knew that I would specialize in algebraic topology, but it turned out that no courses in the subject were available at *mekhmat* at the time, except Boltyansky's course, which I felt was not up to the level – at NYU, I had studied an algebraic topology course based on the deep lecture notes of Lipman Bers, published in 50 copies in a "rotaprint" edition and not included in his Selected Works.

Had I become a third-year rather than a second-year student, I would certainly have connected with the group of third-year students (Fuchs, Turina, Novikov,

Arnold) who were studying the latest research papers on the subject, but this did not happen.

But I did meet Fuchs, Turina, Novikov, Arnold in my first undergraduate years in Moscow, not in one of the lectures or seminars, but in camping expeditions outside the city – in the taiga forests, on white-water kayak trips, in the Ural or Caucasus mountains. At the time, none of us was allowed to leave the country, so the search of new vistas and adventures, sometimes dangerous, took place within Soviet Russia.

The photograph below shows six of us standing on the Mura mountain pass (3500 meters) in the Caucasus in June 1967. The group consisted of four mathematicians and three colleagues from the chemistry faculty (Mochigin, Chapovsky, Uskov). The third chemistry professor, Uskov, does not appear in the picture: he took the photograph.

A. Sossinsky, I. Girsanov, D. Fuchs, E. Mochigin, G. Turina, Yu. Chapovsky (l.-r.)

The picture is beautiful, expressive, and in a sense tragic: Girsanov, Mochigin, Chapovsky were killed in an avalanche in the Ural mountains the same year, and Turina, in a white-water expedition in 1968.

Returning to mathematics and 1958 – I began studying and working in geometric topology under Lyudmila Keldysh, having rejected the advances of Academician P.S. Alexandrov, the head of the Chair of Topology and Geometry, who was willing to become my scientific advisor. For the next few years, my academic carrier continued smoothly, unhampered by the totalitarian regime.

But I have wandered in this digression from the subject matter of this essay, so let me move to the analysis of the behavior of several of my contemporaries listed above.

1 The Unusual Case of Yuri Manin

Manin differs from all the other mathematicians in the list above in that he was the only one with whom I was not on a first-name basis (addressing each other, we both used the second person plural **vy** rather than **ty**), and although I have heard many of his talks, I have never spoken to him at length. Besides, unlike all the other people in the list, he never participated in olympiads and other mathematical contests, he did not take part in math circles, he did not have an inspiring math teacher in high school. He is the perfect example of a Russian youngster "who is smart and honest, cannot perform on the violin or play ice hockey," and so has only one career choice – to do mathematics or theoretical physics. Manin did both.

Not knowing Manin well, I cannot assert that he never expressed, orally or in written form, his opposition to the regime. However, it is clear that he favored democracy and was strongly opposed to the communist bureaucracy – as soon as it became possible, he emigrated (1987) to Germany, traveled all over the world, but would regularly return to Russia to give talks.

2 The Opposite Case: Anatoly Vershik

When I became a mekhmat student, Tolya Vershik was an undergraduate at the Leningrad *matmekh*. Unlike Manin, he participated in olympiads and other math competitions, in math circles, he had excellent high school teachers, and chose mathematics because of his love for the subject itself, rather than because he had no other choice.

Also unlike Manin, he opposed the regime, became an active dissident, published an illegal "anti-Soviet" newspaper, and remained a very vocal opponent of the regime all his life. Even today, Vershik continues to live and work in Saint Petersburg, strongly expressing his attitude to the Putin regime.

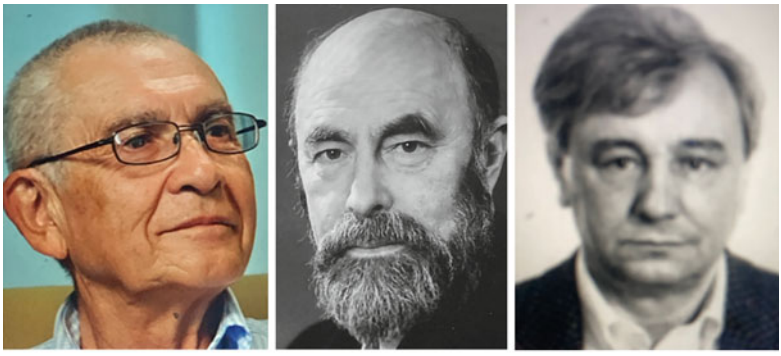
3 The Ambiguous Case of Sergey Novikov

Sergey Novikov comes from a scientific family, his father was the logician and algebraist Academician Petr Novikov, and, more importantly, his uncle was Mstislav Keldysh, President of the Soviet Academy of Sciences and head of the soviet space program. He participated in math Olympiads only once, in seventh grade, "only" got the second prize. His love of mathematics was inherited from his parents and from

his half-brother Leonid Keldysh. Novikov obtained world-class results in algebraic topology and its applications as an undergraduate and was awarded the Fields Medal in 1970, but his uncle's support was not sufficient for him to be allowed to leave the Soviet Union to participate in the ICM in Nice and accept the medal.

Sergey Novikov had very strong family feelings, and regarded the leading students of his mother (Lyudmila Keldysh) as part of the family. I will always be grateful for his help at various moments of my life. I have used the term "ambiguous" in the subtitle above only because some colleagues have a negative attitude to several of his brash statements about the work and personalities of other mathematicians.

As far as I know, while living in the Soviet Union, Novikov did not openly criticize the regime. But when leaving the country became possible (1987), he was one of the first to "choose freedom" and emigrate to the US. In the States, he continued his outstanding work in mathematics and mathematical physics, and placed in the internet various texts about his colleagues and their work that some people consider misleading. Novikov often returns to Russia, where his seminar (which he heads jointly with V. Buchstaber) remains one of the leading research platforms in Moscow.



Yu. Manin A. Vershik S. Novikov

4 The Most Typical Attitude

Here I will not describe the attitude of a concrete mekhmat professor, but that of a "generic" one, the attitude of the vast majority of my colleagues. It was quite simple. The generic professor strongly disapproves of the regime, but expresses his disapproval only among trusted friends, usually at home in the kitchen. At university, he pretends to agree with the authorities – not from lack of courage, but because of the basically correct understanding that one's serious actions against the regime would produce, at best, the loss of one's position, or, at worst, a stay at the Gulag.

5 A Paradoxical Principle

Very soon after my arrival in Moscow, I was contacted by the KGB. Two of its agents attempted – unsuccessfully – to make me an informer. This attempt was repeated by different agents – also unsuccessfully – a year later. A more detailed description of these meetings, as well as of a third meeting in the 1970s, appear in my article "In the other direction" in Golden Years of Moscow Mathematics, AMS Publ. (2007). It was during these meetings that I first applied a new principle concerning how I should talk to different categories of people.

The principle, which is absolutely perpendicular to the most typical attitude described above, is very simple:

Whoever you speak to, be it your best friend or someone you suspect of being a KGB informer, you tell him the truth about yourself and refuse to say anything (good or bad) about any other concrete person.

For instance, I would say:

"Last night I had the opportunity to read *The Cancer Ward*, it is really a great book," instead of saying: "So-and-so gave me *The Cancer Ward* to read, it's a great book."

The logic behind that attitude is that the KGB already knows enough bad things about me anyway, and what I don't want to happen is to give them any information about others.

The application of that principle invariably led to a dead end in any line of questioning, be it by secret service agents in an official meeting or by a prosecutor in a court of law.

6 Fighting the Regime: My Resignation from Mekhmat

I now return to my own relationship with the authorities. Up to the end of the 1960s, things moved ahead without serious difficulties: I passed my diploma examination without any troubles worth reporting, and obtained the diploma (the Russian version of a Western MS degree), was recommended for graduate work by the topology chair, successfully passed the graduate exams, defended my candidate's thesis (equivalent to the Western PhD) in 1965, and obtained the position of Assistant Professor at the mekhmat topology chair. The trouble began later, in 1974, when I was already promoted to the position of Associate Professor.

It was a difficult period of my life, my wife was in hospital seriously ill, I had a heavy teaching load at the university, two children to feed and clothe and take to school. On March 3, 1974, the Communist Party member of the topology chair, responsible for the ideological attitudes of its members, called me aside and told me that it was my turn to take the one-semester Marxism-Leninism course. I replied that it was impossible for me in the coming semester, and that I would think about taking it next year, but my colleague replied that the course has become obligatory, I would automatically lose my job if I did not take it, and that it was really not much trouble at all.

"You just come to the class once a week, you can sit somewhere in the back and read a book or correct papers, choose a topic in Marxist philosophy and copy some (slightly modified) excerpts from approved textbooks, that's all there is to it," he said.

Swallowing my pride, I agreed to take the course. My colleague gave me an application form. I read:

> I, the undersigned |. . . ..| .... | ask the Mechanics and Mathematics Party Organisation to kindly recommend me for a one semester course at the University of Marxism-Leninism beginning in |. . . . . |. . . .|19. . . ..|.

I took a pen to sign, but my hand was physiologically incapable of putting my name on that document! I tore up the form, seized a blank piece of paper and wrote the resignation from my position at the department effective in a month's time. Several people of the topology chair, including the chairman (P.S. Alexandrov), tried to make me change my mind, but succeeded only in part: I agreed to finish my lectures on a novel course called "Analysis on Manifolds" that no one else at Alexandrov's chair could continue.

Thus, I had lost the first battle with the authorities: its result was that I had been forced to exchange my prestigious mekhmat position for a modest editor position at the popular science magazine "Kvant." And yet I felt that I had scored a moral victory. In subsequent months, I had the following poignant experience several times: walking along the corridors of mekhmat, I would meet one of my colleagues, who, after furtively looking around, would come up to me and silently shake my hand. I would also enjoy demonstratively not shaking hands with those of my colleagues who actively supported the wrongdoings of the party bureaucracy.

7 After Solzhenitsyn's Exile

This episode also occurred in 1974, but later in the year, when Solzhenitsyn was exiled from Soviet Russia. Together with Maxim Khomyakov, my former classmate and best friend, we wrote a strongly worded letter in support of the novelist. At the time, Soviet newspapers were publishing letters by eminent personalities and by ordinary people supporting Solzhenitsyn's banishment. In particular, such a letter signed by P.S. Alexandrov and A.N. Kolmogorov appeared in Pravda.

We made a dozen copies of our letter and addressed it to the street number and street of the International Herald Tribune, the addressee being one of the editors, with his name misspelled, and mailed it from mail boxes in different parts of Moscow. But we had underestimated the efficiency of the KGB: none of the letters reached their destination. The result was that Maxim Khomyakov was forced to resign from his job, while I, being briefly unemployed at the time, was left alone.

In this connection, an important little episode occurred in January 1975: walking in the corridors of the department, I saw Kolmogorov coming in my direction from the other end of the hall. I quickly moved toward him and firmly shook his hand.

I clearly remember the expression of relief in his face and the silent sigh that he was unable to suppress just before the handshake. After that we discussed some questions concerning *Kvant* magazine as if nothing had occurred in our lives.

8 Bella Muchnik's People's University

Bella Subbotovskaya-Muchnik was a student, then a graduate student of mekhmat, defended her candidate dissertation (PhD) in mathematical logic in 1963 and, surprisingly, became a school teacher. In 1978, she founded an informal educational institution, often called The Jewish People's University, for those students who had been flunked the Moscow University entrance examinations in the framework of the shamefully famous "Jewish auditoria." This ended tragically in September 1983, when Bella was killed by a truck late at night in what everyone regarded as a KGB assassination.

A lot has been written about this unusual institution, so I will mention only two important facts, not clearly indicated in the literature. First, it was Alexander Vinogradov (more about him below) who was the first to lecture at the People's University and who invited many colleagues to work there, including the present author. Further, in 1979, during a visit to Moscow, Fields medalist John Milnor was invited by Vinogradov and myself to give a lecture to the students, which took place at Bella's overcrowded apartment, with me doing the translation.

9 Aleksandr Esenin-Volpin and the Letter of the 99

EseninVolpin, son of the famous Russian poet Sergey Esenin, was himself a poet, a human rights activist and dissident, as well as a mathematical logician. He obtained his candidate degree from mekhmat MSU in 1949, was arrested several times for "anti-Soviet activities." In 1968, just before the arrival of Nixon to the Soviet Union, he was forcibly placed in a psychiatric institution. This led to the appearance of the now famous "Letter of the 99," which asked, in very mild terms, for his release from that institution, and was immediately published in the West. The number 99 indicates the number of people (mostly mekhmat professors) who signed it. Esenin-Volpin's life story is well documented, so here I will simply say that he eventually emigrated to the U.S.

But I would like to add three remarks that connect this story with the main subject matter of this essay. First, many (if not most) of my colleagues listed at the beginning of this text, in particular Fuchs, Novikov, and Kirillov, signed the letter. Secondly, P.S. Alexandrov and A.N. Kolmogorov did not sign the letter, but wrote their own (even milder) letter asking for Esenin-Volpin's release. Finally, I did not sign the Letter of the 99, although the Wikipedia articles about me say that I did. The fact is that I would have signed it, but it was circulated for only 2 days, during which I was out of Moscow.

10 Yuri Shikhanovich and the Chronicle of Current Events

Like Esenin-Volpin, Yuri Shikhanovich was a dissident, an inmate of political labor camps and doubtful psychiatric institutions, as well as a mathematical logician and a graduate of mekhmat. Perhaps his main achievement in the human rights movement was related to his role in the creation and functioning of the anti-Soviet journal *The Chronicle of Current Events*.

I met Shikhanovich and became friends with him in my graduate years, when he invited me to do the exercise class in his Calculus course at the newly created Department of Mathematical Linguistics of MSU. I knew from the start that he was a dissident and read some issues of his Chronicle, but was never directly involved in human rights activities.

Unfortunately, I was unable to find any serious publication about Shikhanovich other than a short obituary. There is, surprisingly, no Wikipedia article about him, although there is a detailed Wikipedia article about the *Chronicle* (the article contains several inaccuracies). Let me just add that many interesting events in Shikhanovich's life occurred outside the time span of the present essay, and the last years of his life were happy and productive.

11 My Late Friend Sasha Vinogradov (Before Italy)

Unfortunately, I did not get to know Aleksandr Vinogradov in my undergraduate and graduate years – I would have learned a lot of useful mathematics from him, and possibly my research would have proceeded along different lines.

I became friends with Sasha soon after I left the department (in 1974). His seminar in Algebraic Topology and Differential Equations, followed by a seminar in Mathematical Physics, took place on Wednesday afternoons, but in fact the entire day was devoted to joint activities. It began with an outdoor soccer game (in any weather), followed by a shower, lunch at the professor's cafeteria, the two seminars, and was concluded by a walk to the Moscow University metro station. Perhaps the main result of this activity was the book "Smooth Manifolds and Observables," of which I am very proud to be a co-author.

Sasha was strongly opposed to the communist ideology, but his position re the authorities was to avoid any conflict. He emigrated to Italy in 1990, but returned to Soviet Russia every year for a summer school in Differential Equations, again avoiding any conflict with the authorities. Sasha adapted beautifully in Italy – learning Italian well enough to write excellent poetry in the language – created an important seminar and a school of young Italian researchers.

But all this lies outside the time interval of this essay, so I will conclude this section by saying that Alexandre Vinogradov passed away in 2019 in the process of heroically writing the second edition of the book "Smooth Manifolds and Observables" while fighting a terminal cancer.

12 My Friend Mitya Fuchs (Before California)

Dmitry Fuchs became my friend in the early 1960s in the context of several hiking expeditions. I used (not to say abused) this friendship in order to learn contemporary Algebraic Topology from him.

The tragedy of his life was the unexpected death of his first wife. Galya Turina drowned in a white-water kayak expedition in the Urals in 1970. It took him nearly a year to recover from that shock, and it was Izrael Gelfand who brought him back to life by suggesting that he take up painting.

When he returned to mathematics, Fuchs constructed, jointly with Gelfand, the so-called Gelfand-Fuchs cohomology. His work is precisely summarized in Wikipedia articles, his life, in an article written by S.Tabachnikov, myself, and others.

To return to my own relationship with Fuchs and Gelfand, let me mention that in 1974 Gelfand invited Mitya and me to leave the topology chair to work in his Biology Laboratory. Fuchs accepted, while I turned down the invitation.

Later, my wife Elena Efimova in the role of matchmaker invited Mitya to a seat in a play, next to Irina Kloumova, editor at Kvant magazine: Irina eventually became Mitya's second wife.

Mitya Fuchs "chose freedom" in 1990, and is now professor emeritus at the University of California at Davis.

A. Vinogradov B. Subbotovskaia D. Fuchs

13 The Older Generation

I have decided not to write extensively about the great mathematicians of the older generation – P.S. Alexandrov, A.N. Kolmogorov, I.M. Gelfand, A.D. Alexandrov – although I knew all of them quite well. For a good reason. Who am I, being several levels below them as a researcher, to assess the value of their work, and who am I, a person who never experienced life in the Stalin regime, to judge their behavior with respect to the communist authorities?

P. Alexandrov L. Keldysh A. Kolmogorov

14 Conclusion

Rereading this text, I have the feeling that I have written too much about myself, and not enough about the general situation and the attitudes of other, more important, persons. I ask the reader for forgiveness – I have neither the strength, nor the time, to rewrite this essay.

Philosophy Behind Mathematics

Ken'ichi Ohshika

Contents

1 Introduction 882
2 Poincaré: Conventionalism 883
3 Mathematicians in the Vienna Circle I: Background and Hans Hahn 885
4 Mathematicians in the Vienna Circle II: Karl Menger 888
5 Natural Science Versus Mathematics 892
6 Conclusion 894
References 894

Abstract

In this chapter, we look at how philosophical standpoints widely held by mathematicians, conventionalism, and logical positivism were brought about through works of mathematicians. Starting from Poincaré's philosophical ideas, we try to understand philosophical ideas of two mathematicians in the Vienna Circle, Hans Hahn and Karl Menger, through their mathematical works. After that, we also touch upon a problem about the relationship between natural science and mathematics, which the Vienna Circle left unsolved.

Keywords

Vienna circle · Conventionalism · Logical positivism · Poincaré · Hahn · Menger · Popper

K. Ohshika (✉)
Department of Mathematics, Gakushuin University, Tokyo, Japan
e-mail: ohshika@math.gakushuin.ac.jp

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_70

1 Introduction

In studying mathematics on a daily basis, most mathematicians today are aware of neither the philosophical foundations of mathematics nor the philosophical significance which their mathematics may have. Still they have some philosophical tendencies which are brought about from their mathematical work, and at the same time uphold their research activities. Such implicit philosophy appears on the surface only occasionally, but if we look back on the history of mathematics from the end of the nineteenth century to the first half of the twentieth century, we find that there was very fruitful interaction between mathematics and philosophy then. In this chapter, we are going to look at two important philosophical views which are closely related to mathematics, and how they were born through mathematicians' work. The first is "conventionalism," and the second is "logical positivism."

The conventionalism is a term contrived to stand for the philosophical standing of a great and versatile mathematician, Henri Poincaré. As we shall see in the following section, the conventionalism expresses Poincaré's flexible thinking, particularly his quite radical anti-Kantian attitude concerning non-Euclidean geometry and the universe. In the "Vienna Circle," this attitude would be extended to more general geometry and even to axiomatic systems of mathematics. Logical positivism is a philosophical trend dating back to the beginning of the twentieth century. It started and flourished in particular in Vienna before the culture of this city was brutally destroyed by the political disaster leading to the Anschluss. It is the name of Wittgenstein that is usually associated with this term. Still, we should note that logical positivism is a way of thinking somehow commonly shared by Viennese philosophers, scientists and mathematicians, in particular those who participated in the Vienna Circle. After the destruction of intellectual freedom under the Nazi rule, some of the people in the circle had a luck to flee from continental Europe, and started new academic lives in other worlds, mainly in the USA and the UK. These émigrés contributed to the foundations of the modern philosophy in the Anglo-Saxon world, which is referred to as analytic philosophy.

After examining in the first section Poincaré's conventionalism with regard to geometry and space, and his rather Kantian attitude towards mathematical formalism, we shall look at how these two philosophical views, conventionalism and logical positivism, were combined and developed by mathematicians in the Vienna Circle, Hans Hahn and Karl Menger. Both are not only first-rate mathematicians, but also had deep knowledge and insight of philosophy. Through their works, we can understand how the development of modern mathematics freed us from old Kantian apriorism, both in geometry and logic.

In the last part of this chapter, we shall look at a problem which was not fully solved by the conventionalism and the logical positivism espoused by the Vienna Circle: the relation between natural science and mathematics. Another Viennese philosopher, Karl Popper, did pioneering work, criticizing some traits of the Vienna Circle, but the problem remains far from being solved.

2 Poincaré: Conventionalism

We start by looking at Poincaré's conventionalism in geometry and apriorism in logic before turning to the work of mathematicians in the Vienna Circle.

Henri Poincaré (1854–1912) is a versatile mathematician whose works extend to so many branches of mathematics, but to understand his conventionalist standpoint, it is indispensable to look at his work in geometry (and topology). (We note that some of the description here overlaps with Ohshika (2019).) From the days of Euclid up to the eighteenth century, it had been believed that the only possible geometry is Euclidean geometry, a system of geometry where the parallel postulate holds, or in other words, where the sum of interior angles of any triangle is equal to π. If we consider geometry on a sphere, then by defining great circles to be straight lines, we can consider non-Euclidean geometry, where there are no disjoint parallel lines. Spherical geometry was known even to ancient Greeks, and in the modern era, it was studied by Euler comprehensively, but it was regarded as a part of Euclidean geometry. It was through the work of Bolyai and Lobatchevsky that non-Euclidean geometry really began. Their geometry, in which for any straight line ℓ and a point P outside ℓ, there are infinitely many straight lines disjoint from ℓ that pass through P, corresponds to hyperbolic geometry where the sum of interior angles of any triangle is less than π. Beltrami and also Poincaré constructed a model of hyperbolic geometry, realizing it as either the upper half-plane or the interior of the unit disc, making use of Riemann's idea to describe geometry using infinitesimal metrics.

In several of his essays contained in Poincaré (1902, 1908, 1911, 1917), Poincaré talked about the epistemic consequences of the invention of non-Euclidean geometry. According to Kantian epistemology, it is a synthetic a priori judgment that our space is Euclidean and has three dimensions. Poincaré rejected this view by (a) analyzing how we construct the notion of space around us from sensations we get from outside, and then (b) by showing that Euclidean geometry is just a possible choice which looks convenient to us. The first part (a) of Poincaré's idea is shared by positivists such as Mach whose view we shall look at in the next section. Poincaré's real originality lies in part (b), which should be called the geometric conventionalism. He explained convincingly that there can be no possible observations which verify Euclidean geometry, rejecting the non-Euclidean one. It is well known that Einstein was influenced by Poincaré's conventionalism in the process of accomplishing his general relativity.

Poincaré's view on the relation between the real world and geometry, based on his conventionalism, is anti-Kantian and was very revolutionary at the same time. Compared to this, his view on the foundations of mathematics was more intuitionistic, anti-formalistic, and rather close to the Kantian standpoint. We now turn to looking at his view on the latter topic.

For geometry, its axiomatization was already started at its beginning in ancient Greece, in Euclid's "Elements." Also the rules of logic, such as syllogism and the law of excluded middle, were formulated explicitly in ancient Greece, especially by Aristotle. In contrast, in other parts of mathematics, in particular in analysis, most notions, including those using infinity above all, had been used without precise

definitions, relying on intuition. Georg Cantor, the founding father of set theory, formulated the notion of set rigorously, and found that we can define that two sets have the same number of elements, or to be more precise, have the same cardinality, if and only if there is a bijection between them. Making use of this, he showed that there are more than one kind of infinity, in particular that the cardinality of natural numbers (countable infinity) is different from the cardinality of real numbers (uncountable infinity).

Based on Cantor's new notions, Richard Dedekind and Giuseppe Peano tried to axiomatize the natural numbers. While the possibility of mathematical induction had been considered to be derived from some property of the set of natural numbers before their work, in Peano's axioms, it is rather a part of the axiomatic definition of natural numbers. In the same vein, Ernst Zermelo tried to axiomatize set theory. His system of axioms made it possible to save set theory from the famous Russell's paradox, and also includes a controversial one, the axiom of choice.

Poincaré was not so enthusiastic about this trend of axiomatizing mathematics. Analyzing axiomatic approaches attentively, he thought that although they tried to eliminate contradictions by axiomatization, there is no guarantee that this would not introduce a new contradiction. In Poincaré (1917), he says

> Mais s'il a bien fermé sa bergerie, je ne suis pas sûr qu'il n'y ait pas enfermé le loup.
> But if he surely closed his sheep pen, I am not sure if he did not lock up the wolf inside.

As for the axiom of choice, Poincaré took a transcendental position to claim that it should be valid, a position which we can call Kantian. (See Cassinet (1983); Longy (2001) for a detailed analysis.) It seems that what Poincaré had in mind when he talked about the axiom of choice was its use for a countable family of sets whereas what is really problematic is its application to an uncountable family of sets. In fact, a later development of the foundations of mathematics tells us that it is possible to construct a system of axioms without the axiom of choice for uncountable families of sets.

To sum up his view toward mathematical formalism or logicism, he says in Poincaré (1908):

> Si même ils y avaient entièrement réussi, les Kantiens seraient-ils définitivement condamnées au silence? Peut-être pas, car en rduisant la pensée mathématique une forme vide, il est certain qu'on la mutile.
> Even if they had totally succeeded (in axiomatisation), would the Kantians be damned to silence for good? Probably not, for, by reducing the mathematical thinking to an empty form, they would certainly mutilate it.

It is rather amazing that Poincaré took a Kantian position with regard to the foundations of mathematics whereas he was clearly anti-Kantian in geometry. We surmise that this choice has something to do with the facts that his mathematics has a quite geometric flavor even when the subject is neither geometry nor topology, and that the geometric intuition was very important for his research.

3 Mathematicians in the Vienna Circle I: Background and Hans Hahn

As we mentioned in the Introduction, we shall next focus on mathematicians in the "Vienna Circle" (Wiener Kreis), Hans Hahn and Karl Menger in particular, and the trend of philosophy, called logical positivism, which flourished in Vienna before the Anschluss. The University of Vienna is known to have a tradition of philosophical studies on the foundations of science. The name of Ernst Mach (1838–1916), who held a chair at the University of Vienna from 1895 to 1902, comes to mind when we look at the beginning of this tradition. In fact, the people in the Vienna Circle can be regarded as philosophical successors of Mach, and some of them founded the "Ernst Mach Society" (Verein Ernst Mach) and organized public conferences under its name in the 1920s. Mach, as a physicist, is well known in particular for his work on acoustics and optics. He wrote influential works in philosophy, "Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen" and "Erkenntnis und Irrtum : Skizzen zur Psychologie der Forschung" among them. His philosophical position, often dubbed "empirio-criticism" along with that of Richard Avenarius, can be regarded as another departure from orthodox Kantism with Poincaré's conventionalism. At the same time, his philosophy is a precursor of the Viennese logical positivism combined with conventionalism, which is the topic of this section.

In the books which we mentioned above, Mach tried to analyze how our notion of the existence of objects is constructed from senses which we receive from outside. Naïve materialism, as was expounded by Vladimir Lenin in his criticism of the empirio-criticism Lenin (1909), presupposes the existence of things independent of human recognition. In a sense, this position regards the epistemology as being superfluous. In the other extreme, there is a position of solipsists such as George Berkeley. The latter denies the existence of external world, that is, things independent of and outside us, and admits only the existence of minds and ideas, and of course of God. Kant's transcendental idealism, set forth in his main work Kant (1781), is an attempt to justify the existence of things outside minds, by distinguishing "thing-in-itself (Ding an sich)" from its appearance. Mach's philosophy expounded in these books can be regarded as a different way to bridge the gulf between existence and recognition, which rejects both naïve materialism and Kantian thing-in-itself. He starts with our sensations, such as colors, sounds, pressures, spaces, and times, and analyzes how the external world and the existence of things outside us are deduced from them. In this way, he can explain how things outside us turn up without assuming the existence of Kantian thing-in-itself. We note that this argument is quite similar to Poincaré's analysis of our recognition of space, which we explained in the preceding section.

Thus Mach tried to consolidate the foundations of the external world based on our sensations. This can be regarded as a precursor of the modern worldview of physicists after the birth of quantum dynamics, in particular Heisenberg's uncertainty principle. Mach's philosophy had a big influence on Russian Marxists such as Alexander Bogdanov, Anatoly Lunacharsky, and Nikolai Velentinov among others.

This made Lenin, in exile in Geneva and London, write a disparaging criticism of the empirio-criticism Lenin (1909) from the viewpoint of naïve materialism, which can be barely called Marxism: Marx, who had a deep Hegelian philosophical background would not have adhered to such simple-minded materialism. Although this rough work of Lenin, in which he identifies the empirio-criticism with Berkeley's solipsism, has minor value as a philosophical work, it was canonized in Stalin's era and served as heavy shackles on Soviet philosophers (see Valentinov 1968; Bakhurst 2018).

Now we return to the topic of this section and the next: the Vienna Circle. The Vienna Circle was initiated by Hans Hahn (mathematician, 1879–1934), Moritz Schlick (physicist, philosopher) who inherited the chair of Mach from Ludwig Boltzmann, and Otto Neurath (sociologist, economist). (Before this, there had been what should be called the proto-Vienna Circle founded by Hahn, Neurath, and a physicist, Frank Philipp. See Introduction by Karl Menger of Hahn (1980) and Haller (1991).) The circle held seminars, lectures, and above all meetings in Viennese cafés. The existence of first-rate mathematicians, Hahn at the beginning, and Karl Menger and Kurt Gödel later, made the circle easy to absorb the most recent progress in mathematics and to make use of it for philosophical thinking.

In the years during which the Vienna Circle was active, problems concerning the foundations of mathematics, upon which we have touched in the preceding section, were intensely studied. This led to the consolidation of the axiomatic systematization of mathematics, which became common knowledge of mathematicians later. In this field, the impact of Gödel's (first) incompleteness theorem, his work during his Vienna years, is enormous. Another important aspect of mathematics in these years is the debates and antagonism between "formalists" and "intuitionists." Roughly speaking, the former, represented by David Hilbert, claimed that mathematics is just a series of logical manipulations, which obey rules of logic, starting from non-defined terms and axioms, whereas the latter, represented by L. E. J. Brouwer, tried to base mathematics on construction and intuition, and not on logical manipulation, rejecting in particular the law of excluded middle. The standpoint of the Vienna Circle is closer to the formalist, as we shall see soon.

The name of one of the founders of the circle, Hans Hahn, appears in contemporary textbooks on measure theory and on functional analysis which we usually refer to. In measure theory, the Hahn decomposition theorem appears in most textbooks. It states that for any signed measure μ defined on a $\boldsymbol{\sigma}$-algebra $\boldsymbol{\sigma}(X)$ on a set X, there is a decomposition of X into two disjoint $\boldsymbol{\sigma}(X)$-measurable sets, on one of which $\boldsymbol{\mu}$ induces a positive measure and on the other of which $\boldsymbol{\mu}$ induces a negative measure. This work was completed before Hahn came back to Vienna. He returned to Vienna, after taking a position at the University of Czernowitz and then serving for the army, to succeed the chair of his mentor, Gustav von Escherich. In Vienna, Hahn's interest was focused on what is now called "functional analysis." Hahn is regarded as one of the founding fathers of this field of mathematics, and he is known for fundamental results such as the Hahn-Banach theorem and the uniform boundedness principle. It is noteworthy that Hahn used the axiom of choice (or its equivalent) to prove these

theorems. This should be noted as an example of non-triviality and usefulness of the axiom of choice for uncountable families, which Poincaré did not realize.

Hahn's philosophical standpoint is clearly and succinctly expressed in his pamphlet titled "Superfluous entities or Occam's razor" which is reprinted in Hahn (1980). There he declares that his philosophical standpoint is that of empiricism and shows his sympathy toward British philosophers such as John Locke, David Hume, and Bertrand Russell, whereas he puts himself against Plato and German metaphysics, and above all, Kant. He distinguishes two ways of thinking: one is "world-denying" and the other is "world-affirming." According to Hahn, Plato is a typical world-denying philosopher. Distrusting the world of our senses, which is not constant and sometimes deceiving, Plato considered that outside and above this world, there is a world of true being, ideas. He declared that the things which we sense in this world are just projections of ideas. Hahn says that the history of western philosophy from ancient Greek age through medieval theology until French rationalism and German idealism is dominated mainly by the world-denying philosophy, possibly except for Democritus and the Epicureans. For Hahn, it is the British empiricists such as Locke, Hume, and his contemporary Russell that liberated us from this tradition.

William of Occam is a Franciscan in the fourteenth century England. Occam's razor refers to the proposition saying that "entities must not be multiplied beyond necessity," which he repeated in his works and which we can regard as expressing his view of nominalism against mainstream scholasticism of his day. Hahn says that Occam's razor should be used to cut away the world-denying philosophy such as Plato's "ideas" and Kant's "things-in-themselves." Hahn believes, along with Mach, that we should start from observations obtained by our senses as the only foundations of our knowledge. What distinguishes his philosophy from that of Mach is his belief that logical deduction and mathematics are tautological and hence do not add anything new but logically equivalent contents. This understanding of mathematics is shared by most people in and around the Vienna Circle at that time (see for instance Neurath (1973)), with an exception of Karl Menger, whom we are going to talk about in the next section, and who was not totally in agreement with this idea. This view that mathematics is a tautology was first formulated in the form of philosophy by Wittgenstein (1922), and this book was highly praised by Hahn, although there was a subtle difference in the use of the word "tautology" between Wittgenstein and Hahn, Uebel (2005). We note that Karl Popper, whom we shall talk about later, was quite antagonistic towards Wittgenstein and his work, and had quite different views on natural science from Wittgenstein, but still shares this view on mathematics.

The above summary of Hahn's pamphlet may give an impression that his standpoint would be close to that of naïve realism, for instance, that of Lenin, but Hahn's philosophy is not that simple. First, Hahn did not assume things prior to and outside our senses, but started from our observations as the foundations of our knowledge following Mach's philosophy. Secondly, Hahn clearly distinguished the real world from logic-mathematics, and said that there is no unequivocal correspondence between the two.

Hahn died in 1934. This was 4 years before the Anschluss, but Austria already fell into the hands of fascists that year, and the Social Democratic Party of Austria would be banned 2 years after then.

4 Mathematicians in the Vienna Circle II: Karl Menger

Now we turn to the second protagonist of our study of the Vienna Circle, Karl Menger (1902–1985). He studied mathematics at the University of Vienna. (See Golland and Sigmund (2000), (Menger 2002, Sigmund's chapter) for an academic biography of Karl Menger.) A fairly detailed description on Menger's encounter with Hans Hahn and his first original work in mathematics can be found in his own memoir (Menger 1994, Chap IV). Menger was a first year student of the university when Hahn returned to Vienna as a professor in 1921. Menger participated enthusiastically in the seminar organized by Hahn titled "Neueres über den Kurvenbegriff," which dealt with the problem of how to give a mathematically precise definition of curves.

A traditional intuitive understanding of curves was turned into a mathematically precise definition for the first time by Camille Jordan. He defined a curve on a plane to be a continuous image of an interval. Peano constructed in Peano (1890) a continuous image of an interval filling up a plane, which is not injective but finite-to-one. Based on this construction, Hahn proved that more generally, every compact, connected and locally connected metrizable set in a Hausdorff space can be a continuous image of an interval, Hahn (1914). This result of Hahn showed that the definition of curves as continuous images of intervals is too general. On the other hand, if we impose the condition that the map is injective, then we can regard neither a graph nor a figure-eight closed curve as a curve. Therefore, to define mathematically what we intuitively regard as curves, more work would be necessary. Menger, learning this state of difficulty in defining curves, tried hard to solve this problem, and reached a conclusion that curves should be defined by considering neighborhoods of the points lying there and their boundaries. He also used the notion of continuum, that is, compact connected metrizable space, under the restriction that it is embedded in a Euclidean space. We note that the notion of continuum had already been introduced by Cantor. This result of Menger was published as Menger (1929) much later, only in 1929, in which he says:

> Wir nehmen Kurve ein Kontinuum K von der Eigenschaft daß, wenn p ein beliebiger Punkt von K und U_1 eine offene Umgebung von p ist, eine offene Umgebung von P $U_2 < U_1$ existiert, mit deren Begrenzung K keinen zusammenhängenden Teil gemein hat.
>
> We call a continuum K (in a Euclidean space) a curve when it has the property that if p is any point of K and U_1 is an open neighbourhood of p, then there exists an open neighbourhood $U_2 \subset U_1$ of p such that the intersection of the boundary of U_2 and K contains no connected subset (except for singletons). (Phrases in parenthesis are added to explain the context of this definition.)

He also tried to define the dimensions of sets inductively based on this line of idea. First, a zero-dimensional set is defined as follows:

> Wir nennen nulldimensional eine nichtleere Menge, welche keinen zusammenhängenden Teil enthalt.
> We call a non-empty set zero-dimensional when it contains no connected subset (which is not a singleton).

Then an n-dimensional set is defined inductively as follows:

> Wir nennen n-dimensional eine Menge M mit der Eigenschaft, daß zu jedem ihrer Punkte p und zu jeder offenen Umgebung U_1 von p eine offene Umgebung von p $U_2 < U_1$ existiert, mit deren Begrenzung M eine höchstens $(n - 1)$-, mindestens für einen Punkt aber auch tatschlich eine $(n - 1)$-dimensionale Menge gemein hat.
> We say a set M (in a metric space) to be n-dimensional when it has the property that for any point $p \in M$ and any open neighbourhood U_1 of p, there exists an open neighbourhood $U_2 \subset U_1$ of p whose frontier has at most $(n - 1)$-dimensional intersection with K, and at least for some point p, such a neighbourhood actually has $(n - 1)$-dimensional intersection with K.

The mathematical notion of dimension was first defined for linear spaces, as can be seen in the work of Hermann Grassman. Riemann introduced a generalized space equipped with what is now called a Riemannian metric, which is the origin of the notion of manifold, and its dimension was defined to be equal to the dimension of the Euclidian space on which it is modeled. Although there is a phrase as follows in (Poincaré 1917, Chapitre III), where he tried to define dimension for a more general space, based on the notion of cuts, originally due to Dedekind, it is not formulated rigorously. (Menger himself remarked that he found this work of Poincaré after completing his work, in a footnote of (Menger 1929, note 15).) After observing the properties of curves and surfaces using cuts, Poincaré defines an n-dimensional continuum as follows:

> Un continu a n dimensions quand on peut le décomposer en plusieurs parties en y pratiquant une ou plusieurs coupures qui soient elles-mêmes des continus à $n - 1$ dimensions. Le continu à n dimensions se trouve ainsi défini par le continu à $n - 1$ dimensions; c'est une définition par récurrence.
> A continuum has n dimensions when we can decompose it into several parts, by performing one or more than one cuts which are continua of $n - 1$-dimensions themselves. Thus the continuum of n-dimensions is defined by the continuum of $n - 1$-dimensions; this is a definition by induction.

We can easily see an analogy between this definition and Menger's. Still, by "continuum," Poincaré meant something like a submanifold embedded in a Euclidean space or a union of submanifolds intersecting transversely each other. Since Poincaré did not use the theory of general topology developed by Hausdorff, he could not make the definition mathematically rigorous. In fact, it is easy to see that this definition will break down if we apply it to objects with cone-like singularities.

Menger's definition relies on general topology and deals with quite general topological objects, although they must be assumed to lie inside a metric space.

Menger had worked out these results while he was still a student, but they were not completed as a paper immediately since he fell ill with tuberculosis soon after. He wrote his thesis based on this idea after his recovery. After obtaining a doctoral degree, he was offered a job under Brouwer, a Dutch topologist who is famous as a proponent of mathematical intuitionism.

Although his relation with Brouwer got tense during his stay in Amsterdam because of a problem of priority for the definition of dimension, Menger got a serious interest in Brouwer's idea of intuitionism. He returned to the University of Vienna as an associate professor in 1927. Soon after his return to Vienna, Menger wrote two books, one on dimension theory, Menger (1928e), and the other one on curve theory, Menger (1932), which summarize his work on these two subjects started in his student years. At the same time, Menger was invited to join the Vienna Circle by Hahn.

Menger introduced new findings (his own and also others') on the foundations of mathematics to the circle. Besides his works in topology and geometry, Menger himself wrote several papers on logic and the foundations of mathematics during these years while he actively participated in the circle. For instance, in his paper in 1930 (see (Menger 1979, Chap 2)), he gave a clear explanation on mathematical intuitionism, which was propounded mostly by Brouwer at that time, but in a very obscure way. In the 1920s, there were debates between proponents of intuitionism, first advocated by Kronecker, and represented by Brouwer at that time, and those of formalism, represented by Hilbert. The intuitionists, against the formalists, claim that when one proves a theorem, the logical inference alone is not enough, and that mathematical constructions should be accompanied. This in particular implies, in Brouwer's version of intuitionism, the rejection of using the rule of "excluded middle" (i.e., the logical rule that $p \vee \overline{p}$ *is* always true), hence proofs by reduction ad absurdum, and existential propositions. Brouwer further tried to construct a set theory free of the rule of excluded middle in Brouwer (1919).

Brouwer's papers on this topic are reputed to be very difficult to read. Menger deciphered this work of Brouwer and showed in Menger (1928a, b, c, d) that Brouwer's intuitionistic set theory can be interpreted within the framework of ordinary set theory. Also in his public lecture in 1932 (see (Menger 1979, Chap 1)), Menger demonstrated that mathematical intuitionism can be kept within formalistic mathematics only by changing the axioms of logic, once the element of value judgment is removed from the intuitionists' claim. For instance, one can consider axioms of logic such that there are more than two possibilities for the validity of a proposition p, that is, by admitting the possibility that neither p nor $\overline{p}$ holds. By changing axioms in this way, some theorems which were already proved using reduction ad absurdum might not be proved any more, but the formalistic foundations of mathematics is preserved.

Menger's standpoint expressed in these works can be regarded as an extension of conventionalism to logic and the foundations of mathematics. As we saw in Sect. 2, Poincaré, who was the founder of conventionalism with regard to space and time,

took rather Kantian view for the foundations of mathematics. It should also be emphasized that Gödel, who was 4 years younger than Menger, was also a student of Hahn, and that his first incompleteness theorem, evidently one of the most important work in the foundations of mathematics, was first presented in Menger's colloquium. Menger not only had first-hand information on the fast developments of the foundations of mathematics but also wrote himself several papers on this topic.

Another philosophically important aspect related to Menger's mathematical work is that he dealt with very general geometric objects compared to his predecessors, above all Poincaré. Even in his first work, he considered very general objects, continua, to define dimensions. As was explained above, only objects Poincaré had in mind when he tried to define dimension are manifolds or a union of manifolds with the same dimension intersecting transversely.

Later, from the 1940s on, Menger developed the theory of statistical (or probablisitic) metric space, Menger (1942, 1951a, b, c). For two points p, q in a space X, instead of an ordinary distance function $d(p, q)$ which defines a structure of a metric space, he considers a monotone non-decreasing function F_{pq} defined on the set of real numbers which takes values in [0,1], with the properties $F_{pq}(t) = 0$ for every $t < 0$ (the positivity), $F_{pq} = F_{qp}$ (the symmetry) and $F_{pq} * F_{qr} \leq F_{pr}$ (a substitute of the triangle inequality), where * denotes the convolution. (The last condition using convolution is originally due to Wald (1943).) The intuitive meaning of this function is that $F_{pq}(t)$ is the probability that the distance between p and q is less than or equal to t. In particular if $F_{pq}(0) > 0$ for $p \neq q$, then this means that p, q are indistinguishable with the probability $F_{pq}(0)$. As this observation shows, this theory of Menger is a precursor of the fuzzy set theory.

In Menger (1979, Chap 20), Menger explains the philosophical significance of this work. If we try to construct our notion of space, as did Mach and Poincaré, including its geometric/topological property, starting from our senses, then we need to take into account the fact that if two points are too close then we cannot distinguish them. Poincaré also noticed this and repeatedly described it in his essays Poincaré (1902, 1908, 1911, 1917, 2002). In Poincaré's description which appeared in (Poincaré 1911, Chap 3), in our senses, we may perceive that $A = B$, $B = C$, $A < C$, that is, pairs of points A, B and B, C are undistinguishable for us, whereas we can distinguish A from C. This rather confusing expression can be interpreted clearly using Menger's probabilistic metric: $F_{AB}(0) > 0$, $F_{BC}(0) > 0$, whereas $F_{AC}(0) = 0$.

Poincaré's geometric conventionalism made it possible to consider non-Euclidean spaces, or more generally Riemannian manifolds as models of our space. Still if we look at them microscopically, for instance, if we look at their tangent spaces, they have a Euclidean geometry. Therefore, we can call Poincaré's conventionalism "global conventionalism." Menger's work above shows that we can also consider "local conventionalism" or even "infinitesimal conventionalism." Modern theoretical physics is founded on general relativity and quantum mechanics. It was Riemann's work extending the geometric objects enormously and Poincaré's geometric conventionalism combined with Mach's positivism that paved the way for general relativity. On the other hand, there had not been a new notion of space which supports quantum mechanics. Menger was aware of this situation, and tried to

construct a model of space which is compatible with quantum mechanics. Later there appeared many geometric theories which support the worldview of quantum mechanics, but it should be remembered that Menger tried to construct a geometric model relying on his deep philosophical thinking well ahead of his time.

5 Natural Science Versus Mathematics

By choosing a rather formalistic position towards mathematics, people in the Vienna Circle consolidated the foundations of mathematics, and overcame Kantian mindset both for geometry and logic. On the other hand, they tended to project the integrity and logical strictness of mathematics onto natural science, trying to eliminate imprecise, non-mathematical elements from science. (This is not exactly Menger's attitude though.) Naturally, this was not readily acceptable for natural (or even social) scientists in general.

There is a typical picture of a mathematician put against natural science in a passage of Brecht's piece "Lebens des Galilei" (Brecht 1952, 1998), where Galilei tries to demonstrate his finding of satellites of Jupiter (via his telescope) to court scholars of Florence.

> PHILOSOPHER: Thank you my boy. I am afraid it is not as simple as that. Before we use your celebrated instrument, Mr. Galilei, we beg the favour of a discussion - subject: can these moons exist?
> MATHEMATICIAN: Of a formal discussion.
> GALILEO: I strongly suggest you take a look first. You will find that a discussion would have wasted your time.
> ANDREA: *indicating stool*: Here please.
> MATHEMATICIAN: Surely! You know your proposition that there is no support in the heavens in contradiction to the wisdom of the ages? . . .

This is from the English adaptation performed in New York, based on the translation by Laughton. In the Berlin version Brecht (1998), the corresponding passage is a bit different from the English version as follows. The last line by the mathematician, a bit longer than the New York version, contains geocentrism:

> DER PHILOSOPHER Danke mein kind. Ich fürchte, das alles ist nicht ganz so einfach. Herr Galilei, bevor wir Ihr berühmtes Rohr applizieren, möchten wir um das Vergnügen eines Disputs bitten. Thema: Können solche Planeten existieren? / DER MATHEMATIKER Eines formalen Disputs./ GALILEI Ich dachte mir, Sie schauen einfach durch das Fernrohr und überzeugen sich?/ ANDREA Hier bitte./ DER MATHEMATIKER Geweiß, Geweiß.- Es ist Ihnen natürlich bekannt, daß nach der Ansicht der Alten Sterne nicht möglich sind, die um einen anderen Mittelpunkt als die Erde kreisen, noch solche Sterne, die im Himmel keine Sttze haben?

A mathematician is thus depicted, alongside a philosopher, as someone who starts a discussion within the framework of the "wisdom of the ages," ignoring the

evidences before their eyes. Of course this is not the role played by mathematicians in modern age, but there remains a doubt among natural scientists that what is important for mathematicians is theories, not facts.

Logical positivism is clear in pointing out the tautological character of mathematics, but it did not answer the question of why natural laws can be aptly described by mathematics. Karl Popper's "philosophy of science," which can be found in his first book published in Vienna (Popper 1935, 1959) took an approach quite different from the Viennese logical positivists. In the book, he tries to solve the problem of justifying the validity of natural laws. Positivists tried to solve this problem by introducing the "inductive methods," which says that by accumulating many observations and experiments, and by induction, we have a scientific law. This is obviously logically untenable, for we have only finitely many observations and experiments whereas any scientific law is a universal proposition. Popper's solution to this problem is very clear: he claims that scientific laws are nothing but hypotheses which are not negated by observations and experiments. He also gave an answer to the problem of demarcation between a scientific theory from a nonscientific one. Popper claims that a scientific theory must be falsifiable, that is, it must be possible to conceive a result of an experiment or an observation which would falsify the theory if it happened. As we can easily see, mathematics was not regarded as a scientific theory by Popper. As we mentioned before, he shared with the Vienna Circle the view that mathematics is a tautology.

There is an issue in Popper's view of a scientific theory that it did not reflect what really happened in the process of scientific discoveries, in particular when really new theories such as general relativity or quantum theory appeared. This type of criticism was put forward, among others, by Kuhn, in his very famous epoch-making work, Kuhn (1962). Another issue in Popper's view is that although he regarded a scientific theory as a system, he did not take into account the effect of the theory on the observations/experiments. This second issue is more important for mathematicians.

Scientific theories are frequently formulated in a form of mathematical formulae. Although Newton invented for himself the mathematical theory which he needed for his theory of dynamics, in the modern era, natural scientists usually make use of mathematics which already exists to formulate their theories. Even such a great physicist as Einstein would not have been able to formulate his general relativity without relying on the differential geometry originated by Riemann. Since observations and experiments are always theory-laden, they are affected by what kind of mathematics we have in our hands.

If the Vienna Circle had existed for longer than it really does, and if they could go on to discuss such things with Popper or other philosophers, mathematicians, and physicists, then they could have dug into such problems, and a new view on the relationship between mathematics and science might have been born. Unfortunately, after the Anschluss, they fled to far-flung parts of the world, and this prospect was never realized.

6 Conclusion

In works belonging to recent philosophical trends such as structuralism and post-structuralism, modern mathematical notions and arguments appear frequently. Some of them are serious applications of mathematics, but there are some in which the significance of introducing mathematics seems to be hard to justify. For instance, Lacan talked about automata in his famous lecture on "la lettre volée" and Borromean rings in his lectures titled "R. S. I." (see Lacan (1974–1975, 1978)). It is fairly easy to realize the relevance of the notion of automaton in the first, whereas in the second, it is rather dubious if Lacan really understood the mathematics which he was trying to use (knot theory).

The book written by Bricmont and Sokal (1997), which reveals inexact, or worse, nonsensical use of mathematics by French post-modern thinkers, had a devastating effect among mathematicians and scientists on the reputation of fashionable trends of French philosophy. Although their analysis is well documented and adequate in many ways, what they took as examples are only parts of contemporary philosophy, even if we restrict our attention to the French one. The present author believes that there still remain both philosophical problems in mathematics and mathematical problems in philosophy, and serious discussions as were done in the Vienna Circle are still possible between mathematicians and philosophers and should be desirable for both.

Acknowledgments This research was supported by the JSPS Fund for the Promotion of Joint International Research 18KK0071.

References

Bakhurst D (2018) On Lenin's materialism and empiriocriticism. Stud East Eur Thought 70(2): 107–119

Brecht C Bertold; trans Laughton (1952) Gallileo. Grove Press

Brecht B (1998) Leben des Galilei. Suhrkamp Verlag

Bricmont J, Sokal A (1997) Impostures intellectuelles. Odile Jacob

Brouwer LEJ (1919) Intuitionistische Mengenlehre. Jahresber Dtsch Math-Ver 28:203–208

Cassinet J (1983) La position d'Henri Poincaré par rapport à l'axiome du choix, à travers ses ecrits et sa correspondance avec Zermelo (1905-1912). History and Philosophy of Logic 4(1–2):145–155

Golland L, Sigmund K (2000) Exact thought in a demented time: Karl Menger and his Viennese mathematical colloquium. Math Intell 22(1):34–45

Hahn H (1914) Mengentheoretische charakterisierung der stetigen kurve. Sitzungsberichte Wien 123:2433–2489

Hahn H (1980) Empiricism, logic, and mathematics, Vienna circle collection, vol 13. D. Reidel Publishing Co., Dordrecht/Boston, philosophical papers, Edited by Brian McGuinness, With an introduction by Karl Menger, Translated from the German by Hans Kaal

Haller R (1991) The first Vienna circle. Springer, Dordrecht

Kant I (1781) Critique of pure reason. The Cambridge edition of the works of Immanuel Kant. Cambridge University Press (1991 Edition)

Kuhn T (1962) The structure of scientific revolutions. University of Chicago Press

Lacan J (1974–1975) R. S. I. Séminaire XXII, preprint available at http://staferla.free.fr/S22/S22.htm

Lacan J (1978) Le moi dans la théorie de Freud et dans la technique de la psychanalyse. Seuil
Lenin VI (1909) Materialism and Empirio-criticism. Progress Publishers (1970 Edition), Moscow
Longy F (2001) Mathématiques et intuitions: Zermelo et Poincaré face à la théorie axiomatique des ensembles et à l'axiome du choix. Philosophia Scientiae 5(2):51–87
Menger K (1928a) Bemerkungen zu Grundlagenfragen. I. Jahresber Dtsch Math-Ver 37:213–226
Menger K (1928b) Bemerkungen zu Grundlagenfragen. II. Jahresber Dtsch Math-Ver 37:298–302
Menger K (1928c) Bemerkungen zu Grundlagenfragen. III: Über Potenzmengen. Jahresber Dtsch Math-Ver 37:303–308
Menger K (1928d) Bemerkungen zu Grundlagenfragen. IV: Axiomatik der endlichen Mengen und der elementargeometrischen Verknüpfungsbeziehungen. Jahresber Dtsch Math-Ver 37:309325
Menger K (1928e) Dimensionstheorie. Leipzig: B. G. Teubner. iv, 318 S
Menger K (1929) Zur Dimensions- und Kurventheorie. Monatsh Math Phys 36:411–432
Menger K (1932) Kurventheorie. Herausgegeben unter Mitarbeit von G. Nöbeling. VI + 376 S. 1 Tafel. Leipzig, B. G. Teubner
Menger K (1942) Statistical metrics. Proc Natl Acad Sci U S A 28:535–537
Menger K (1951a) Ensembles flous et fonctions aléatoires. C R Acad Sci Paris 232:2001–2003
Menger K (1951b) Probabilistic geometry. Proc Natl Acad Sci U S A 37:226–229
Menger K (1951c) Probabilistic theories of relations. Proc Natl Acad Sci U S A 37:177–180
Menger K (1979) Selected papers in logic and foundations, didactics, economics, Vienna circle collection, vol 10. D. Reidel Publishing Company, Dordrecht/Boston/London
Menger K (1994) Reminiscences of the Vienna circle and the mathematical colloquium, Vienna circle collection, vol 20. Kluwer Academic Publishers, Dordrecht, with a preface by Lionello Punzo, Edited and with an introduction by Louise Golland, Brian McGuinness and Abe Sklar
Menger K (2002) Selecta Mathematica, vol 1. Springer, Wien
Neurath O (1973) Wissenschaftliche Weltauffassung: Der Wiener Kreis. In: Neurath M, Cohen RS (eds) Empiricism and sociology. Springer, Dordrecht, pp 299–318
Ohshika K (2019) Poincaré's geometric worldview and philosophy. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer, Cham, pp 435–450
Peano G (1890) Sur une courbe, qui remplit toute une aire plane. Math Ann 36(1):157–160
Poincaré H (1902) La science et l'hypothèse. Flammarion, Paris, 284 S. 12^{mo}
Poincaré H (1908) Science et méthode. Flammarion, Paris
Poincaré H (1911) La Valeur de la Science. Flammarion, Paris
Poincaré H (1917) Dernières pensées. Flammarion, Paris
Poincaré H (2002) Scientific opportunism – L'opportunisme scientifique. An anthology. Compiled by Louis Rougier. Edited by Laurent Rollet. Birkhäuser, Basel
Popper K (1935) Logik der Forschung. Zur Erkenntnistheorie der modernen Naturwissenschaft. Springer Verlag, Wien
Popper K (1959) The logic of scientific discovery. Hutchinson & Co
Uebel TE (2005) Learning logical tolerance: Hans hahn on the foundations of mathematics. Hist Philos Logic 26(3):175–209
Valentinov N (1968) Encounters with Lenin. Oxford University Press, New York
Wald A (1943) On a statistical generalization of metric spaces. Proc Natl Acad Sci U S A 29:196–197
Wittgenstein L (1922) Tractatus Logico-Philosophicus. K. Paul, Trench, Trubner & Co., Ltd, London

Some Disillusioned Thoughts About School and Teaching

Valentin Poénaru

Abstract

This is an outline of my experience, from philosophy, to physics and then to mathematics, and my personal view on today's important philosophical problems. This leads me to some critical thoughts concerning high-school teaching, particularly in France, related to religious or political fanaticism.

Keywords

High-school teaching · High-school programs · Religion · AMS classification: 97D30

Since a few years, I sometimes write, occasionally, nonmathematical texts too and call them literary, mostly in English. Occasionally, I might also write in French. So, it happens that sometimes I am a writer. But I must confess that, to my regret, I never managed to create, out of nothing, characters who can jump out of the written prose, with a life of their own, like Cervantes, Dostoevski, or Balzac were able to do. The only thing I can do is to put down on paper impressions or fleeting thoughts, which sometimes cross my mind in a night of insomnia. Because it can happen that I may have insomnia, usually of a mathematical kind, but then also, of a literary nature. And in that case, the story unfolds all by itself, with me as a mere spectator, and words get organized all by themselves. And when there is a matter of maths, then there are no words at all, but rather abstract images, with their own life.

Note. The first version of this essay was written in French, a first part in October 2014, in Bures-sur-Yvette, France, and a second part in February 2015, in Vuda Point Fiji.

V. Poénaru (✉)
IHES, Bures-sur-Yvette, France

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_68

When I grew up, as a child or a teenager, in that Balkan country where I happen to have been born, I hated the school where I had to go. With rare exceptions, elementary school teachers, or high-school teachers, were never endeared by me, neither then nor now, neither there nor here.

There were exceptions, of course. For a number of years, in high school which in Romania lasted from the age of 10–18, I had a remarkable history teacher. He was the son of a priest; we were in an Orthodox country where priests, unless they are monks, have to be married. And that particular priest, the father of my teacher, had been himself a professor of ancient languages at the Theological faculty, which existed there before Communism. And this father had insisted that his son, at home, should speak living languages, old Greek and the Latin of Cicero, in addition to Romanian, of course.

The lectures of my teacher were brilliant, nothing to do with the official programs concocted by the high authorities of the Ministry of education, of that country. And what he told us was so interesting that even the worst pupils listened to him, fascinated. I love to think that, had we not lived in a communist country, my teacher would have made a big university career.

It is that particular teacher, and also my father, who gave me a big taste for history, which I practice as a hobby, when I need to relax. And, in that particular domain I am today knowledgeable enough to be able to identify the sources of my teacher from long ago, and even sometimes to disagree with him, concerning those things he told us. When it comes to reading history, I can do that in several languages. And it is from an Italian author, actually a former professor of nuclear engineering at the university of Milano, and also former boss of the Italian version of NSF, who later became a remarkable historian, that I learned something which I cannot resist telling you now. Toward 1900, Charles Péguy, that big hero of both Catholics and the French left, tells a friend: “Yes, the war will certainly come, and then the first thing we should do is to shoot Jaurès.” As you may know, Jaurès was shot the very first day of the war, in August 1914. Not by Péguy, of course, but then Péguy certainly suggested it big. And it is not here in France that these things were ever taught in school.

Mario Silvestri, my nuclear expert turned historian, wrote the best pages I have ever read on Fascism, Italian, or from elsewhere, and also on the mechanisms of the big economic crisis from 1929 to 1933. And he is also one of the rare historians to give a good account of the impact of science and technology on history.

To change topic, maybe not quite, I was always puzzled by how close friends can have very different memories concerning things they went through together, and have such radically different judgments about the same events. Already when I was seven, in elementary school, Mircea became a close friend, and he stayed so until he died of lung cancer, in Paris, in 1987. And for him, our elementary school, which we went through together, was a beautiful thing, very important for his future career as medical doctor and biologist. But for me, it was a black jail: That is how I felt about it. In my memory, the worst moments of my life were that school (4 years), the military service which I did under the shadow of Stalin, and my first marriage. Of

course, there were other unpleasant moments in my life, but those left the biggest mark in my memory.

I certainly disliked school, but I was very far from being a bad pupil. Year after year, I had the best grades in all the subjects. But I rarely did my homework, neither at home nor in school, and I very often missed the lessons. Surreptitiously, together with a chum who incidentally also became a mathematician, we left the school building and went to the boxing school instead.

But I know how to manage, and here is a typical example. One day, there was a rumor in my high school class that for once our teacher of Romanian was going to control our homework. I never did that homework at home, and so during the break I wrote quickly a little text. When the teacher started checking, he noticed that I was presenting something much shorter than the other pupils. Of course, I had not had the time to do something longer. And then some friendly voices started mumbling: "He just wrote it, Ha!Ha!Ha!" So my teacher asked me to read aloud my effort. I expected some punishment, but I was actually congratulated for my elegant prose. In fact, when necessary I can write fast and well. I was even offered once a fancy job as editorial writer for a weekly magazine. But I am not in the mood of telling you now about that. I am just a mathematician, a bit of a physicist too, but neither a journalist nor a historian, although history interests me a lot.

The only thing that I learned in that high school was to live in society, together with other people of my age. But concerning the big philosophers, mathematics or physics, I had to learn all that alone, not in school. I still remember that when I was 13 or 14 my maths teacher realized that I knew already calculus, and then he got very cross. He said that I did that just in order to challenge him, and he started disliking me big. It is certainly not under the guidance of my various maths teachers that I became a mathematician. Actually, maths as such did not interest me much then. It is physics I wanted to understand, and mathematics was needed there.

I got to physics via philosophy. And then, why philosophy, and how does one get from philosophy to physics?

My father loved to take long walks with me in the nearby beautiful parks and talk to me about all sorts of things. And when I was about 9 or 10, there was an issue which bothered me a lot and I told him about it. How is it possible that something moves, when at any given moment it stays still? It seemed to me then that this meant that movement was impossible, and this thought tormented me. My father was very puzzled by my worries, which he did not understand; he did not know what to tell me.

But at home, there was a big library, the books of my parents, filled with interesting things. I started to read the Greek philosophers, and very soon I found out that Zeno of Elea, 24 centuries before, had already shared my worries. That was the beginning of my philosophical adventures. And for a while at least, my great hero was Immanuel Kant, the Kant of the "Critique of pure reason" (*Kritik der reinen Vernunft*).

There were also other books in my parental library. I found physics books, and although I could not understand exactly what they said – they were written in the

language of mathematics, opaque to me then – I still grasped that they could get much farther than the philosophers.

There is here, in this story, a matter of personal taste too, of course. Both physics and mathematics I had to learn all by myself, maths in order to be able to understand physics. And then, with their strange beauty, these two topics marveled and excited me, in a breathtaking way. And I have continued to be fascinated by them, ever since. Things concerning life I found much less attractive, and the questions of organizing human society even much less so.

Actually, moral philosophy and ethics I found frankly boring. I know, of course, quite well that we are directly concerned by the organization of society and interaction between human beings; our everyday life depends on that. But then so is good plumbing, of course. There are all sorts of things, essential for our survival, but they have nothing to take away your breath with excitement, nothing to marvel you, nor to clinch your thirst for universality and the absolute. And, intellectually speaking, I am not at all attracted by those things; they leave me ice cold.

I have just described you my personal road to philosophy, from philosophy to physics, and then quite naturally from physics to mathematics. But being brilliant in maths in school was never an attractive thing for me.

All these things notwithstanding, there were still some good things in that Balkan high school of mine. One had to attend classes only half-day, sometimes in the morning, sometimes in the afternoon, depending on the years. So, I had plenty of free time. I never did my homework, but I read all the books I wanted to read, and very importantly, I could think, think, and think.

I find that nothing is more important than free thought, when you can play around with mathematical objects or philosophical ideas, without any preestablished plan or schedule, just wander through the magic forest of abstract thought. And this way, I became the person I am.

Here in France, through my children, I discovered a school even more repressive than mine, although I come from the third world of the Balkans and here we are in civilized land. No question of free time, no possibility to read what you want, or even to think freely. No question of letting your mind explore that far away land of ideas.

And then, the ridiculous way in which various topics are put together in school, here. Why the Devil should history and geography be lumped together? Geography fits well with geology or geophysics, but not with history. Of course, historical atlases are useful, but that is not geography. I suspect that some big brain, some high civil servant in the Ministry of education, had the stupid idea that space and time correspond to geography and history, and hence the two should go together. Or, maybe, there is here some ultranationalistic idea about ancestral blood and sacred soil of our mother country (theirs, not mine!).

And then, why lump physics and chemistry together, rather than maths and physics, like in the more intelligent educational systems of Germany or Russia? All this, I think, is not unrelated with the lack of visibility surrounding theoretical physics in this land, which lasted so long. Until late, well into the first half of the twentieth century, big topics like Boltzmann's statistical mechanics or Maxwell's great synthesis electricity-magnetism-optics were totally ignored here in France. Of

course not really by everybody: There was also Henri Poincaré, the solitary genius. But then, he was a mathematician, and mathematics has always fared very well in this country.

Then, what can one say about how quantum theory was received here in France? This all-important breakthrough was totally ignored by the university, until the 1950s. Between the two wars, there were important voices claiming that it was just some Germanic metaphysical speculation. My friend Pierre Cartier told me that toward 1950, when as a young student he attended a physics course, in Paris, the professor started by saying: "Ladies and gentlemen, the so-called atomic hypothesis will never be mentioned in my course." Five years after Hiroshima, atoms were still hypothetical for this disciple of Marcellin Berthelot (not only the well-known chemist, but also minister of education, who banned atoms from high school in the beginning of the last century).

Concerning physics in high school, that is total disaster. A friend of mine, a famous physicist, told me that while around 1950, physics taught in high school was at least 40 years old; now it is about a 150.

To crown all this, why should philosophy in high school go together with the political-ideological doddering of those guys who call themselves the intellectuals? Concerning this, I still remember that at a party, after a painting exhibit of my friend Hermann Braun-Vega, a lady who had heard that I was a mathematician turned to me saying: "But you are a scientist, you sir you are not an intellectual!"

And I do not talk here about the teaching of biological sciences in high school, because I am less competent there. When I went to high school, in Romania, that was in Stalin's time, and biology was completely controlled by a certain Lysenko. Very few people still know about him today. Anyway, genetics was banned, and its practitioners went to jail. So, when I escaped from Romania, in August 1962, I was totally ignorant in biology. And I can tell you now a little funny story. In the Summer of 1964, I attended a big topology conference, in Cambridge, England, organized by my friend Christopher Zeeman, soon to become my cousin by marriage, and a bit later, Sir Christopher too. And one day, he invited me to a little lunch party with Francis Crick and an Anglican bishop. And my ignorance was such that I had never heard about DNA and I totally ignored who Crick was. The bishop was, actually, not much better. He kept calling him Mister Watson Crick, to the big annoyance of Crick and the great amusement of Christopher. But there I stood and did not understand a thing.

Coming back now to those intellectuals from above, the most important figure of the Enlightenment was Kant, who had started his career as a scientist, as a great astronomer actually, but I will not tell you here about what he achieved. Then, at a certain point, Kant started pondering, why and how did Newton succeed in understanding the world, while for more than 2000 years the philosophers hat failed to do so. And out of this came the Critique of pure reason, already mentioned, from which much of what is interesting and serious in philosophy proceeds. And if one looks at things, before and after that, a lot of philosophers were scientists too. I will quote here, at random, Descartes, Pascal, Leibniz, Mach, Poincaré, Bertrand Russel, Ludwig Wittgenstein, the members of the Vienna Circle *(der Wiener Kreis*), Popper,

and Tom Kuhn; the list is long. And then, some towering physicists of the twentieth century were imbued with philosophy, like Einstein, Bohr, Heisenberg, or Schrödinger. And I could add to the list the mathematicians Hermann Weyl and Kurt Gödel.

Why then lump together philosophy, in the classroom lessons, with the political doddering of a character like Sartre, the man who treated his conscience problems on par with the liver ailments?

Can one talk seriously about philosophy today without Gödel's undecidability and without the Turing machine and its cousins?

I will tell you now some personal opinions, for which I take full responsibility. I can only see two really important philosophical issues today: on the one hand, the foundations of mathematics, the big open gap between mathematical truth and what one can prove (see here Gödel and company), and on the other hand, the interpretation of quantum mechanics, the problem of measurement, and its incidence on the existence of objective reality (see here John Bell and company). Those two problems touch deeply the fundamental relation between thought, i.e., what goes on in our mind, and what is outside, physical reality and mathematical reality, both of them, to me, with the same ontological status. And, connected with all this is also the ontological status of the laws of nature.

Coming back to those intellectuals, with their blatant lack of depth and knowledge as soon as they leave their very narrow horizon, they should better learn from Voltaire, who from his English exile brought back in his luggage Newton's *Principia.* And he gave it to his girlfriend the marquess Emilie du Châtelet, who produced a beautiful translation, the first of its kind in France, with beautiful comments. And this was an important thing, but I do not intend to deliver you here a lesson in the history of science.

And I should not forget here that little gang of intellectuals called by some, rightly so, "the impostors." I am thinking here, of course, of Sokal's hoax and the "Science Wars." These impostors tell all kinds of phony things concerning the scientific topics I have been mentioning here. And they present their stuff bathing in a very heavy, highly undigestible, scholarly pedantic, sauce. They believe that if you quote Gödel or quantum uncertainty, this is good excuse for telling, afterward, all kinds of phony stuff.

Coming back to my starting point, I do not like school, but at least I had the good luck to be submitted to it in the third world Balkan country where I come from, and that gave me what I needed most, much much free time.

The pages you just read were written toward October 2014, but then a bit later a bloody terrorist attack was unleashed against a Paris satyrical magazine; that was in January 2015, and that triggered in my mind the text which follows now.

Religion can be a very serious challenge today; I am thinking about the fundamentalist and radical Islam, but before we had the challenge of Communism and then, even earlier in time, when the Catholic creed was intolerant and uncompromising, that was a challenge too.

Concerning Islam, in its days of glory, the Ottoman empire was relatively tolerant. My "tolerant," to be taken with quotation marks, refers to the general

climate of yore, when religious wars were creating bloody havoc in Germany, France, and England. But then, when the Ottoman empire was, later on, in the throes of death, then it became really murderous, like any wild beast in such throes, and then we had those Bulgarian atrocities, much bemoaned by Gladstone, or the 1915 Armenian genocide.

Of course, in their long history, the Ottoman Turks did other murderous deeds, but to get something comparable to our present Islamic terrorism, one has to go back 800 years, to the times of the Old Man of the Mountain and to his assassins, stuffed with hashish. I believe the same Arab word covers both things.

And it is only the radical Islam which, these days, assassinates satyrical draftsmen in France or Denmark or elsewhere. It is radical Islam which pronounces death sentences against writers like Salman Rushdie.

From our present vantage point, the other traditional religions have quieted down long ago. Of course, that is not the case with political movements like Fascism, Nazism, or Communism, which somehow replaced them. But I will not talk about them now.

You will have probably already guessed that I am very fond of Enlightenment, and concerning our present topic, here are two typical heroes of it. Toward 1700, John Locke wrote his book "On Tolerance," where he preaches great tolerance in religious matters, with one exception: no tolerance for the fanatic Catholics, very intolerant then. At least, this is how they were in these times; they calmed down big, since.

And few years later, Voltaire came with his famous "Ecrasez l'infâme !" meaning religious fanaticism.

As far as I am personally concerned, I do not believe in God; more exactly, I fail to understand what the whole issue of His existence or inexistence means. But would I be a believer, which I am not, I would not do what the good Christians daily do, pray God for my health or my little problems; I would think that is sacrilege. Besides, believers tell me about Creation, the creation of the world, something I cannot understand; I am already at pains making sense of the existence of the world, so when it comes to its creation . . .

Comes to my mind a comment by Steve Weinberg, the great physicist, in one of his essays for the New York Review of Books; he says something like this: There are many people who do horrible things, but when a nice person does that, then religion is somewhere there. Yes, at least under its radical, fundamentalist, and uncompromising form, religion is a problem.

I have talked here a lot about Islam and Catholic creed, but I was born in an Orthodox country, and when I was a young man I got to know Orthodoxy quite well. I had then strong links with Orthodox monks, something which earned me, some years ago, an invitation to a theology conference, at the Vatican. But since I thought I was out of place there, I refused very politely.

Coming back to Orthodoxy, one of its basic ideas is the existence of a vertical direction, toward transcendence, an idea which still appeals to me, and then also a horizontal direction, this world. And I was always told that this is what interests, mostly, the Catholics.

All this might be so, but then Orthodoxy is quite arch-conservative and can be even more antisemitic than Catholicism. And then, Romanians are too ultra-nationalistic, believing that one should always do as the grand-grand-fathers have done, so Romania never knew Humanism, Renaissance, Reformation, Baroque, Scientific Revolution, or Enlightenment. The country stayed too long outside of the European civilization, and that is still very feelable today.

To come back to those bloody events from 2015, when the French minister of education suggested that one should introduce more courses of civic education in school, I do not know whether I should laugh, weep, or get angry. Why the Devil make school even more boring! We do not need more courses on civic morality, pure annoying doddering, but rather courses on science, something which could interest and excite the young minds, thirsty for the absolute, and develop their curiosity too.

History teaches us that there is nothing better than science, for cutting that branch on which the fundamentalist, uncompromising religious fanaticism sits. Remember that while the Vatican did some long due excuses for Galileo, it never did so for Giordano Bruno. I believe this is because Bruno's idea of an infinite universe, with infinitely many inhabited worlds, puts dynamite under the belief in a special unique destiny for humanity, which God the Son came specially to save, on our modest planet. And Bruno conceived this idea as a direct consequence of his reading of the breakthrough of Nicolas Copernic. The Church can never forgive this kind of ideas.

One hears a lot of talk these days about Revolution. But how many people know that the origin of the word is the Latin verb "revolvere," and the noun is "revolutio," which means circular movement, hence a backward one, certainly not a linearly forward one. Remember that Copernic talks about revolution of celestial bodies. And during many centuries, all European revolutions, starting with the Hussites of Bohemia, all the way to the puritan revolution in England, were made not for a "shining future" (the French leftists say "l'avenir radieux"), but in the name of a mythical past, for going back to Paradise on Earth. All this in the name of God and for the salvation of the eternal soul. Of course, this was no longer true for the next revolutions, starting with the American and the French one.

And this is because, in the meanwhile, came the TRUE revolution, the scientific one, and that brushed away all that talk about the God on the image of Whom man is made, or about the eternal soul and its salvation. Of course one can be scientifically minded and not cut the bridge with transcendence, with the individualistic mysticism. But that kind of religious spirit is totally foreign to the traditional, highly organized, and structured religion, too often fundamentalist and uncompromising. There is no contradiction between the kind of individual and individualistic religion I was talking about before, and "laïcité," a French word which I may translate as secular spirit, something which humanity needs so much today.

A little prentice, here. In the 1920s, when Kemal Atatürk decided to modernize his country, Turkey, and get it out of underdevelopment, he identified religion as the main thing in the way. And since the word "laïcité" did not exist in Turkish, he imported it from French.

Just before ending, I feel I still have to tell you some personal ideas of mine. I think there were only two important revolutions, the scientific one and the

technological-industrial one, each coming in several successive waves. All the other revolutions might have had, here and there, some reasonable, pragmatic ideas, but most times they were beds of Procust, brutal attempts to impose a rigid grid on top of a complex and complicated humanity. That always brought a lot of suffering and misery. Nothing appealing to me, here.

I love scientific revolutions, but I dread like hell political ones, and everything they come with. About a 100 years ago, the Russian philosopher Nicolas Berdiaev wrote a book "The new Middle Ages." And there, he writes that if once people have cherished Utopias, now the time has come when those Utopias might fall on our heads, and, in that case, God have mercy on us. That is what he said.

Section III

Influences in Twentieth- and Twenty-First-Century Mathematical Practice

Bharath Sriraman and Kyeonghwa Lee

Influences in Twentieth- and Twenty-First-Century Mathematical Practice

Bharath Sriraman and Kyeonghwa Lee

Contents

1 Introduction 910
2 Mathematical Practice 910
3 Mathematics as a Cultural Practice 911
4 Historiography of Mathematics: Who Has the Prerogative? 912
5 From Natural Philosophy to the Present 913
References 914

Abstract

What characterizes twentieth-century mathematics? Or twenty-first-century mathematics? As opposed to mathematical practices? The former is too broad and subject to different answers depending on the individual or the group being addressed. Hilbert's 23 problems published in 1900 would probably be mentioned because they influenced generations of mathematicians in the ensuing 123 years. Many names could also be mentioned such as Hardy, Ramanujan, Kolmogorov, Arnold, Poincaré, Grothendieck, the Bourbaki, the Langlands program, and others as shaping the agenda of twentieth- and twenty-first-century mathematics. Mathematical practice, on the other hand, is a broader arena than mathematics per se. It is broader because it includes not only mathematicians making original contributions to the field but also others that use mathematics in other contexts. These can vary from applications of mathematics in the real world, revisiting historical sources, (auto) biographical and metamathematical accounts, to making mathematics more accessible through textbooks. Others could simply view mathematics as a recreational activity (e.g., Bhaskara's Lilavati, Victorian

B. Sriraman (✉)
University of Montana, Missoula, MT, USA
e-mail: sriramanb@mso.umt.edu

K. Lee
Seoul National University, Seoul, South Korea
e-mail: khmath@snu.ac.kr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_137

Mathematics, etc.). In this section of the Handbook, a selection of chapters are found that serve as examples of mathematical practice in a broader sense. The perspective conveyed in this section provides a balance between Western and non-Western contributions to mathematical practice.

Keywords

Twentieth-century mathematics · Twenty-first-century mathematics · Historiography of mathematics · Victorian mathematics · Indian mathematics · Greek mathematics · Ethics · Financial mathematics · Postmodernism

1 Introduction

What characterizes twentieth-century mathematics? Twenty-first-century mathematics? As opposed to mathematical practices? The former is too broad and subject to different answers depending on the individual or the group being addressed. Hilbert's 23 problems published in 1900 would probably be mentioned because they influenced generations of mathematicians in the ensuing 123 years. Many names could also be mentioned such as Hardy, Ramanujan, Kolmogorov, Arnold, Poincaré, Grothendieck, the Bourbaki, the Langlands program, and numerous others as shaping the agenda of twentieth- and twenty-first-century mathematics.

In the twentieth century, among the external factors that influenced mathematics, one could say that pure mathematics (topology, algebraic geometry, and analysis) were essentially a continuation of mathematics of the nineteenth century; the influences came essentially from Riemann, Klein, Poincaré, and the Italian school of algebraic geometry. Mathematical logic was influenced by philosophy. The influences to mathematical physics came primarily from physics: quantum mechanics and relativity. Applied mathematics certainly had external influences: physics, the wars of the twentieth century, industry, medicine, computer science, etc. Discrete mathematics emerged in its own right.

2 Mathematical Practice

Mathematical practice, on the other hand, is a broader arena than mathematics per se. It is broader because it includes not only mathematicians making original contributions to the field but also others that use mathematics in other contexts. These can vary from applications of mathematics in the real world, revisiting historical sources, (auto) biographical and metamathematical accounts, to making mathematics more accessible through textbooks. Felix Klein's classical book *Elementary Mathematics from a Higher Standpoint* is an example of a textbook aimed at making mathematics accessible to school teachers over a century ago. It still remains a classic today. Textbooks (and teachers) can inspire people toward the study of mathematics. Others could simply view mathematics as a recreational

activity (e.g., Bhaskara's Lilavati, Victorian Mathematics, etc.). In this section of the Handbook, a selection of chapters serve as examples of mathematical practice in a broader sense. The perspective conveyed provides a balance between Western and non-Western contributions to practice. The contributions explore Greek, Indian, Levant, Javanese, Victorian, and Western contributions to mathematical practices culminating in twenty-first-century financial mathematics. The arrangement reveals a progression of thought and practice starting with the Greeks, suggesting mathematics as a cultural practice across time and place.

3 Mathematics as a Cultural Practice

Diophantine arithmetic is regarded as a difficult area of mathematics (and mathematical practice). Questions are easy to pose but difficult to answer. Modern number theory textbooks relegate Pell equations toward the back even though the history of mathematical practice in number theory suggests otherwise. Negrepontis, Farmaki, and Brokou provide a reinterpretation of Plato's philosophy in terms of periodic anthyphairesis. They focus on one of the most difficult and profound books of Euclid's Elements, namely, Book X, and examine palindromically periodic anthyphairesis through the lens of *Politicus*. One learns of the remarkable contributions of Theaetetus in his work on the general Pell equation. This chapter also points to connections and similarities to work on Pell equations done by Brahmagupta.

Ancient Indian mathematics has been subject to all sorts of interpretations, appropriations, and misappropriations in the literature ranging from speculations on the computational nature of Vedic mathematics to nationalistic discourses in twenty-first-century India. The view of ancient Indian mathematics today stands in stark contrast to denigrating views during colonial times expressed by the likes of W.W. Rouse Ball who wrote:

> The Hindoos, like the Chinese, have pretended that they are the most ancient people on the face of the earth, and that to them all sciences owe their creation, But it is probable that these pretensions have no foundation; and in fact no science or useful art (except a rather fantastic architecture and sculpture) can be definitely traced back to the inhabitants of the Indian peninsula. . .. (Ball 1908, p. 146)

Dani in his chapter on Ancient Indian mathematics provides a detailed review of the Sulbasutras. The Sulbasutras are a striking example of mathematical practices embedded in cultural contexts and provide us a unique window to the mathematics in India from the millennium BCE. Engagement with performance of yajnas (sacrificial fire worship) with elaborate and intricate structures led to knowledge of geometry, in particular the statement of what is called Pythagoras theorem, well ahead of the latter, and also several interesting constructions. There are four major Sulbasutras, Baudhayana, Manava, Apastamba, and Katyayana, that have been studied, to varying degrees, from an historical perspective, starting from about the 1870s, with Thibaut's translation and editing of Baudhayana Sulbasutra. Many authors, from

around the world, and from India, have contributed to the study over the years. Knowledge on the topic nevertheless remains patchy and limited, with many unanswered questions. In this overall context, Dani's chapter aims to present a comprehensive review of the mathematical contents of the Sulbasutras, together with a critical evaluation of various claims that have been made about them.

Mathematics as a cultural practice is also revealed in Karjanto and Beauducel exploration of the intricate relationship between number theoretic congruence formulas and the Javanese calendar, especially how this type of arithmetic permeates the daily lives of Javanese in Indonesia. The Javanese calendar which combines the *Saka* Hindu, lunar Islamic, and western Gregorian calendars is used for cultural and spiritual purposes.

4 Historiography of Mathematics: Who Has the Prerogative?

Who has the prerogative in writing the historiography of mathematics? Is it mathematicians or historians of mathematics? History of mathematics is a typical undergraduate course taken by students in North America and other parts of the world. Burke's account of the mathematics of three ancient cultures Thebes, Babylon, and Alexandria brings episodes from the history of measurement – some of which are found in textbooks today. These cultures have also been a legitimate area of study within the community of historians of mathematics who were more attuned to the context of the cultures that created this mathematics. Colonization played a role in creating "access" to western historians to artifacts in the greater Levant. Mathematicians, on the other hand, are able to telescope forward from origins and changes of notation into research advances in the field of mathematics. One could trivialize the history of mathematics as simply being the "wait" for the right notation to come along which encapsulates cumbersome attempts of the past in a more efficient manner. But is it okay to trivialize the mathematics of the past? Narratives in textbooks shape the worldviews of students on the origins, nature, and trajectory of the development of mathematics. Issues of precedence, the influence of colonial narratives such as the deplorable quote (one of many) by Ball earlier in this introduction, can portray the development of mathematics by nonwestern cultures in a negative or skewed light.

Blasjö argues that "presentist historiography is not the natural outlook of the mathematician, but rather a direct antithesis of the mathematician's most fundamental reason for studying history. To study past mathematics precisely as it appeared to active researchers at the time is not foreign to the mathematician, but a direct corollary of the mathematician's core conviction that only those with first-hand research experience in a field of mathematics truly understand it." In other words, the hope is that more mathematicians today become interested in shaping the narrative of mathematics instead of casting it aside as the purview of historians.

5 From Natural Philosophy to the Present

No account of practices in twentieth- and twenty-first-century mathematical practices would be complete without acknowledging the impact of Newton. It was only in the second half of the nineteenth century that expressions such as "scientist" and "mathematician" began to replace "natural philosopher." Britton and Sriraman provide a chapter on Newton's Opticks where they describe his insights into developing a theory of color through his experiments on light. Donnelly, Dunkum, and McCoy delve into the original source of enumerating integer sequences with alternate parity entries, through an 1839 paper by Olry Terquem. Their chapter makes remarkable connections between Terquem's "forgotten" problem, the Fibonacci sequence, a one-player game of numbers, continuant polynomials (an extended version of), the Euclidean Algorithm, and Bezout's Lemma.

Does mathematics have/make a social impact? Ernest's chapter examines the interaction of ethics with mathematical practice. His chapter analyzes resistances from the community of mathematicians that simply said that mathematics can be pursued for its own sake. To wit, G.H. Hardy studied pure mathematics unapologetically (!) because it had little or no applications to real life during his time period. In fact in Victorian times, pure mathematics was mostly viewed as a recreational activity with various periodicals tailored toward circles of the gentry, who enjoyed doing mathematics for its own sake. Bertrand Russell started out with the infamous Mathematical Tripos in Cambridge and pursued it for its own sake despite his disenchantment in the way it was taught. Russell not only turned to philosophy but also wrote about the ills of war and the social responsibility of Science. Similarly, Ernest reviews ideas of social responsibility since mathematicians increasingly are in the background of media communications given the data-saturated world we find ourselves in – and media increasingly shapes worldviews of the population.

One finds numerous personal accounts from mathematicians on events that shaped their worldviews and practices. Norbert Wiener's accounts of how he learned mathematics is an astonishing example of traveling to different places to learn firsthand from the giants of the twentieth century – from Hardy and Russell in Cambridge, and from Hilbert and Landau in Göttingen (Wiener 1953, 1956). It should be noted that Wiener also took a public stand against the militarization of science and the need for ethical considerations in Science. Such (auto) biographical accounts have great value for the community at large. They reveal the formative experiences and the openness of the mathematical community in the past. In a similar vein, Lifyland provides a personal account of influences in a mathematicians' (his) professional journey through mathematics. It reveals the interplay between practices and the development of a working philosophy of mathematics.

Zalamea surveys Grothendieck's "known" work (1949–1991), before his seclusion in the Pyrenees. This chapter offers an insight into the massive contributions of Grothendieck in addition to the dialectics between abstract categorical constructions and structured, concrete models. Zalamea writes that "the abstract and the concrete, the universal and the particular, categorical reason and mathematical heart, become all necessary to *nurture mathematical imagination*." His chapter ends with the

insight into Grothendieck's work as a *back-and-forth between types and archetypes*, elevated to a method.

Sinclair, on the other hand, examines mathematical practices from a postmodernist viewpoint that is in concordance with how human societies were actually shaped. It moves way from a linear, patriarchal depiction of mathematics to viewing how mathematical ideas can be used to generate new ways of imagining relations "that pertain to existence (who we are, as people, collectives, societies) and to experience." Numerous mathematical examples are used to "transform normative relations in order to produce new ones – in particular, new spatial imaginaries for disrupting systems of oppression."

Finally, the ubiquity of mathematics with the world of finance is explored in depth by Davis starting from its prehistoric origins to the use of Brownian motion and stochastic calculus to price stocks. Cutting-edge mathematical ideas coupled with advances in AI are continually applied in the world of finance, and the world at large. Mathematical practice is prevalent everywhere today both overtly and covertly – unimaginable in 1900.

Taken as a whole, this section of the Handbook provides an insight into a highly diverse set of current practices in mathematics shaped by our shared past culminating in the present.

References

Ball WWR (1908) A short account of the history of mathematics. Macmillan, London/New York
Wiener N (1953) Ex-prodigy: my childhood and youth. MIT Press, Cambridge, MA
Wiener N (1956) I am a mathematician. Gollancz, London

The Pell Equation in the Pythagoreans, Theaetetus, and Hindu Mathematics

Stelios Negrepontis, Vassiliki Farmaki, and Marina Brokou

Contents

1 Introduction 917
2 Geometric Anthyphairesis 921
2.1 Definitions 921
2.2 The Logos Criterion for Periodic Anthyphairesis 922
3 The Anthyphairetic Basis of Plato's Philosophy: An Intelligible True Being Is the Philosophical Analogue of a Dyad in Periodic Anthyphairesis (*Parmenides*, *Sophistes*) 925
3.1 The One of the Second Hypothesis of the *Parmenides* Consists of a Dyad <One and Being> in a Philosophical Analogue of Periodic Anthyphairesis 926
3.2 The Division and Collection of the Platonic Idea "the Angler" in the *Sophistes* 218b-221c 928
3.3 The Division and Collection of the Platonic Idea "the Sophist" in the *Sophistes* 234e-236d and 264b-268d 931
3.4 The Cyclical-Periodic Nature of "Collecting/Comprehending in One" 935
4 Deciphering the Mathematical Statement of the Weak Theaetetus' Periodicity Theorem from the Philosophical Statement of the Theorem in the *Theaetetus* 147d-148b and Its Imitations by Plato in the *Politicus* 283a-287b and the *Philebus* 16c, 23a-25e 937
4.1 *Theaetetus* 147d3-148b2 939
4.2 The Mathematical Interpretation of the *Theaetetus* 147d3-148b2 Passage, Describing the Weak Theaetetus' Periodicity Theorem in Semi-Philosophical Terms 940
4.3 The Mathematical Statement of Theaetetus' Theorem Reported by Plato in the *Theaetetus* 147e2-148b3 in Semi-Philosophical Terms 945
4.4 "peiro mimoumenos" (πειρῶ μιμούμενος) in *Theaetetus* 148d4-5: The Plan to Imitate Theaetetus' Periodicity Theorem in Platonic Philosophy 945
4.5 Philosophical Imitation of the Weak Theaetetus' Periodicity Theorem in the *Philebus* 16c, 23b-25e: Every Intelligible Being Is the Mixture of the Anthyphairetic Infinite and the Anthyphairetic Finite 946

S. Negrepontis (✉) · V. Farmaki · M. Brokou
Department of Mathematics, Athens University, Athens, Greece
e-mail: snegrep@gmail.com; vfarmaki@math.uoa.gr; marinabrk@gmail.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_72

4.6 Philosophical Imitation of the Weak Theaetetus' Periodicity Theorem in the *Politicus* 283a-287b: Every Intelligible Being Is the Product of a Double Measurement, First Against the Opposite (Excluding a Pleasure, Which Is a Philebean Infinite, Hence a Finite), Secondly Against the Geometric Mean of These Opposites (a Generation, Hence an Infinite) 949
4.7 The Equivalence of the Three Passages Reveals the Statement of the Weak Theaetetus' Periodicity Theorem 952
4.8 The Philebean Finite Compared to the Pythagorean Finite 952
5 The Reconstruction of the Proof of the Weak Theaetetus' Periodicity Theorem, Employing Solely Tools from Book X of Euclid's *Elements* (the Apotome Lines (Propositions X.73-107), the Binomial Lines (Propositions X.36-72), and Their Conjugacy (Propositions X.112-114), and a Pigeon Hole Argument 953
5.1 Apotome Lines 953
5.2 Binomial Lines and Their Conjugacy with the Apotome Lines 958
5.3 The Realization of the Importance of the Conjugacy Propositions of Book X in the Modern Era 960
5.4 The Role of the Assumed Line ("προτεθείσα," "ἐκκειμένη") of Apotome and the Binomial Lines 962
5.5 The Transformation of the Anthyphairetic Relations of a to b for $a^2 = Nb^2$ for Some Non-Square Number to Conjugation Relations 962
5.6 The Recursive Definition of Generalized Side and Diameter Numbers 966
5.7 The Proof of the Weak Theaetetus' Periodicity Theorem**:** The Anthyphairesis of a to b for $a^2 = Nb^2$ for Non-Square Number N is Periodic 966
6 The Definition of the Intelligible Being "the Statesman" in the *Politicus* Is a Philosophical Imitation of the Strong Theaetetus' Palindromic Periodicity Theorem ... 967
6.1 The First Ten Steps of the Division and Collection (258b-267c) 967
6.2 The Need to Continue the Division (267c-268d and 274e-277c) 968
6.3 The Myth of Palindromic Periodicity with Kronos Era and Zeus Era (268d-274e) 968
6.4 The Division for the Statesman Must Be Continued According to the Myth of Palindromic Periodicity (276a1-7) 969
6.5 The Last Ten Division Steps and the Logos Criterion for the Statesman (267c-268d and 274e-277c, 289e-305e) 972
6.6 Logos in the Division and Collection of the Statesman 976
6.7 The Definition of the Statesman as a Philosophical Palindromically Periodic Anthyphairesis (*Politicus* 258c-267c, 289e-305e) 976
7 The Reconstruction of the Proof of the Strong Theaetetus' Palindromic Periodicity Theorem, Employing Methods Solely from the Theaetetean Book X of Euclid's *Elements* and the Pigeonhole Argument at the *Politicus* 272d-e 979
7.1 The Strong Theaetetus' Palindromic Periodicity Theorem 979
7.2 The Reconstruction of Theaetetus' Palindromic Periodicity Theorem 980
7.3 Plato's Argument for Going from the Kronos Era to the Zeus Era (*Politicus* 272d6-e6) Appears to be a Philosophical Version of the Pigeonhole Argument Employed in the Proof 983
8 The Pell Equation for $N = 2$ by the Pythagoreans 984
9 A Reconstruction of Theaetetus' Proof of the General Case of the Pell Property 988
9.1 The Tools for the Proof of the Pell Property in Book X of the *Elements* 988
9.2 The Crucial Step for a Reconstruction of Theaetetus' Proof of the General Pell Property, Employing the Palindromic Periodicity Theorem and Proposition X.97 990
9.3 The Reconstruction of Theaetetus' Proof of the General Case of Pell Property ... 991

9.4 The Third Purpose of Book X of the *Elements* Is the Extension of the Quadratic Anthyphairetic Periodicity to the Canonical Solids by Means of Application of Areas, in an Attempt to Counter the Successes of the Non-Anthyphairetic Approach by Archytas and Eudoxus 993
10 The Problem of Pell in Hindu Mathematics 999
10.1 Reception of the Hindus' Solution to the Problem of Pell in Europe 999
10.2 Brahmagupta's *Bhavana*/Composition 1000
10.3 The Unknown Origins of *Cakravala* 1005
10.4 A Simple Key Proposition, Deduced from the Reconstructed Proof of the Theaetetus' Periodicity Theorem, Shows Its Connection with *Cakravala* 1006
10.5 The Initial Step of the *Cakravala* Process 1007
10.6 The Third Step of the Bhaskara's II *Cakravala* Process for $N = 61$ and Its Relation to the Reconstructed Theaetetus' Method 1008
10.7 The General *Bhaskara II Cakravala* Step 1010
10.8 On the Cyclical Meaning of the Term *Cakravala* 1014
10.9 The *Cakravala* Method Appears to Have Similarities with the Reconstructed Theaetetus' Conjugacy Method 1015
11 Summary 1016
Bibliography 1018

Abstract

The reinterpretation of Plato's philosophy in terms of periodic anthyphairesis, in fact of palindromically periodic anthyphairesis in the *Politicus,* and the reading of Book X of *Euclid's Elements* under the new light, reveal deep mathematical contributions by Theaetetus, including a proof of the general Pell equation. Fascinating similarities of Theaeteus' reconstructed proofs with the Hindus' solution of the problem of Pell are noted.

Keywords

Plato · Theaetetus · Book X of Euclid's *Elements* · Palindromically periodic anthyphairesis · Pell equation · Cakravala

1 Introduction

Our study of Plato's work has shown that the basic entity in Plato's philosophical system, the intelligible Being, is a philosophical imitation of a dyad in periodic anthyphairesis. (The terms mentioned in the introduction will be defined in the main text below.) (The simple mathematics of anthyphairesis are reviewed in Sect. 2.) This becomes evident, among others, in the One of the second hypothesis of the *Parmenides* and in the Divisions of the Angler and the Sophist in the *Sophistes* (Sect. 3).

Thus, we should not be greatly surprised to find that Plato's philosophical theories are hiding and may reveal deep mathematical secrets. Especially so, as according to his own admission, Plato's later work was especially influenced by the mathematical discoveries of Theaetetus, as reported in the *Theaetetus* 147c-148d. We trace the imitation of which Plato is talking ("πειρῶ μιμούμενος," 148d4-5), in the *Philebus* 16c, 23a-25e, according to which an intelligible Being is the mixture of an anthyphairetic infinite/

incommensurable and an anthyphairetic finite/commensurable, and in the *Politicus* 283a-287b, according to which intelligible Beings are generated by the double measurements of an entity against its opposite, provided this is not a pleasure, and is thus finite, and against the geometric mean of the two, an infinite generation; and we recognize that both these are philosophical statements closely imitating Theaetetus' periodicity theorem, stated in semi-philosophical language in *Theaetetus* 147c-148d: "if a, b are lines commensurable in power only, in particular if $a^2 = Nb^2$ for some non-square number N, then the anthyphairesis of a to b is *periodic*" (Sect. 4).

More amazing, the elaborate definition of the Statesman, occupying the whole *Politicus*, is an imitation not only of periodic anthyphairesis, as was the case with the Divisions in the *Sophistes* (examined in Sect. 3), but of a *palindromically periodic* one, revealing with this imitation Theaetetus had established the Theorem: "if $a^2 = Nb^2$, N non-square, then the anthyphairesis of a to b is *palindromically periodic*," an important result that in modern era was gradually discovered by the work of Fermat, Wallis, and Brouncker (seventeenth century), and Lagrange and Euler (eighteenth century) (Sect. 6).

Armed with Plato's revelations we next attempt to reconstruct Theaetetus' proof. We find the complete set of tools in Book X of the *Elements*, a work known to be by Theaetetus. Already in 1572 Bombelli realized that Book X is devoted to the *conjugation* (established in Propositions X.112-114 of the *Elements*) between *apotome* (defined in X.73 and studied in Propositions X.73-107) and *binomial* lines (defined in X.36 and studied in Propositions X.36-72), a conjugation so striking that made him introduce the complex numbers with the name of *binomio*. It is thus the ideal tool for the proof of the palindromic periodicity of quadratic anthyphairesis. Once we know, thanks to Plato, that Theaetetus had indeed proved the palindromic periodicity theorem and the tools of Book X, then the reconstruction of the proof is one-way, essentially obligatory, simply consisting in the realization that the anthyphairetic remainders γ_n of the anthyphairesis of a to b, for $a^2 = Nb^2$, N non-square, are simple apotome lines of the form $\gamma_n = (-1)^n(q_n b - p_n a)$, and in the transformation of the anthyphairetic relations ($\gamma_{n-1} = I_n\gamma_n + \gamma_{n+1}$) into relations of conjugation $\varphi_n(I_n b + \varphi_{n+1}) = b^2$, necessarily with the help of the reduced remainders φ_n defined by $\gamma_n\, \varphi_{n+1} = \gamma_{n+1}\, b$.

The one basic argument yet needed for the proof and absent from Book X, the *pigeonhole principle*, is again gratefully supplied by Plato in precisely the point needed, namely when the first half of the period in the anthyphairetic division goes to its second palindromic half, in Plato's terms at the moment of transfer from the Kronos era to the Zeus era (*Politicus* 272d-e). We emphasize that the anthyphairetic interpretation of Plato's philosophy is what, despite his immersion in Greek philosophy and mathematics, Weil lacked, deciding, as he did, against an anthyphairetic use of the conjugation of Book X (Sects. 5 and 7).

Theaetetus had a strong motivation to deal with the Pell problem, since the case $N = 2$ was a venerated Pythagorean discovery not unnoticed by Plato (*Politeia* 546c), closely connected with their discovery of the periodicity of the anthyphairesis $a^2 = 2b^2$ (Sect. 8); and once Theaetetus obtained the palindromic periodicity theorem, the proof of the Pell property is within easy reach. But Theaetetus' preoccupation with the Pell problem can be traced in Book X: The classification of

apotome lines in six classes in X.84/85 makes sense only in relation to the Pell property. Thus, (i) the reconstructed Theaetetus' proof of periodicity in fact involves the Pell property, since the number λ_{n+1}, employed in the definition of m_{n+1} and φ_{n+1}, is the absolute value of the Pell number of the remainder γ_n; (ii) the first and the fourth apotome, in the classification X.84/85 of the apotome lines in the *Elements*, are apotomae with *positive Pell number* (corresponding to the *even remainders* of the anthyphairesis of a to b for $a^2 = Nb^2$, N non-square, and b the assumed "one foot" line), the second and fifth with *negative Pell number* (corresponding to the *odd remainders*); (iii) more impressive and relevant, the first and second apotome lines, in the classification X.84/85 of the apotome lines in the *Elements*, are those with *square Pell number*, the fourth and fifth with *non-square Pell number*; and (iv) Propositions X.97 produces an apotome with *square Pell number* from any apotome line. These are essentially the simple additional tools, available in Book X, needed to prove the Pell property (Sect. 9).

In conclusion, on the basis of our arguments that

(a) Plato's intelligible Beings are imitations of dyads in periodic anthyphairesis in his whole work
(b) Plato's intelligible Beings in the *Theaetetus*, *Politicus*, and *Philebus* are imitations of Theaetetus' theorem on the palindromically periodic anthyphairesis of lines incommensurable in length and commensurable in power
(c) The proof of Theaetetus' theorem employs the conjugacy of the quadratic lines apotome and binomial in the Theaetetean Book X of the *Elements*
(d) That there are strong indications that Book X is concerned with the proof of the general Pell property

we conclude that Theaetetus had a proof not only of the palindromic periodicity theorem, but also of the general property, a simple consequence of the palindromic periodicity theorem.

In the final Sect. 10 on the Hindus' contribution to the solution of the Pell problem we include the intriguing and fascinating evidence of the similarities between the reconstructed Theaetetus' proof of the palindromic periodicity of quadratic anthyphairesis and the Pell equality and the methods developed by the Hindus. We want to emphasize that we have no way of knowing if the origin of the methods developed by the Hindu mathematicians are in some way connected with Theaetetus' reconstructed proofs.

The Hindus' method of solving the problem of Pell has come in two stages.

First Brahmagupta in the seventh century AD, introduced the *bhavana* method, namely the composition rule of dyads, given a non-square natural number N:

$$(q,p)*(q',p') = (qq' + Npp',\ qp' + q'p),$$

used mostly for the square case ($q = q'$, $p = p'$); the *bhavana* * operation is multiplicative on the absolute value of the Pell number $q^2 - Np^2$ of the dyads (q, p).

The Brahmagupta rule of composition appears close to an arithmetized version of the geometric rule of "multiplying" simple apotomae or binomial lines, where $a^2 = Nb^2$:

$$(qb - pa)(q'b - p'a) = ((qq' + Npp')b - (qp' + q'p)a)b,$$

and which, for the square case, is closely connected with Proposition X.97 of the *Elements*.

The *bhavana* rule of squaring has been employed by Brahmagupta for the solution of the problem of Pell for those non-square N for which one can find a dyad (q, p) with Pell number equal to -4 or $+4$, or -2 or $+2$, or -1. These "global" reductions and shortcuts computationally are simple enough to verify, but their mathematics certainly rely on the palindromic nature of the periodicity of the quadratic anthyphairesis, since from a part of the period one can find the whole, and hence solve the Pell problem. The most interesting Brahmagupta reduction is the one when the Pell number is -4, because it appears to rely not only on the palindromicity but also on a "square root" argument related to the quite interesting Proposition X.91 of the *Elements*.

The second *cakravala*/circle method has been mostly advanced by Bhaskara II in the twelfth century AD, although it existed for at least a century before him. For a non-square number N, the basic step of *cakravala* is given by the Brahmagupta-type composition

$$(q, p)*(m/\lambda, 1/\lambda) = ((mq + Np)/\lambda, (q + mp)/\lambda),$$

where λ is the absolute value of the Pell number of the dyad (q, p), and m is a natural number to be determined by the *kuttaka* indeterminate method and an extremal condition.

Weil has presented the *cakravala* method, stripped of the mathematically unnecessary method of *kuttaka*, and unencumbered by ingenious shortcuts. In this "pure" form the *cakravala* method appears as the arithmetized form of the reconstructed Theaetetus conjugacy method presented in Sect. 5. Indeed, for a non-square number N, the *cakravala*

(a) Invariably starts with the dyad $(q_1 = m_1, p_1 = 1)$, or else, if N is nearer to $(m_1 + 1)^2$, with the dyad $(q_2 = m_1 + 1, p_2 = 1)$; these dyads appear as the arithmetized version(s) of the initial anthyphairetic remainder(s) $\gamma_1 = a - m_1 b$, or else $\gamma_2 = (m_1 + 1)b - a$, respectively, of the anthyphairesis of a to b, for $a^2 = Nb^2$.

(b) Proceeds inductively from the given dyad (q_n, p_n), whose Pell number ${q_n}^2 - N{p_n}^2$ has absolute value equal to λ_{n+1}, to the next dyad (q_{n+1}, p_{n+1}), with a Brahmagupta-type composition of the given dyad with the dyad $(m_{n+1}/\lambda_{n+1}, 1/\lambda_{n+1})$:

$$(q_n, p_n)*(m_{n+1}/\lambda_{n+1}, 1/\lambda_{n+1}) = \left(q_{n+1}, p_{n+1}\right),$$

where m_{n+1} is the greatest number satisfying the rule $m_n + m_{n+1} =$ multiple of λ_{n+1} and $m_{n+1} \leq m_1$; this procedure appears close to the arithmetized version of the geometric rule $\gamma_n \varphi_{n+1} = \gamma_{n+1} b$, namely

$$(-1)^n(q_n b - p_n a)\varphi_{n+1} = (-1)^{n+1}\left(q_{n+1}b - p_{n+1}a\right)b,$$

where φ_{n+1} has been determined by the conjugation relation $\varphi_n(I_n b + \varphi_{n+1}) = b^2$, into which the anthyphairetic relation $\gamma_{n-1} = I_n\gamma_n + \gamma_{n+1}$ has been transformed and thus shown by conjugacy to have the form $\varphi_{n+1} = (a - m_{n+1}b)/\lambda_{n+1}$, where m_{n+1} is determined by the rule $m_n + m_{n+1} = I_n\lambda_{n+1}$, according to the Theaetetus conjugacy method, as reconstructed in Sect. 5.

(c) Employs the term "*cakravala*," unmistakably indicating periodicity (and not just iteration).

This appears to refer to the periodicity of the geometric anthyphairesis of a to b, for the non-square number N, according to Theaetetus' periodicity theorem.

2 Geometric Anthyphairesis

2.1 Definitions

Let us recall the *geometrical concept of anthyphairesis* of a dyad of opposite magnitudes, the greater a and the smaller b, $a > b$:

$a = I_0 b + \gamma_1$,	with $\gamma_1 < b$,
$b = I_1\gamma_1 + \gamma_2$,	with $\gamma_2 < \gamma_1$,
...	
$\gamma_{n-1} = I_n\gamma_n + \gamma_{n+1}$,	with $\gamma_{n+1} < \gamma_n$,
$\gamma_n = I_{n+1}\gamma_{n+1} + \gamma_{n+2}$,	with $\gamma_{n+2} < \gamma_{n+1}$

...

Here

$$I_0, I_1, I_2, I_3, \ldots, I_n, I_{n+1}, \ldots,$$

denotes the sequence of successive quotients (natural numbers), and

$$a > b > \gamma_1 > \gamma_2 > \ldots > \gamma_n > \gamma_{n+1} > \ldots$$

the sequence of successive remainders (magnitudes) of the anthyphairesis of a to b.

If the anthyphairesis of a to b is *finite*, the last remainder necessarily divides the immediately previous remainder and is the greatest common divisor of a and b (Propositions VII.1, 2 of *Euclid's Elements* for natural numbers, Proposition X.3 for magnitudes).

In case a remainder never divides the immediately previous remainder, the anthyphairesis is said to be *infinite*. If the anthyphairesis of a to b is infinite, then a and b are *incommensurable* (Proposition X.2 of *Euclid's Elements*).

2.2 The Logos Criterion for Periodic Anthyphairesis

2.2.1 The Theaetetean Theory of Ratios of Magnitudes

Aristotle's *Topics* 158-9, a passage important for the history and reconstruction of Greek mathematics, informs us that before the theory of ratios of magnitudes that dominated Euclidean geometry, and which is Eudoxus' theory majestically presented in Book V of the *Elements*, there was a previous theory based on the

Definition If a, b are homogeneous magnitudes and c, d are homogeneous magnitudes then
$a/b = c/d$ if and only if $\mathrm{Anth}(a, b) = \mathrm{Anth}(c, d)$
(where $\mathrm{Anth}(a, b)$ denotes the sequence of successive *quotients* of the anthyphairesis of a to b).

A theory of ratios based on this pre-Eudoxian definition, If it were meant to be a theory as general as the Eudoxian theory, would need Eudoxus' condition (Def. 4 in Book V of the *Elements*), and so would be historically untenable. Knorr (1975) reconstructs such a theory which was quite adequate for the purposes of quadratic incommensurabilities, and for which, as can be shown, there is no need to an appeal to Eudoxus' condition. Parenthetically, it is worth mentioning that as carefully noted by Knorr, the crucial Theorem 6, p 338, the analogue of the fundamental Eudoxian Proposition V.8 of the *Elements*, makes essential use of the Eudoxus' condition (V, Def. 4). There is however evidence that this pre-Eudoxian theory applied only to a limited class of pairs of magnitudes, chiefly the lines that are rational with respect to an assumed line, in the sense of definition 3 of Book X of the *Elements*—hence the term "alogos" ("ἄλογος") for those ratios whose square is not commensurable. One anonymous comment in *Scholia in Euclidem* V.22 (eds: Heiberg and Stamatis 1977) alludes to a restricted, in fact quadratic, pre-Eudoxian theory of magnitudes, and another, *Scholia in Euclidem* X.2 (eds: Heiberg and Stamatis 1977), criticizes the general Eudoxian theory, clearly from a Platonic point of view and in favor of the earlier restricted quadratic theory of ratios of magnitudes.

It is generally accepted that the theory of ratios based on anthyphairesis is due to Theaetetus.

2.2.2 The Logos Criterion for Periodic Anthyphairesis

A natural consequence of Theaetetus' definition of equality of ratios for magnitudes is the following

Proposition (Logos Criterion for the Periodicity of an Anthyphairesis)

Let a, b be two line segments, with $a > b$, and anthyphairesis given as follows:

$a = I_0\, b + \gamma_1$, with $b > \gamma_1$,
$b = I_1\, \gamma_1 + \gamma_2$, with $\gamma_1 > \gamma_2$,
…
$\gamma_{n-1} = I_n\, \gamma_n + \gamma_{n+1}$, with $\gamma_n > \gamma_{n+1}$,
$\gamma_n = I_{n+1}\, \gamma_{n+1} + \gamma_{n+2}$, with $\gamma_{n+1} > \gamma_{n+2}$,
…
$\gamma_{m-1} = I_m\, \gamma_m + \gamma_{m+1}$, with $\gamma_m > \gamma_{m+1}$,
$\gamma_m = I_{m+1}\, \gamma_{m+1} + \gamma_{m+2}$, with $\gamma_{m+1} > \gamma_{m+2}$,
…
…….

and assume that there are indices n<m such that

$\gamma_n/\gamma_{n+1} = \gamma_m/\gamma_{m+1}$ ("Logos Criterion").

Then the anthyphairesis of a to b is eventually periodic, in fact

$$\mathrm{Anth}(a, b) = [I_0, I_1, \ldots, I_n, \text{period } (I_{n+1}, I_{n+2}, \ldots, I_m)].$$

Proof. By the Theaetetean definition of analogy, $\mathrm{Anth}(\gamma_n, \gamma_{n+1}) = \mathrm{Anth}(\gamma_m, \gamma_{m+1})$. But $\mathrm{Anth}(a, b) =$

$$\left[I_0, I_1, \ldots, I_n, \mathrm{Anth}\left(\gamma_n, \gamma_{n+1}\right)\right] =$$

$$\left[I_0, I_1, \ldots, I_n, I_{n+1}, I_{n+2}, \ldots, I_m, \mathrm{Anth}\left(\gamma_m, \gamma_{m+1}\right)\right] =$$

$$\left[I_0, I_1, \ldots, I_n, I_{n+1}, I_{n+2}, \ldots, I_m, \mathrm{Anth}\left(\gamma_n, \gamma_{n+1}\right)\right] =$$

$$\left[I_0, I_1, \ldots, I_n, I_{n+1}, I_{n+2}, \ldots, I_m, I_{n+1}, I_{n+2}, \ldots, I_m, \mathrm{Anth}\left(\gamma_m, \gamma_{m+1}\right)\right] =$$

…=
$\mathrm{Anth}(a, b) = [I_0, I_1, \ldots, I_n, \text{period } (I_{n+1}, I_{n+2}, \ldots, I_m)]$.

Proposition If a, b, c, d are lines, such that $ad = bc$, then $\mathrm{Anth}(a, b) = \mathrm{Anth}(c, d)$.

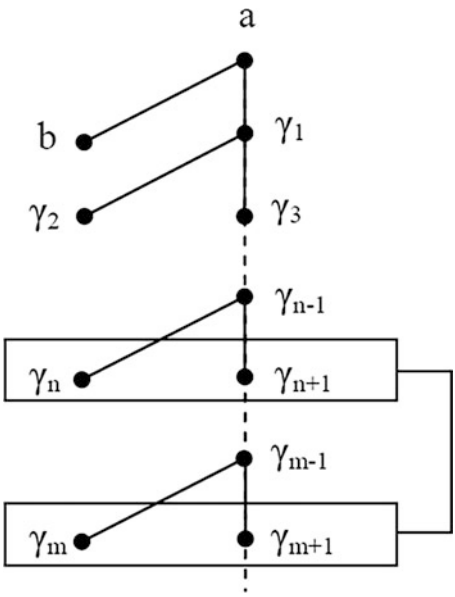
Proof Let
$a = I_0 b + \gamma_1$, $\gamma_1 < b$.
Hence $ad = I_0 bd + \gamma_1 d$. From hypothesis, $bc = I_0 bd + \gamma_1 d$.

By Proposition I.44-45 of the *Elements*, there is line δ_1, such that $\gamma_1 d = \delta_1 b$.
Hence $bc = I_0 bd + \delta_1 b$. Hence $c = I_0 d + \delta_1$. We note that $\delta_1 < d$, since $\gamma_1 < b$.
Hence the relation $c = I_0 d + \delta_1$, $\delta_1 < d$ is an anthyphairetic relation.
Thus $\mathrm{Anth}(c, d) = [I_0, \mathrm{Anth}(d, \delta_1)]$, $\mathrm{Anth}(a, b) = [I_0, \mathrm{Anth}(b, \gamma_1)]$,
and $\gamma_1 d = \delta_1 b$, as in the first step. The proof is now completed by induction.

Note. If we take as the definition of analogy the Theaetetean definition in terms of equality of anthyphairesis, then the proposition is a tautology.

2.2.3 An Abbreviated Representation of the Logos Criterion

We note that the criterion is stated solely in terms of the remainders, and no use is made in its statement of the quotients; thus, the criterion can be adequately represented in the following abbreviated, Zeno-like, form:



2.2.4 Example. The Anthyphairesis of *a* to *b* with $a^2 = 19b^2$

We exhibit the Logos Criterion with the first incommensurability that Theodorus did not prove in his lesson to Theaetetus and his companion in *Theaetetus* 147c-d.

(i) *The anthyphairesis of a to b*

$$a = 4b + \gamma_1, \gamma_1 < b \text{ (whence } \gamma_1 = a - 4b),$$
$$\text{(and } b = 2\gamma_1 + \gamma_2, \gamma_2 < \gamma_1 \text{(whence } \gamma_2 = 9b - 2a)),$$

$$\gamma_1 = \gamma_2 + \gamma_3, \gamma_3 < \gamma_2 \text{ (whence} \gamma_3 = 3a - 13b),$$
$$\text{(and } \gamma_2 = 3\gamma_3 + \gamma_4, \gamma_4 < \gamma_3 \text{(whence } \gamma_4 = 48b - 11a)),$$

$$\gamma_3 = \gamma_4 + \gamma_5, \gamma_5 < \gamma_4 \text{ (whence } \gamma_5 = 14a - 61b),$$
$$\text{(and } \gamma_4 = 2\gamma_5 + \gamma_6, \gamma_6 < \gamma_5 \text{(whence } \gamma_6 = 170b - 39a)),$$

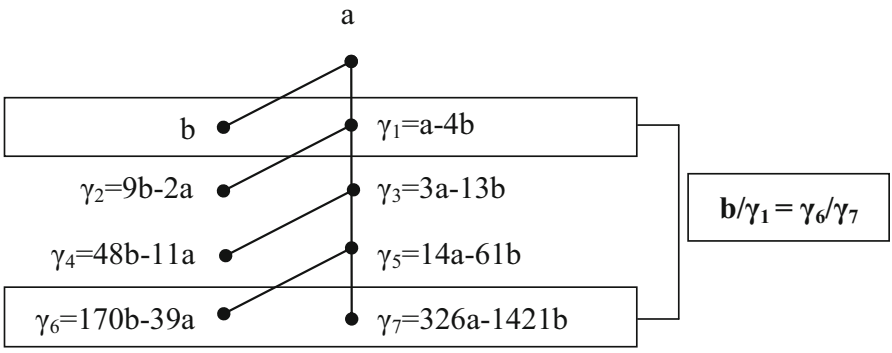
$$\gamma_5 = 8\gamma_6 + \gamma_7, \gamma_7 < \gamma_6 \text{ (whence } \gamma_7 = 326a - 1421b).$$

(ii) *Verification of the Logos Criterion* $b/\gamma_1 = \gamma_6/\gamma_7$.
In fact, $b \cdot \gamma_7 = b \cdot (326a - 1421b) = \gamma_1 \cdot \gamma_6 = (a - 4b) \cdot (170b - 39a)$.
The Logos Criterion results in the full knowledge of the ratio a/b:

$$\text{Anth}(a, b) = [4, \text{period } (2, 1, 3, 1, 2, 8)].$$

By Proposition X.2, of course it is concluded that a, b are incommensurable.

(iii) *Abbreviated, Zeno-like, representation of the Logos Criterion*



3 The Anthyphairetic Basis of Plato's Philosophy: An Intelligible True Being Is the Philosophical Analogue of a Dyad in Periodic Anthyphairesis (*Parmenides*, *Sophistes*)

According to our interpretation, an *intelligible (Platonic true) Being* consists of a dyad of parts satisfying the *philosophical analogue of a dyad in periodic anthyphairesis* (The first publications on this interpretation can be found in Negrepontis (1999) and AVGI (2000). Details on the anthyphairetic interpretation of the second hypothesis of the *Parmenides* have appeared in Negrepontis n.d.; of the Divisions in the *Sophistes* and *Politicus* in Negrepontis 2019, and of the *Philebus* in Negrepontis 2005. In the present work we have included, in upgraded form, those parts which are needed to make our arguments self-contained.). The paradigmatical true Being is the One of the second hypothesis in the *Parmenides*, consisting of the dyad <One, Being> (studied in Negrepontis n.d.), or of the dyad <Being, Not-Being>, such as <Just, Unjust>, in the *Sophistes* and *Politicus* (studied in Negrepontis 2012, 2018), in a philosophical analogue of geometric anthyphairesis, for which the exact analogue of the Logos Criterion for periodicity holds. The dialectic meaning of Logos is the Logos Criterion, the cause of periodic anthyphairesis.

3.1 The One of the Second Hypothesis of the *Parmenides* Consists of a Dyad <One and Being> in a Philosophical Analogue of Periodic Anthyphairesis

The paradigmatical true Being is the One of the second hypothesis in the *Parmenides* 142b1-155e3, consisting of the dyad

<One, Being>, or the dyad <Being, Not-Being> in the *Sophistes*, in the philosophical analogue of periodic anthyphairesis. The anthyphairetic interpretation of a platonic intelligible true Being has been obtained by our analysis of the dyad <One, Being> in the second hypothesis of the *Parmenides*, and of *Philebus* 15a-17a (Negrepontis n.d.).

3.1.1 The Infinite Anthyphairesis of the Dyad One and Being

In the first part, *Parmenides* 142d9-143a3, it is shown that the dyad < One, Being > is in infinite anthyphairetic division:

$$\begin{aligned}
&\text{One} = \text{Being} + \text{One}_1, \text{Being} > \text{One}_1,\\
&\text{Being} = \text{One}_1 + \text{Being}_1, \text{One}_1 > \text{Being}_1,\\
&\ldots\\
&\text{One}_n = \text{Being}_n + \text{One}_{n+1}, \text{Being}_n > \text{One}_{n+1},\\
&\text{Being}_n = \text{One}_{n+1} + \text{Being}_{n+1}, \text{One}_{n+1} > \text{Being}_{n+1},\\
&\ldots
\end{aligned}$$

Thus there is then an infinite multitude of remainders-parts of the anthyphairetic division:

$$\text{One} > \text{Being} > \text{One}_1 > \text{Being}_1 > \ldots > \text{One}_n > \text{Being}_n > \ldots$$

3.1.2 The Logos Criterion and the Anthyphairetic Periodicity of the *Parmenides* Dyad <One, Being>

We outline below the argument that shows that the dyad <One, Being> is not only in infinite anthyphairesis, but in fact in periodic one. In the sequel, the passage 144c2-d4 is concerned with the presence of the One in the Being, setting the dilemma whether the presence be of a part or of a whole of the One part in each part of the Being.

As noted by previous scholars, the passage has similarities to the passage 130e4-131e7 in the introduction of the *Parmenides*, posing the question (the "dilemma of participation") whether the presence of the One, be as part or as whole, in each of the sensibles that participate in the One.

However, the "whole" horn of the dilemma is rejected in both 144c-d and 130e-131e, but the "part" horn *only* in the 130e-131e, while fully accepted in 144c2-d4. Because of this crucial difference, we are led to conclude that the

"part" horn of the dilemma must have been found acceptable somewhere in the dialogue in between. And indeed, as it has not been noted by previous scholars, the place where the "part" horn is accepted is found to be the passage 138a3-7, where it is stated that the presence of the One in the parts of the Being is realized by a circle of "touchings" ("ἅψεις"). A "touching," as explained in the passage 148d5-149d7, occurs between parts of the dyad <One, Being>, successive in their order of generation, and the touchings generate numbers according to the formula

$$\text{Number of parts} - \text{units} = \text{touchings} + 1;$$

whose origin is clearly musical with terms-chords and musical intervals in place of parts-units and touchings, respectively. This correspondence suggests the interpretation that a "touching" in 138a3-7 is meant to be a ratio of consecutive parts of the anthyphairesis of One, Being.

In view of the earlier statement that the division of the One to Being is anthyphairetic, this interpretation is tantamount to the statement that the anthyphairesis of One to Being satisfies the "Logos" Criterion: there is an index k such that

$$\text{One}/\text{Being} = \text{One}_k/\text{Being}_k,$$

and thus finally the dyad <One, Being> satisfies a philosophical analogue of periodic anthyphairesis.

This interpretation is strengthened by the realization that the *Philebus* 15a1-c3 sets a whole-part dilemma, precisely as in the *Parmenides* 144c2-d4, and that this is followed by *Philebus* 15d4-5, a veritable replica of 138a3-7, with "logoi" ("λόγοι")-ratios running in circles in place of "touchings" in circles.

3.1.3 Equalization of Parts, in Consequence of Periodicity

The equalization of the One with the Being now proceeds smoothly, in the passage 144d4-e3 as follows:

$$\text{the } \textit{number} \text{ of parts of the One} =$$

$$\text{logoi between successive parts (from One on)} + 1,$$

a *finite* number, since the "logoi" will be different till a complete circle-period of ratios is achieved, and there will be no other ratios. The periodicity then implies that the number of parts of the One = the number of parts of the Being.

The equalization ("ἐξισοῦσθον") of One and Being rests on the above equality.

3.1.4 The Self-Similar Property of One and Many

The two basic features of the One of the second hypothesis are summarized by the statement that the One of the second hypothesis is both One and Many, in the sense that

The One is divided anthyphairetically into Many, in fact an infinite multitude of, parts.

Each part of the Many is equalized by the periodicity of the anthyphairesis, and thus the One of the second hypothesis is indeed a One in the sense of self-similarity, namely that each part is the same as the whole (144e8-145a2).

3.1.5 The One of the Second Hypothesis in the *Parmenides* is the Paradigmatical True Intelligible Being

The One of the second hypothesis is declared, in 155d3-6, to be a true intelligible Being, a Platonic Idea.

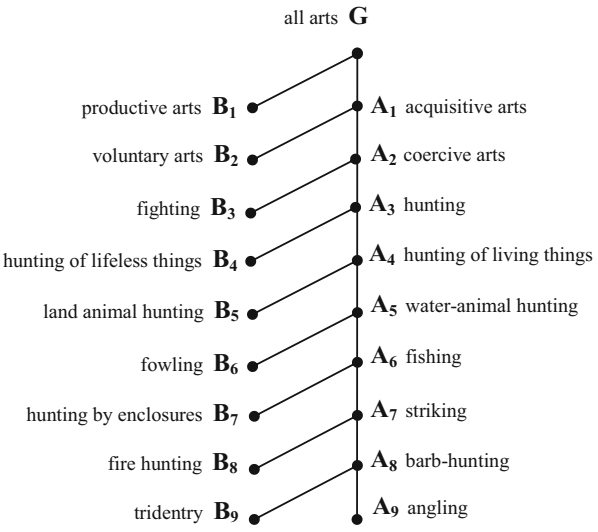
The One of the second hypothesis, although infinite in structure, nevertheless is fully knowable, possesses "ἐπιστήμη" 155d6, since full knowledge of the One is achieved, at a finite stage, when a full period of the Logoi is completed.

3.2 The Division and Collection of the Platonic Idea "the Angler" in the *Sophistes* 218b-221c

3.2.1 The Division of the Angler

The method of Division and Collection, also called "Name and Logos" (cf. *Theaetetus* 201e2-202b5, *Sophistes* 218c1-5, 221a7-b2, 268c5-d5) is exemplified at the start of the *Sophistes* 218b-221c by the definition of the Being Angler. In the scheme below, the binary division process which leads to the Angler is reproduced.

The Division of the Angler (*Sophistes* 218b-221c)



3.2.2 Logos Criterion of the Angler

The description of the Logos Criterion of the Angler is contained in the *Sophistes* 220e2-221c3 passage (which are presented below divided into two parts [A] and [B] for the sake of convenience) (The emphasis and italics in the present quote and in all future quotes are due to the authors.):

> [A] "Stranger: Then of striking which belongs to barb-hunting [A_8],
> that part which proceeds *downward from above ("ἄνωθεν εἰς τὸ κάτω"),*
> is called, because tridents are chiefly used in it, tridentry [B_9],
> I suppose. . ..
> Stranger: The kind that is characterized by the opposite sort of blow,
> which is practised with a hook and strikes,...
> and proceeds *from below upwards ("κάτωθεν εἰς τοὐναντίον ἄνω")*,
> being pulled up by twigs and rods.
> By what name, Theaetetus, shall we say this ought to be called?
> Theaetetus: I think our search is now ended and we have found the very thing we set before us a while ago as necessary to find.
> Stranger: Now, then, you and I are only agreed about the name of angling, [A_9]"
> (220e2-221b1)
> [B] "but *we have acquired also a satisfactory "Logos"* of the thing itself.
> *For* **(γὰρ)**
> of art as a whole, half was acquisitive,
> and of the acquisitive, half was coercive,
> and of the coercive, half was hunting,
> and of hunting, half was animal hunting,
> and of animal hunting, half was water hunting,
> and of water hunting [A_5]
> *the whole part from below ("τὸ κάτωθεν τμῆμα ὅλον")* was fishing, [A_6]
> and of fishing, half was striking,
> and of striking, half was barb-hunting, [A_8]
> and of this [A_8]
> the part in which the blow is pulled
> *from below upwards ("τὸ περὶ τὴν κάτωθεν ἄνω")* was angling. [A_9]" (221b1-c3).

(Plato, *Theaetetus* and *Sophistes*, trans: Fowler 1921), with modifications by the authors).

In [A], the opposing relation of tridentry to angling is carefully explained: all fishing with a hook is divided into

> Tridentry (=fishing with a trident), which is described as
> fishing with a hook with an art that proceeds **from above downwards**, and
> Angling (=fishing with a rod), which is described as
> fishing with a hook with an art that proceeds **from below upwards**.
> We have pushed Division all the way to the angling; thus, we have certainly found "the name" of angling.

But now, in [B], it is claimed that "the Logos" of the angling has been found, too.

The justification—the proof—that we have indeed found the Logos, too, is contained in the remainder of [B], since this remaining part of [B] starts with a "for" ("gar"), and this justification can be seen to consist of:

(i) An accurate recounting of all the division-steps, abbreviated in the sense that of the two species into which each genus is divided, only the one that contains the Angler is mentioned, while the species opposite to it is omitted
(ii) A reminder that the last species, the angling, is characterized as the part of its genus that proceeds "from below upwards"
(iii) The ONLY new information (since (i) and (ii) are repetitions of things already contained in the Division and in [A]), concerns the species of fishing, three steps before angling, and which informs us for the first time that this species is "the whole part from below" of its genus.

Since this is an abbreviated account, there is no explicit information on the species opposite to "fishing"—namely "fowling," but since "fishing" was described not simply as "the part from below" of its genus, but emphatically as "the whole part from below," it follows that the opposite species—"fowling"—must be characterized as "the (whole) part from above" of the same genus. In fact, there can be no other justification for the presence of the term "whole" in the description of "fishing" with a view to arguing that we have obtained "Logos," except to indicate and imply this description for its opposite species, "fowling."

We recall that the part of [B] from the word "for" ("γὰρ") on is explicitly a justification of the claim that we have succeeded in finding the "Logos" of the angling. We may then ask: What is the "Logos" of the Angler that reasonably results from such a justification? There can really be only one answer: The "Logos" we are looking for is the equality of the "philosophic ratio" of tridentry to angling, namely the equality of the ratio "of from above downwards to from below upwards," to the ratio "of fowling to fishing."

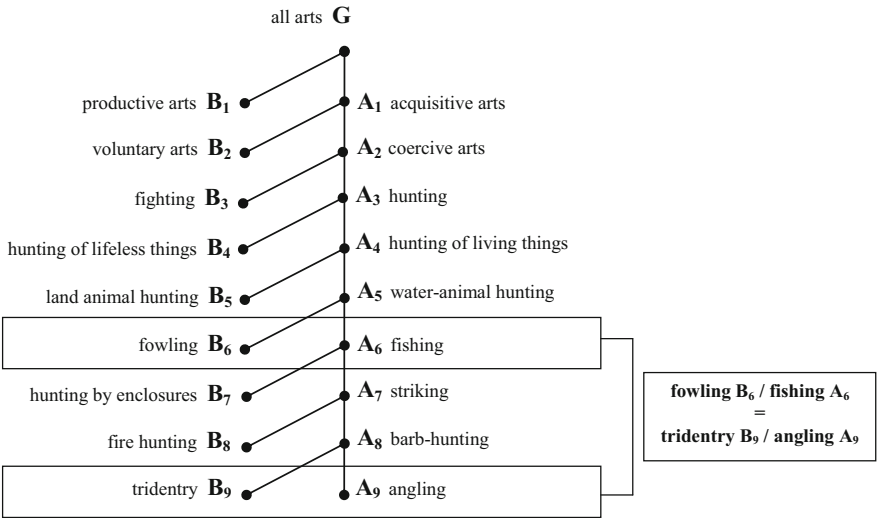
Thus

fowling B_6/fishing A_6 = tridentry B_9/angling A_9 =
from above downwards/from below upwards.

Since the species tridentry and angling form a pair of opposite species, and the species fowling and fishing form another pair of opposite species in the division scheme for the Angler, the resulting "Logos" bears a most uncanny similarity to the Logos Criterion for the periodicity of the anthyphairesis of geometric magnitudes, and of geometric powers in particular.

3.2.3 The Division and Collection of the Angler

The Division and Collection of the Angler thus takes the following form:



Thus the Division and Collection of the Angler consists of the Division, and of the Logos, described in b, analogous to the Logos Criterion for periodicity for geometric anthyphairesis established in Sect. 2.

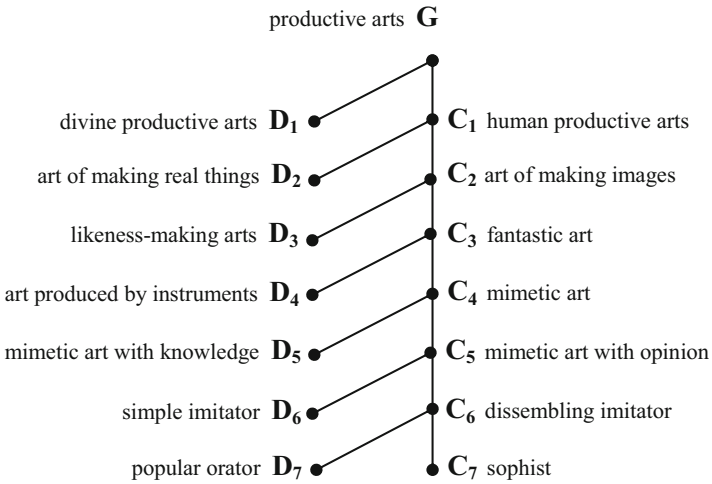
3.3 The Division and Collection of the Platonic Idea "the Sophist" in the *Sophistes* 234e-236d and 264b-268d

It now appears that the Division and Collection of a Platonic Idea—and the Angler is certainly a lowly paradigm of a Platonic Idea—is very much like the anthyphairetic division and the Logos Criterion of a geometric "power." When Socrates expressed, in the *Theaetetus* 145c7-148e5, his exhortation to imitate the geometric situation, it seems that he meant a much closer imitation that anybody has thus far suspected! But before proceeding on to wide-ranging conclusions, it would be prudent to examine whether the Division and Collection of the Sophist, in the *Sophistes* 264b-268d, has the same kind of structure and, in particular, whether there is a similar type of "Logos." So we shall take a look at the Division and Collection of the Sophist.

3.3.1 The Division of the Sophist

The Division for the Sophist follows the same pattern as the Division of the Angler, starting with a Genus, in this case "all the productive arts," proceeding by binary division of each Genus to two species, where the next Genus is that species of the previous step in the division that contains the entity to be defined, in this case the Sophist, and ending with the division-step that produces the Sophist as a species. The whole division scheme is as follows:

Division of the Sophist (*Sophistes* 264b-268d)



3.3.2 The Fundamental Analogy in the Divided Line of the *Politeia* 509d-510b and the Logos Criterion-Collection of the Sophist

We will make essential use of the fundamental analogy in the Divided Line of the *Politeia* 509d-510b (Plato 1969):

the opinable/the knowable = the likeness/that of which it is a likeness.

The Logos Criterion for the Sophist is now immediately recognized:

The art of making real things D_2/the art of making images C_2 = mimetic art with knowledge D_5/mimetic art with opinion C_5.

3.3.3 The Next Two Logoi of the Division and Collection of the Sophist Analogous to the Geometrical Ratios of Periodic Anthyphairesis

The definition of the Sophist presents an additional feature that ties it even closer to the mathematical model: There are two Logoi after the Logos Criterion, and if the mathematical anthyphairetic model is indeed followed, we would expect to have two further equalities of Logoi, namely the logos of the third step should be equal to the ratio of the sixth division-step, and the logos of the fourth division-step should be equal to the ratio of the seventh and final division-step.

(c1) *The equality of the third and sixth ratios of the Division of the Sophist*. In the third division-step, the Genus "humanly produced images" is divided into two species,

"Stranger: I see the likeness-making art ("εἰκαστικὴν") as one part of [the image-making art]. This is met with, as a rule, whenever anyone produces the imitation by following the proportions of the

original in length, breadth, and depth, and giving, besides, the appropriate colors to each part." (235d6-e2);
"Stranger: Now then, what shall we call that which appears, because it is seen from an unfavorable position, to be like the beautiful, but which would not even be likely to resemble that which it claims to be like, if a person were able to see such large works adequately? Shall we not call it, since it appears, but is not like, an appearance?
Theaetetus: Certainly." (236b);
"Stranger: And to the art which produces appearance, but not likeness ("φάντασμα ἀλλ' οὐκ εἰκόνα"), the most correct name we could give would be "fantastic art" ("φανταστικὴν"), would it not?
Theaetetus: By all means.
Stranger: These, then, are the two forms of the image-making art ("εἰδωλοποικῆς") that I meant, the likeness-making ("εἰκαστικὴν") and the fantastic ("φανταστικήν")." (236c3-7);
"Stranger: We must remember that there were to be two parts of the image-making class ("εἰδωλουργικῆς"), the likeness-making ("εἰκαστικόν") and the fantastic ("φανταστικὸν")" (266d8-9).

In the sixth division-step, described in 267e7-268a8, the Genus opinionated-imitator is divided into the two species, "the simple" ("ἁπλοῦν") and "the dissembling" ("εἰρωνικὸν") opinionated-imitator, the simple imitators being those who

> "think they know that about which they have only opinion,"

while the "dissembling" are those who

> "because of their experience in the rough and tumble of arguments, strongly suspect and fear that they are ignorant of the things which they pretend before the public to know."

Thus the simple imitators do not distort their opinion, but rather express a likeness of their opinion, while the dissemblers distort and disguise their opinion behind a false appearance. Therefore we have

> *Proposition*. The ratio of the third step in the Division of the Sophist,
> namely **the ratio of the likeness-making arts to the phantastic arts**,
> is equal to
> the ratio of the sixth step in the Division of the Sophist,
> namely **the ratio of the simple imitator to the dissembling imitator**.

(c2) *The equality of the fourth and seventh ratios*. In the fourth division-step, described in 267a1-b3, the Genus "fantastic arts" is divided into two species, as follows:

"Stranger: Let us, then, again bisect the fantastic art.
Theaetetus: How?
Stranger: One kind is that produced by instruments, the other that in which the producer of the appearance offers himself as the instrument.
Theaetetus: What do you mean?
Stranger: When anyone, by employing his own person as his instrument, makes his own figure or voice seem similar to yours, that kind of fantastic art is called mimetic."

Thus, in mimetic art the instrument of imitation is the imitator himself, while in the nameless opposite art the instrument of imitation is other than the imitator.

In the seventh division-step, described in 268a9-c4, the dissembler is divided into the demagogue and the sophist, the demagogue being

> "one who can dissemble in long speeches in public before a multitude,"

while the sophist is someone

> "who does it in private in short speeches and forces the person who converses with him to contradict himself."

Thus, he who listens to a dissembler is deceived, if that dissembler is a demagogue, and contradicted not by himself but by another instrument of deceit (namely the demagogue himself), while if the dissembler is a sophist, the listener is forced by the sophist to himself become the instrument of deceit.

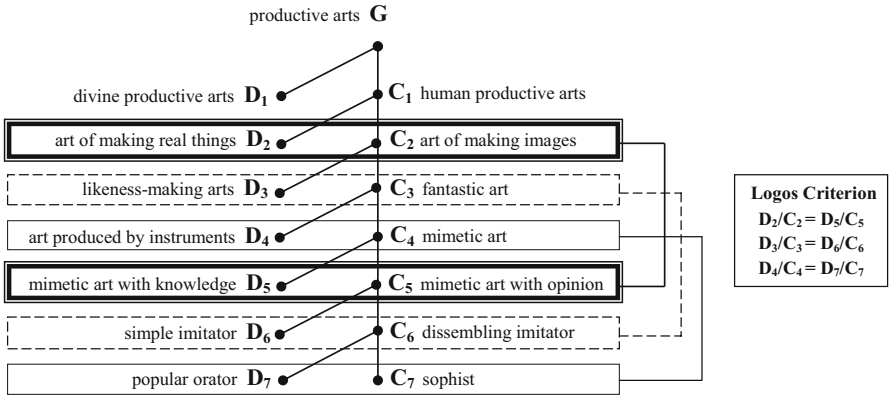
Thus we have:

> *Proposition.* The ratio of the fourth step in the Division of the Sophist,
> namely the ratio of the imitator who uses other instruments of imitation to the imitator who is himself the instrument of imitation
> is equal to
> the ratio of the seventh step in the Division of the Sophist,
> namely the ratio of the demagogue,
> whose listener is contradicted and deceived not by himself but by another, to the sophist,
> whose listener is forced to be contradicted and deceived by himself.

Propositions (c1) and (c2) provide powerful additional evidence in favor of the anthyphairetic interpretation of Division and Collection.

3.3.4 The Division and Collection of the Sophist

The complete Division and Collection of the Sophist (*Sophistes* 264b-268d) can thus be summarized in the following scheme:



Thus, we have full confirmation for the anthyphairetic meaning of Logos that we proposed in the Division and Collection of the Angler. This time, the Logos is not of the simple-minded, obvious type that appeared in the definition of the Angler (from above downwards/from below upwards), but of a more sophisticated and philosophical kind, playing a central role in the dialectics of the *Politeia* (The correlation of the *Politeia* ratio in the divided line with the Division and Collection of the Sophist in the *Sophistes* is hinted at by Proclus in *eis Politeian* 1,290,7–10).

3.4 The Cyclical-Periodic Nature of "Collecting/Comprehending in One"

There are several descriptions of the method of Division and Collection in Platonic dialogues, specifically in the *Sophistes*, *Politicus*, *Phaedrus*, and *Philebus*; in every one of these descriptions, Collection is described by an expression that denotes that it has a *cyclical-periodic nature*. Here is a list of such descriptions, indicating, in each of them, the term showing cyclical nature-periodicity.

[1]
Eleatic Stranger. Look sharp, then; it is now our business not to let the beast get away again
(ΞΕ. Ἄγε δή, νῦν ἡμέτερον ἔργον ἤδη τὸν θῆρα μηκέτ' ἀνεῖναι),
for we have almost *comprehended* (*περιειλήφαμεν*) him
into a kind of *encircling net of the devices* (*ἐν ἀμφιβληστρικῷ τινι*)
we employ *"logoi"* about such subjects (τῶν *ἐν τοῖς λόγοις* περὶ τὰ τοιαῦτα ὀργάνων). *Sophistes* 235a10-b2

The Sophist (the beast) must be defined by catching, comprehending, and encircling him in a net of "logoi." *Circularity* most clearly indicated by the terms *"περιειλήφαμεν" and "ἐν ἀμφιβληστρικῷ τινι."* The term *"περιειλήφαμεν"* is the same as the term used in Book XIII of *Euclid's Elements* to describe the *comprehension* of each of the five canonical solids in a sphere (e.g., Proposition XIII.15: Κύβον συστήσασθαι καὶ *σφαίρᾳ περιλαβεῖν*).

[2]
[*Eleatic Stranger*] The *division of species in genera* (*Τὸ κατὰ γένη διαιρεῖσθαι*) . . .
and of *many kinds* differing from one another (καὶ πολλὰς ἑτέρας ἀλλήλων)
but *circularly contained in one greater form* (*ὑπὸ μιᾶς ἔξωθεν περιεχομένας*)
and again of one idea (καὶ μίαν αὖ) through many wholes (δι' ὅλων πολλῶν) into one (ἐν ἑνὶ).
Sophistes 253d1- e3

The nature of Division and Collection is first Division and then cyclical Collection into One, cyclicity indicated somewhat ambiguously by the term "*περιεχομένας*."

[3]
Eleatic Stranger.Then let us try again (Πάλιν τοίνυν ἐπιχειρῶμεν);
let us *divide in two* (*σχίζοντες διχῇ*) the class we have taken up for discussion
and proceed always by way of the right-hand part of the thing divided (πορεύεσθαι κατὰ τοὐπὶ δεξιὰ ἀεὶ μέρος τοῦ τμηθέντος),

clinging close to the company to which the sophist belongs (ἐχόμενοι τῆς τοῦ σοφιστοῦ κοινωνίας),
until, having *circularly stripped* (*περιελόντες*) him of all common properties,
and left him only his own peculiar nature (τὴν οἰκείαν λιπόντες φύσιν ἐπιδείξωμεν μάλιστα μὲν ἡμῖν αὐτοῖς). *Sophistes* 264d10-265a1

This passage precisely describes the Division, as it is presented in the case of the Angler and the Sophist (cf. Sects. 3.2, 3.3 above), and indicates the circularity by the term "*περιελόντες*."

[4]
Eleatic Stranger. Why did *we go around in circle* (*περιήλθομεν ἐν κύκλῳ*)
making a great many *divisions* in vain? (πάμπολλα *διοριζόμενοι* μάτην)
Younger Socrates. For my part, I thought nothing that was said was in vain (μάτην οὐδὲν τῶν ῥηθέντων ἔδοξε ῥηθῆναι). *Politicus* 283b1-c2

Explicit statement that Division is circular and periodic: *"περιήλθομεν ἐν κύκλῳ,"* seemingly in vain, but in fact not so.

[5]
[*Eleatic Stranger*.] the *dissimilarities* of all sorts (τὰς δὲ αὖ παντοδαπὰς *ἀνομοιότητας*), when seen in multitudes (ὅταν ἐν πλήθεσιν ὀφθῶσιν),
he must find it impossible to be discouraged or to stop
(μὴ δυνατὸν εἶναι δυσωπούμενον παύεσθαι) until (πρὶν ἂν)
one *circularly confines* (*περιβάληται*) all the things which are related to each
(σύμπαντα τὰ οἰκεῖα) *into one circle of similarity* (*ἐντὸς μιᾶς ὁμοιότητος*),
circularly enclosing (*ἕρξας*) them, on the basis of their essential nature (οὐσίᾳ) of some genus. *Politicus* 285a4-b6

The circularity of Division and Collection here is strongly indicated by the terms *"περιβάληται," "ἐντὸς μιᾶς ὁμοιότητος," and "ἕρξας."*

[6]
[*Eleatic Stranger*.] By far our first and most important object should be to exalt
the method itself of ability to *divide in kinds* (*κατ' εἴδη. . . διαιρεῖν*),
and therefore, if a *"logos"* (*λόγον*), even though it be very long, makes the hearer better able to discover the truth, we should accept it eagerly and should not be offended by its length, or if it is short, we should judge it in the same way.
And, moreover, anyone who finds fault with *the length of "logoi"* (*λόγων μήκη*)
in our divisions and *of the circular periods* (*καὶ τὰς ἐν κύκλῳ περιόδους*)
must not merely find fault with the speeches for their length and then pass them quickly and hastily by. . . *Politicus* 286d6-287a6

The periodicity is described by the terms *"λόγων μήκη," "τὰς ἐν κύκλῳ περιόδους."*

[7]
Socrates. We say that *One and Many* (*ἓν καὶ πολλὰ*) are identified always, both now and in the past, by *"logoi"* (*ὑπὸ λόγων*) *running circularly around* (*περιτρέχειν*)
everywhere *in every intelligible being* (*καθ' ἕκαστον τῶν λεγομένων ἀεί*).
Philebus 15d4-8

In every intelligible Being, the Many, generated by anthyphairetic division, become One by the circularity, periodicity (*περιτρέχειν*) of the "logoi" resulting in the equalization and Collection of the Many into One.

[8]
and one way,
dividing the left-hand part (ὁ μὲν τὸ ἐπ' ἀριστερὰ *τεμνόμενος* μέρος),
continued to divide (*πάλιν τοῦτο τέμνων*),
did *not return to the previous parts* (*οὐκ ἐπανῆκεν πρὶν ἐν αὐτοῖς*)
discovered a sort of left-handed love ἐφευρὼν (ὀνομαζόμενον σκαιόν τινα ἔρωτα),
which it very justly reviled (ἐλοιδόρησεν μάλ' ἐν δίκῃ),
but the other way,
leading us to the right-hand part of madness,
discovered a love having the same name as the first, but divine, which it held up to view and praised as the author of our greatest blessings. *Phaedrus* 266a3-b1

Phaedrus 264e-266c. The Division and Collection of the right Love is periodic by its opposition to the explicit nonperiodicity of the left-sinister Love. The right, good eros, is the one who is not bad, like the left, its badness consisting in not having periodic division. Nonperiodicity is indicated by the terms: "*οὐκ ἐπανῆκεν πρὶν ἐν αὐτοῖς*."

[9]
is able to divide beings in kinds (κατ' εἴδη τε διαιρεῖσθαι τὰ ὄντα) and
to *circularly comprehend* (*περιλαμβάνειν*) each of them under one idea (μιᾷ ἰδέᾳ δυνατὸς).
Phaedrus 273d8-e4

Circularity indicated, as in [1], by the term *"περιειλήφαμεν."*

4 Deciphering the Mathematical Statement of the Weak Theaetetus' Periodicity Theorem from the Philosophical Statement of the Theorem in the *Theaetetus* 147d-148b and Its Imitations by Plato in the *Politicus* 283a-287b and the *Philebus* 16c, 23a-25e

Our aim is to understand Theaetetus' contribution and to show that it is stated in mathematical terms as follows:

> *weak Theaetetus' Periodicity theorem*: if a, b are lines, such that $a^2 = Nb^2$ for a non-square number (or more generally, such that a, b are commensurable in square only), then the anthyphairesis of a to b is periodic.

We will not be content only for a direct examination of the passage *Theaetetus* 147d-148b, but will also look at (some of) the imitations announced in 148d5.

In the beginning paragraphs of the dialogue *Theaetetus* (145d6, e9), written in memory of the deceased in battle Theaetetus, the question "what constitutes knowledge of an intelligible Being?" is set by Socrates; soon after (147d3-148b2) Plato

describes a mathematical discovery by Theaetetus (and his friend) and he makes clear that this discovery by Theaetetus is of paramount importance for Plato's philosophy, in particular for the question set above, and that therefore Theaetetus' mathematical discovery is worthy of imitation in his philosophy, hence Socrates' urge to Theaetetus "try to imitate" (πειρῶ μιμούμενος, 148d5).

Our main concern in the present section, and the subsequent Sect. 6, is to decide what is the exact nature of Theaetetus' mathematical discovery. Plato's statements on Theaetetus' mathematical discovery and contribution are couched in philosophical language, not providing an unambiguous mathematical statement, and unfortunately we have no other ancient sources on that matter. It is clear that we will have a chance of success only if we reach a deep understanding of Plato's philosophy and of how Plato has proceeded with the imitation, in the hope that understanding the philosophical imitation will throw light in and decipher the mathematical original. This is no mean task, because the "imitation" turns out to extend throughout the whole trilogy *Theaetetus-Sophistes-Politicus*, dedicated to the knowledge of the intelligibles, and even beyond, to the work of Plato in his last years including notably the *Philebus*.

In the present Sect. 4 we will examine three crucial passages.

The *first* (in Sects. 4.1–4.4) is the description by Plato of the mathematical discovery of Theaetetus, following a lesson on some quadratic incommensurabilities by Theodorus, described in the *Theatetus* 147d3-148b2 in semi-philosophical language, according to which

> every power is infinite/incommensurable and is collected into One.

The *second* (in Sect. 4.5) is the principal philosophical statement in the *Philebus* 16c, 23a-25e, according to which

> every intelligible Being is the mixture of
> the two Philebean principles of Infinite and Finite, interpreted as
> Anthyphairetic Infinite/Incommensurability and
> Anthyphairetic Finite/Commensurability, respectively.

The *third* (in Sect. 4.6) is the difficult but illuminating passage in the *Politicus* 283a-287b, according to which

> every intelligible Being is the product of two measurements,
> the one against its opposite, from which pleasure is to be excluded, and
> the other against their mean.

A comparison of these key passages confirms that the late period of Plato's philosophy has been heavily influenced from the mathematical work of Theaetetus, and we are led to the following *weak form* of Theaetetus' Theorem on Quadratic Incommensurabilities.

Weak periodicity form of Theaetetus' Theorem. If N is a non-square natural number and a, b are lines such that $a^2 = Nb^2$, then the anthyphairesis of a to b is periodic.

In the next Sect. 5 we will reconstruct the proof of the periodicity theorem, employing tools solely in the Theaetetean Book X of the *Elements*, in Sect. 6 we will study in detail Plato's imitation of the strong form of Theaetetus' theorem on palindromic periodicity in the *Politicus*, and in Sect. 7, we will reconstruct the strong palindromic periodicity theorem, using essentially the same tools from Book X.

4.1 *Theaetetus* 147d3-148b2

The "distributive" interpretation of the plural "hai dunameis" ("αἱ δυνάμεις") in 147d7-8 opens the way for a meaningful understanding of this crucial passage. But why was Plato interested for Theaetetus' mathematical discovery? It is not simply to eulogize Theaetetus; he has something more significant in mind.

4.1.1 Socrates Asks Theaetetus What Is the Nature of Knowledge of an Intelligible Being 145d6, e9

Socrates, with a characteristic ironic understatement (μικρὸν δέ τι ἀπορῶ, 145d6), sets the basic question for the trilogy *Theaetetus, Sophistes, and Politicus*,

What constitutes knowledge (ἐπιστήμη) of an intelligible Being (ἐπιστήμη ὅτι ποτὲ τυγχάνει ὄν, 145e9).

4.1.2 Theaetetus Realizes that Socrates' Question Is Similar to a Mathematical Discovery He Had Made Recently 147c7-d1

After some abortive attempts, Theaetetus realizes that Socrates' question on the acquisition of knowledge of intelligible Beings (ἐπιστήμη) is similar, relevant to his (and his companion's) treatment of the mathematical problem, conceived during Theodorus' lesson on quadratic incommensurabilities.

Theaetetus. It seems easy just now, Socrates, as you put it;
but you are probably asking the kind of thing that came up among us recently
(οἷον καὶ αὐτοῖς ἡμῖν ἔναγχος εἰσῆλθε)
when your namesake, Socrates here, and I were talking together. 147c7-d1

4.1.3 Theaetetus' Definition of "dunamis" 148a6-b1

Theaetetus. Those lines γραμμαὶ which square (τετραγωνίζουσι) the square (ἰσόπλευρον) number
we called length (μῆκος), and
those which square the non-square (ἑτερομήκη) number
we called powers/"dunameis" (δυνάμεις). 148a6-b1

4.1.4 Theodorus' Lesson 147d3-6

Theodorus here was drawing some figures (ἔγραφε) for us
about powers/"dunameis", about the three square feet and five square feet

showing [ἀποφαίνων] these "dunameis" are not commensurable in length (μήκει οὐ σύμμετροι) with the one foot line, and so, selecting (προαιρούμενος) each one in its turn up to the seventeen square feet "dunamis"
and at that he somehow stopped (πως ἐνέσχετο). 147d3-6

4.1.5 Theaetetus Mathematical Discovery 147d7-e1

Now we had the idea (ἡμῖν οὖν εἰσῆλθέ τι τοιοῦτον),
since the "dunameis" were shown (ἐφαίνοντο) to be
infinite in multitude (ἄπειροι τὸ πλῆθος),
to try (πειραθῆναι) *to collect them into one (συλλαβεῖν εἰς ἕν),*
by which we could henceforth call all the "dunameis". 147d7-e1

4.1.6 Explaining Further Theaetetus' Mathematical Discovery 148b1-2, d5-6

because (ὡς)
[the dunameis] are not commensurable in length (μήκει μὲν οὐ συμμέτρους) with
the others [namely, the lengths],
but [are commensurable] only in the areas of the planes which they have the power (δύνανται) to form. 148b1-2
Socrates. you *comprehended (περιέλαβες)* them [the "dunameis"],
though they were many (πολλὰς οὔσας), *in one kind (ἑνὶ εἴδει).* 148d5-6
(Plato, *Theaetetus*, trans: Fowler 1921, with modifications by the authors)

4.2 The Mathematical Interpretation of the *Theaetetus* 147d3-148b2 Passage, Describing the Weak Theaetetus' Periodicity Theorem in Semi-Philosophical Terms

4.2.1 The Distributive Plural Reading of "dunameis" in 147d7-8: *Every* "dunamis" Separately, *Not All* "dunameis" Collectively, Possesses Division Ad Infinitum and Collection into One

Our starting point is our attempt to make some sense of how Theodorus' lesson resulted in that

"the "dunameis" were shown (ἐφαίνοντο) to be infinite in multitude (ἄπειροι τὸ πλῆθος)" (147d7-8, in Sect. 4.1.5 above).

4.2.2 The Traditional Reading of 147d7-8 and Its Rejection

We look at Burnyeat's (Burnyeat 1978) treatment:

"The key sentence is 147d 7-e 1, which I render as follows:
*Since the dunameis were turning out to be unlimited in number, it occurred to us to attempt to collect them up into a single way of speaking [*i.e*., a formula or definition] of all these dunameis together.*
Theaetetus is recounting the thoughts suggested to himself and his companion *by and during Theodorus' lesson*, and the idea that there is an endless series of whole number squares (or sides of such squares) *would hardly need to be prompted by a process as protracted as Theodorus' lesson*.

> That there are an indefinite, perhaps infinite, number of squares with incommensurable sides, on the other hand, is precisely the hypothesis that would suggest itself as Theodorus proceeded from case to case proving more and yet more examples of incommensurability, perhaps by a method which could be endlessly reapplied.
> Therefore, it is likely that, in context, "all these dunameis" refers to squares with incommensurable sides rather than to squares generally."

Thus Burnyeat correctly rejects the interpretation that

> the expression "infinitely many powers" in 147d7-8 should mean "infinitely many non-square numbers,"

as this infinity would be already known, trivial anyway, and not prompted by Theodorus' lesson.

But the suggestion of Burnyeat that

> the expression "infinitely many powers" in 147d7-8 should mean "infinitely many incommensurable sides,"

is open to the same objection, namely that it "*would hardly need to be prompted by a process as protracted as Theodorus' lesson,"* since the incommensurability of the power for $N = 2$, known to the Pythagoreans a long time ago and for this reason expressly omitted from Theodorus' lesson, is seen to imply the *incommensurability of the infinitely many powers,* namely for each N of the form $N = 2 \cdot n^2$, with $n = 1$, 2,. . . Thus Burnyeat's interpretation should be equally rejected.

4.2.3 The Distributive Plural Interpretation of "dunameis" in 147d7-8

Once these two interpretations are rejected, it is hard to think of any other interpretation of an infinite set of "dunameis" distinguished by Theodorus' lesson.

We have already argued in (Negrepontis 2018) that the only interpretation of the sentence in question that makes sense is by switching to a *distributive plural,* and render the passage *Theaetetus* 147d7-e1 as follows:

> Because, as a result of Theodorus' lesson, it was shown that *every power* was divided into an infinite multitude of parts ("ἄπειροι τὸ πλῆθος") (147d7-8), Theaetetus and his companion had the idea and in fact succeeded in *collecting* [the infinitely many parts of] *every power* into One ("συλλαβεῖν εἰς ἕν").

Distributive Plural of Dunameis Before and After 147d7-8 Throughout Passage 147d3-148b2

We find initial support for this distributive interpretation in the observation that in fact the plural about powers is distributive *throughout passage 147d3-148b2*:

in the occurrence of "dunameis" in 147d4-5 (in item 3.1.4 above),
just before 147d7-8

> [Theodorus] showing [ἀποφαίνων] these "dunameis" are not commensurable in length (μήκει οὐ σύμμετροι) with the one foot line

since it means of course that every "dunamis" separately is not commensurable in length with the one foot line and not all the "dunameis" together; and,

in the occurrence of "dunameis" in 148b1-2 (in item 3.1.6 above),

just after 147d7-8

> [the dunameis] are not commensurable in length (μήκει μὲν οὐ συμμέτρους) with the others [namely, the lengths],
> but [are commensurable]only in the areas of the planes which they have the power (δύνανται) to form,

since it means of course that every "dunamis" separately is incommensurable or commensurable to the one foot, and not all the "dunameis" together.

It is natural then to conclude that the plural about "dunameis" in the crucial in-between occurrence 147d7-8 (in Sect. 4.1.5 above) is distributive as well.

Schematically

[Theodorus] showing [ἀποφαίνων] these *"dunameis"* are not commensurable in length (μήκει οὐ σύμμετροι) with the one foot line 147d4-5	Distributive plural of "dunameis" before 147d7-8
"the *"dunameis"* were shown (ἐφαίνοντο) to be infinite in multitude (ἄπειροι τὸ πλῆθος)" 147d7-8	What plural of "dunameis" at 147d7-8?
[the *"dunameis"*] are not commensurable in length (μήκει μὲν οὐ συμμέτρους) with the others [namely, the lengths], but [the *"dunameis"*are commensurable] only in the areas of the planes which they have the power (δύνανται) to form 148b1-2	Distributive plural of "dunameis" after 147d7-8

Support for Distributive Reading of "dunameis" in 147d7-8 from *Scholia in Platonem*

A commentary to the *Theaetetus* passage 162e6-7 on Theodorus, found in the anonymous work *Scholia in Platonem* (abbreviated SIP), runs as follows:

> If we accept the judgement of the many as the dominant one in geometry,
> we would be ridiculous to claim
> (καὶ γὰρ εἰ τὴν τῶν πολλῶν κρίσιν λάβοιμεν ἐπὶ γεωμετρίας κυρίαν, γελοῖοι ἂν ὦμεν),
> that maginitudes are *incommensurable* to each other
> (ἀσύμμετρα λέγοντες ἀλλήλοις μεγέθη),
> and that *the finite line is divisible* ad infinitum
> (καὶ τὴν πεπερασμένην εὐθεῖαν διαιρετὴν εἶναι εἰς ἄπειρον,
> and *the like* (καὶ τὰ τοιαῦτα).

The following reading of the commentary is suggested:

Theodorus, when proving that each of the "dunameis" (of 3,5,. . .,17 feet) and the one foot line (the magnitudes) are *incommensurable* to each other, as in statement 147d3-6, employed,

not any considerations of probability

(which would be worthless and make him ridiculous),

but logically impeccable proofs ("ἀπόδειξιν δὲ καὶ ἀνάγκην," 162e4-5) involving the *division* ad infinitum of each of these finite lines,

so that every one of these lines, each power, is *infinite in multitude*, as in statement 147d7-8.

This reading leads to a distributive interpretation of "hai dunameis" in 147d7-8.

Support for Distributive Reading of "dunameis" in 147d7-8 from the *Anonymi Commentarius in Platonis Theaetetum*

The *Anonymi Commentarius in Platonis Theaetetum* (*ACIPT*), possibly Eudorus, first century BC (cf. Tarrant (Tarrant 1985)), lends strong support to the distributive interpretation. Indeed, commenting on the infinity arising from the "powers infinite in multitude," and their collection achieved by Theaetetus and his companion, in the *Theaetetus* 147d-148b passage, he makes the following statements:

[1]. an infinity (τὸ ἄπειρον) in general must be collected (περιλαμβάνειν καὶ ὁ[ρίζ]ειν), if possible (καθ' ὅσον ἐνδ[έ]χεται) *ACIPT* 37, 3-12;

[2]. a line admits of the infinite either by (infinite) increase or (infinite) division (αἱ γραμμαὶ ἐπιδέχονται τὸ ἀόριστον, εἴ τις αὐτὰς ἢ αὔξοι ἢ διαιροῖ) *ACIPT* 36,45-48;

[3]. the collection of an infinity in a line can be achieved by passing over to numbers (ὁρίζονται δ' ὑπὸ τῶν ἀριθμῶν) *ACIPT* 36,48-37,3; (μετέβησαν οὖν καὶ ἐπὶ τούτων ἐπὶ τοὺς ἀριθμούς, ἵνα περιορίσωσι καθολικῶι τινι) *ACIPT* 42, 30-33.

Note: Collection renders 'perilambanein' (as in the *Theaetetus* 148d6) (26,11; 37,5; 37,10; 37,29; 37,43; 37,46; 45,48; 46,39), periorizein (42,32) and horizein (37,1; 37,11).

[4]. a line infinite by increase cannot be collected by numbers (Ὁ ἀριθμὸς ἄπειρός ἐστιν κατὰ τὸ αὔξεσθαι· οὐ δύναται οὖν τις τὸ ἐπ' ἄπειρον προϊὸν περιλαβεῖν) *ACIPT* 37, 39-44;

Thus collection can be achieved only for a line divided into infinite parts

[5]. the collection of an infinity in a line can be achieved by passing over to numbers, because numbers are commensurable to each other (ἦλ]θ[ον] οὖν ἐπὶ τὸ[ν ἀριθμ]ὸν διὰ τὸ [.]κ[.....]ρν τῶι πάντα[ς] τοὺς ἀρι[θμ]ο[ὺ]ς συμμ[έτ]ρους [εἶ]να[ι] πρ[ὸς] ἀλλή[λους]) *ACIPT* 26,13-18;

[6]. the collection of an infinity in a line is achieved by passing from the less clear magnitudes to the more clear numbers (ὥσπερ δὲ οἱ περὶ Θεαίτητον μετέβησαν ἐπὶ τοὺς ἀριθμοὺς ὡς σαφεστέρους) *ACIPT* 32,1-4; (πρῶτον μὲν οὖν ἀπὸ τῶν ἀσαφεστέρων ἐπὶ τὰ σαφέστερα δεῖ μεταβαίνειν ὡς ἀπὸ τῶν μεγεθῶν μετέβησαν ἐπὶ τοὺς ἀριθμούς) *ACIPT* 44,50-45,3; namely from incommensurables to commensurables, that is to "as number to number" (ἔχει γὰρ λόγον, ὃν ἀριθμὸς πρὸς ἀριθμόν) *ACIPT* 41,40-42,18.

Note. Exactly the same point is made in the *Scholion* X.39 to the *Elements*.

> In analogy to the description for cubes: "they came to the cubes," as "the solid itself is commensurable to the solid; for it has ratio, as number to number, while the sides are incommensurable."

The interpretation of these statements is the following:
By statement [2] we conclude that

> the plural in the expression "the "dunameis" infinite in multitude", as already suspected, is distributive and "every "dunamis" infinite in multitude" is meant;

by statement [4] and the fact that Theaetetus and his companion succeeded in collecting the infinite in multitude "dunameis," we conclude that

> every "dunamis" is infinite in power by division ad infinitum;

by statements [3], [5], and [6] we conclude that

> the infinity of every "dunamis" is due to incommensurability; and,

by statements [3], [5], [6], we conclude that

> 'collection into One' concerns the infinite multitude of parts of a "dunamis" brought about by Theodorus' incommensurability proofs, and is made possible because a "dunamis" is not just any line incommensurable in length, but a line whose square is commensurable, namely a Theaetetean mixture of commensurability and incommensurability.

Support for Distributive Reading of "dunameis" in 147d7-8 from the Definition of a Platonic Idea as "dunamis" in the *Sophistes* 247d-248d

In the *Sophistes* 247d-248d Plato defines the Platonic Idea as a "dunamis." The meaning of this definition will be explained briefly. According to this definition the fundamental property of a Platonic Idea is that, in interaction, active or passive, with an opposite element, forming an indefinite dyad, it has the "power," in fact the "power to communicate" with each other (252d3, 253a8, 253e1, 254b8, 254c5), in fact the "power for equalization with each other" (257b).

Irrespective of the precise meaning of the definition, Plato imitates here the term used in the *Theaetetus* 147d-148d for the mathematical "dunamis." In the *Sophistes* the Division and Collection is a method for the acquisition of knowledge of a Platonic Idea, thus of a "dunamis"; it is certainly the case that every "dunamis" separately is Collected into One. This suggests again that in the *Theaetetus* the Collection of the "dunameis" (who were collected into One because they appeared to be infinitely many) is to be understood distributively, for each power separately.

4.3 The Mathematical Statement of Theaetetus' Theorem Reported by Plato in the *Theaetetus* 147e2-148b3 in Semi-Philosophical Terms

We are now able to interpret the passage 147e2-148b3 that describes in semi-philosophical terms Theaetetus' mathematical contribution. There are two main problems in producing a mathematical statement that makes us understand Theaetetus' mathematical discovery and contribution and reaching an understanding of Theaetetus' mathematical theorem:

(i) The kind of infinity in Sect. 4.1.5
(ii) The mathematical meaning of "collection," "comprehension" of the power(s) in One in Sect. 4.1.5 and Sect. 4.1.6.

Since, according to the clarification, Sect. 4.2.2 above, Theodorus' proofs were showing that every *power*, namely every line a such that $a^2 = Nb^2$ for some non-square number N, was *infinite* in multitude, namely was divided into an infinite multitude of parts, Theaetetus (and young Socrates) tried to collect ("συλλαβεῖν εἰς ἕν" [147d8], in Sect. 4.1.5), or to comprehend ("ἐνὶ εἴδει περιέλαβες" [148d6] in Sect. 4.1.6), the infinite multitude of these parts into One, namely according to our interpretation of Collection in Sect. 3, to equate all parts, something achieved by anthyphairetic periodicity.

Thus the mathematical statement of Theaetetus' theorem must be as follows:

Weak Theaetetus' periodicity theorem. If N is a non-square number, b is the one foot line, and $a^2 = Nb^2$, then the anthyphairesis of a to be is periodic.

4.4 "peiro mimoumenos" (πειρῶ μιμούμενος) in *Theaetetus* 148d4-5: The Plan to Imitate Theaetetus' Periodicity Theorem in Platonic Philosophy

In view of the importance for our work to determine with certainty the theorem that Plato tells us that Theaetetus proved, we will not be content with the arguments interpreting the passage itself that describes the theorem, but we will look carefully into the philosophical imitations of this theorem that Plato, as he tells us, employed (cf. *Theaetetus* 148d4-5); specifically we will look at two impressive imitations, in the *Philebus* 16c, 23b-25e (Sect. 4.5) and in the *Politicus* 283a-287b (Sect. 4.6).

4.5 Philosophical Imitation of the Weak Theaetetus' Periodicity Theorem in the *Philebus* 16c, 23b-25e: Every Intelligible Being Is the Mixture of the Anthyphairetic Infinite and the Anthyphairetic Finite

The most striking philosophical imitation of the weak Theaetetus' periodicity theorem is the central theme in the *Philebus* 16c, 23b-25e:

Every intelligible Being is the mixture of the the two Philebean principles, Infinite and Finite.

A dyad of lines that are incommensurable but have commensurable squares is a dyad of lines with infinite anthyphairesis with squares with finite anthyphairesis, thus a mixture of anthyphairetic infinite and anthyphairetic finite; and,

such a dyad has, according to the weak Theaetetus' periodicity theorem, periodic anthyphairesis, and is thus a geometric entity whose philosophic analogue is an intelligible Being, as we have seen in Sect. 2.

Thus to establish that the central *Philebus* thesis is indeed a philosophic imitation of the weak Theaetetus' periodicity theorem, we simply have to prove that the Philebean Infinite is the anthyphairetic Infinite and the Philebean Finite is the anthyphairetic Finite. This will be done in Sect. 4.5.1 below.

4.5.1 The Philebean Infinite Is the Philosophical Analogue of Infinite Anthyphairesis and of Incommensurability, and the Philebean Finite Is the Philosophical Analogue of Finite Anthyphairesis and of Commensurability (*Philebus* 23b-25e)

We first examine Plato's description of the *Philebus principle of the Finite*:

> And the things which do not admit of more and less and the like, but do admit of all that is opposed to them (τούτων δὲ τὰ ἐναντία πάντα δεχόμενα) —first *equality* and the equal (τὸ ἴσον καὶ ἰσότητα), then the *double* (τὸ διπλάσιον), and *anything which is a ratio of number to number* (πᾶν ὅτιπερ ἂν πρὸς ἀριθμὸν ἀριθμὸς) or measure to measure, all these might properly be assigned to the class of *the finite* (τὸ πέρας). 25a6-b2
> The class of the *equal* (ἴσου) and *double* (διπλασίου), and everything which puts an end (παύει) to the differences between opposites and makes (ἀπεργάζεται) them *commensurable* (σύμμετρα) and harmonious (σύμφωνα) by the introduction (ἐνθεῖσα) of number (ἀριθμὸν). 25d11-e2

(Plato, *Philebus*, trans: Fowler 1925, with modifications by the authors)

A comparison of these passages with Propositions 1-8 of Book X of the *Elements* immediately implies that the *Philebean principle of the Finite* is the philosophic analogue of *finite anthyphairesis* and thus, by means of Proposition X.2 of the Elements, of *commensurability.*

Proclus, in *Scholia in Euclidem* 7,1-5 (Morrow 1970) fully confirms the identification of Philebean Finite with commensurability:

> And if the Finite were absent (*τοῦ δὲ πέρατος ἀναιρεθέντος*),
> there would be no (*οὐκ ἄν ποτε*) commensurability (*συμμετρία*) or equality of ratios (κοινωνία λόγων), no identity of kinds (ταυτότης εἰδῶν) and equality (ἰσότης) in mathematics (*ἐν τοῖς μαθήμασιν ἐφαίνετο)*, nor anything else that belongs in the column of the better (ὅσα τῆς ἀμείνονός ἐστι συστοιχίας), nor any knowledge of such matters (οὐδ' ἂν ἐπιστῆμαι τῶν τοιούτων ἦσαν), nor any fixed and precise concepts (οὐδὲ καταλήψεις μόνιμοι καὶ ἀκριβεῖς). 7,1-5

We next examine Plato's description of the *Philebus principle of the Infinite*:

> For wherever they are present, they do not allow *each (of the dyad) to be a quantity* (ποσὸν ἕκαστον) to exist; they always (ἀεὶ) introduce in every instance an opposition—*(the dyad) more emphatic than that which is quieter, or* vice versa (σφοδρότερον ἡσυχαιτέρου καὶ τοὐναντίον) —and thus they create *the (dyad) more and less* (τὸ πλέον καὶ τὸ ἔλαττον), thereby doing away with *fixed quantity* (ποσὸν). 24c2-6
> For, as I said just now, if they did not abolish *quantity* (ποσὸν), but allowed it and *measure* (τὸ μέτριον) to make their appearance in the abode *of the (dyad) more and less* (μᾶλλον καὶ ἧττον), *the (dyad) emphatically and gently* (σφόδρα καὶ ἠρέμα), those latter would be banished from their own proper place. When once they had accepted *quantity* (τὸ ποσόν), they would no longer be *(the dyad) hotter or colder*; for *(the dyad) hotter and colder* (θερμότερον. . .ψυχρότερον) are always (ἀεὶ) progressing and never stationary; but *quantity* (ποσὸν) is at rest and does not progress. By this reasoning *(the dyad) hotter and its opposite* are shown to be *infinite* (ἄπειρον). 24c6-d7
> we ought to collect all things that are *scattered* (διέσπασται) and *split up* (διέσχισται) to be put in the class of *the infinite* (τὸ τοῦ ἀπείρου γένος). 25a1-3

An Infinite is a dyad of the general form "the (dyad) More and Less" (μᾶλλόν τε καὶ ἧττον) (25c5-d2), as for example the (dyad) Hotter and Colder (θερμοτέρου καὶ ψυχροτέρου, 24a7-8), or the (dyad) Emphatic and Gentle (τὸ σφόδρα καὶ ἠρέμα, 24c1-2). The first observation is that *the Philebean infinite is a dyad of opposites*.

The dyad More and Less is Infinite if it renews itself at every stage, ad infinitum (ἀεὶ), so that it becomes divided into an infinite multitude of parts. But if, on the other hand, the process stops then each part of the dyad turns into quantity/number, and becomes an element of the class Finite (Here are some passages, by Pythagoreans and neo-Pythagoreans, that employ "ποσὸν" (quantity) and "ἀριθμὸς" (number) as equivalent: Xenocrates, *Fragment* 87; Theon Smurneus 19,15-18; Nicomachus, *Introductio Arithmetica* (1,2;1,7,1,1-2; 2,22; 2,23,1,1-6), *Harmonicum enchiridion* (4,1,37-38; 6,1,1-6), and *Theologoumena arithmeticae* (20,1-21,2); lamblichus, *Commentaria in Nicomachi Geraseni Pythagorei introductionem arithmeticam* 1,49,1-3; Iamblichus, *De communi mathematica scientia* 7,33-61; Proclus, *Comments to Euclid* 35,21-36,3 (Procli Diadochi 1873); *Anonyma Scholia to Euclid's Elements* V.9, V.10, X.39.). But we have already seen that an element of

the Finite is a commensurable pair of finite anthyphairesis. It is now clear that *the Philebean Infinite* is a dyad in *infinite anthyphairesis*, and thus, by Proposition X.2 of the *Elements*, *incommensurable*.

Proclus, in *Scholia In Euclidem* 6,19-22 (Morrow 1970), fully confirms the identification of Philebean Infinite with incommensurability:

> If there were no Infinite (*τῆς μὲν ἀπειρίας οὐκ οὔσης*),
> all magnitudes would be commensurable (*τά τε μεγέθη πάντα σύμμετρα ἂν ἦν*) and
> there would be nothing irrational or ratioless (*οὐδὲν ἄρρητον οὐδὲ ἄλογον*),
> features that are thought to distinguish geometry from arithmetic (οἷς δὴ δοκεῖ διαφέρειν τὰ ἐν γεωμετρίᾳ τῶν ἐν ἀριθμητικῇ). 6,19-22

4.5.2 Every Intelligible Being Is the Mixture of the Infinite and the Finite (*Philebus* 16c, 23c-d, 25b)

According to the *Philebus*, each of the Platonic Ideas is a *mixture* of the Philebean *Infinite* and the Philebean *Finite* (Some scholars, e.g., Cherniss (Cherniss 1945), Ryle (Ryle 1966), do not think that "ta aei legomena einai" refer to the intelligible Beings; this position is to be understood within a general view that, Plato's repeated dithyrambic descriptions of Division and Collection not withstanding, downgrades the method.)

Socrates. One which is easy to point out, but very difficult to follow for through it *all the inventions of art* have been brought to light. See this is the road I mean.
Protarchus. Go on what is it?
Socrates. A gift of gods to men, as I believe, was tossed down from some divine source through the agency of a Prometheus together with a gleaming fire; and the ancients, who were better than we and lived nearer the gods, handed down the tradition that
all the things which are ever said to exist ("τῶν ἀεὶ λεγομένων εἶναι") *consist of one and many and have inherent in them the finite and the infinite* ("πέρας δὲ καὶ ἀπειρίαν ἐν αὐτοῖς σύμφυτον ἐχόντων"). *Philebus* 16c
Note. "a Prometheus" is most likely a reference to Pythagoras, whose philosophy is based on the principles of the Infinite and the Finite (Philolaus, *Fragments* 1, 2, 6 and Aristotle, *Metaphysics* 987a13-28). But see also Sect. 4.7 below.

Socrates. We said that God revealed in the universe two elements, *the infinite and the finite*, did we not?
Protarchus. Certainly.
Socrates. Let us, then, assume these as two of our classes, and a third, made *One by combining ("ἕν τι συμμισγόμενον") these two*. *Philebus* 23c-d

Socrates. Well, what shall we say is the nature of *the third class, made by combining ("τὸ μεικτὸν") these two? Philebus* 25b

Thus, we will be able to discover the nature of Platonic Ideas and Beings, if we understand the meaning of the terms "Philebean Finite," "Philebean Infinite," and "Philebean mixture" of Infinite and Finite.

4.6 Philosophical Imitation of the Weak Theaetetus' Periodicity Theorem in the *Politicus* 283a-287b: Every Intelligible Being Is the Product of a Double Measurement, First Against the Opposite (Excluding a Pleasure, Which Is a Philebean Infinite, Hence a Finite), Secondly Against the Geometric Mean of These Opposites (a Generation, Hence an Infinite)

4.6.1 The "double measurement" in the *Politicus* 284e2-8

According to the *Politicus* 284e2-8, numbers, magnitudes, and qualities are subject to the *double measurement*, as follows:

> *Eleatic Stranger.* We should evidently divide the art of measurement (*μετρητικήν*) into two parts in accordance with what has been said.
> *One part* is all the arts which measure (*μετροῦσιν*) number, length, depth, breadth, and thickness against their opposites (*πρὸς τοὐναντίον*);

(Plato, *Statesman*, trans: Fowler 1921, with modifications by the authors)

Thus a number, a magnitude, or a quality, say x, is measured firstly against its opposite, say y; measurement against the opposite is about dyads, such as length and shortness, excess and deficiency, greatness and smallness, the greater and the less, or the great and the small (283c-e), of a magnitude or number or quality against its opposite (284e).

> *the other* is those which measure them in relation to the moderate (τὸ μέτριον), the fitting (τὸ πρέπον), the opportune (τὸν καιρὸν), the needful (τὸ δέον), and all the other standards that are situated *in the mean* (*τὸ μέσον*) between the extremes (τῶν ἐσχάτων). *Politicus* 284e2-8

Secondly, x is measured against the *mean* of x and y.

4.6.2 The First Measurement Must Not Be Pleasure (286c8-d6)

Not all opposite dyads x, y are worthy of a double measurement. In fact if the dyad of opposites (*x*, *y*) is a *pleasure* (ἡδονὴ) (and pain), then such a pair is unworthy of a second measurement.

> **Eleatic Stranger.** What I have to say, then, is that you and I, remembering what has just been said, must praise or blame the brevity or length of our several discussions,
> not [only] by comparing their various lengths with one another (μὴ πρὸς ἄλληλα τὰ μήκη κρίνοντες),
> but with reference to that part of the art of measurement (κατὰ τὸ τῆς μετρητικῆς μέρος), which we said before must be borne in mind; I mean the standard of fitness (πρὸς τὸ πρέπον).
> **Younger Socrates**. Quite right.
> **Eleatic Stranger.** But we must not measure all lengths by fitness (πρὸς τὸ πρέπον), either.
> For we shall *not in the least want to measure by fitness* (*οὐδὲν προσδεησόμεθα*) a length (μήκους) whose first measurement has produced (ἁρμόττοντος) *pleasure* (*πρὸς τὴν ἡδονὴν*,
> except, perhaps, as a secondary consideration (πάρεργόν τι). 286c8-d6

4.6.3 Every Pleasure (ἡδονὴ) Is a Philebean Infinite (*Philebus* 27e,31a,41d); Thus the First Measurement Is a Philebean Finite

We might be mystified why Plato rejects a dyad of opposites that represents a pleasure. It will begin to make sense if we realize that pleasure is a philosophical term for the Philebean anthyphairetic Infinite: Every pleasure (ἡδονὴ), or more precisely every dyad (pleasure and pain) (ἡδονὴ καί λύπη), is an Infinite.

Socrates. Is *pleasure and pain* a *finite*, or are they among the things which *admit of more and less*?
Philebus. Yes, they are among those which *admit of the more*, Socrates; for *pleasure* would not be absolute good if it were not *infinite in multitude and in the more*.
Philebus 27e

Socrates. Let us, then, remember . . . that *pleasure* was itself *infinite* and belonged to the class which, in and by itself, has not and never will have either beginning or middle or end. *Philebus* 31a

Socrates. And have we not also said and agreed and settled something further?
Protarchus. What?
Socrates. That both *pleasure and pain* admit of *the more and less* and are of the class of *the infinite*.
Protarchus. Yes, we have said that, certainly. *Philebus* 41d

That "the dyad x, y produces *pleasure*" is equivalent, as explicitly stated in the Platonic dialogue on pleasures *Philebus* 27e, 31a, 41d, with "the dyad x, y is an *Infinite*." Thus the condition on the first measurement is that the dyad x, y *is not a Philebean Infinite*, hence necessarily it is a *Philebean Finite*.

4.6.4 The Second Measurement Is x Against the Geometric Mean of the Dyad x and y

We interpret "the mean (x, y)," in the description of the second measurement in Sect. 4.6.1, to mean the philosophical analogue of the *geometric mean* of x and y.

This choice is strongly supported in the *Elements*, and specifically in the Theaetetean Book X of the *Elements*. It is remarkable that in the *Elements* there are 738 occurrences of the term "*μέσου*" (and its variants), and these exclusively and without exception refer to the "geometric mean" (of two numbers or of two magnitudes), while NONE of them refers to the arithmetic or any other mean, or any other sense; of them, the great majority 624 occur in the Theaetetean Book X (This remark is credited to A. Bassiakou (Bassiakou 2004)). The alternative term used "μέτριον" (283e, 284a, c, d, e) has normally a meaning equivalent to "*μέσον*."

4.6.5 The Second Measurement Is an Intelligible Generation, a Philebean Infinite

In the *Parmenides* 142c-143a, the infinite anthyphairetic intelligible division of the dyad One, Being, outlined in Sect. 3, is described, with an insisted repeated use of words referring to generation, such as "*γίγνεσθαι*" (142d5), "*γί[γ]νεται*" (142e4), "*γίνηται*" (142e6), and "*γιγνόμενον*" (143a1). Proclus in *Platonic Theology* 3,91, 9-24, comments on this frequent occurrence, as referring to "the progress/division of the intelligible multitude" (*περὶ τῆς προόδου τοῦ νοητοῦ πλήθους*).

A similar description of the Infinite intelligible anthyphairetic division as generation can be seen is the *Philebus*: "generation in substance" (*γένεσιν εἰς οὐσίαν*, *Philebus* 26d-8), "generated substance" (*γεγενημένην οὐσίαν*, 27b7-9).

The same insistent use of words related to generation, such as "*ὄντως γιγνόμενον*" (283e3-7), "*κατὰ τὴν τῆς γενέσεως ἀναγκαίαν οὐσίαν*" (283d7-9), *"πρὸς τὴν τοῦ μετρίου γένεσιν"* (284b9-c1, d5-6), refers to the intelligible infinite anthyphairetic division of the second measurements against the mean in the *Politicus*.

Thus the condition on the second measurement is that the dyad x, mean (x, y) is a *Philebean Infinite*.

4.6.6 Deciphering the Meaning of the Double Measurement: The Philosophical Double Measurement of the *Politicus* 283a-287b Is the Philosophical Imitation of the Theaetetus' Condition of a Dyad of Lines Incommensurable in Length Only

In order to understand the meaning of these two measurements, let us assume that x and y are lines, and $m = \text{mean}(x, y)$ is the geometric mean of x and y, given by Propositions VI.13 and II.14, namely m is a line such that $x/m = m/y$.

Let r be any "assumed line" ("προτεθείσα εὐθεῖα"), as in Definition X.3.

By Proposition II.14, there are lines a, b, such that $xr = a^2$, $yr = b^2$.

By Proposition VI.1, $\boldsymbol{x/y = a^2/b^2}$.

On the other hand, mr is the geometric mean of the squares a^2, b^2, since, by Proposition VI.1, $a^2/mr = xr/mr = x/m = m/y = mr/yr = mr/b^2$.

But the geometric mean of a^2, b^2 is obviously ab.

Hence $ab = mr$.

Therefore, $\boldsymbol{x/m = a/b}$,

since, again by Proposition VI.1, $x/m = xr/mr = a^2/ab = a/b$.

Thus the philosophical double measurement of the *Politicus* x against y a Finite, and x agaist m an Infinite is revealed to be the exact imitation, philosophical analogue of the condition of the weak Theaetetean theorem a against b incommensurable and a^2 against b^2 commensurable.

4.6.7 The Double Measurement Is the Cause of all Division and Collection in an Intelligible Being (*Politicus* 283a-287b)

Each of the Angler, the Sophist, the Statesman, the Weaving is an intelligible Being. An intelligible Being is knowable to humans as an Art, an entity with ontological status inferior to that of an intelligible Being, in that divisions are more spread out and Unity and Collection is less pronounced than in the intelligible Being itself. All Arts, and therefore all intelligible Beings become known to the human soul by the method of *Division and Collection* (284e-285c), equivalently "Division according to Kinds" (286d-287a), equivalently "Name and Logos" (285c-286b). The method of Division and Collection is the philosophical analogue of periodic anthyphairesis.

The cause of Division and Collection, and eventually of all intelligible Beings, is the fundamental condition of the double measurement (284a5-7, 284b7-c1, 286b9-10).

4.7 The Equivalence of the Three Passages Reveals the Statement of the Weak Theaetetus' Periodicity Theorem

Schematically:

Theaetetus 147d-148b	*Philebus* 16c, 23b-25e	*Politicus* 283a-287b
If a, b are lines, such that $a^2 = Nb^2$ for a non-square number, or more generally such that a, b commensurable in square only	The mixture of the Philebean Infinite, the Incommensurable, and the Philebean Finite, the Commensurable	The double measurement, against the opposite, excluding pleasure, hence a Finite, and against the mean, a generation in substance, hence an Infinite
then the infinitely many parts are collected into One	produces every intelligible Being, namely every philosophical analogue of a dyad in periodic anthyphairesis, and Collection of the Many into One	produces every intelligible Being, or every Division and Collection

In view of the fact, described in Sect. 3, that an intelligible Being is the philosophical analogue of a dyad in periodic anthyphairesis, whose parts are equalized and collected into One precisely because of anthyphairetic periodicity, we recognize that

> each of the *Philebus* 16c, 23b-25e and the *Politicus* 283a-287b
> is a philosophical analogue, philosophical imitation,

of the following

> *weak Theaetetus' Periodicity theorem*: if a, b are lines, such that $a^2 = Nb^2$ for a non-square number N (or more generally, such that a, b are commensurable in square only), then the anthyphairesis of a to b is periodic.

We have insisted on the details showing the philosophical imitation and analogy, so that the statement of the mathematical original, the weak Theaetetus' periodicity theorem, would be established with certainty and exactness.

4.8 The Philebean Finite Compared to the Pythagorean Finite

It should be pointed out that according to the Pythagoreans, too, the two principles of Beings is the Infinite and the Finite (cf. Philolaus, *Fragments* 1, 2, 6, and Aristotle, *Metaphysics* 987a13-28). But, while the Infinite, in both Plato and the Pythagoreans, is anthyphairetic in content, the Finite for the Pythagoreans is the Gnomon. The evolved corresponding Platonic concept is Logos, as may also be seen from a comparison of (the possibly nongenuine) *Fragment* 11 by Philolaus (knowledge is achieved by Gnomon) with the passage 546b-c of the *Politeia* on geometric number

(where Gnomon is in effect replaced by Logos). But Plato in the *Philebus*, not content with such replacement, insists that the principle of Finite has anthyphairetic content as well. Plato is able to modify the original Pythagorean scheme *only because* he has now in his possession the Theaetetus' periodicity theorem.

5 The Reconstruction of the Proof of the Weak Theaetetus' Periodicity Theorem, Employing Solely Tools from Book X of Euclid's *Elements* (the Apotome Lines (Propositions X.73-107), the Binomial Lines (Propositions X.36-72), and Their Conjugacy (Propositions X.112-114), and a Pigeon Hole Argument

Once the statement of (the weak form of) Theaetetus' theorem has been determined (in Sect. 4), the next step, to be attempted in the present Sect. 5, is to find out how Theaetetus was able to prove this theorem. We know that Book X of the *Elements* is generally attributed to Theaetetus, and it is natural to try to find in that Book the possible tools for a reconstruction of the proof of Theaetetus' theorem.

The reconstruction of the proof of the weak Theaetetus' periodicity theorem rests on the apotome lines (examined in Sect. 5.1), on the conjugacy of the apotome lines with the binomial lines (examined in Sect. 5.2), on the inductive transformation of the anthyphairetic relations of the quadratic equation $a^2 = Nb^2$ for non-square N into conjugacy relations (in Sect. 5.5), and on a pigeonhole argument.

5.1 Apotome Lines

The first concept in Book X that will turn out to be relevant in our quest for tools needed for the proof of the Theaetetus' periodicity theorem is the concept of an *apotome* line.

5.1.1 The Definition of the Apotome Line

Definition

Apotome Line (Proposition X.73 of the *Elements*)

An *apotome* line $\gamma = \zeta - \eta$ *with assumed line* r is equal to the difference of a *whole* line ζ, an *annex* line η, with $\zeta > \eta$, such that ζ, η are incommensurable, ζ^2, η^2 commensurable to the square r^2 of the *assumed line r.*

The study of the apotome lines occupies a part of Book X of considerable significance, Propositions X.73-107, 35 in total, consisting of (almost) six hexads, as follows:

X.73-78, defining six classes of subtractional irrational lines, one of which (X.73) is the *apotome* line, and showing the "alogia," incommensurability in square, of

each of these lines (these "alogoi" lines will be related to apotome lines in the hexad X.91-96 below);
X.79-84, showing the uniqueness of representation of each of these lines;
X.84/85, defining the six classes of apotome lines and,
X.85-90, constructing each of these lines;
X.91-96, constructing the "square-root" of each of these apotome lines defined in X.84/85 and corelating it to a line defined in X.73-78;
X.97-102, constructing the "square" of each of the lines defined in X.73-78, and corelating it to an apotome line defined in X.84/85; and,
X.103-107, a curtailed hexad, showing invariance under commensurability of each of the lines defined in X.73-78.

5.1.2 Simple Apotome Lines

The only place in Book X where apotome lines are expressly *constructed* is in the hexad of Propositions X.85-90. We notice that the apotome lines constructed there are somewhat less general than what one might expect from the definition of an apotome line. In order to explain this difference we introduce the following definition.

Definition

A line γ is a *simple apotome* if there are a non-square number N, lines a and b such that $a^2 = Nb^2$, and natural numbers p, q, λ, such that either $\lambda\gamma = pa - qb$ or $\lambda\gamma = qb - pa$.

A simple apotome is *integral* if $\lambda = 1$, *fractional* if $\lambda > 1$.

The canonical choice for the assumed line of a simple apotome is the line b, corresponding to the one foot line of *Theaetetus* 147d5, or the line a, or any other line r commensurable in power to a, namely such that a^2, r^2 commensurable areas.

Note. In the *Elements* the definition of a simple apotome is *not* explicitly given, but it is clear that Book X is preoccupied with simple apotomes, since all six kinds of apotome lines constructed in Propositions X.85-90, the only construction of apotome lines in Book X, are simple apotome lines.

Indeed, for the construction of the line in Propositions X.85-87, we let N be a non-square number, such that $N = q^2 - p^2$ (a method for finding such numbers is described in Lemma 1 in X.28/29), lines a, b such that $a^2 = Nb^2$, and construct the apotome line $qb - a$ (with respect to assumed line b in X.85, a in X.86, and r incommensurable to both a and b in X.87).

For the construction of the line in Propositions X.88-90, we let N be a non-square number, such that $N = q^2 + p^2$ (a method for finding such numbers is described in Lemma 2 in X.28/29), lines a, b such that $a^2 = Nb^2$, and construct the apotome line $a - qb$ (with respect to assumed line a in X.88, b in X.89, and r incommensurable to both a and b in X.90).

We will return in Sect. 9 to some more detailed remarks on the apotome lines in X.85-90.

It must be emphasized that the general definition of an apotome line (in X.73) allows for lines of the form $\zeta - \eta$, where $\zeta^2 = Nb^2$, $\eta^2 = Mb^2$, for some two (different)

non-square numbers N, M. Such an apotome line indeed appears in Proposition XIII.17 of the *Elements*: The side s of the face of the canonical 12hedron is the line $1/3(a - a')$, where $a^2 = 15d^2$, $a'^2 = 3d^2$, and d the diameter of the sphere in which the 12hedron is comprehended.

Thus Theaetetus in Book X is interested in the general apotome lines, so that a line such as the one arising in the study of the 12hedron will be included in the definition of apotome lines, and this appears to be the sole reason for adopting the general definition, but his central interest, as attested by all his explicit constructions in Book X, appears to be the *simple* apotome lines arising from the lines a, b, where $a^2 = Nb^2$, for some non-square number N.

5.1.3 All the Anthyphairetic Remainders of the Anthyphairesis of $a^2 = Nb^2$, for Some Non-square Number N, Are Simple Apotome Lines

The following proposition connects the concept of an apotome line with the study of the anthyphairesis of a to b wnen $a^2 = Nb^2$, and is the first serious indication that the concept is introduced in Book X precisely for the purpose of studying the quadratic antyphairesis.

Proposition If N is a non-square number, a, b lines such that $a^2 = Nb^2$, then every anthyphairetic remainder of the anthyphairesis of a to b is a simple apotome.

Notation. For two *incommensurable* (homogeneous) magnitudes a, b, with $a > b$, we denote the infinite *anthyphairesis* of a to b by

$$
\begin{aligned}
&a = I_0 b + \gamma_1, && \text{with } b > \gamma_1,\\
&b = I_1\gamma_1 + \gamma_2, && \text{with } \gamma_1 > \gamma_2,\\
&\gamma_1 = I_2\gamma_2 + \gamma_3, && \text{with } \gamma_2 > \gamma_3,\\
&\dots\\
&\gamma_n = I_{n+1}\gamma_{n+1} + \gamma_{n+2}, && \text{with } \gamma_{n+1} > \gamma_{n+2},\\
&\dots,
\end{aligned}
$$

the infinite sequence of successive *quotients* of the anthyphairesis of a to b by

$$I_0, I_1, I_2, I_3, \dots, I_{2n}, I_{2n+1}, \dots, \text{and}$$

the infinite sequence of successive *remainders* of the anthyphairesis of a to b by

$$a > b > \gamma_1 > \gamma_2 > \dots > \gamma_{2n} > \gamma_{2n+1} > \dots.$$

Definition *The generalized side-diameter numbers p_n, q_n*

$$
\begin{aligned}
&p_1 = 1, && q_1 = I_0,\\
&p_2 = I_1, && q_2 = I_1 I_0 + 1,\\
&\dots && \dots\\
&p_{2n-1} = I_{2n-2}\, p_{2n-2} + p_{2n-3}, && q_{2n-1} = I_{2n-2}\, q_{2n-2} + q_{2n-3},\\
&p_{2n} = I_{2n-1}\, p_{2n-1} + p_{2n-2}, && q_{2n} = I_{2n-1}\, q_{2n-1} + q_{2n-2},\\
&\dots && \dots
\end{aligned}
$$

Note. The (classical) side and diameter numbers are identified to the generalized side and diameter numbers of the anthyphairesis of the diagonal (diameter) to side of a square.

Indeed the anthyphairesis of diameter to side of a square is $I_0 = 1, I_1 = I_2 = I_3 = \ldots = 2$.

Then $p_1 = 1$, $q_1 = 1$, and,

by the recurrence relation of the generalized side-diameter numbers,

$$p_{n+2} = 2p_{n+1} + p_n, q_{n+2} = 2q_{n+1} + q_n.$$

We assume by induction, for the term n, that

$$p_{n+1} = p_n + q_n, q_{n+1} = 2p_n + q_n.$$

Then by the recurrence relations of the side-diameter numbers

$$\begin{aligned} p_{n+2} &= 2p_{n+1} + p_n = p_{n+1} + p_{n+1} + p_n = p_{n+1} + (p_n + q_n) + p_n \\ &= p_{n+1} + (p_n + q_n + p_n) = p_{n+1} + (2p_n + q_n) = p_{n+1} + q_{n+1}, \text{ and} \end{aligned}$$

$$\begin{aligned} q_{n+2} &= 2q_{n+1} + q_n = q_{n+1} + q_{n+1} + q_n = q_{n+1} + 2p_n + q_n + q_n \\ &= q_{n+1} + 2(p_n + q_n) = q_{n+1} + 2p_{n+1}, \end{aligned}$$

hence, for the next term $n + 1$, we have

$$p_{n+2} = p_{n+1} + q_{n+1},$$

$$q_{n+2} = 2p_{n+1} + q_{n+1}.$$

Note. A major difference at this stage between the classical side and diameter numbers and the generalized ones is that the recursive rule defining the generalized side and diameter numbers is far inferior to the classical rule, in that at the stage n they are defined in terms of the *previous two stages*, $n-1$ and $n-2$, and not just in terms of the immediately previous one $n-1$. The recursive relations analogous to the side and diameter numbers in terms of the previous stage only will be obtained in Sect. 5.6 below.

The generalized side and diameter numbers are useful in expressing the anthyphairetic remainders in terms of the initial magnitudes a and b.

Lemma ***(anthyphairetic remainders in terms of the generalized side and diameter numbers).*** For all $n = 1, 2, \ldots$

$$\gamma_{2n-1} = p_{2n-1}\, a - q_{2n-1}\, b$$

$$\gamma_{2n} = -p_{2n}\, a + q_{2n}\, b,$$

$$\ldots$$

Proof The proof is by induction.

$$\text{For } n = 1$$

By the anthyphairetic relation $a = I_0 b + \gamma_{1,}$ it follows

$$\gamma_1 = p_1 a - q_1 b, \text{hence}$$

$$p_1 = 1, q_1 = I_0.$$

By the anthyphairetic relation $b = I_1 \gamma_{\ 1} + \gamma_{2,}$ it follows

$$\gamma_2 = b - I_1 \gamma_1$$

$$= b - I_1 (a - I_0 b)$$

$$= -I_1 a + (1 + I_0 I_1) b, \text{hence}$$

$$\gamma_2 = -p_2 a + q_2\, b, \text{with } p_2 = I_1, q_2 = 1 + I_0 I_1.$$

Inductive step
If $n > 1$, assume that

$$\gamma_{2n-3} = p_{2n-3}\, a - q_{2n-3}\, b,\ \gamma_{2n-2} = -p_{2n-2}\, a + q_{2n-2}\, b,$$

then we need to see the same relations holding true for γ_{2n-1} and $\gamma_{2n.}$
By the anthyphairetic relation $\gamma_{2n-3} = I_{2n-2}\, \gamma_{2n-2} + \gamma_{2n-1}$ it follows

$$\gamma_{2n-1} = \gamma_{2n-3} - I_{2n-2}\, \gamma_{2n-2}$$

$$= (p_{2n-3}\, a - q_{2n-3}\, b) - I_{2n-2}\, (-p_{2n-2}\, a + q_{2n-2}\, b)$$

$$= (I_{2n-2}\, p_{2n-2} + p_{2n-3})\, a - (I_{2n-2}\, q_{2n-2} + q_{2n-3})\, b,$$

thus, by setting

$$p_{2n-1} = I_{2n-2}\, p_{2n-2} + p_{2n-3}, q_{2n-1} = I_{2n-2}\, q_{2n-2} + q_{2n-3},$$

we have that

$$\gamma_{2n-1} = p_{2n-1}\, a - q_{2n-1} b.$$

By the anthyphairetic relation $\gamma_{2n-2} = I_{2n-1}\, \gamma_{2n-1} + \gamma_{2n,}$ it follows

$$\gamma_{2n} = \gamma_{2n-2} - I_{2n-1}\, \gamma_{2n-1}$$

$$= (-p_{2n-2}\, a + q_{2n-2} b) - I_{2n-1}\, (p_{2n-1} a - \mathrm{q}_{2n-1} b)$$

$$= -(I_{2n-1}\, p_{2n-1} + p_{2n-2})\, a + (I_{2n-1}\, q_{2n-1} + q_{2n-2})\, b,$$

thus, by setting

$$p_{2n} = I_{2n-1}\, p_{2n-1} + p_{2n-2}, q_{2n} = I_{2n-1}\, q_{2n-1} + q_{2n-2},$$

we have that

$$\gamma_{2n} = -p_{2n}\, a + q_{2n}\, b.$$

Finally, the *uniqueness* of the generalized side and diameter numbers is a direct consequence of the incommensurability of a, b.

e.g., if $\gamma_{2n-1} = p_{2n-1}\, a - q_{2n-1}\, b = pa - qb$,
then $(p_{2n-1} - p)\, a = (q_{2n-1} - q)\, b$,
and if $p_{2n-1} - \mathrm{p} > 0$
then ($q_{2n-1} - q > 0$, thus) $a/b = (q_{2n-1} - q)/(p_{2n-1} - p)$,
in that case a, b would be commensurable magnitudes.

5.2 Binomial Lines and Their Conjugacy with the Apotome Lines

5.2.1 The Definition of Binomial Lines

Definition
Let r be an assumed line, ζ, η lines with $\zeta > \eta$, incommensurable, and such that ζ^2, η^2 are commensurable to r^2. Then (given in Proposition X.36 of the *Elements*) the line $\zeta + \eta$ is called *line of two names* ("ἐκ δύο ὀνομάτων"), or *binomial*.

The *greater part* of the apotome is the line ζ and the *smaller part* is the line η.

Propositions X.36–72 are exclusively about binomial lines, forming about one-third of Book X of the *Elements*.

X.36–41 is the hexad for binomial lines corresponding	to the hexad X.73–78,
X.42–47	to X.79–84,
X.47/48	to X.84/85,
X.48–53	to X.85–90,
X.54–59	to X.91–96,
X.60–65	to X.97–102 and
X.66–72	to the pentad X.103–107.

5.2.2 Simple Binomial Lines

The definition of a simple binomial line is entirely analogous to the definition of a simple apotome line, and will not be given explicitly. The constructions of binomial lines in Book X, X.48–53 are completely analogous to the ones for apotomes in X.85–90, and they produce simple binomial lines.

5.2.3 Question

Thus the binomial lines are exactly as important for Book X as the apotomes. We have found a connection of the apotomes with the remainders of the quadratic anthyphairesis a to b, for $a^2 = Nb^2$, N non-square number. The question then is: What is the role of the binomials in Book X?

5.2.4 The Role of the Binomial Lines in Book X as Conjugate to the Apotome Lines

The question on the role of the binomial lines is answered by the appearance, toward the end of Book X, of the three fundamental and truly miraculous Propositions X.112–114 on the *conjugacy* between the lines of two names and the apotomae lines.

Proposition X.112.
If r is a rational line, ζ is a line of two names, with names $\zeta_1 > \zeta_2$, and η is a line such that $\zeta \cdot \eta = r^2$, then η is an apotome with names $\eta_1 > \eta_2$, and $\zeta_1/\eta_1 = \zeta_2/\eta_2$ is a commensurable ratio.

Proposition X.113.
If r is a rational line, ζ is an apotome, with names $\zeta_1 > \zeta_2$, and η is a line such that $\zeta \cdot \eta = r^2$, then η is a line of two names with names $\eta_1 > \eta_2$, and $\zeta_1/\eta_1 = \zeta_2/\eta_2$ is a commensurable ratio.

Proposition X.113 *(Construction of conjugate to an apotome).*
Let BD $[\zeta]$ = BC − CD $[\zeta_1 - \zeta_2]$ be an *apotome* with assumed line A $[r]$, and
KH $[\eta]$ be a line with rectangle BD·KH $[(\zeta_1 - \zeta_2)\cdot\eta = r^2]$ equal to A^2.
Then KH is a *binomial* line KH = KF + FH $[\eta = \eta_1 + \eta_2]$, such that
the terms of KH are commensurable with the terms of BD and
in the same ratio (namely BC/KF=CD/FH $[\zeta_1/\eta_1 = \zeta_2/\eta_2 = \mu/\nu]$ commensurable), and
KH has the same order as BD.

(namely, if BD is an apotome line (of class 1,2,3, 4,5,6 wrt A,
then KH is a binomial line of class 1,2,3, 4,5,6, respectively, wrt A; and,
the side of a square equal to $BC^2 - CD^2$ is commensurable to BC
iff the side of a square equal to $KF^2 - FH^2$ is commensurable to KF.
The line KH is the *conjugate* to the apotome line BD.).

Proposition X.114.
If ζ is a line of two names, with names $\zeta_1 > \zeta_2$, η is an apotome with names $\eta_1 > \eta_2$, $\zeta_1/\eta_1 = \zeta_2/\eta_2$ is a commensurable ratio, and $\eta\cdot\zeta = r^2$,
then r is a rational line.

Note that

if $\gamma = qb - pa$ is a simple apotome wrt b, where N is a non-square number and $a^2 = Nb^2$, we set $\delta = qb + pa$, and note that $\gamma \cdot \delta = q^2b^2 - p^2a^2 = (q^2 - Np^2)b^2 = \lambda b^2$.

Set $\lambda\gamma^{-1} = qb + pa$.

Then $\gamma \cdot \gamma^{-1} = b^2$, namely the line γ^{-1} is the "inverse" to the apotome line γ. Informally we might think that $\gamma/b \cdot \gamma^{-1}/b = 1/1$, namely the ratio (essentially, the modern real number) γ^{-1}/b is the inverse of γ/b.

5.3 The Realization of the Importance of the Conjugacy Propositions of Book X in the Modern Era

5.3.1 The Insightful Connection of Bombelli of the Binomial Line with the Complex Number

Even before any careful examination of Book X, we are immediately stricken by the basic contents of the Book, no less than Bombelli was impressed in 1572. A. Weil (Weil 1984, p. 202) describes beautifully the connection:

> "The first concept, motivated by the work of the Italian algebraists of the sixteenth century on equations of degree 3 and 4, had been introduced by *Bombelli* in Book I of his Algebra of 1572, in close imitation of Euclid's theory of irrationals of the form $a + \sqrt{b}$ in his Book X ('*binomial*' in the terminology of Campanus' Latin translation of Euclid; accordingly, *binomio* is Bombelli's word for a complex number $a + \sqrt{-b}$). Further developments belong to the history of algebra and analysis rather than to number theory; suffice it to say that Euler played a decisive role in extending to complex numbers the main operations of analysis."

What impressed Bombelli in 1572 (Bombelli 1572) were the principal concepts of Book X, the apotome (introduced in X.73) and the binomial line (introduced in X.36), and their conjugacy in Propositions X.112–114, a conjugacy very similar to the relation between a complex number and its conjugate.

5.3.2 Euler Employed, for his Theory of Quadratic Continued Fractions, Conjugation of Quadratic Expressions Very Similar to the Theaetetus Conjugacy of Book X

The conjugation of quadratic expressions, very similar to the conjugacy of the apotome and binomial lines, has been used effectively for the examination of the continued fraction expansion of the square roots of non-square natural numbers by Euler (in the middle of the eighteenth century). Euler, in computing the continued fraction of, say, $\sqrt{54}$ formed the remainder/"apotome" $\sqrt{54}–7$, where $7 = m_1 =$ the integral part of $\sqrt{54}$. In order to proceed to the next step, he formed the conjugate "binomial" $\sqrt{54} + 7$, and applied the equivalent of X.113, to find $(\sqrt{54}–7) \cdot (\sqrt{54} + 7) = 5$. Thus the inverse of $\sqrt{54} + 7$ is $(\sqrt{54} + 7)/5$. The great advantage of conjugacy is that it allows for the trivial computation of both the next quotient I_1, the integral part of $(\sqrt{54} + 7)/5$, $I_1 = 2$, and the next remainder/"apotome" $(\sqrt{54} + 7)/5–2 = (\sqrt{54}–3)/5$.

But there is no evidence that Euler was in any way aware of Theaetetus conjugacy.

5.3.3 Heiberg's Failure to Realize the Importance of the Conjugacy Propositions

It is perhaps not surprising that the importance of the conjugacy Propositions of Book X has been lost by the erudite philologist and historian but nonmathematician Heiberg:

> Heiberg considers that this proposition and the succeeding ones are interpolated, though the interpolation must have taken place before Theon's time. His argument is that X.112-115 are nowhere used. (Heath 1908, 1926, p 246)

5.3.4 Weil's Disappointing Oversight to Realize Theaetetus' Motivation in Introducing the Conjugacy Propositions of Book X

Weil (Weil 1984, p 15) realizes the importance of the presence of conjugacy in Book X, but at the end dismisses the possibility that these conjugacies were used in the manner that they inspired Bombelli:

> "Not only is Euclid himself well aware of
> the relation $(\sqrt{r} + \sqrt{s})\,(\sqrt{r} - \sqrt{s}) = r - s$
> but even
> the identity $1/(\sqrt{s} + \sqrt{r}) = \sqrt{r}/(r - s) - \sqrt{s}/(r - s)$
> may be regarded as the essential content of prop. 112 of that book.
> *Unfortunately*, Euclid's motivation in Book X seems to have been
> the wish to construct a general framework for the theory of regular polygons and polyhedra, and *not*, as modern mathematicians would have it, an algebraic theory of quadratic fields.
> So we are left to speculate idly whether, in antiquity or later, identities involving square roots may not have been used, at least heuristically, in arithmetical work."

Knorr (Knorr 1983, p 60) provides further information on Weil's view. Knorr, after himself dismissing the role of Propositions X.112–114,

> "Certain potentials of the subject matter are entirely missed, for instance, insight into how products of the form $(a \pm b\sqrt{N})\,(c \pm d\sqrt{N})$ bear on the finding of solutions to integral relations of the form $x^2 - Ny^2 = \pm m$";

he then confirms Weil's view at footnote (Knorr 1983, p 66, fn. 53):

> "At the time of his course of lectures on the history of the Pell equation (Institute for Advanced Study, 1978–79), Professor A. Weil expressed to me his disappointment over Euclid's *utter failure* to perceive this potentially fruitful development of the theory of Book X."

The intimate connection of Plato's philosophy with periodic anthyphairesis, including the theorems proved by Theaetetus, has been partially presented in Sects. 3 and 4 above, but, as we shall see in the next Sect. 6, Plato was so deeply impressed by *the palindromic periodicity of quadratic anthyphairesis* established by Theaetetus that he made palindromic periodicity the dominant theme in the *Politicus*.

It seems to us clear that Weil's disappointing oversight to realize the true role of the conjugation of the lines of apotome and binomial lines in the Theaetetean Book X of the *Elements* was the result of a lack of a deeper understanding on his part of the intimate connection of Plato's philosophy with periodic anthyphairesis, and of the deep influence that Theaetetus exercised on Plato's later philosophy (especially on the trilogy *Theaetetus-Sophistes-Politicus* and the *Philebus*). This oversight is

especially disappointing since Weil was in fact steeped in Greek Mathematics and Plato's philosophy.

5.4 The Role of the Assumed Line ("προτεθείσα," "ἐκκειμένη") of Apotome and the Binomial Lines

Before proceeding with the reconstruction of the proof of Theaetetus' weak periodicity theorem we make a simple remark on the role of the *assumed line*.

For the same apotome line γ, say $\gamma = pa - qb$ or $\gamma = qb - pa$, the external classification of apotome lines (X.84/85) distinguishes three cases: The assumed line is either (commensurable to) a, or (commensurable to) b, or some other line c incommensurable to a and to b but commensurable to them in square and thus rational.

It is clear that the *internal* structure of an apotome line is not affected at all by the choice of the assumed line. Why then is there a need for the assumed line? Its role can only be *external*, i.e., in the compatibility of the apotome line with other apotome or binomial lines, for the purpose of performing simple algebraic or geometric operations. Thus we compose a line with its conjugate, or square a line, or take the square root of a line, or take a commensurable "perturbation" of the line, only if we have lines with the same assumed line. In fact, these operations are all applications of Proposition I.44–45 and the assumed line takes the role of the "fourth" proportional.

We consider further the role of the assumed line in Sect. 9 as well.

5.5 The Transformation of the Anthyphairetic Relations of a to b for $a^2 = Nb^2$ for Some Non-Square Number to Conjugation Relations

Once we have established

(a) That Plato states unequivocally that Theaetetus has proved the quadratic periodicity theorem (and imitates in the *Politicus* the stronger palindromic one, as we shall see in Sect. 6)
(b) That the theory of the apotome and binomial lines and their conjugacy, the ideal tools for proving the quadratic (palindromic) periodicity theorem, is fully developed in Book X,

then the reconstruction of the proof of quadratic periodicity theorem is *one way*, essentially obligatory, consisting of the natural transformation of the anthyphairetic relations into conjugate relations. This is precisely the content of the next.

Theorem (Theaetetus) (The transformation of the anthyphairetic relations of a to b for $a^2 = Nb^2$ for some non-square number to conjugation relations).

Let N be a non-square number, a, b lines with $a^2 = Nb^2$, and let the anthyphairesis of a to b be denoted as follows.

$a = I_0 b + \gamma_1$, with $b > \gamma_1$,
$b = I_1\gamma_1 + \gamma_2$, with $\gamma_1 > \gamma_2$,
$\gamma_1 = I_2\gamma_2 + \gamma_3$, with $\gamma_2 > \gamma_3$,
...
$\gamma_n = I_{n+1}\gamma_{n+1} + \gamma_{n+2}$, with $\gamma_{n+1} > \gamma_{n+2}$,
...

Then there are natural numbers
m_n, with $m_1 =$ the unique number such that ${m_1}^2 < N < (m_1 + 1)^2$, namely I_0 and λ_n, with $\lambda_1 = 1$, such that setting

$$\varphi_n = (a - m_n b)/\lambda_n, N - {m_n}^2 = \lambda_n\lambda_{n+1}, m_{n+1} = \lambda_{n+1}I_n - m_n,$$

we have

$$b^2 = \varphi_n\left(I_n b + \varphi_{n+1}\right), \text{ the conjugacy relations, and}$$

$$\gamma_1 = \varphi_1, \text{and } \gamma_n\varphi_{n+1} = \gamma_{n+1}b, \text{for } n = 1, 2, \ldots.$$

Proof *Step 1* (*The transformation of the second anthyphairetic relation into a conjugation relation*).

The *first anthyphairetic relation* is then

$$a = m_1 b + \gamma_1, \text{with } \gamma_1 < b.$$

$\gamma_1 = a - m_1 b$, an apotome; set $\varphi_1 = \gamma_1$. Set $\lambda_2 = N - {m_1}^2$.

(a) Setting φ_2 by $\gamma_2 b = \gamma_1\varphi_2$, we find that $b^2 = \varphi_1(I_1 b + \varphi_2)$, and $\lambda_2 I_1 > m_1$
(b) Setting $m_2 = \lambda_2 I_1 - m_1$, we find that $\varphi_2 = (a - m_2 b)/\lambda_2$
(c) Setting $\lambda_3 = 1 + I_1(m_1 - m_2)$, we find that $N - {m_2}^2 = \lambda_2\lambda_3$

Proof of Step 1:

(a) The second anthyphairetic relation is $b = I_1\gamma_1 + \gamma_2$.
We want to transform this anthyphairetic relation into a conjugation relation.
From it we obtain $b^2 = \gamma_1 I_1 b + \gamma_2 b$.
In order to factor $\gamma_1 = \varphi_1$ we employ Propositions I.44–45 to construct a line φ_2, such that

$$\gamma_2 b = \gamma_1\varphi_2;$$

indeed we then have

$$b^2 = \gamma_1 I_1 b + \gamma_1 \varphi_2 = \varphi_1 (I_1 b + \varphi_2).$$

This transformed anthyphairetic relation is now a conjugation relation, allowing for the use of X.113.

Since $\varphi_1 = a - m_1 b$, the line $a + mb$ is a binomial line, an N-binomial, and

$$\varphi_1 \cdot (a + m_1 b) = a^2 - {m_1}^2 b^2 = Nb^2 - {m_1}^2 b^2.$$

Set $N - {m_1}^2 = \lambda_2$.
Hence $I_1 b + \varphi_2 = (a + m_1 b)/\lambda_2$, hence $\lambda_2 \varphi_2 + \lambda_2 I_1 b = a + m_1 b = S_2$.
Note that $a > m_1 b$ and $\lambda_2 I_1 b > \lambda_2 \varphi_2$;
since $\gamma_2 b = \gamma_1 \varphi_2$ and $\gamma_1 > \gamma_2$, hence $b > \varphi_2$;
thus $\lambda_2 I_1 b > S_2/2 > m_1 b$, hence $\lambda_2 I_1 > m_1$.

(b) We set

$$m_2 = \lambda_2 I_1 - m_1.$$

We find immediately that

$$\varphi_2 = (a - m_2 b)/\lambda_2,$$

a fractional N-apotome.

(c) We note that $N - {m_2}^2 = (N - {m_1}^2) + ({m_1}^2 - {m_2}^2) = \lambda_2 + (m_1 - m_2) \cdot (m_1 + m_2) =$
$= \lambda_2 + (m_1 - m_2) \cdot \lambda_2 I_1 = \lambda_2 (1 + I_1 (m_1 - m_2)) = \lambda_2 \lambda_3$.

Step 2 (*The transformation of the third anthyphairetic relation into a conjugation relation*).

(a) Setting φ_3 by $\gamma_3 b = \gamma_2 \varphi_3$, we find that $b > \varphi_3$, $b^2 = \varphi_2 (I_2 b + \varphi_3)$, and $\lambda_3 I_2 > m_2$
(b) Setting $m_3 = \lambda_3 I_2 - m_2$, we find that $\varphi_3 = (a - m_3 b)/\lambda_3$
(c) Setting $\lambda_4 = \lambda_2 + I_2 (m_2 - m_3)$, we find that $N - {m_3}^2 = \lambda_3 \lambda_4$

Proof of Step 2:

(a) We transform the third anthyphairetic relation $\gamma_1 = I_2 \gamma_2 + \gamma_3$ into a conjugation relation.
From it we obtain $\gamma_1 b = I_2 \gamma_2 b + \gamma_3 b$.
In order to factor γ_2 we employ Propositions I.44–45 to construct a line φ_3, such that

$$\gamma_3 b = \gamma_2 \varphi_3;$$

we then have

$$b > \varphi_3 \text{ and}$$

$$\gamma_1 b = I_2\gamma_2 b + \gamma_2\varphi_3 = \gamma_2(I_2 b + \varphi_3).$$

From it we obtain $\gamma_1 b^2 = \gamma_2 b(I_2 b + \varphi_3)$.
Recalling that $\gamma_2 b = \gamma_1\varphi_2$, we have $\gamma_1 b^2 = \gamma_2 b(I_2 b + \varphi_3) = \gamma_1\varphi_2(I_2 b + \varphi_3)$, hence

$$b^2 = \varphi_2(I_2 b + \varphi_3).$$

This transformed anthyphairetic relation is now a conjugation relation, allowing for the use of X.113.
Since $\varphi_2 = (a - m_2 b)/\lambda_2$, then

$$\varphi_2\cdot(a + m_2 b) = \left(a^2 - {m_2}^2 b^2\right)/\lambda_2 = \left(Nb^2 - {m_2}^2 b^2\right)/\lambda_2 = \left(N - {m_2}^2\right)b^2/\lambda_2$$
$$= \lambda_2\lambda_3 b^2/\lambda_2 = \lambda_3 b^2.$$

Hence $I_2 b + \varphi_3 = (a + m_2 b)/\lambda_3$, hence $\lambda_3\varphi_3 + \lambda_3 I_2 b = a + m_2 b = S_3$.
Note that $a > m_2 b$ and $\lambda_3 I_2 b > \lambda_3\varphi_3$, since $\gamma_3 b = \gamma_2\varphi_3$ and $\gamma_2 > \gamma_3$, hence $b > \varphi_3$; thus $\lambda_3 I_2 b > S_3/2 > m_2 b$, hence $\lambda_3 I_2 > m_2$.

(b) We set

$$m_3 = \lambda_3 I_2 - m_2.$$

We find immediately that

$$\varphi_3 = (a - m_3 b)/\lambda_3,$$

a fractional *N*-apotome.

(c) We note that $N - {m_3}^2 = (N - {m_2}^2) + ({m_2}^2 - {m_3}^2) = \lambda_2\lambda_3 + (m_2 - m_3)\cdot(m_2 + m_3) =$
$=\lambda_2\lambda_3 + (m_2 - m_3)\cdot\lambda_3 I_2 = \lambda_3(\lambda_2 + I_2(m_2 - m_3)) = \lambda_3\lambda_4.$

General step (*the transformation of the general anthyphairetic relation into a conjugation relation*).
We assume that for some natural number n, line φ_n, and natural numbers m_n, λ_n, λ_{n+1} have been defined, such that

$$\varphi_n = (a - m_n b)/\lambda_n, N - {m_n}^2 = \lambda_n\lambda_{n+1}.\text{Then}$$

(a) setting φ_{n+1} by $\gamma_{n+1}\, b = \gamma_n\varphi_{n+1}$,
we find that $b > \varphi_{n+1}$, $b^2 = \varphi_n(I_n b + \varphi_{n+1})$, and $\lambda_{n+1} I_n > m_n$

(b) setting $m_{n+1} = \lambda_{n+1} I_n - m_n$,
we find that $\varphi_{n+1} = (a - m_{n+1} b)/\lambda_{n+1}$; and,
(c) setting $\lambda_{n+2} = \lambda_n + I_n(m_n - m_{n+1})$,
we find that $N - {m_{n+1}}^2 = \lambda_{n+1}\lambda_{n+2}$.

The *Proof* is entirely analogous to the proof of Step 2, shown just above. The details are omitted.

Note. The transformation of the anthyphairetic relations into conjugacy relations is realized with the introduction of the "reduced apotomes" φ_n.

As mentioned already, the introduction of these reduced apotomes provide a one-way, essentially obligatory, transition from the anthyphairetic relations to the conjugacy relations, so that Theaetetus' conjugacy theory can be applied.

It should be pointed out that the definition of φ_{n+1}, via the equality $\gamma_{n+1} b = \gamma_n \varphi_{n+1}$, an elementary application of Proposition I.44 of the *Elements*, does not appear explicitly in Book X; and that the reduced apotomes play a basic role in the proof of the palindromic periodicity theorem in Sect. 7 and in the proof of the general Pell property in Sect. 9 below. It is, however, quite remarkable that, as we shall see in Sect. 10, the reduced apotomes make their unmistakable reappearance in the *cakravala* method of dealing with the Pell problem by the Hindu mathematician Bhaskara II.

5.6 The Recursive Definition of Generalized Side and Diameter Numbers

Proposition $(-1)^n (q_n b - p_n a)((a - m_{n+1} b)/\lambda_{n+1}) = (-1)^{n+1}(q_{n+1} b - p_{n+1} a) b$,
$\lambda_{n+1} q_{n+1} = m_{n+1} q_n + N p_n$,
$\lambda_{n+1} p_{n+1} = q_n + m_{n+1} p_n$.

These are the general recursive formulas really analogous to the recursive formulae for the side and diameter numbers for $N = 2$, in the sense that the recursion is not from the previous two steps, but only from the previous one.

5.7 The Proof of the Weak Theaetetus' Periodicity Theorem: The Anthyphairesis of *a* to *b* for $a^2 = Nb^2$ for Non-Square Number *N* is Periodic

From Theorem 4.5 just proved, and from the fact that the set of pairs (λ, m) of natural numbers, with λ bounded above by N and m by m_1, is finite, it follows, by an application of *the pigeonhole principle*, that there are indices k and s, with $k < s$, such that $\varphi_k = \varphi_s$. The equality $\varphi_k = \varphi_s$ clearly implies the equality $\gamma_k/\gamma_{k+1} = \gamma_s/\gamma_{s+1}$, and this is exactly the Logos Criterion for eventual periodicity of the anthyphairesis.

Note. It would be highly desirable, but there is no trace of the combinatorial pigeonhole argument to be found in Book X; however, a more careful application of the pigeonhole argument for the proof of a stronger Theaetetus theorem on

palindromic periodicity can be clearly associated with a passage in Plato's *Politicus*, as we shall see in Sects. 6 and 7 below.

6 The Definition of the Intelligible Being "the Statesman" in the *Politicus* Is a Philosophical Imitation of the Strong Theaetetus' Palindromic Periodicity Theorem

The third dialogue in the trilogy on Plato's theory of knowledge is the *Politicus*. The express purpose of the dialogue is the definition of the art of statesmanship by the method "Name and Logos," equivalently "Division and Collection," involving steps with binary division and the Logos Criterion, in agreement with the method employed in the *Sophistes*. We have seen in Sect. 3 that in the *Sophistes* an intelligible Being, like the Angler or the Sophist, becomes fully known by the definition of an Art, the Art of angling or the Art of sophistry, respectively, employing the method of Division and Collection, equivalently of "Name and Logos," which we interpreted as a philosophical analogue of periodic anthyphairesis.

6.1 The First Ten Steps of the Division and Collection (258b-267c)

The aim of the *Politicus* is to obtain the knowledge ("ἐπιστήμη") of the intelligible Being the Statesman, by the method of Division and Collection, the method employed in the *Sophistes*. In the first part of the dialogue, 258b-267c, the first ten steps of the Division are obtained. The initial genus G of all arts is successively divided in the binary manner familiar from the *Sophistes*. It is clear that the same interpretation of the Division as philosophical analogue of periodic anthyphairesis holds here as well.

Thus,

G divided into [practical B_1, scientific ("γνωστική") A_1], (258b-259d)

A_1 divided into [appraise, determine ("κρίνειν") B_2, command A_2], (259d-260c)

A_2 divided into [someone else's commands B_3, one's own commands A_3], (260c-261b)

A_3 divided into [lifeless herd B_4, herd of living beings (=animals) A_4], (261b-d)

A_4 divided into [herd of a single animal ("μονοτροφίαν") B_5, herd of many animals A_5], (261d-e)

A_5 divided into [herd of aquatic animals B_6, herd of land animals A_6], (264b-d)

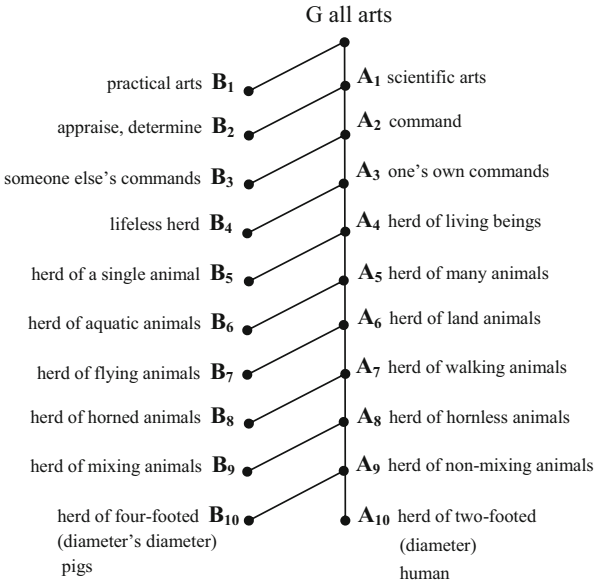
A_6 divided into [herd of flying animals B_7, herd of walking animals A_7], (264e-265b)

A_7 divided into [herd of horned animals B_8, herd of hornless animals A_8], (265b-d)

A_8 divided into [herd of mixing animals B_9, herd of nonmixing animals A_9], (265d-e)

A_9 divided into [herd of four-footed (diameter's diameter) pigs B_{10}, herd of two-footed (diameter) (human) A_{10}], (265e-266d).

Here is the figure summarizing these ten initial steps of the Division.



Note that after A_4 we are dealing with

the arts of feeding an animal herd ("ἀγελαιοτροφικὴ," 261b-d);

what is really being divided is the nature of the herd, narrowing it down, by successive binary division and exclusion, till we get to

the arts of commanding a human herd (A_{10}, 265e-266d).

6.2 The Need to Continue the Division (267c-268d and 274e-277c)

The division steps described in Sect. 6.1 is by no means the end of our quest for the definition of the Statesman, as one might think. This is so because A_{10} consists of the arts of commanding a human herd but we are interested to define the statesmanship, namely the optimal art of commanding a human herd.

This means that there is still ahead of us a second process of successive binary divisions, in all of which the herd is now fixed as the human herd, and what is being successively narrowed down, by binary division and exclusion, is the nature of commanding this human herd ("ἀγελαιοκομικὴ," 275e), till we get to the optimal commanding.

6.3 The Myth of Palindromic Periodicity with Kronos Era and Zeus Era (268d-274e)

According to this myth 268d-274e, there are two eras in the world, the Kronos era and the Zeus era.

The world, to be thought of as a spherical entity, rotates in a certain direction during the Kronos era. After a long time, when all the possible stages of the Kronos era are exhausted ("τὸ γήινον ἤδη πᾶν ἀνήλωτο γένος" "πάσας τὰς γενέσεις ἀποδεδωκυίας"), it is necessary to reverse the direction ("πάλιν ἀνέστρεφεν") and to go from the Kronos era to the Zeus era (272d6-e6).

When the Zeus era is in danger of losing its direction and God perceives that it is in trouble, and that "it might founder in the tempest of confusion and sink in the infinite sea of dissimilarity ("κηδόμενος ἵνα μὴ χειμασθεὶς ὑπὸ ταραχῆς διαλυθεὶς εἰς τὸν τῆς ἀνομοιότητος ἄπειρον ὄντα πόντον"), he … reversed ("στρέψας") whatever had become unsound and unsettled in the previous period when the world was left to itself, set the world in order, restored it and made it immortal ("ἀθάνατον") and ageless ("ἀγήρων")," thus returning (273c4-e4). In this way, we are back again to the Kronos period and completing a full period. The change from the Zeus era to the Kronos era is in order to avoid the bad, nonperiodic infinite, something achieved by the Logos Criterion.

The world moves forever repeating this period, alternating the Kronos era with the Zeus era.

The relation of the two eras is especially interesting. It is schematically described as follows: In the Zeus era (in which era we are presently) men come from nonexistence (a) into existence (b), then they grow (c), then they get a beard (d), then their beard becomes white (e), then they grow old and are about to die (f). If at this time the world happens to switch to the Kronos era, then the human will go from stage (f) to (e) then to (d) then to (c) then to (b) and then they pass into nonexistence (a) (270d-271a).

Summarizing, according to the myth, the world is periodic, and the period consists of two half periods, the Kronos era followed by the Zeus era, in other words it is palindromically periodic.

6.4 The Division for the Statesman Must Be Continued According to the Myth of Palindromic Periodicity (276a1-7)

The first ten division steps of the definition of the Statesman, corresponding to the Kronos era, must be completed with the last ten division steps of the Zeus era, in **palindromic** to the Kronos era manner, with **Logos circularly** comprehending both of them (276a). Thus every Division and Collection has palindromically periodic (and thus infinite) form.

The passage 276a1-7 below is crucial for our interpretation of the *Politicus*, because in this passage Plato makes perfectly clear that *the palindromic periodicity of the myth must be applied for the Division and Collection of the Statesman*. The passage 276a1-7 is the following:

		Younger Socrates:. . .but *how (τίνα τρόπου)* would *the next division (ἡ μετὰ τοῦτο διαίρεσις)* be made? Stranger: *In the same kinds (κατὰ ταὐτὰ)*
as (καθ' ἅπερ) we divided (διηρούμεθα) before (ἔμπροσθεν) the art of *feeding herds* in		
	those that go on foot [A_7]	
and the winged, [B_7]		
	and the unmixed breeds [A_9]	
	and the hornless, [A_8]	
		we must *divide* *the art of tending herds* *in arts similar to (τοῖς αὐτοῖς)*
those (τούτοις),		
comprehending (περιειληφότες) in the logos (ἐν τῷ λόγῳ) both (ὁμοίως)		
		the present [Zeus] era (τὴν ἀγελαιοκομικὴν τήν τε νῦν)
and *the Kronos era.* *(καὶ τὴν ἐπὶ Κρόνου βασιλείαν)*		

We may describe the application of the myth of the two eras as describing in symbolic and philosophical manner the palindromic periodicity of the definition of the Statesman, an imitation by Plato of the Theaetetus' palindromically periodic quadratic anthyphairesis:

	The second half of the division (*ἀγελαιοκομική*, where the commanded herd is the human herd, held fixed, and the nature of the commanding herdsman is the variable) corresponds to the **Zeus** era,
while the first half (*ἀγελαιοτροφική*, where the nature of the commanded herd is the variable) corresponds to the **Kronos** era,	
in the myth of the two eras (Sect. 6.3). Thus (276a) is precisely the point in the *Politicus*, where the method of Division and Collection, namely the method of philosophical periodic anthyphairesis, is explicitly connected to *palindromicity*, since	
the Kronos era	
	and the Zeus era

(continued)

<table>
<tr><td colspan="2">are in palindromic relation in the myth of world ‘aneilixis’.</td></tr>
<tr><td>In this passage
the first half of the process,
with first ten division steps
G divided into B_1, A_1,
A_1 divided into B_2, A_2,
. . .,
A_9 divided into B_{10}, A_{10},
is associated explicitly with
the Kronos era,</td><td></td></tr>
<tr><td></td><td>and, also explicitly, it is stated that
the definition of the Statesman
must be completed,
as in the myth of the two eras,
with its second half,
with last ten division steps, say
A'_{11} divided into B'_{10}, A'_{10},
A'_{10} divided into B'_9, A'_9,
. . .,
A'_2 divided into B'_1, A'_1,
corresponding to
the Zeus era,
palindromic to</td></tr>
<tr><td>the Kronos era</td><td></td></tr>
<tr><td colspan="2">the whole comprising a full period, cycle,
achieved by the Logos Criterion $B''_1/A''_1 = B_1/A_1$.</td></tr>
</table>

(There are also the many steps A_{11}, A_{12},. . ., . . ., A'_{12}, A'_{11} which are not examined in detail, but are given collectively in groups, in the form of **“contingent causes”** (287b-289a). We will not be concerned with this part of the definition of the Statesman, as it is of no importance for our reconstruction.

This is the place to be reminded of that *all descriptions* of the method of (Division and) Collection in the Platonic dialogues exhibit *the cyclical, periodic nature* of the method (as noted in Sect. 3 above). In particular, the Platonic Division and Collection is not like the Aristotelian division into a genus and differentia, clearly a strictly finite division, but, having a periodic form, is an infinite division. In accordance with the Divisions and Collections of the *Sophistes*, this cycle must be realized by the *Logos Criterion*. This takes the following form:

The last genus of the Zeus era, denoted A'_1, will be divided into

the art of the strategy B''_1 and of the statesmanship A''_1,
for which the *Logos Criterion* indeed holds $B''_1/A''_1 = B_1/A_1$.

The new element in the *Politicus* Division and Collection, introduced in Sects. 6.3 and 6.4, in relation to the *Sophistes*, is not only periodicity, as in the *Sophistes*, but in addition the presence of *palindromicity within the period*; in other words we expect that the anthyphairesis will have a period, but one presenting a symmetry with respect to the middle of the period. In the language suggested by the myth of Sect. 6.3, the division

steps within the period will be divided into two parts, the first half to be the Kronos era, and the second half the Zeus era. We then expect that the definition of the Statesman must exhibit the palindromicity of these two eras. We show in Sects. 6.5 and 6.6 below that this is precisely the plan suggested and realized by Plato for defining the Statesman.

6.5 The Last Ten Division Steps and the Logos Criterion for the Statesman (267c-268d and 274e-277c, 289e-305e)

The last ten steps $A'_{10},\ldots,A'_1$ in the Zeus era and the Logos Criterion in the definition of the Statesman are described in 289e-305e.

6.5.1 The Game (277e)

Plato has no intention to give the Zeus steps explicitly and in succession, so these steps are not given in their correct order, in which they appear in the definition of the Statesman, but in the form of a *game*, reminding of a grammatical children's game, as explicitly stated in 277e. To find the right order and the Logos Criterion and thus complete the definition of the Statesman is going to be our task.

6.5.2 The Rules of the Game

However, Plato sets precise rules which enable us to complete this definition.

Rule 1. There is a *one to one correspondence between the already defined genera* B_j *and the palindromically analogous to them genera* B'_j *to be defined for* $j = 1, 2,\ldots,10$ *(276a, 277e-279a)*

To every division step (B_j, A_j) of the Kronos era, namely of "agelaiotrophike," where some type of herd B_j, governed and led, is *rejected,*

there is a unique division step (B'_j, A'_j) of the Zeus era, namely of "agelaiokomike," where some "similar" type of governing and leadership B'_j of the human herd is rejected, for $j = 1,..,10$.

The "positive" genera A'_j are *not given*, but the method of Division and Collection implies that in principle each A'_j is uniquely defined, as relative complement of B'_j in A'_{j+1}.

Rule 2. Ordering of government systems (291c-303b)

There are six types of conventional government, namely
government (I) by one, or by few, or by many, and
government (II) with laws, or without laws,
in every possible combination in each of the two criteria (I) and (II).
Thus
Tyranny is government by one, without laws,
Oligarchy by few without laws,
Anarchy by many without laws,
Democracy by many with laws,
Aristocracy by few with laws, and
Monarchy by one with laws.

(Plato does not use the term Anarchy for the government by many without laws in the *Politicus*, but he does use the term elsewhere, e.g., in the *Laws* 942a-d).

There is a quite long discussion on these types of government in the *Politicus* 291c-303b with the following conclusions:

(2a) There is an ordering of these six types of government as follows:
All six known types of government, all governments without true knowledge, will be eventually rejected; of them the worst is Tyranny (government by one, without laws), then Oligarchy (government by few, without laws), then Anarchy/ Democracy without laws, and then, in reverse /palindromic order, government with laws: first Democracy, then Aristocracy, then Monarchy.

This ordering is so, because

(i) in the absence of government by knowledge, it is better to have laws, than to be governed without laws, and

(ii) (without laws the many govern better than the few and the few better than the one, while in the presence of laws the situation is precisely the reverse.

(2b) All these conventional types of government are to be rejected, because they are not based on true knowledge. Hence each of them (possibly more than one at a time) will be rejected at some A′j.

(2c) All types of government with laws are to be rejected because laws, and in fact everything written, is *lifeless*, as opposed to true knowledge, which has life and soul. This is strongly supported by Plato's similar rejection, in the final part of the *Phaedrus* 274c5-277a5, of the written word, in general, as an image and as not truly alive.

Rule 3. Ordering of the genera B'_j to be rejected (303d-e)

In the Zeus era, a relatively better type of government will *be rejected later,* namely will appear as B'_j, with smaller index j, while a relatively worse type of government will appear as B'_k with greater index k.

Rule 4. The three arts to be rejected last (303e-305e)

The arts of judgeship, rhetoric, and strategy are the relatively highest arts of government, and are thus the last three to be rejected; they will then be the genera B'_2, B'_1, B''_1, respectively.

6.5.3 The Search for the Ten Palindromic Correspondences

We now apply the Rules set by Plato in order to find out the palindromic nature of the Division for the Statesman.

[*Division step rejecting* $\mathbf{B'_4}$ *palindromic to* $\mathbf{B_4}$]

This is definitely the most interesting palindromic correspondence. In a famous passage in the *Phaidrus* 274c-277a Plato argues against the written word, as lifeless. Evidently it is for this reason that Plato introduces the highly artificial division step into herds of living entities and lifeless herds B_4. We realize that this is done precisely in order to reject as its palindromical analogue all governments that are based not on true intelligible knowledge, but on written laws.

	at $\mathbf{B'_4}$ every government with written laws (including Democracy D, Aristocracy Ar, Monarchy M) must be rejected as being, according to *Phaidrus* 274c5-277a5, *lifeless* government, in *palindromic* fashion to
the rejection at $\mathbf{B_4}$ of the *lifeless herds*	

[*Division step rejecting* $\mathbf{B'_3}$, $\mathbf{B'_6}$, $\mathbf{B'_7}$ *palindromic to* $\mathbf{B_3}$, $\mathbf{B_6}$, $\mathbf{B_7}$, respectively]
It is fairly clear by Rule 1 that

	at $\mathbf{B'7}$ the heralds, whose patron god Hermes, the herald of the Gods, is flying, are to be rejected as governors (290a-c) in *palindromic* fashion to
the rejection at $\mathbf{B_7}$, of the herds of flying animals;	

	at $\mathbf{B'_6}$ *the merchants and shipmasters* are rejected as governors (289e-290a), in *palindromic* fashion to
the rejection at $\mathbf{B_6}$, of *the herds of aquatic animals*; and	

	at $\mathbf{B'_3}$ *the priests and diviners* are to be rejected as governors (290c-291a) in *palindromic* fashion to
the rejection at $\mathbf{B_3}$	

[*Division step rejecting* $\mathbf{B'_{10}}$ *palindromic to* $\mathbf{B_{10}}$]

Some argument is needed to convince us that the rejection of the sophists as governors in 291b-c, and again in 303b-c (of the sophists' sophists, this second time to be exact), is palindromically related to the rejection of the pigs as herd at B_{10} (265e-266c). Here Plato is multiply playful: In the first place the distinction of the pig from human is that it is four-footed; this is playfully described, in geometric language, as "diameter's diameter." We observe that the queer expression *"diameter's diameter"* has some similarity of form with the equally queer expression *"sophists' sophists."*

But we would be rather at a loss to find any more significant association; unless we note that while a true dialectician is preoccupied with "the diameter itself" (*Politeia* 510d5-511a1), a partless Being, it is the *sophists* who, according to *Meno* 85b4, call *"diameter"* the straight line that joins one edge of a square with its opposite edge, a divisible magnitude. The meaning of this distinction is possibly the following: Sophists occupy themselves with magnitudes and in general with

sensibles, while the dialectician occupies himself with the idea of the diameter, something that is rather the logos of the diameter to the side of a square. In any case the *Meno* suggests that there is a connection between the diameter and the sophists; hence there is also indeed a connection between "*diameter's diameter"* the four-footed pig and "*sophists' sophists,"* hinting of course that the worst sophists are like pigs. We are clearly ready, granted a certain playfulness, to accept that, by Rule 1,

	the sophists' sophists, rejected as governors at $\mathbf{B'_{10}}$, is palindromically related with
the herd of pigs, described playfully as "diameter's diameter," rejected at $\mathbf{B_{10}}$	

[*Division step rejecting* $\mathbf{B'_8}$, $\mathbf{B'_9}$ *palindromic to* $\mathbf{B_8}$, $\mathbf{B_9}$, respectively]

It follows necessarily, by the rules 2 and 3, that the two types of conventional government without laws, namely the lowest Tyranny and next lowest Oligarchy must be associated with the two lowest remaining free positions in the Zeus era, namely B'_9, B'_8, respectively. It becomes immediately clear why Plato likens twice (in 291a-b, 303c-d) the worse classes of conventional politicians with Centaurs, creatures of mixed breed, and Satyrs, horned creatures! Thus

	at $\mathbf{B'_9}$ the *Tyrants,* likened to *Centaurs*, creatures of mixed breed, must be rejected, as palindromically related to
herds of *animals with mixed breed* already rejected at $\mathbf{B_9}$, and	
	at $\mathbf{B'_8}$ the Oligarchs, likened to Satyrs, horned creatures, must be rejected, as palindromically related to
herds of horned animals already rejected at **B8**	

[*Division step rejecting* $\mathbf{B'_5}$ *palindromic to* $\mathbf{B_5}$]

It is not immediately obvious from the *Politicus* why Democracy without laws, Anarchy, should be palindromic to the rejected herd at B_5, and thus be placed in the position of the rejected part in B'_5. However, this is well explained in the *Nomoi* (Laws) (Plato 1967 & 1968):

	Democracy without laws at B'_5 identical with the Anarchy described in the *Nomoi* 942a5-d2, as "κατὰ μόνας δρᾶν," i.e., as a government where each person acts singly, is palindromically related to

(continued)

the rejected herd at B_5, described as "μονοτροφία," i.e., as herd consisting of a single animal	

[*Division step rejecting* $\mathbf{B'_2}$, $\mathbf{B'_1}$ *palindromic to* $\mathbf{B_2}$, $\mathbf{B_1}$, respectively]

We are left with the arts of judging, rhetoric, and strategy, for which Plato sets Rule 4, that they will be rejected last (303e-305e). Of them

	at $\mathbf{B'_2}$ *the Judges* are rejected as governors (305b-c) in *palindromic* fashion to
the rejection at $\mathbf{B_2}$, of *the art of* "κρίνειν"(260a-c)	

	at $\mathbf{B'_1}$ *the Rhetoricians* are rejected as governors (304c-e) in *palindromic* fashion to
the rejection at $\mathbf{B_1}$, of *the* practical *arts*	

This completes the quest for the palindromic correspondences between the first ten steps of the Kronos era with the last ten steps of the Zeus era. The art of strategy is still not rejected. It is left to be rejected at the last step when the Logos Criterion must be formed.

6.6 Logos in the Division and Collection of the Statesman

The very last division is the division of A'_1 into the practical art of the strategy B''_1 and the gnostical art of the statesman A''_1 (304e-305a).

The Logos Criterion consists in equating

the initial logos:

practical arts B_1/ scientific arts A_1

with

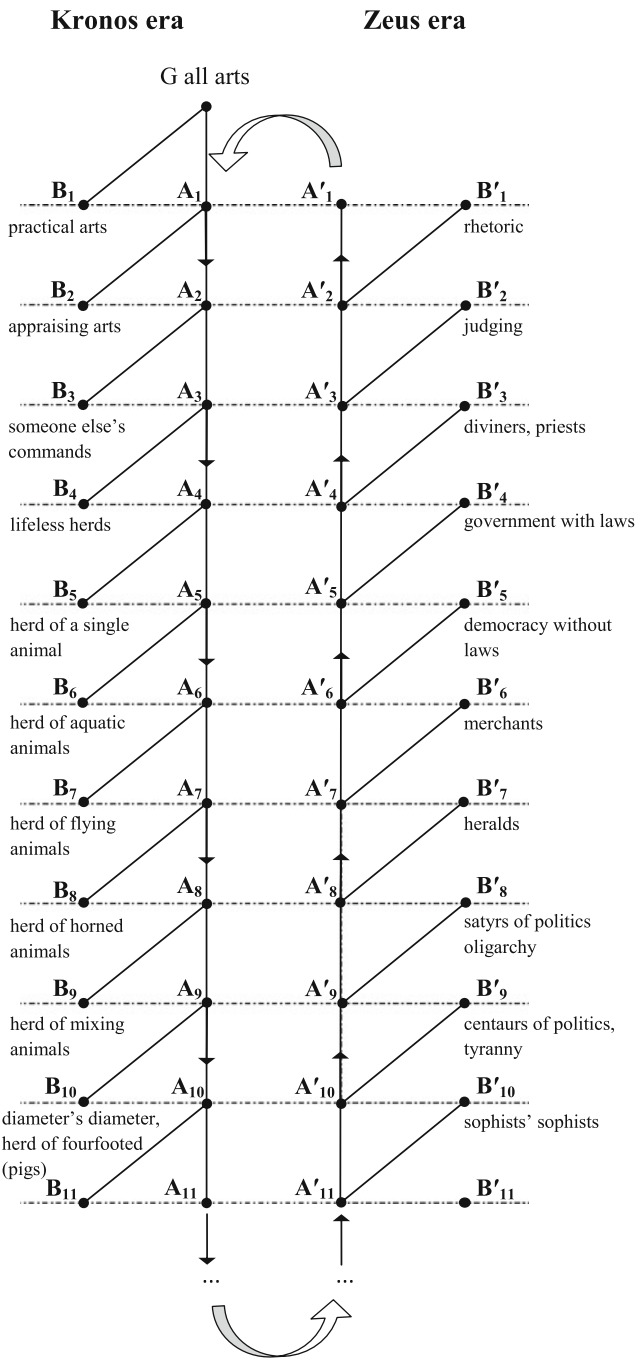
the final logos:

the practical art of strategy B''_1/the scientific art of statesmanship A''_1.

Thus $B_1/A_1 = B''_1/A''_1$, and periodicity is insured.

6.7 The Definition of the Statesman as a Philosophical Palindromically Periodic Anthyphairesis (*Politicus* 258c-267c, 289e-305e)

We are now ready to collect all the pieces of the Division and Collection of the Statesman. The following figure incorporates the information needed for the definition of the Statesman as a philosophic palindromically periodic anthyphairesis in the *Politicus*.

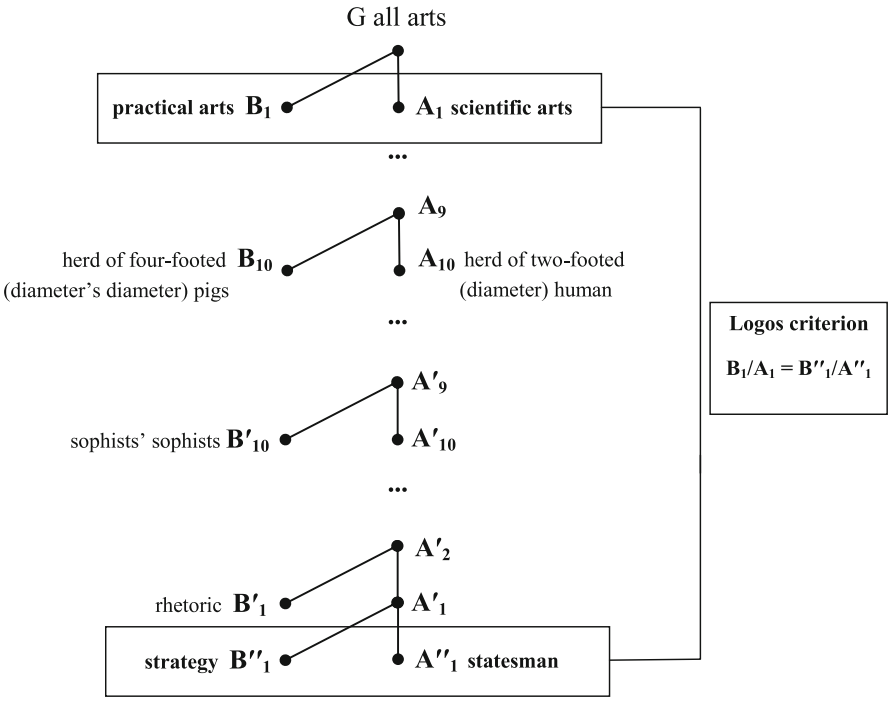

Kronos era
Zeus era
G all arts
B1
A1
A′1
B′1
practical arts
rhetoric
B2
A2
A′2
B′2
appraising arts
judging
B3
A3
A′3
B′3
someone else's commands
diviners, priests
B4
A4
A′4
B′4
lifeless herds
government with laws
B5
A5
A′5
B′5
herd of a single animal
democracy without laws
B6
A6
A′6
B′6
herd of aquatic animals
merchants
B7
A7
A′7
B′7
herd of flying animals
heralds
B8
A8
A′8
B′8
herd of horned animals
satyrs of politics oligarchy
B9
A9
A′9
B′9
herd of mixing animals
centaurs of politics, tyranny
B10
A10
A′10
B′10
diameter's diameter, herd of fourfooted (pigs)
sophists' sophists
B11
A11
A′11
B′11
...
...

The figure of the Division of the Statesman is to be read as follows:

start the division with the left column,
read from up G downwards to A_{10},
consisting of the ten first binary division steps of the *Kronos era*
the initial genus G is divided into B_1 and A_1,
A_1 is divided into B_2 and A_2,
...,
A_9 is divided into B_{10} and A_{10};
continue with the contingent causes
A_{10} to A_{11} to . . .to A'_{11},
the summary steps in the division steps, and
complete the division with the right column,
read from below A'_{11} upward to A'_1,
consisting of the ten last binary division steps of the *Zeus era*
A'_{11} is divided into B'_{10} and A'_{10},
A'_{10} is divided into B'_9 and A'_9,
...,
A'_2 is divided into B'_1 and A'_1, and
A'_1 is divided into B''_1 and A''_1.

Finally the *Logos Criterion* $B_1/A_1 = B''_1/A''_1$, shown in the figure below,

results in the definition of Statesman as the philosophical analogue not simply of periodic, but of *palindromically periodic* anthyphairesis.



It is clear that the purpose of the *Politicus* is to exhibit that the perfect definition of an intelligible Being is by means of a philosophical analogue of a palindromically periodic anthyphairesis. There can be no doubt that the whole of *Politicus* is a philosophical imitation of the strong Theaetetus' theorem on the palindromic periodicity of the anthyphairesis of quadratic incommensurability $a^2 = Nb^2$ for non-square number N. This will be examined in the next Sect. 7.

7 The Reconstruction of the Proof of the Strong Theaetetus' Palindromic Periodicity Theorem, Employing Methods Solely from the Theaetetean Book X of Euclid's *Elements* and the Pigeonhole Argument at the *Politicus* 272d-e

7.1 The Strong Theaetetus' Palindromic Periodicity Theorem

The Division and Colletion for the Statesman in the *Politicus* is clearly a continuation of the examples of Division and Collection that were given in the *Sophistes*, principally for the Angler and for the Sophist, which were shown (in Sect. 3) to be philosophical imitations of a dyad in periodic anthyphairesis.

Furthermore, it has been shown in Sect. 4 that the theorem Plato has (a) reported in the *Theaetetus* 147c-148d that Theaetetus has proved and (b) imitated philosophically in the *Politicus* 283a-287b and the *Philebus* 16c, 23a-25e, is the Theorem: "if $a^2 = Nb^2$ for non-square number N, or more generally if a, b are commensurable in power only lines (namely commensurable in power a^2 to b^2 and incommensurable in length a to b), then the anthyphairesis of a to b is periodic," and in Sect. 6, that a reconstruction of the proof can be obtained with tools exclusively from the Theaetetean Book X of the *Elements* PLUS a pigeonhole argument.

But now prompted by the extremely elaborate philosophical imitation of the periodic palindromicity theorem in the *Politicus*, described in Sect. 6, we have no choice but to admit that in fact Plato is forcefully declaring that in fact Theaetetus has proved the following:

> *Strong Theaetetus' Palindromic Periodicity Theorem.* If a and b are two lines, such that $a^2 = Nb^2$ for a non-square number N, or more generally, incommensurable in length only, then the anthyphairesis of a to b is *palindromically periodic*.

This of course is essentially the important

> *Theorem* Fermat (de Fermat 1679), Brouncker, Wallis, Euler (Euler 1767), Lagrange (Lagrange 1769a, 1769b, 1770).
> If N is a non-square natural number, then the continued fraction of $\sqrt{N}$ is periodic, and in fact palindromically periodic. In fact,there are numbers n and m, $k_1, k_2, \ldots, k_{n-1}, (k_n)$, such that the sequence of successive quotients of the continued fraction of $\sqrt{N}$ has the form
> $[m, \text{period } (k_1, k_2, \ldots, k_{n-2}, k_{n-1}, (k_n), k_{n-1}, k_{n-2}, \ldots, k_2, k_1, 2m)]$,

which was established gradually during the seventeenth and eighteenth century, with contributions by Fermat, Brouncker, Wallis, Euler (who gathered important experimental evidence on the periodicity of the continued fractions of particular quadratic irrationals), and Lagrange (who proved the periodicity). Modern proofs of this result can be found in Hardy and Wright 1938; Weil 1984; Kahane 1985; Fowler 1987.

We have no great difficulty to provide a reconstruction of the proof of the strong Theaetetus' palindromic periodicity theorem with the same tools that we have found in Book X. of the *Elements* and which have been used to prove the weak Theaetetus' periodicity theorem, in Sect. 5. This is done in Sect. 7.2 below.

We emphasize that our reconstruction is based solely on tools from Book X of the *Elements*, but assumes a historical relevance only because of the compelling imitation by Plato with which he constructs the whole of *Politicus*. It is inconceivable that Plato based his later philosophy on something mathematical that was not resting in absolutely solid ground (Recall Plato's passage on the importance of necessity and mathematical proof in geometry (examined in Sect. 4.2.2):

ἀπόδειξιν δὲ καὶ ἀνάγκην οὐδ' ἡντινοῦν λέγετε ἀλλὰ τῷ εἰκότι χρῆσθε,
ᾧ εἰ ἐθέλοι Θεόδωρος ἢ ἄλλος τις τῶν γεωμετρῶν χρώμενος γεωμετρεῖν,
ἄξιος οὐδ' ἑνὸς μόνου ἂν εἴη. *Theaetetus* 162e4–7).

The unmistakable and necessary consequence of our analysis is that, despite the lack of any explicit ancient mathematical evidence, Theaetetus' palindromic periodicity theorem was in possession of the Academy.

Of course the reconstruction of the strong palindromic periodicity theorem makes heavy use of the pigeonhole principle. But, as we have noted in Sect. 5, there is no trace of such principle in Book X of the *Elements*, in fact in any ancient mathematical source. Thus a basic argument, the pigeonhole principle, needed for the proof is missing from Book X.

However, it is remarkable that Plato makes a clear reference to the pigeonhole principle, and in fact precisely at the point needed: In the *Politicus* 272d6-e6, namely when he wants to explain why a soul makes the transition fom the Kronos era to the Zeus era, namely from the first half of the period in the anthyphairetic division to its second palindromic half, Plato explains this transition by means of an argument that seems pure pigeonhole: The soul is forced to go to the Zeus era, to the second palindromic semi-period, because "the time of the generations was accomplished," "the Kronos type generations have all been used up," "the soul had fulfilled all its generations," and "there were as many seeds fallen on the earth as prescribed"—essentially many ways of saying that "all pigeonholes were filled" (Sect. 7.3).

7.2 The Reconstruction of Theaetetus' Palindromic Periodicity Theorem

We next proceed with the proof of the full theorem on palindromic periodicity, described in philosophical imitation by Plato in the *Politicus* (The original proof is contained in the unpublished manuscript Negrepontis 2003).

Definition (Inverse of the conjugate φ_n*) Define the line ω_n by the equality

$$\varphi_n{}^* \cdot \omega_n = b^2$$

(where we have set $\varphi_n{}^* = (a + m_n\ b)/\lambda_n$, which implicitly has been used in Sect. 5 anyway).

Since $\varphi_n{}^*$ is a line of two names, it follows from Book X conjugacy (*Proposition X.112*) that ω_n is an apotome. In fact we can obtain more detailed information.

Proposition (Properties of ω_n).

(i) The line ω_n is an apotome, and in fact

$$\lambda_{n+1}\ \omega_n = a - m_n\ b;$$

(ii) $\omega_n < b$
(iii) $\omega_1(\varphi_1 + 2\ m_1 b) = b^2$
(iv) $\omega_{n+1}(I_n b + \omega_n) = b^2$

Proof All statements follow easily by the definition of ω_n and the periodicity theorem, proved in Sect. 5; the details are left for the reader.

We order the elements of the two sequences (φ_n) and (ω_n) in one sequence as follows:

$$\varphi_1, \omega_1, \varphi_2, \omega_2, \ \ldots, \varphi_n, \omega_n, \ \ldots.$$

From *the pigeonhole principle*, using the periodicity theorem and the above Proposition, there exists *a first element* in the sequence so ordered, such that all the previous elements of the sequence are pairwise distinct, and the element in question coincides with a previous element.

There are two cases to examine:

Case I: The elements φ_1, ω_1, φ_2, ω_2,. . ., φ_{k-1}, ω_{k-1} are pairwise distinct and the element φ_k coincides with a previous element. Then we claim that $\omega_{k-1} = \varphi_k$.
Case II: The elements φ_1, ω_1, φ_2, ω_2,. . ., φ_{k-1}, ω_{k-1}, φ_k are pairwise distinct and the element ω_k coincides with a previous one. Then we claim that $\varphi_k = \omega_k$.

Proof for Case I. There are three subcases:

(a) If the apotome line φ_k coincides with the apotome line φ_1, then we have a contradiction [Proof similar to (b) below]
(b) If the apotome line φ_k coincides with the apotome line φ_j for some j with $1 < j < k$, then we have a contradiction.

[we are to obtain a contradiction from (b).

From the supposition it follows that the two conjugate lines of two names are equal, i.e., that $\varphi_k^* = \varphi_j^*$, and since $\varphi_k^* \cdot \omega_k = \varphi_j^* \cdot \omega_j = b^2$ (by the definition of φ_k^*), we obtain that $\omega_k = \omega_j$. But we have $\omega_j(I_{j-1}b + \omega_{j-1}) = b^2$ and $\omega_k(I_{k-1}b + \omega_{k-1}) = b^2$ (both by Proposition (iv)). Hence $I_{k-1}b + \omega_{k-1} = I_{j-1}b + \omega_{j-1}$; since (on account of Proposition (ii)) both sides of the equation represent the anthyphairetic division of the same whole line with respect to the same line b it follows that the remainders are equal, i.e., $\omega_{k-1} = \omega_{j-1}$, a contradiction, since the elements $\varphi_1, \omega_1, \varphi_2, \omega_2, \ldots, \varphi_{k-1}, \omega_{k-1}$ have been assumed to be pairwise different. Hence φ_k is not equal to φ_j for some j with $1 < j < k$.]

(c) If the apotome line φ_k coincides with the apotome line ω_j, for some j with $j < k-1$, then we have a contradiction. [Proof similar to (b) above.]

Therefore, since φ_k is equal to one of the elements $\varphi_1, \omega_1, \varphi_2, \omega_2, \ldots, \varphi_{k-1}, \omega_{k-1}$, the only possibility left, by (a), (b), and (c), is to have $\omega_{k-1} = \varphi_k$.

Case II is dealt with similarly and we obtain $\varphi_k = \omega_k$.

We now complete the proof by showing the palindromicity in Case I.

From the fact, already proved, that $\omega_{k-1} = \varphi_k$, and the relations $\varphi_k(I_k b + \varphi_{k+1}) = b^2$ (Theorem of Sect. 5.5) and $\omega_{k-1}(I_{k-2}b + \omega_{k-2}) = b^2$ (Proposition of Sect. 7.2. (iv)), we deduce that

$$I_k b + \varphi_{k+1} = I_{k-2}b + \omega_{k-2};$$

since (on account of Theorem of Sect. 5.5 and Proposition of Sect. 7.2) both sides of this equation represent the anthyphairetic division of the same whole line with respect to the same line b, it follows that the two remainders and the two quotients are equal, i.e.,

$$I_k = I_{k-2} \text{ and } \varphi_{k+1} = \omega_{k-2}.$$

Continuing in the same way we obtain that

$$I_{k+1} = I_{k-3} \text{ and } \varphi_{k+2} = \omega_{k-3}.$$

Still continuing in the same way we obtain that

$$I_1 = I_{2k-3} \text{ and } \varphi_{2k-2} = \omega_1.$$

But we have that
$\varphi_{2k-2}(I_{2k-2}b + \varphi_{2k-1}) = b^2$ (Sect. 5) and
$\omega_1(2m_1 b + \varphi_1) = b^2$ (Proposition of Sect. 7.2. (iv));
hence

$$I_{2k-2}b + \varphi_{2k-1} = 2\, m_1 b + \varphi_1.$$

Since (on account of Theorem of Sect. 5.5) both sides of this equation represent the anthyphairetic division of the same whole line with respect to the same line b, it follows that the two remainders and the two quotients are equal, i.e.,

$$I_{2k-2} = 2\, m_1 \text{ and } \varphi_{2k-1} = \varphi_1.$$

From the equality $\varphi_{2k-1} = \varphi_1$ (Logos Criterion (Theorem of Sect. 5.5)),
we deduce
the *periodicity of the anthyphairesis*, and
from the equalities $I_{2k-2} = 2\, m_1$, and $I_1 = I_{2k-3}$, $I_2 = I_{2k-4}$,. . ., $I_k = I_{k-2}$, and $I_{k+1} = I_{k-3}$,
we deduce
the *palindromicity in the period of the anthyphairesis*, as required.
Case II is dealt with similarly.

The general commensurability in power only requires only minor modifications left to the reader. The proof of the theorem is complete.

7.3 Plato's Argument for Going from the Kronos Era to the Zeus Era (*Politicus* 272d6-e6) Appears to be a Philosophical Version of the Pigeonhole Argument Employed in the Proof

For when the time of all these [Kronos era] generations was accomplished (ἐτελεώθη)	
	and the change to[Zeus era] was to take place (μεταβολὴν ἔδει γίγνεσθαι)
and all the earth-born [Kronos era] genus had already been used up (ἀνήλωτο), since every soul (ἑκάστης τῆς ψυχῆς) had fulfilled (ἀποδεδωκυίας) all its generations (πάσας τὰς γενέσεις), by falling into the earth (εἰς γῆν πεσούσης) as many seeds (τοσαῦτα σπέρματα) as prescribed to it (ὅσα ἦν ἑκάστῃ προσταχθὲν), then	
	the heldsman (κυβερνήτης) of the universedropped the tiller and *withdrew* *(ἀπέστη)* to his place of outlook, and fate (εἱμαρμένη) and innate desire (σύμφυτος ἐπιθυμία) made the universe (κόσμον) turn backwards (πάλιν ἀνέστρεφεν). 272d6-e6

Thus for a soul a change from the Kronos era to the Zeus era becomes necessary when the (definitely finite) number of generations of the soul in the Kronos era and

the seeds prescribed to the soul in this era have all been used up; this is precisely the pigeonhole argument.

Not only is Plato in the *Politicus* 272d6-e6 describing, in philosophical-mythological terms a *pigeonhole principle*, but he does so at the point of passing from the Kronos era to the Zeus era, in other words at the point of passing from the first semiperiod to the second and palindromically symmetric to the first. It is remarkable that this is precisely the point, in our reconstructed proof of the strong Theaetetus' palindromically periodic theorem, where the mathematical pigeonhole principle is employed.

Sections 8, 9, 10 The Ancient History of the Pell Problem
We now have all the information we need from Plato and from Book X of the *Elements* to deal with the Pythagorean proof of the Pell property for $N = 2$ (Sect. 8), the reconstructed poof of the Pell property for any non-square number N by Theaetetus (Sect. 9), and the solution of the problem of Pell by the Hindu mathematicians (Sect. 10).

8 The Pell Equation for $N = 2$ by the Pythagoreans

The Pythagoreans discovered the incommensurability of the diameter to the side of a square, by proving that the anthyphairesis of the diameter to the side is infinite, in fact with sequence of quotients the infinite sequence [1, 2, 2, 2,...], and applying Proposition X.2 of the *Elements*. They next introduced the double sequence of the *side p_n and diameter q_n numbers* defined *recursively* by

$$p_1 = q_1 = 1,$$

$$p_{n+1} = p_n + q_n, q_{n+1} = 2p_n + q_n \text{ for } n = 1, 2, \ldots,$$

whose anthyphairesis has as sequence of quotients a finite initial segment of the infinite anthyphairesis of the diameter to the side of a square. According to the ancient authors recounting the side and diameter numbers, Theon Smyrneus (Theon Smyrnaeus, ed.: Hiller 1878) 44,18-45,8, Iamblichus, *Comments to Nikomachus* (Iamblichus, ed.: Pistelli 1894) 92,23-93,6, and mainly Proclus, *Commentary to Politeia* (Procli Diadochi, ed.: Kroll 1899–1901) 2, 24,16–25,13 and 2, 27, 1–29,4, the Pythagoreans verified and then showed that these numbers satisfy the Pell Diophantine equation

$$q_n{}^2 = 2p_n{}^2 + (-1)^n.$$

Some scholars (such as Freudenthal 1953) support, but more recently others are either ambivalent (such as Fowler 1994) or have challenged (such as Unguru 1991; Acerbi 2000) the thesis that the Pythagoreans had obtained a real proof by induction, or in fact any kind of proof of the Pell property. In our work (Negrepontis and

Farmaki 2019, Chap. 8), reading the sources meticulously, we have some novel arguments showing that the Pythagoreans had indeed an inductive proof, but without the principle of mathematical induction.

The Pythagoreans noted that each rational diameter by itself is not a true diameter, because its square is not equal to double the square of its side (something impossible for numbers, by the incommensurability of the diameter to the side of a square, as noted by Proclus 2,27,1-4), but equal to double the square minus 1 or plus 1; however, they found encouragement into the fact that the square of the diameter is double the square minus 1 or plus 1 in an alternating (ἐναλλάξ, Theon 44,14, Iamblichus 92,23) manner. In consequence, in order to ensure exact diameter type equality to side (ἀπίσωσις,92,27; ἰσότης, 93,1; ἀπισώσει, 93,3 in Iamblichus; ἰσότητα, 45,5 in Theon) they were led to form the sum of the square of a rational diameter together with (*μετὰ*) the square of the next diameter, which sum is then exactly equal to the double (in Proclus 2, 25,11) of the square of the side and (*καὶ*, a connective different from *μετὰ*) the square of the next side. Proclus describes this basic *heuristic* idea of the Pythagoreans, as follows:

> ἐὰν δὲ λάβωμεν ἁπάσας τὰς ἀπὸ τῶν τοιούτων διαμέτρων, ἔσονται διπλάσιαι ὄντως, ***ὧν***
> ἑκάστη μονάδι μείζων ἢ ἐλάσσων διπλασίου·
> οἷον ἡ ἐννέα ***μετὰ*** τοῦ μθ′ τῆς τοῦ κε′ ***καὶ*** δ′.
>
> If we take together all the diameters, their squares are really double,
> [as, for example, if we receive the square 9 of the diameter 3 *together with* (***μετὰ***) the square 49 of the diameter 7, they are really double of the square 25 of the side 5 *and* (***καὶ***) the square 4 of the side 2]
> *of which* (hon) squares of diameters
> each is greater or less of the double square of the side by a unit. 2,25,9–12

As with all good heuristic ideas, this one is really a step in *analysis*, going backward in the chain of deduction: In order to prove what is required, namely the Pell property of the side and diameter numbers, it is enough to prove that

> the square of the diameter *together with (μετὰ)* the square of the next diameter is really double the square of the side *and (καὶ)* the square of the next side.

Proclus tells us that the Pythagoreans were encouraged by this method:

> διὸ καὶ οἱ Πυθαγόρειοι ἐθάρρησαν τῇ μεθόδῳ. *Proclus* 2,25,12–13

Thus the Pythagoreans realized that it is enough to prove that

> for any *natural numbers* a, b, the square of $a + 2b$ *together with* ***(μετὰ)*** the square of a equals double the square of $a + b$ *and* ***(καὶ)*** of b.

But in fact the Pythagoreans carried the analysis one step further, going one more step bacward in the chain of deduction. As with Propositions II.4–8, the statements

are geometric, about lines, and so the Pythagoreans stated that it will be enough to prove that

> for any *lines* a, b, the square of $a + 2b$ *together with* ***(μετὰ)*** the square of a equals double the square of $a + b$ *and* ***(καὶ)*** of b.

Thus the initial heuristic idea of the Pythagoreans led them to the discovery, by analysis, of the geometric Proposition II.10 of the *Elements*:

for lines a, b, $(a + 2b)^2 + a^2 = 2((a + b)^2 + b^2)$.

Now it is the turn of synthesis/proof.

From Proposition II.10, by reversing the analytic steps, we are to obtain the required statement about the side and diameter numbers. However before going to the side and diameter numbers, the Pythagoreans, wishing to prove that II.10 is really a geometric statement, not equivalent to, but stronger than, the corresponding arithmetical one, they interposed (προετίθεσαν, 2,27,11) the *elegant theorem* (*θεώρημα γλαφυρὸν*, 2,27,12), which they proved geometrically (γραμμικῶς, linearly, 2,27,17), clearly turning II.10 into the statement of a (geometric) inductive step (if a is diameter and b side, then $a + 2b$ is diameter and $a + b$ side):

> διὰ τὸ θεώρημα [Proposition II.10] ἔσται
> τὸ ἀπὸ τῆς ΑΔ *μετὰ* τοῦ ἀπὸ τῆς ΔΓ διπλάσιον τοῦ ἀπὸ τῆς ΑΒ *καὶ* ἀπὸ τῆς ΒΔ.
> *ὧν*
> τὸ ἀπὸ τῆς ΔΓ διπλάσιον τοῦ ἀπὸ τῆς ΑΒ·
> καὶ λο[ιπὸν] *ἄρα* τὸ ἀπὸ τῆς ΑΔ τοῦ ἀπὸ τῆς ΒΔ διπλάσιον· 2,27,24-2,28,3,
>
> By the theorem [Proposition II.10] we have that
> the square of ΑΔ together with (*μετὰ*) the square of ΔΓ is
> double the square of ΑΒ and (καὶ) the square of ΒΔ,
> of which (*ὧν*)
> the square of ΔΓ is double the square of ΑΒ,
> hence *ἄρα* the square of ΑΔ is double the square of ΒΔ.

The wording of this argument is truly remarkable, because it preserves intact the linguistic peculiarity with which the Pythagorean heuristic idea was conceived (*μετὰ*, καὶ, *ὧν*), thus confirming fully the analytical nature of the argument, and revealing at the same time that the Pythagorean heuristics was for the purpose of an inductive argument.

T*he statement of the elegant theorem*:

> ἡ μὲν διάμετρος προσλαβοῦσα τὴν πλευράν, ἧς ἐστιν διάμετρος, **γίνεται πλευρά**,
> ἡ δὲ πλευρὰ ἑαυτῇ συντεθεῖσα καὶ προσλαβοῦσα τὴν διάμετρον τὴν ἑαυτῆς **γίνεται διάμετρος** 2, 27, 13–16
>
> the diameter, having received (προσλαβοῦσα) the side, of which it is diameter, becomes side (γίνεται πλευρά), and
> the side, being composed with itself (ἑαυτῇ συντεθεῖσα) and having received (προσλαβοῦσα) the side's diameter becomes diameter **(γίνεται διάμετρος)**.

In the final step of synthesis, Proposition II.10, in an arithmetic form, immediately deduced from the geometrical, is used for the proof of every inductive step of the Pell equation, stated in a manner linguistically identical with the statement of the elegant theorem (so that the latter serves as the model for identical language for every inductive step in the proof of the Pell equation). The first inductive step is as follows:

> λαβέτω οὖν ἡ διαμετρικὴ μονὰς πλευράν, ὅ ἐστι μονάδα (καὶ γὰρ αὕτη διάμετρος ἦν), **γίνεται** δυὰς καὶ **πλευρὰ** τετράδος·
> ἡ δὲ πλευρικὴ μονὰς ἑαυτῇ συντεθεῖσα λαβέτω διάμετρον ἄλλην μονάδα (ἦν γὰρ καὶ ἡ διαμετρική), **γίνεται** τριὰς [καὶ **διάμετρος**]
> καὶ ποιεῖ τὸν ἐννέα, μονάδι μείζονα τοῦ ἀπὸ δυάδος· 2,28,17–21
>
> The diametric unit receiving side . . . becomes side, and
> The side unit composed with itself and receiving diameter. . .becomes [diameter].

(the second step 2,28,22–24; the third step 2,28,27-29,1)

Thus the Pythagoreans proved the following:

Proposition P(*n*) Let n be a given natural number. The n-th side number p_n and diameter number q_n satisfy the Pell Diophantine equation

$q_n^2 = 2p_n^2 + (-1)^n$.

The Pythagorean proof, reconstructed according to above arguments, now runs smoothly as follows:

Proposition P(1) clearly holds.

We obtain a proof Proof (2) for Proposition P(2), by *applying the arithmetical version of Proposition II.10* with $a = p_1$, $b = q_1$, using the recursive definition, and using the fact that Proposition P(1) holds.

We next obtain Proof (3) for Proposition P(3) by applying Proposition II.10 with $a = p_2$, $b = q_2$, using the recursive definition, and using the fact that Proposition P (2) holds. It is clear that continuing in this way we can obtain Proof (n) for Proposition P(n).

The number of steps of these separate proofs Proof (n) is going to infinity as n is going to infinity (and thus ad infinitum, καὶ ἀεὶ οὕτως, 2,29,4).

So the ancients had a proof Proof (n) for Proposition P(n) for every n, but they had no proof for all n simultaneously, namely no proof for the Proposition $(\forall n)$P(n), because this proof would require an infinite number of steps.

We conjecture that the Pythagorean proof of the Pell property for $N = 2$ marked the *Birth of Induction*. It was *unfinitized* induction, Greeks never finitized mathematical induction (Plato's *Parmenides* 149a7-c3 has been proposed (Wedberg 1955 and Acerbi 2000) as a perfect specimen of ancient proof by mathematical induction; however our analysis in (Negrepontis and Farmaki 2021, Sect. 3.6) shows that Plato means this to be an argument involving only finite induction, since there are only finitely many "ἄψεις"/logoi in every true Being, forming a period.).

9 A Reconstruction of Theaetetus' Proof of the General Case of the Pell Property

Theaetetus had a strong motivation to study the Pell property for every non-square number, because of the Pythagorean discovery of the Pell property for $N = 2$ (Sect. 8). Once a proof of the palindromic periodicity theorem has been obtained (Sect. 7), the proof of the general case of the Pell property, namely for every non-square natural number, is straightforward.

9.1 The Tools for the Proof of the Pell Property in Book X of the *Elements*

There are strong internal signs that the Pell problem is a central concern in Book X.

One such sign is the classification of the apotome lines in six classes in X.84/85, which for the simple apotome lines is intimately related with whether their Pell number is square or not.

The second sign is Proposition X.97 stating that the square of every apotome line can be made equal to a rectangle one of whose sides is a rational line and the other an apotome with square Pell number.

A natural extension of X.97 that the product of two (simple) apotome lines is equal to a rational line times an (simple) apotome line with Pell number the product of the Pell numbers of the two initial lines, not stated explicitly in the *Elements*, is all that is needed on top of the palindromic periodicity theorem, to provide a complete proof of the Pell equation.

9.1.1 The Pell-Related Classification of the Apotome Lines in Six Classes in X.84/85

The classification of all apotome lines in six kinds in the definitions X.84/85, which has played no role for the proof of the palindromic periodicity, is according to two criteria.

The first criterion is *external* and concerns the commensurability or incommensurability of the assumed line in relation to the whole part or to the annex part of the apotome. This criterion does not have to do with the internal structure of the apotome, since *the same apotome line* can have assumed line commensurable to the whole part, or to the annex part, or to neither of the two. This criterion is introduced for making the apotome lines compatible to each other, by having the same/commensurable assumed line.

The second criterion is *internal*, has to do with the structure of the apotome line, and can be described as follows: If g–h is an apotome line then we form the right triangle with hypotenuse g and one vertical side h, and we define π to be the second vertical side: $\pi^2 = g^2 - h^2$. The criterion for classification is whether the lines π and the whole g are commensurable or not. At a formal level it is a rather opaque criterion, but it is immediately illuminated if we apply it on the simple apotome lines, which, as we have seen with the constructions in Propositions X.85–90 and

with the proof of the palindromic periodicity theorem (Sects. 5 and 7), are the main object of interest in Book X. A simple apotome line γ has the form either $\gamma = (qb - pa)/\lambda$ or $\gamma = (pa - qb)/\lambda$, where p, q, λ are natural numbers, with $a^2 = Nb^2$, where N a non-square number.

We introduce the following crucial **Definition**. The *Pell number* of a simple apotome line γ, denoted by Pell(γ), is the number $(q^2 - Np^2)/\lambda^2$.

For $\gamma = (qb - pa)/\lambda$, Pell($\gamma$) is a positive rational number, for $\gamma = (pa - qb)/\lambda$ negative.

It becomes evident that the second, internal criterion for the classification of the simple apotome lines is intimately related with the Pell problem, in fact can now be stated *solely in terms of the Pell number of the apotome* and in fact with whether the *Pell number of the apotome line is square or not*.

Classification of the simple apotome lines according to X.84/85.

γ first apotome	either $\gamma = (qb - pa)/\lambda$	Pell(γ) square and b assumed line [X.85]
	or $\gamma = (pa - qb)/\lambda$	$-$Pell(γ)·N square and a assumed line
γ second apotome	either $\gamma = (qb - pa)/\lambda$	Pell(γ) square and a assumed line [X.86]
	or $\gamma = (pa - qb)/\lambda$	$-$Pell(γ)·N square and b assumed line
γ third apotome	either $\gamma = (qb - pa)/\lambda$	Pell(γ) square and assumed line not commensurable to b, or to a [X.87]
	or $\gamma = (pa - qb)/\lambda$	$-$Pell(γ)·N square and assumed line not commensurable to b, or to a
γ fourth apotome	either $\gamma = (qb - pa)/\lambda$	Pell(γ) positive non-square and b assumed line
	or $\gamma = (pa - qb)/\lambda$	$-$Pell(γ)·N non-square and a assumed line [X.88]
γ fifth apotome	either $\gamma = (qb - pa)/\lambda$	Pell(γ) positive non-square and a assumed line
	or $\gamma = (pa - qb)/\lambda$	$-$Pell(γ)·N non-square and b assumed line [X.89]
γ sixth apotome	either $\gamma = (qb - pa)/\lambda$	Pell(γ) positive non-square and assumed line not commensurable to b or to a
	or $\gamma = (pa - qb)/\lambda$	$-$Pell(γ)·N non-square and assumed line not commensurable to b, or to a [X.90]

If $\gamma = (qb - pa)/\lambda$, then $\pi^2 = (q^2b^2 - p^2a^2)/\lambda^2 = ((q^2 - p^2N)/\lambda^2)b^2 = \text{Pell}(\gamma)b^2$, hence

qb and π are commensurable if and only if $\text{Pell}(\gamma)$ is a square number.

If $\gamma = (pa - qb)/\lambda$, then $\pi^2 = (p^2a^2 - q^2b^2)/\lambda^2 = ((p^2N - q^2)/\lambda^2)b^2 = -\text{Pell}(\gamma)b^2 = -\text{Pell}(\gamma)a^2/N = (-\text{Pell}(\gamma)\cdot N/N^2)a^2$ hence,

pa and π are commensurable if and only if $-\text{Pell}(\gamma)\cdot N$ is a square number.

The simple verifications of equivalence of the definitions X.84/85 with the corresponding conditions in the table above.

The table shows that the classification of the (simple) apotome lines γ is most closely related to the Pell number $\text{Pell}(\gamma)$, and in fact to whether $\text{Pell}(\gamma)$, or the modified expression $-\text{Pell}(\gamma)\cdot N$, is a square or non-square number, in case γ is of the form $qb - pa$, or of the form $pa - qb$, correspondingly.

We note that an even remainder $\gamma_{2n} = q_{2n}b - p_{2n}a$ of the quadratic anthyphairesis a to b, $a^2 = Nb^2$, N non-square, is

a first apotome if its Pell number $q^2 - Np^2$ is a square number,

a fourth apotome otherwise.

9.1.2 The Pell-Related Proposition X.97 of the *Elements*

Proposition X.97, stating that the "square" of every apotome line is a first apotome line, is crucial for the proof of the Pell property.

Proposition X. 97 The square on an apotome applied to a rational straight line produces as breadth a first apotome.

[If γ is an apotome with respect to the assumed line b and $\gamma^2 = \chi\ b$, then χ is a first apotome]

Proposition X.97 reveals its true content when applied to simple apotome lines related to the quadratic incommensurability $a^2 = Nb^2$, N non-square, namely to the apotomae of the form either $\gamma = (pa - qb)/\lambda$ or $\gamma = (qb - pa)/\lambda$, for which the assumed line is the line b. The proof of Proposition X.97 in fact shows that for those lines γ,

$\text{Pell}(\chi) = (\text{Pell}(\gamma))^2$.

The close connection of Proposition X.97 with the Pell property is evident.

9.2 The Crucial Step for a Reconstruction of Theaetetus' Proof of the General Pell Property, Employing the Palindromic Periodicity Theorem and Proposition X.97

The principal application of Proposition X.97 is with respect to the reduced remainders $\varphi_n = (a - m_nb)/\lambda_n$.

Proposition If φ_n is a reduced remainder in the initial semi-period of the period of the anthyphairesis of a to b, and ω_n is its corresponding conjugation line in the final semiperiod, then

$$\text{Pell}\,(\varphi_n) = -\lambda_{n+1}/\lambda_n, \text{Pell}\,(\omega_n) = -\lambda_{\text{n}}/\lambda_{n+1}, \text{and}$$

$$\text{Pell}\,(\varphi_n \cdot \omega_{\text{n}}/b) = 1$$

Proof By direct computation

$$\varphi_n = (a - m_n b)/\lambda_n,$$

hence Pell $(\varphi_n) = -(N - m_n{}^2)/\lambda_n{}^2 = -\lambda_n\,\lambda_{n+1}/\lambda_n{}^2 = -\lambda_{n+1}/\lambda_n$, and

$$\omega_n = (a - m_n b)/\lambda_{n+1},$$

hence Pell $(\omega_n) = -(N - m_n{}^2)/\lambda_{n+1}{}^2 = -\lambda_n\,\lambda_{n+1}/\lambda_{n+1}{}^2 = -\lambda_n/\lambda_{n+1}$.

Thus by Proposition X.97

$$\text{Pell}\,(\varphi_n{\cdot}\omega_n/b) = \text{Pell}\,(\varphi_n)\ \text{Pell}\,(\omega_n) = (-\lambda_{n+1}/\lambda_n)(-\lambda_n/\lambda_{n+1}) = 1.$$

But note that line $\varphi_n.\omega_n/b$ is *not* a solution to the Pell problem yet, because it is a fractional N-apotome, not an integral one.

9.3 The Reconstruction of Theaetetus' Proof of the General Case of Pell Property

Proposition

$$\text{Pell}\,\left(\gamma_{n+1}\right) = \text{Pell}\,(\gamma_n)\ \text{Pell}\,\left(\varphi_{n+1}\right).$$

Proof By definition $\gamma_{n+1}\ b = \gamma_n\ \varphi_{n+1}$, namely $\gamma_{n+1} = \gamma_n\ \varphi_{n+1}$/b. The conclusion follows by a natural extension of Proposition X.97 to a product of two apotome lines, rather than to the square of an apotome line, an extension not mentioned explicitly in Book X.

Proposition

$$\text{Pell}\,(\gamma_n) = \text{Pell}\,(\varphi_1)\ \text{Pell}\,(\varphi_2)\ldots\text{Pell}\,(\varphi_n).$$

If $a^2 = Nb^2$, then Anth$(a, b) = [m_1, \text{period}(I_1, I_2,\ldots, I_n, (I_{n+1}), I_n,\ldots,I_2, I_1, 2m_1)]$.

Then, with middle γ_{n-1} present, we have

$a = m_1 b + \gamma_1$	$\varphi_1 = \gamma_1$
$b = I_1\gamma_1 + \gamma_2$	$\varphi_2\,\gamma_1 = \gamma_2 b$

(continued)

$\gamma_1 = I_2\gamma_2 + \gamma_3$	$\varphi_3\ \gamma_2 = \gamma_3 b$
...	...
$\gamma_{n-2} = I_{n-1}\ \gamma_{n-1} + \gamma_n$	$\varphi_n\ \gamma_{n-1} = \gamma_n b$
$\gamma_{n-1} = I_n\ \gamma_n + \gamma_{n+1}$	$\varphi_{n+1}\ \gamma_n = \gamma_{n+1} b$
$\gamma_n = I_n\ \gamma_{n+1} + \gamma_{n+2}$	$\omega_n\ \gamma_{n+1} = \gamma_{n+2} b$
...	...
$\gamma_{2n-3} = I_3\ \gamma_{2n-2} + \gamma_{2n-1}$	$\omega_3\ \gamma_{2n-2} = \gamma_{2n-1} b$
$\gamma_{2n-2} = I_2\ \gamma_{2n-1} + \gamma_{2n}$	$\omega_2\ \gamma_{2n-1} = \gamma_{2n} b$
$\gamma_{2n-1} = I_1\ \gamma_{2n} + \gamma_{2n+1}$	$\omega_1\ \gamma_{2n} = \gamma_{2n+1} b$
$\gamma_{2n} = 2m_1\ \gamma_{2n+1} + \gamma_{2n+2}$	

Hence, the Pell number of the integral N-apotome $\gamma_{2n+1} = p_{2n+1}a - q_{2n+1}b$, equal by definition to the number $(N p_{2n+1}{}^2 - q_{2n+1}{}^2)$, is equal to 1, since

$$\mathrm{Pell}\left(\gamma_{2n+1}\right) =$$

$$\mathrm{Pell}(\omega_1\ \gamma_{2n}/b) = \mathrm{Pell}(\omega_1)\ \mathrm{Pell}(\gamma_{2n}) =$$

$$\mathrm{Pell}(\omega_1)\ \mathrm{Pell}(\omega_2\ \gamma_{2n-1}/b) = \mathrm{Pell}(\omega_1)\ \mathrm{Pell}(\omega_2)\ \mathrm{Pell}(\gamma_{2n-1}) =$$

$$\mathrm{Pell}(\omega_1)\ \mathrm{Pell}(\omega_2)\ \mathrm{Pell}(\omega_3\ \gamma_{2n-2}/b) = \mathrm{Pell}(\omega_1)\ \mathrm{Pell}(\omega_2)\ \mathrm{Pell}(\omega_3)\ \mathrm{Pell}(\gamma_{2n-2}) =$$

$$\ldots$$

$$\mathrm{Pell}(\omega_1)\ \mathrm{Pell}(\omega_2)\ \mathrm{Pell}(\omega_3) \ldots \mathrm{Pell}(\omega_\mathrm{n})\ \mathrm{Pell}\left(\varphi_{n+1}\right)\ \mathrm{Pell}(\varphi_n) \ldots \mathrm{Pell}(\varphi_3)\ \mathrm{Pell}(\varphi_2)\ \mathrm{Pell}(\varphi_1) =$$

$$(\mathrm{Pell}(\omega_1)\mathrm{Pell}(\varphi_1))\ (\mathrm{Pell}(\omega_2)\mathrm{Pell}(\varphi_2))\ (\mathrm{Pell}(\omega_3)\mathrm{Pell}(\varphi_3)) \ldots (\mathrm{Pell}(\omega_\mathrm{n})\mathrm{Pell}(\varphi_\mathrm{n}))\ \mathrm{Pell}\left(\varphi_{n+1}\right)$$
$$= 1{\cdot}1{\cdot}1{\cdot} \ldots {\cdot}1{\cdot}(-1) = -1.$$

If we continue with one more period, the Pell number of the last remainder of the second period will clearly be +1.

If the middle term γ_{n-1} is not present, then the middle(-1) term in the product for the Pell number of the last remainder of the first period is missing, and the last remainder of this period, already has Pell number + 1.

Note. We might conclude by saying that Tannery's (1882) and van der Waerden's (1976) intuition was right, that the Greek geometers had somehow proved the general Pell property, but they definitely did not possess the decisive arguments, namely an understanding of the deep mathematical secrets hidden in Plato's philosophy and the Pell content of some key Propositions and the classification of apotome lines in Book X.

9.4 The Third Purpose of Book X of the *Elements* Is the Extension of the Quadratic Anthyphairetic Periodicity to the Canonical Solids by Means of Application of Areas, in an Attempt to Counter the Successes of the Non-Anthyphairetic Approach by Archytas and Eudoxus

We have argued that the principal two purposes of Theaetetus in Book X is the proof of the theorem on palindromic periodicity and the proof of the Pell problem.

9.4.1 The Question on the Role of the Pythagorean Application of Areas in Book X

There are still several Propositions in Book X that have not been used as tools for the proof of these two results. This is essentially the hexad of Propositions X.91–96 and its conjugate X.54–59, distinguished by the fact that they employ the *Pythagorean application of areas, always in defect*, together with several Propositions X.17–21, 33–36, that are preparatory for the two hexads.

The question is: What is the purpose of the introduction of the Pythagorean application of areas in Book X? We will outline our interpretation as this is not the main subject of our present work. It will be the subject of a separate more detailed publication.

9.4.2 In Order to Answer the Question in 8.4.1 We Concentrate on the Only Use of the Application of Areas Introduced in Book X, the Study of the Canonical 20heron in Book XIII, Especially in Propositions XIII.11 and 16 of the *Elements* by Means of Proposition X.94

It seems reasonable to try to find the answer to the Question set in 8.4.1 by analyzing the only use of the group of propositions of Book X that employ application of areas. It is only Proposition X.94 employed for the proof of Proposition XIII.16 on the canonical solid 20hedron, with use of XIII.11.

Definition (Proposition X.76) A line Ω is a *minor* line
if there are two lines Φ, Ψ, such that
$\Phi > \Psi$, $\Omega = \Phi - \Psi$,
$\Phi^2 + \Psi^2$ is a *rational* area (namely commensurable to b^2), and
$\Phi\Psi$ is a *medial* area (namely equal to a rectangle with sides commensurable in square).

Proposition XIII.16 The side, edge of the 20hedron is a *minor* with respect to the diameter of the sphere in which the 20hedron is comprehended.

Proposition XIII.16 is a simple consequence of Propositions XIII.11 and X.94.

Proposition XIII.11
Let Ω be the side of a canonical 5gon, comprehended by a circle of radius b, and let γ be the simple apotome $\gamma = (5b - a)/4$, with $a^2 = 5b^2$.

Then

$$\Omega^2 = \gamma \cdot 2b.$$

Proposition X.94. If the line γ is is a *fourth (simple) apotome* and a rational line r, such that $\Omega^2 = \gamma \cdot r$, then Ω is a *minor* line.

Here is an *outline of the proof* of Proposition X.94.

We employ the special Pythagorean application of areas in defect, a special form of the application of areas, as follows.

[1] Let γ be a *(simple) apotome*, $\gamma = qb - pa$, where q, b are rational numbers, and a, b are lines commensurable in square,

namely there are natural numbers M, N, such that $a^2/b^2 = M/N$.

[2] We employ *Pythagorean application of areas in defect with data qb and pa*/2.

Since $qb > pa$, it follows that $qb/2 > pa/2$, and hence an application of areas in defect with these data is possible.

[3] We form *the circle* with diameter qb, we take the height $pa/2$, and

form *the right triangle*, with hypotenuse the diameter qb,

and third point the point where the height $pa/2$ meets the circumference.

[4] We set Φ to be one of the sides of the right triangle, so that $\Phi^2 = (qb - x)qb$, and

Ψ the other side, so that $\Psi^2 = qbx$.

We then verify all required conditions,

$(\Phi - \Psi)^2 = \gamma \cdot qb$,

$\Phi^2 + \Psi^2 = (qb)^2$, a rational area,

$\Phi\Psi = qb \cdot pa/2$, a medial area.

Proof of XIII.11 We verify that γ is a fourth apotome, and we apply the fundamental Proposition X.94 on γ.

9.4.3 The Special Application of Areas Introduced by Theaetetus in Book X, with Givens qb and pa/2, in Notation of Sect. 9.4.2, Is Such That the Anthyphairesis of the Solution x to b Is Eventually Periodic

The following proposition clarifies the role played by the *special* Pythagorean application of areas appearing in Proposition X.94,

namely a Pythagorean application of areas in defect with data two lines qb and $pa/2$ that are not in arbitrary relation, as in a general application of areas, but are *commensurable in square*.

The effect of this condition is to produce an equation of the form

$$Ax^2 + Cb^2 = Bxb,$$

with natural numbers A, B, C as coefficients.

Finally, such an application of areas produces an anthyphairesis of x to b that is eventually periodic.

Proposition If γ is a *fourth (simple) apotome*, $\gamma = qb - pa$,

where q, p are rational numbers, and a, b are lines commensurable in square, namely there are natural numbers M, N, such that $a^2/b^2 = M/N$, then the *Pythagorean application of areas in defect* with data qb and $pa/2$

(a) leads to the construction of a line x, such that

$$x(qb - x) = p^2a^2/4 = \left(p^2M/4\,N\right)b^2, \text{namely}$$

$$x^2 = qxb - \left(Mp^2/4\mathrm{N}\right)\cdot b^2;$$

(b) the *discriminant* Δ of this application of areas is equal to

$$\Delta = \left(q^2 - M/4Np^2\right)b^2; \text{and}$$

(c) the anthyphairesis of x to b is *eventually periodic*.

We note that because of the commensurability in square of a, b, the application of areas has *natural numbers* as coefficients.

(corresponding, by setting $\omega = x/b$,

to a second degree equation $\omega^2 - q\omega + Mp^2/4N = 0$ with integer coefficients).

The *proof of the proposition* is a straightforward consequence of

the invariance of the discriminant under anthyphairetic substitution (no use is made of the conjugation theory) with a pigeonhole argument. The details are omitted.

9.4.4 The Interpretation of the *Anonyma Scholia X.135, 185* to the *Elements*

In turn the proposition, establishing the periodicity of anthyphairesis, produced by this special application of areas, provides a satisfactory interpretation of the hitherto mysterious philosophically phrased *Scholia X.135, 185* to the *Elements*.

Scholion in Euclid's Elements X.135, 1–9.
It is worthy of admiration (Θαυμάζειν ἄξιον), that
the ruling, controlling power (κρατητικὴ δύναμις) of the triad (τῆς τριάδος)
finitises (ἀφορίζει) the alogos power (ἄλογον δύναμιν) and
reaches (διήκει) till the extremes (μέχρι τῶν ἐσχάτων),
because each kind of alogia (ἀλογίας) is *finitised (ἀφορίζεται)* by some mean (μεσότητος),
one by the geometric (γεωμετρικῆς) mean, the other by the arithmetical (ἀριθμητικῆς) mean,
the other by the harmonic/musical (μουσικῆς) mean;
and it seems that the substance of the soul sittting upon (ἐπιβατεύουσα) the substance of the magnitudes according to the "logoi" of the soul (κατὰ τοὺς ἐν αὐτῇ λόγους)
finitises (ὁρίζειν) everything indefinite (ἀόριστον) and the infinite of the alogia (τὴν τῆς ἀλογίας ἀπειρίαν)
by means of these three bonds (δεσμοῖς).

Scholion in Euclid's Elements X.185
holding together (συνέχουσα) its scattered and divided quality (τὸ σκεδαστὸν αὐτῆς)

According to these *Scholia*, phrased in the language of platonic philosophy, some power controls and *finitizes* the anthyphairetic infinity present in each of the 13 (1 medial, 6 additive, 6 subtractive) *alogoi lines*, by means of the three bonds (of addition, subtraction, and geometric mean). The finitization can only refer to some form of *anthyphairetic periodicity.*

The only anthyphairetic periodicity present in the alogos line minor is the anthyphairetic periodicity produced by the special Pythagorean application of areas in defect provided by X.94, according to the proposition above.

9.4.5 The Answer to the Question in 8.4.1: The Purpose of the Introduction of Application of Areas in Book X Is the Extension of the Quadratic Anthyphairetic Periodicity to the Canonical Solids

Since, according to Proposition XIII.16, the side, edge of the 20hedron is a minor wrt the diameter of the sphere in which the 20hedron is comprehended; it follows that the finitizing power of periodic anthyphairesis is present and dominates the canonical solids. This is precisely our interpretation of the purpose of the introduction of application of areas in Book X.

9.4.6 The Answer Given in Sect. 9.4.5 Should Be Understood in Context to the Successes of the Non-Anthyperetic Approach by Archytas and Eudoxus in Various Basic Questions

Despite a claim to the contrary in the *Theaetetus* 148b2 ("καὶ περὶ τὰ στερεὰ ἄλλο τοιοῦτον"), Theaetetus failed to extend the quadratic, plane periodic anthyphairesis to a cubic stereometry, since in fact it does not hold. Theaetetus' anthyphairetic method of proof on quadratic incommensurabilities does not extend to cubic incommensurabilities.

Soon after, an alternative, antagonistic *non-anthyphairetic approach to incommensurabilities* took place, proposed by *Archytas* and his student Eudoxus.

Archytas succeeded in achieving arithmetical, non-anthyphairetic proofs, not only of quadratic incommensurabilities, first shown in terms of periodic anthyphairesis by Theaetetus, but also of cubic incommensurabilities, as Theaetetus could not do. The proofs were based on Proposition VII.27 (if a, b are relatively prime numbers, then a^n, b^n are relatively prime for any number n as well), and its derivative Proposition VIII.14, on the existence of an arithmetical mean proportional for the plane case, and two *arithmetical mean proportionals* for the cubic incommensurabilities.

Archytas' incommensurability proofs were criticized by Plato in the *Philebus* 16c5-17a5 as "eristic," "non-dialectical" proofs.

Soon after, *Eudoxus,* a student of Archytas, developed a *non-anthyphairetic approach to the theory of ratios of magnitudes*, replacing the limited anthyphairetic theory by Theaetetus, exposed in Book V of the *Elements*, foreshadowing Dedekind

cuts. Eudoxus theory was sharply criticized in *Anonymon Scholion X.2* to the *Elements* by circles close to Plato.

The new theory of ratios by Eudoxus made possible *Archytas'* ingenuous *non-anthyphairetic approach to duplication of the cube* by means of constructing two mean proportionals (as initially conceived by Hippocrates).

Archytas' duplication of the cube was clearly criticized by Plato in the *Politeia* 527d-528d (Plato 1969) and by Plutarch, in *Quaestiones Convivales* 718 B8-F4.

(In this section we follow Negrepontis 2019, pp. 83–86, where the reader can find more details).

All these successes by Archytas and Eudoxus have put Theaetetus and Plato in the defensive, and in need of some success of the anthyphairetic approach in solid geometry. Propositions X.94 and XIII.16, interpreted as in Sects. 9.4.4 and 9.4.5, are such results.

9.4.7 Passage *Meno* 86e-87b Shows Plato's Strong Interest for Theaetetus' Special Application of Areas in Defect in Book X

The *Meno* passage 86e-87b, generally considered mysterious, was shown in Negrepontis and Farmaki 2019, Sect. 7, to go from an application of areas in defect with data a and m, with $a/2 > m$, into a circle with diameter a and an inscribed (*right*) *triangle* with hypotenuse a and height $m/2$ (and not an isosceles one, as claimed there), establishing a close relation of the *Meno* passage with Propositions X.18 and X.33, preparatory for Proposition X.94. Plato is of course interested in Theaetetus' introduction of application of areas, since via them a finitization is achieved by means of periodic anthyphairesis.

9.4.8 Reappraisal of the Old Dispute About Geometric Algebra: Pre-Euclidean Application of Areas Is Mostly About Periodic Anthyphairesis and Secondarily Geometric Algebra

Our interpretation of the purpose of the application of areas in Book X forces a rather drastic reappraisal of the dispute on geometric algebra. This is the principal pre-Euclidean application of areas, and while Theaetetus is certainly practicing some geometric algebra, nevertheless his main interest is the periodic anthyphairesis of the application of areas. The purer occurrence of geometric algebra in the application of areas must perhaps wait until periodic anthyphairesis is no longer a main concern, e.g., in Apollonius.

9.4.9 The Purposes of Book X of the *Elements*

Book X is thus revealed to have a most exciting purpose: Supplying the tools for the statement and proof of the palindromic periodicity of the anthyphairesis of the quadratic irrationals and of the closely related solution of the Pell problem, and still further attempting to extend the anthyphairetic periodicity to space and the canonical solids. These purposes were hitherto unsuspected, as can be seen by opinions expressed by authors who have studied Book X, such as:

van der Waerden (1950, p 172), who writes:

> "Book X does not make easy reading. . .The author succeeded admirably in hiding his line of thought".

Mueller (1981, pp. 270–271), who writes:

> "One would, of course, prefer an explanation that invoked a clear mathematical goal intelligible to us in terms of our own notions of mathematics and which, under analysis, would lead univocally to the reasoning in book X. Unfortunately, book X has never been explicated successfully in this way nor does it appear amenable to explication of this sort. Rather, book X appears to be expedient for dealing with a particular problem and at the same time a mathematical blind alley";

Taisbak (Taisbak 1982, p 58), who must be credited with demystifying the mathematical structure of Book X., writes:

> "We are prepared to face the possibility that there was no other point than to entertain us with good logic";

Knorr (Knorr 1983, p 60), who writes:

> "The true merit of Book X, and I believe it is no small one, lies in its being a unique specimen of a fully elaborated deductive system of the sort that ancient philosophies consistently prized";

The first two purposes (palindromic periodicity, Pell) show that these and such statements must be rejected.

On the other hand, Weil thinks that its purpose is not to prove the strong palindromic periodicity theorem, but *only* to be useful for Book XIII. But Weil's thesis does not take into account the imitation by Plato of the Theaetetus' palindromicity Theorem in the *Politicus*, and our reconstructed proof. Weil's thesis is also rejected by the *Scholia X.135, 185* to the *Elements.*

The purposes of Book X are then the following:

(a) The proof of the strong Theaetetus' parindromic periodicity theorem

(tools needed include apotome X.73–78, binomial X.36–41, conjugacies X.112–114, uniqueness X.79–84, and invariance under commensurability X.103)

(b) The solution of the Pell problem for every non-square N

(tools needed include the classification for apotome X.84/85, constructions X.85–90, X.97 square of an apotome, and possibly X.91)

(c) The proof that some key three-dimensional, space constructions, such as the constructions of canonical polyhedra in Book XIII, have some connections with periodicity, and can be considered finitized, such as

(i) The side of the 20hedron is a "minor line" with respect to the diameter of the sphere in which the 20hedron is comprehended, defined in X.76, constructed in Proposition X.94

(ii) The side of the 12hedron is a "double" apotome (hence the need for a definition of the apotome more general than that of the simple apotome). The details of (c) (i) can be found in Negrepontis, Farmaki, Kalisperi, 2021.

10 The Problem of Pell in Hindu Mathematics

In this final section we will examine the nature of the Hindus' contribution to the solution of the Pell problem (A relevant preliminary study can be found in Brokou 2014). This contribution appears to have come in two stages, first by Brahmagupta in the seventh century AD who developed the *bhavana*, a technique of composition or squaring, and secondly by Bhaskara II in the twelfth century who was associated with the *cakravala* method, practiced with the auxiliary technique of *kuttaka* (Accounts of the method have been given by Singh (1983), Srinivasiengar (1988), Bhanu Murthy (1992), Emch et al. (2005), and Bhattacharyya (2009)). However, according to Weil (1984), *kuttaka* is entirely unnecessary, while various global and local shortcuts, while ingenious, nevertheless obscure the basic periodic nature of the *cakravala* method. The pure form of *cakravala* shows strong similarities to our reconstructed Theaetetus' periodicity proof of quadratic anthypairetic periodicity. We want to emphasize that we have no way of knowing if the origin of the methods developed by the Hindu mathematicians are in some way connected with Theaetetus' reconstructed proofs.

10.1 Reception of the Hindus' Solution to the Problem of Pell in Europe

We can imagine the astonishment of the European mathematicians with the circulation of Colebrooke's book in 1817. Pierre de Fermat in 1657 had solved the problem of Pell specifically for $N = 61$ and, with a letter to Frenicle had challenged him to solve, claiming "misleadingly, or rather mischievously" that he chose small numbers "pour ne vous donner pas trop de peine" (so you don't have too much pain) (Weil 1984, p 97), now it turned out that the very case of $N = 61$ was known to Bhaskara II at about 1150 (Colebrooke 1817, pp. 176–178)!

Hankel, a mathematician and historian of Mathematics, (Hankel 1874) would hail *bhavana* and *cakravala* as "the climax of all Hindu science," a method which "is beyond all praise; [and] is certainly the finest thing achieved in the theory of numbers before Lagrange."

> "der Glanzpunkt ihrer gesammten Wissenschaft ... Diese Methode ist über alles Lob erhaben; sie ist sicherlich das Feinste, was in der Zahlentheorie vor Lagrange geleistet worden ist." (Hankel 1874, pp. 200–202)

However, a little later, Tannery (Tannery 1882, p 325) hypothesized that the Hindu indeterminate analysis had Greek origin, and that with the aid of the Greek rules for approximation one could "by simple steps" pass to the Indian method. Weil writes (Weil 1984, p 19):

> Tannery seems to have... entertained the hope that the solution would some day re-appear with the lost books of Diophantus; so far no evidence for this has been forthcoming.

> Equations of the type $x^2 - Ny^2 = 1$ do occur in Diophantus (e.g. in *Dioph*.V.9 and 1 l), but it is a rational solution that is asked for, even when accidentally a solution in integers is obtained (e.g. in *Dioph*.V.9, where N is of the form $m^2 + 1$, giving $y = 2\ m$, $x = 2\ m^2 + 1$).

Van der Waerden (1976) argues in favor of the Greek origin of Brahmagupta's astronomy:

> "The main subject of Brahmagupta's treatise is astronomy. His astronomical system is based upon the epicycle hypothesis, which is a Greek invention. As one of his tools, Brahmagupta uses a table of sines. The Greeks had tables of chords, which can easily be transformed into tables of sines.
> Some eighty years before Brahmagupta, Varaha Mihira (see Selenius 1975, p 173) presented excerpts from five Siddhantas, one of which, the "Romaka-Siddhanta" was based on the solar and lunar theory of Hipparchos, while another Siddhanta is ascribed to "Paulisa the Greek" (Shukla 1954, p 10). Hence the astronomy and the mathematical tools of the Hindus in the time of Brahmagupta and before were largely derived from Greek sources.
> This makes it even more probable that the Hindu methods of solving Pell's equation also go back to Greek sources."

The arguments by Tannery and van der Waerden were not considered convincing.

We saw in Sect. 5 that Weil did not believe that the conjugation Propositions X.112–114 of Book X of the *Elements* were used for the proof of the periodicity of the quadratic continued fractions/anthyphairesis, and that therefore the Hindu contribution should be considered original. Weil further commented on the "experimental," as he thought, nature of the method of the Hindus, without any proofs, marking nevertheless a "no mean achievement."

> *Firstly,* the facts for which we have just given proofs may have been known to the Indians only experimentally; there is nothing to indicate whether they had proofs for them, or even for part of them.
> For the Indians, of course, the effectiveness of the *cakravala* could be no more than an experimental fact, based on their treatment of a great many specific cases, some of them of Considerable complexity and involving (to their delight, no doubt) quite large numbers.
> As we shall see, Fermat was the first to perceive the need for a general proof, and Lagrange the first to publish one.
> Nevertheless, to have developed the *cakravala* and to have applied it successfully to such difficult numerical cases as $N = 61$ or $N = 67$ had been no mean achievement.
> (Weil 1984, p 24)

10.2 Brahmagupta's *Bhavana*/Composition

10.2.1 Brahmagupta's *Bhavana*/Composition in Colebrooke 1817

Brahmagupta introduces the subject of the resolution of the *varga-prakṛti* (rendered as "affected square" by Colebrooke 1817, or "square-nature" by Weil 1984 p 21, to refer to the general Diophantine equation $q^2 = Np^2 + \lambda$, where N: non-square integer > 1, λ integer and (q, p) a pair of desired integer solutions – cf. Raghavan 1997, p 1008) by presenting an interesting rule (*bhavana* in Sanskrit), according to

which two pairs of already known solutions not necessarily of the same λ are composed and as a result a new pair of solutions is produced, having $\lambda\lambda'$ as the constant term:

> RULE §39–40:
> A root [is set down] two-fold: and [another, deduced] from the assumed square multiplied by the multiplier, and increased or diminished by a quantity assumed.
> The product of the first [pair], taken into the multiplier, with the product of the last [pair] added, is a "last" root.
> The sum of the products of oblique multiplication is a "first" root.
> The additive is the product of the like additive or subtractive quantities.
> The roots [so found], divided by the [original] additive or subtractive quantity, are [roots answering] for additive unity. (Colebrooke 1817, p 363)

The correspondence between the terminology of Coleberooke translating Brahmagupta and the one we are going to use is presented below:

N: "multiplier" (gunaka in the Sanskrit text)
q: "last root"
p: "first root"
λ: "additive or subtractive" (ksepa, praksepa, or praksepaka for Brahmagupta)

10.2.2 Brahmadupta's *Bhavana*/Composition in Modern Notation

Definition. If N is a non-square number and (q, p) a dyad of natural numbers, the *Pell number* of (q, p) with respect to N is $\lambda = q^2 - Np^2$

We note that the Pell number of a dyad is either positive or negative.

The Pell number of the dyad (q, p) coincides with the Pell number Pell(γ) of the simple apotome $qb - pa$ if $qb > pa$ or $pa - qb$ if $pa > qb$.

*Definition of the Brahmagupta's general bhavana/composition rule * for non-square N.* If (q, p) is a dyad with N-Pell number $\lambda = q^2 - Np^2$ and (q', p') a dyad with N-Pell number $\lambda' = q'^2 - Np'^2$, then the dyad $(q'' = qq' + Npp', p'' = qp' + pq')$ has N-Pell number $\lambda'' = \lambda\lambda'$. We write $(q, p)*(q', p') = (qq' + Npp', qp' + pq')$.

*Definition of the Brahmagupta's square bhavana/composition rule * for non-square N.* For the special case where $q' = q$, $p' = p$ and $\lambda' = \lambda$, the dyad resulting from the rule takes the form $q'' = q^2 + Np^2, p'' = 2pq$, with N-Pell number $\lambda'' = \lambda^2$. We write $(q, p)*(q, p) = (q^2 + Np^2, 2pq)$.

10.2.3 The Close Connection of Brahmagupta's General *Bhavana*/Composition and Square *Bhavana* Rule with the Composition and Squaring of Simple Apotome Lines in Book X and Especially Proposition X.97 of the *Elements*

It is difficult to see how Brahmagupta came up with the *bhavana* definition

$$(q,p)*(q',p') = (qq' + Npp', qp' + pq').$$

But the definition becomes natural if we consider it as the arithmetical version of the extended to non-squares geometric Proposition X.97 of the *Elements* for simple apotomae lines:

$$(qb - pa)(q'b - p'a) = ((qq' + Npp')b - (qp' + pq')a)b.$$

The *bhavana*/composition rule plays a basic role in Bhaskara's *cakravala* process, as we shall see in the next section.

10.2.4 Some Examples Worked Out by Bragmagupta

[1]. The Pell problem for $N = 83$: to find natural numbers q, p, such that $q^2 - 83p^2 = 1$.
Brahmagupta starts with the dyad (9, 1).
Then he composes the dyad (9, 1) with itself according to square *bhavana* (Sect. 10.2.2), and finds
(9, 1) * (9, 1) = (164, 18), whose Pell number is 164^2–$83 \cdot 18^2 = 4$.
Then, dividing by 4, the dyad (82, 9), has Pell number $82^2 - 83 \cdot 9^2 = 1$, solving the equation $q^2 = 83p^2 + 1$. Therefore, setting $q = 82$, $p = 9$, he obtained $q^2 - 83\mathrm{p}^2 = 1$.

[2]. The Pell problem for $N = 92$: to find natural numbers q, p, such that $q^2 - 92p^2 = 1$.
Normally the initial step would be the dyad (9, 1), since 9 is the greatest number whose square 81 is less than $N = 92$, but, thanks to a shortcut (explained in Sect. 10.2.5 below), since $10^2 - 92 = 8 < 11 = 92$–9^2, Brahmagupta goes to the dyad (10, 1), whose Pell number is 10^2–$92 \cdot 1^2 = 8$.
Brahmagupta composes the dyad (10, 1) with itself according to square *bhavana* (9.2.2), and finds
(10, 1) * (10, 1) =
(192, 20), with Pell number 192^2–$92 \cdot 20^2 = 8^2 = 64$.
Dividing by 8^2 we get 24^2–$92 \cdot (5/2)^2 = 1$.
We have arrived at a dyad with Pell number 1, but this is not an integral, but a fractional, dyad, (24, 5/2) that has Pell number 1.
Brahmagupta again composes the dyad (24, 5/2) with itself according to square *bhavana* (9.2.2), and finds
(24, 5/2)*(24, 5/2) =
$(24^2 + 92 \cdot (5/2)^2, 2 \cdot (5/2) \cdot 24) = (576 + 575, 120) = (1151, 120)$.
Therefore, setting $q = 1151$, $p = 120$, he obtained $q^2 - 92p^2 = 1$.

10.2.5 Some Reductions by Brahmagupta Employing Square *Bhavana*/Composition

If, for some non-square number N, a dyad can be found with Pell number either
-4, or $+4$, or -2, or $+2$, or -1,
then by repeated use of Brahmagupta's square rule,

it is possible to find a dyad with Pell number +1, thus solving the Pell problem.

We will verify the Brahmagupta's reduction rules, by employing Proposition X.97.

[1] *Reduction from Pell(d) = − 4 to Pell(e) = 4, employing Proposition X.97.*

Let $d = pa - qb$ be an integral simple apotome, with $a^2 = Nb^2$, N non-square, and with assumed line b, such that $\text{Pell}(d) = q^2 - p^2N = -4$.

We note that the line d is a fifth apotome (in the sense of the definitions X.84/85 of the *Elements*).

Set f be the line such that $d^2 = fb$.

By Proposition X.97 of the *Elements*, we have,

since $p^2N + q^2 = 2q^2 + 4 = 2(q^2 + 2)$, that

$$d^2 = (pa - qb)(pa - qb) = \left(\left(p^2N + q^2\right)b - 2pqa\right)b$$
$$= \left(2\left(q^2 + 2\right)b - 2pqa\right)b, \text{hence}$$
$$f = 2\left(q^2 + 2\right)b - 2pqa, \text{and}$$
$$\text{Pell}(f) = (\text{Pell}(d))^2 = \left(p^2N + q^2\right)^2 - 4p^2q^2N = \left(p^2N - q^2\right)^2 = 16.$$

Set $e = f/2$.

Then e is an integral simple first apotome, with $\text{Pell}(e) = 4$.

[2] *Reduction from Pell(d) = −2, to Pell(e) = 1, employing Proposition X.97.*

Let $d = pa - qb$ be an integral simple apotome, with $a^2 = Nb^2$, N non-square, and with assumed line b, such that and $\text{Pell}(d) = q^2 - p^2N = -2$.

We note that the line d is a fifth apotome (in the sense of the definitions X.84/85 of the *Elements*).

Set f be the line such that $d^2 = fb$.

By Proposition X.97 of the *Elements*, we have,

since $p^2N + q^2 = 2q^2 + 2 = 2(q^2 + 1)$, that

$$d^2 = (pa - qb)(pa - qb) = \left(\left(p^2N + q^2\right)b - 2pqa\right)b$$
$$= \left(2\left(q^2 + 1\right)b - 2pqa\right)b, \text{hence}$$
$$f = 2\left(q^2 + 1\right)b - 2pqa, \text{and}$$
$$\text{Pell}(f) = (\text{Pell}(d))^2 = \left(p^2N + q^2\right)^2 - 4p^2q^2N = \left(p^2N - q^2\right)^2 = 4.$$

Set $e = f/2$.

Then e is an integral simple first apotome, with $\text{Pell}(e) = 1$.

[3] *Reduction from Pell(d) = 4 to Pell(e) = 1, employing Proposition X.97 and its generalization for different apotomae.*

Let $d = qb - pa$ be an integral simple apotome, with $a^2 = Nb^2$ and with assumed line b such that $\text{Pell}(d) = q^2 - p^2N = 4$.

We note that the line d is a first apotome (in the sense of the definitions X.84/85 of the *Elements*).

Set g be the line such that $d^2 = gb$.

By Proposition X.97 of the *Elements*, we have,

since $p^2N + q^2 = 2q^2 - 4 = 2(q^2 - 2)$, that

$$\begin{aligned} d^2 = (qb - pa)(qb - pa) &= \big(\big(q^2 + p^2N\big)b - 2pqa\big)b \\ &= \big(2\big(q^2 - 2\big)b - 2pqa\big)b, \text{hence} \end{aligned}$$

$$g = 2\big(q^2 - 2\big)b - 2pqa, \text{and}$$

$$\text{Pell}(g) = (\text{Pell}(d))^2 = \big(q^2 + p^2N\big)^2 - 4p^2q^2N = \big(p^2N - q^2\big)^2 = 16.$$

Set $h = g/2$.

Then h is an integral simple first apotome, with $\text{Pell}(h) = 4$.

Employing the generalization of the Proposition X.97 for the apotomae d and h, we have

$$\begin{aligned} dh &= (qb - pa)\big(\big(q^2 - 2\big)b - pqa\big) \\ &= \big(\big(q\big(q^2 - 2\big) + Np{\cdot}pq\big)b - \big(q{\cdot}pq + p\big(q^2 - 2\big)\big)a\big)b = \\ &= \big(\big(q\big(q^2 - 2\big) + \big(q^2 - 4\big)q\big)\big)b - \big(q^2p + p\big(q^2 - 2\big)\big)a\big)b \\ &= \big(q\big(2q^2 - 6\big)b - p\big(2q^2 - 2\big)a\big)b = \\ &= \Big(4q\big(q^2 - 3\big)^2b - 4p\big(q^2 - 1\big)a\Big)b. \end{aligned}$$

$$\begin{aligned} \text{Pell}(dh/b) &= q^2\big(2q^2 - 6\big)^2 - Np^2\big(2q^2 - 2\big)^2 \\ &= q^2\big(2\big(q^2 - 3\big)\big)^2 - \big(q^2 - 4\big)\big(2\big(q^2 - 1\big)\big)^2 = \ldots = 16. \end{aligned}$$

Set $e = q(q^2–3)^2b - p(q^2 - 1)$a; then e is a simple first apotome, with $\text{Pell}(e) = 1$.

The remaining *bhavana* rules are shown in a similar way.

10.2.6 Brahmagupta's Reduction Rules and Palindromicity

Although all the Brahmagupta's Reduction rules follow essentially from Proposition X.97, it is nevertheless not easy to explain the mathematical reason that makes these rules work.

The rules for the dyads with Pell numbers +4, or −2, or 2, or −1 appear to rely on the *palindromic* periodicity, since the first semiperiod determines completely the full period, exactly because of palindromicity.

The rule from a dyad with Pell number -4 to a dyad with Pell number 4 is interesting because a Pell number -4 can occur early in the anthyphairesis of $a^2 = Nb^2$ (e.g., as we shall see below the third step of the anthyphairesis for $N = 61$ has Pell number -4, while the length of the whole period is 11). The argument of the case [1] in Sect. 10.2.5, together with the rule from Pell number 4 to Pell number 1, shows that the whole quadratic anthyphairesis is determined by the time we find a dyad with Pell number -4.

Thus the Brahmagupta reduction rules, for a non-square number N, rely on the palindromic periodicity of the anthyphairesis of a to b, with $a^2 = Nb^2$.

10.3 The Unknown Origins of *Cakravala*

The greatest known contribution to the *cakravala* method is due to Bhaskara II, a Hindu mathematician of the twelfth century AD. But he is not the originitator of this method.

Bhaskara II in his *Lilavati* and *Bija-ganita* (1150 AD) ascribed the solution of *varga-prakrti* to earlier writers (Selenius 1975, p 168).

Indeed, the "principle of composition" or *bhavana* goes back to Brahmagupta's *Brahmasphuta-siddhanta* (628 AD) (Colebrooke 1817, p 363). As for the *cakravala* method, its first known occurrence is found in the work of *Acarya Jayadeva*, who, according to Selenius (1975, p 168):

> "lived about 1000 or earlier, at least 100 years before *Bhaskara II.* His work is quoted and explained with illustrations in the Sundari [Shukla 1954, p.1-4], written in 1073 by Srimad Udayadivakara as a commentary on the *Laghu-Bhaskariya* of Bhaskara I (629)."

However, Bhaskara II does not explicitly mention Jayadeva (Dutta 2010, p 156), who, was anyway an "otherwise unknown author," as pointed out by Weil (1984, p 22). Thus, the originator of such a momentous mathematical discovery, as the *cakravala* method (that in principle is powerful enough to solve the general Pell problem), is considered as not yet known:

> With this, one is still far short of the general solution; since Colebrooke it has been known that such a one is to be found in the work of Bhaskara (twelfth century; Col. pp.170–184; cf. Datta and Singh, 1oc.cit. pp. 161–172), but an
> almost identical presentation has now been found, attributed to an otherwise unknown author Jayadeva, in a commentary of the eleventh century (cf. K. S. Shukla, Ganita 5 (1954), pp. 1–20). Both authors describe the method as "the cyclic process" (cakravala, from cakra, a wheel). Its true *originator remains unknown*. (Weil 1984, p 22)

Dutta (2010), agreeing with Weil that the originator of *cakravala* is not known, writes on the subject:

> The discovery [of cakravala] took place sometime during the 7th−11th century... However, the true originator is not known. (p 148)

> The *Bijaganita* of Bhäskara II contains – apart from basic results of algebra - a lucid exposition of the *kuttaka, bhavana* and the *cakravala,* and mentions several other indeterminate equations. Bhäskara II attributes the *cakravala* to ancient authors without specifying any name. At the end of his treatise, Bhäskara II makes a general acknowledgement of the prolific algebra works of Brahmagupta, Sridhara and Padmanabha as his chief sources. There is therefore a strong possibility that the lost treatises of Sridhara and/or Padmanabha too contain an account of the *cakravala* algorithm. The history of cakravala became more mysterious with the discovery (in 1954) of a work dated 1073 CE which quotes verses from a preceding algebraist Jayadeva describing the cakravala algorithm (along with the bhavana). Curiously, Jayadeva is not mentioned by Bhäskara. (p 156)
> In 1954, K.S. Shukla discovered the text *Sundari* (1073 CE) of Udaydivakara, a commentary on the *Laghu-Bhaskariya* of Bhaskara and found that it contains quotations from a hitherto unknown mathematician called Jayadeva on the *bhavana* and the *cakravala.* He published the verses in [Shul]. (p 159)

10.4 A Simple Key Proposition, Deduced from the Reconstructed Proof of the Theaetetus' Periodicity Theorem, Shows Its Connection with *Cakravala*

Before examining the *cakravala* method we will state and prove a simple proposition that follows from the reconstructed anthyphairetic proof of the Theaetetus' periodicity theorem (Sect. 5). It will be the key to clarify the link between the reconstructed Theaetetus' proof and the Bhaskara's II approach.

Proposition Let N be a non-square number, $a^2 = Nb^2$, and
γ_n, I_n, λ_n, m_n as in Sect. 5. Then
[1] $\lambda_n + m_n > m_1$, hence $(\lambda_n + m_n)^2 > N$;
[2] if $(m_n + \lambda_n)^2 - N < N - {m_n}^2$, then $I_n = 1$, and hence $m_{n+1} = \lambda_{n+1} - m_n$.

Proof [1] Since $m_1 \geq m_{n+1}$,
we have $m_n + m_1 \geq m_n + m_{n+1} = I_n \lambda_{n+1} \geq \lambda_{n+1}$,
[using *Theaetetus' conjugation* (Sect. 5),
or the weaker statement that $m_n + m_{n+1} = a$ multiple of λ_{n+1}]
hence $\lambda_n\lambda_{n+1} > \lambda_n\lambda_{n+1} - \lambda_2 = (N - {m_n}^2) - (N - {m_1}^2) = {m_1}^2 - {m_n}^2 = (m_1 - m_n)(m_1 + m_n) \geq (m_1 - m_n)\lambda_{n+1}$,
hence $\lambda_n > m_1 - m_n$.
[2] We have $m_{n+1} + m_n = I_n \lambda_{n+1}$
[again employing *Theaetetean conjugation*]
$(\lambda_n + m_n) + m_n < 2\ \lambda_{n+1}$.
[Elementary computation employing $N - {m_n}^2 = \lambda_n\lambda_{n+1}$.
By the arithmetical form of Proposition II.4 of the *Elements*,
it follows from the assumption that

$$(m_n + \lambda_n)^2 - N = {m_n}^2 + {\lambda_n}^2 + 2\, m_n\lambda_n - N < N - {m_n}^2, \text{hence}$$

$$\lambda_n(\lambda_n + 2m_n) = {\lambda_n}^2 + 2\, m_n\, \lambda_n < 2\left(N - {m_n}^2\right) = 2\, \lambda_n\, \lambda_{n+1}, \text{hence}$$

$$\lambda_n + 2m_n < 2\, \lambda_{n+1}, \text{hence}$$

$$(\lambda_n + m_n) + m_n < 2\, \lambda_{n+1}.]$$

We now employ [1]: $m_{n+1} \leq m_1 < m_{n+1} + \lambda_{n+1}$
to obtain

$$m_{n+1} + m_n < 2\, \lambda_{n+1}.$$

It now follows that

$$I_n = 1, \text{and } m_{n+1} = \lambda_{n+1} - m_n.$$

Corollary 1 If $(m_1 + 1)^2 - N < N - {m_1}^2$,
then $I_1 = 1$, $\gamma_2 = b - \gamma_1 = b - (a - m_1 b) = (m_1 + 1)b - a$ and $m_2 = \lambda_2 - m_1$.

Corollary 2 (*characterization of* m_{n+1}). m_{n+1} is the greatest natural number m, such that $m \leq m_1$ and m is of the form $-m_n + k{\cdot}\lambda_{n+1}$ for some natural number k.

10.5 The Initial Step of the *Cakravala* Process

Weil 1984, pp 23–24: cakravala for non-square N invariably starts with a dyad (q, 1) such that q^2 is the closest square to N below or above. Thus in effect cakravala starts with the first or the second anthyphairetic remainder of the anthyphairesis of a to b, $a^2 = Nb^2$.

It is very important to make sure that the initial step of the *cakravala* method for a non-square number N always corresponds to the first remainder of the anthyphairesis of a to b, with $a^2 = Nb^2$.

10.5.1 Weil: The Initial Step of *Cakravala* Is Invariably Anthyphairetic

This calls for several observations.
"Secondly, in order to carry out the *cakravala* [for the non-square number N], one must have a starting point; *invariably* they [the Hindus] choose as such the triple $[(q, p_1 = 1; \lambda_2)]$ for which $[q^2]$ is the closest square to N, above or below.
it is easy to see that $[\lambda_2]$ is then $<2\sqrt{N}$. (Weil 1984, p 23)

[the words and symbols in brackets are added by the authors]

10.5.2 The Starting Dyad of the *Cakravala* Method

We set $a^2 = Nb^2$ and denote by γ_n the nth anthyphairetic remainder of the anthyphairesis of a to b.

It is clear that if q^2 is the closest square to N *below*, then $a - qb = \gamma_1$

(e.g., $N = 67$, $a - 8b = \gamma_1$).

But in case q^2 is the closest square to N *above*, then Bhaskara II takes the dyad $(q, 1)$, (e.g., $N = 61$, (8,1)). This appears to be a violation of the classical anthyphairesis (since, e.g., for $N = 61$, $a - 7b = \gamma_1$). The natural way would be to start with the pair (7, 1), whose Pell number is $7^2–61\cdot1^2 = -12$. However, Bhaskara II had a shortcut that allowed him to go automatically to the next step of the anthyphairesis, by noting that 61 is closer to 8^2 than to 7^2: $8^2–61 = 3 < 12 = 61–7^2$. In such a case the shortcut, described and shown in Sect. 10.4, Corollary 1, allows them to proceed immediately to the second stage with the dyad (8, 1). But in fact there is no violation of the classical anthyphairesis, only a shortcut. Thus, if q^2 is the closest square to N *above* then $qb - a = \gamma_2$, and the Bhaskara II dyad $(q, 1)$ is in fact anthyphairetic.

In order to solve the Pell problem for $N = 61$, Bhaskara II starts with the dyad (8, 1), which has positive Pell number $8^2 - 61\cdot1^2 = 3$.

If N is closest to the square number below, the first step of the anthypairesis is employed

($N = 67$ with initial triple (8, 1, −3), where $a^2 = 67b^2$, $\gamma_1 = 8a - b$)),

if N is closest to the square number above, the second

($N = 61$, with initial triple (8, 1, 3), where $a^2 = 61b^2$, $\gamma_1 = 7a - b$, $\gamma_2 = 8b - a$).

Thus Bhaskara II is careful to start with a triple corresponding to an anthyphairetic pair.

10.6 The Third Step of the Bhaskara's II *Cakravala* Process for *N* = 61 and Its Relation to the Reconstructed Theaetetus' Method

Before examining the general step of the *cakravala* method, it will be instructive to consider the third step of the *cakravala* method by Bhaskara II for the case of $N = 61$.

10.6.1 The Third Step of the Bhaskara's II Cakravala Process for N = 61

According to the Bhaskara's II *cakravala* process, we compose, with the Brahmagupta's *bhavana* process, the dyad (8, 1), with $\lambda = 8^2–61\cdot1^2 = 3$, with the dyad $(m_3/3, 1/3)$, where m_3 is to be determined by the *kuttaka* method:

$(8, 1)*(m_3/3, 1/3) = ((61 + 8\, m_3)/3, (m_3 + 8)/3)$.

The Diophantine equation for the "smallest root" is $3P = M + 8$.

Bhaskara II employs *kuttaka* method, Bachet-Bezout method (Bachet (1612)), based on the Euclidean algorithm Propositions VII.1 & 2 of *Elements*, an application that does not appear explicitly in Greek Mathematics, in order to find at least one integer solution. The *kuttaka* method is explained below in Sect. 10.7.2.

One solution by inspection, omitting *kuttaka*, is $M_0 = 1$, $P_0 = 3$.

Then every integral solution of the equation has the form:

$$M = 3k + 1,$$

$$[P = k + 3]$$

where $k = 0, \pm 1, \pm 2, \ldots$.

We find m_3 to be the solution m the makes $N - m^2$ minimum, the m such that $N > m^2$ and m greatest. It is clear that $k = 2$, $m_3 = 7$ is the required solution.

We find $p_3 = k + 3 = 2 + 3 = 5$, and $q_3 = 39$. Then Pell(39, 5) $= -4$

Now Bhaskara II follows the old *bhavana* rules by Brahmagupta, going from the dyad with Pel number -4 to a dyad with Pell number +4, and then further to the dyad ($q = 1\ 766\ 319\ 049$, $p = 226\ 153\ 980$) with Pell number +1.

By *bhavana*/composition of this pair with itself ad infinitum, we get an infinite multitude of pairs with Pell number 1.

10.6.2 The First Three Steps of the Reconstructed (in Sect. 5) Theaetetus' Method for $a^2 = 61b^2$

Computation of γ_1: $m_1 = 7$, $\gamma_1 = a - 7b = \varphi_1$

$$\text{Pell}(\gamma_1) = -12 = -\lambda_2$$

Computation of φ_2 by means of the conjugation form of anthyphairesis, $\varphi_1(I_1 b + \varphi_2) = b^2$.

We find that $\varphi_1(a + 7b) = a^2 - 7^2 b^2 = 61b^2 - 49b^2 = 12b^2$, hence

$$I_1 b + \varphi_2 = (a + 7b)/12 = b + (a - 5b)/12.$$

Thus $\varphi_2 = (a - 5\text{b})/12$.

$$\text{Pell}(\varphi_2) = -36/12 \cdot 12 = -1/4 = -\lambda_3/\lambda_2.$$

Computation of γ_2 by means of the formula $\gamma_1\ \varphi_2 = \gamma_2 b$:

$$\gamma_1\ \varphi_2 = (a - 7b)\ (a - 5b)/12 = \left(a^2 + 35b^2 - (12ab)\right)/12 = ((96b - 12a)/12)b = (8b - a)b.$$

Thus $\gamma_2 = 8b - a$.

$$\text{Pell}(\gamma_2) = \lambda_3 = 3.$$

Computation of φ_3 by means of the conjugation form of anthyphairesis, $\varphi_2(I_2 b + \varphi_3) = b^2$.

We find that $\varphi_2(a + 5b) = (61b^2 – 25b^2)/12 = 3b^2$, hence

$$I_2 b + \varphi_3 = (a + 5b)/3 = 4b + (a - 7b)/3.$$

Thus $\varphi_3 = (a - 7b)/3$.

$$\text{Pell}\,(\varphi_3) = -4/3.$$

Computation of γ_3 by means of the formula $\gamma_2\,\varphi_3 = \gamma_3 b$:

$$\gamma_2\,\varphi_3 = (8b - a)\,(a - 7b)/3 = \left(15ab - \left(a^2 + 56b^2\right)\right)/3 = ((15a - 117b)/3)b = (5a - 39b)b.$$

Thus $\gamma_3 = 5a - 39b$.

$$\text{Pell}\,(\gamma_3) = -4.$$

The close connection of the two approaches is evident.

10.7 The General *Bhaskara II Cakravala* Step

10.7.1 Description of the *Cakravala* Process by Bhaskara II/Colebrooke

Colebrooke (Colebrooke 1817, p. 175) quotes Bhaskara II in the extract below that describes the *cakravala* process:

> §83–86 Rule for the cyclic method:
> Making the "least" and "greatest" roots and additive, a dividend, additive and divisor,
> let the multiplier be thence found.
> The square of that multiplier being subtracted from the given coefficient, or this coefficient
> being subtracted from that square, (so as the remainder be small;)
> the remainder, divided by the original additive, is a new additive;
> which is reversed if the subtraction be [of the square] from the coefficient.
> The quotient corresponding to the multiplier [and found with it] will be the least root:
> whence the "greatest" root may be deduced.
> With these, the operation is repeated, setting aside the former roots and additive.
> This method mathematicians call that of *the circle* (Colebrooke 1817, p 175)

10.7.2 The General *Cakravala* Process with *Kuttaka* by Bhaskara II and Its Relation to the Reconstructed Theaetetus' Method

The starting point is a (q, p) N-dyad, with absolute value of its Pell number $q^2 - Np^2$ the natural number λ. As we have seen the initial choice of this N-dyad is not a random one, but an anthyphairetic, corresponding to the anthyphairetic remainder γ_1 or γ_2 of the anthyphairesis of a to b, where $a^2 = Nb^2$. We will prove that the resulting N-dyad is the one corresponding to the next anthyphairetic remainder. Thus we assume that $q = q_n, p = p_n$, where $\gamma_n = (-1)^n(q_n b - p_n a)$. In such a case $\lambda = \lambda_{n+1}$, the absolute value of

$$q_n{}^2 - Np_n{}^2.$$

The basic step of the Bhaskara's II *cakravala* process is a Brahmagupta-type composition with the dyad $(m/\lambda_{n+1}, 1/\lambda_{n+1})$, as follows:

$$(q_n, p_n)^* (m/\lambda_{n+1}, 1/\lambda_{n+1}) = ((mq_n + Np_n)/\lambda_{n+1}, (q_n + mp_n)/\lambda_{n+1}),$$

where m is a natural number to *be determined*.

The number m is determined by solving the Diophantine equation for the "smallest root," namely

$$\lambda_{n+1} P = Mp_n + q_n.$$

Bhaskara II employs first the *kuttaka* method, in order to find at least one integer solution, say P_0, M_0, so that we have $\lambda_{n+1} P_0 = M_0 p_n + q_n$. As already mentioned, the *kuttaka* method is based on the Euclidean algorithm (Propositions VII.1 & 2 of *Elements*), and constitutes an ingenious application that does not appear explicitly in Greek mathematics and is equivalent to the Bachet-Bezout method.

Then

$$\lambda_{n+1}(P - P_0) = (M - M_0)p_n$$

and employing Proposition VII.19 of Euclid's *Elements* we obtain

$$(P - P_0)/(M - M_0) = p_n/\lambda_{n+1}.$$

Since, the generalized side and diameter numbers p_n, q_n are relatively prime, it immediately follows that the numbers p_n, λ_{n+1} are relatively prime, and hence it follows, from the fundamental Proposition VII.22 of Euclid's *Elements*, that

$$P - P_0 = k\, p_n,$$

$$M - M_0 = k\, \lambda_{n+1}$$

for k any integer. Thus the general solution of the Diophantine equation is

$$P = P_0 + k\, p_n,$$

$$M = M_0 + k\, \lambda_{n+1}$$

for $k = 0, \pm 1, \pm 2, \ldots$.

Now let m be the *greatest* number that satisfies

$N - m^2 > 0$, $\lambda_{n+1}\, p = mp_n + q_n$.

(for simplicity we avoid the absolute values, a "local" shortcut employed by Bhaskara II, as Weil points out)

> ("Thirdly, the Indian prescription for choosing $\boldsymbol{x}$ within its congruence class modulo m is not quite what we have indicated above, since the rule is to make $N - \boldsymbol{x}^2$ 'small' (i.e., in actual practice, as small as possible), but, as the context shows, *in absolute value*; in other words,

one should substitute $\boldsymbol{x}_1 = \boldsymbol{x} + |\boldsymbol{m}|$ for x if ${x_1}^2 - N$ happens to be less than $N - x^2$. It can be shown that this *has merely the effect of abbreviating the procedure somewhat* when that is the case." (Weil 1984, p 24)).

But we now see, by Proposition of Sect. 10.5, that the number m satisfies

$$N - m^2 > 0,$$

$$\lambda_{n+1} p = m\, p_n + q_n,$$

but the immediately greater natural number m satisfying $\lambda_{n+1} p = m p_n + q_n$, namely the number $m + \lambda_{n+1}$, fails to satisfy $N - (m + \lambda_{n+1})^2 > 0$.

This proves that the *cakravala* number m coincides with the anthyphairetic number m_{n+1}, and thus the N-dyad produced by the *cakravala* process is precisely the dyad (p_{n+1}, q_{n+1}), corresponding to the anthyphairetic remainder γ_{n+1}.

10.7.3 *Cakravala*, Without *Kuttaka*, According to Weil 1984

But Weil shows that the *kuttaka* method is not an essential tool of the *cakravala* method, and it can be avoided (This was in fact shown earlier by Krishnaswamy Ayyangar (Krishnaswamy 1929–30).).

Weil: The Knowledge of the Rules

$m_{n+1} + m_n =$ multiple of λ_{n+1} and $m_{n+1} \leq m_1 < m_{n+1} + \lambda_{n+1}$,
which are used anyway in Bhaskara's II kuttaka approach to cakravala, makes kuttaka unnecessary

That being so, the *kuttaka* is of course *not needed* at the first step, since $q_0 = 1$;
but it is also *not needed* in any further operation,
since, if notations are as above, we have $p' - q'x = -qm'$,
so that at the next step one has to take $x' = -x \pmod{m'}$.
Strangely enough, this does not seem to have been noticed by any of our Indian authors (nor even by their later commentators, down to the sixteenth century); they make no mention of it, and invariably refer to the *kuttaka* for the choice of x, even though their abundant numerical evidence could easily have convinced them that *this was unnecessary.*
(Weil 1984, pp. 23–24)

Kuttaka Is Not Necessary for the *Cakravala* Process, According to Weil

We provide below the details of Weil's simple proof that the use of *kuttaka* is not necessary for the determination of the crucial number m_{n+1}. Both the *kuttaka* method and Weil's argument rely on the formulae $m_n + m_{n-1} = T_n \lambda_n$, a multiple of λ_n (1), and $m_n \leq m_1 < m_n + \lambda_n$ (2), which were derived in Sect. 5 and in Sect. 10.4.

Let N be a non-square number.

As we have seen the method of *cakravala* starts "invariably" with

$$\lambda_1 = 1, m_1, \text{such that}, {m_1}^2 < N < (m_1 + 1)^2, \text{and } q_1 = m_1, p_1 = 1.$$

[Consider above remarks, for the seemingly different choice e.g. for $N = 61$]

Inductively assume that for 1,2,. . .,*n*, λ_n, m_n, q_n p_n are natural numbers such that

$$m_n + m_{n-1} = T_n\lambda_n, \text{a multiple of } \lambda_n \ (1),$$

$$m_n \leq m_1 < m_n + \lambda_n \ (2),$$

$$\lambda_n q_n = m_n q_{n-1} + Np_{n-1}, \lambda_n p_n = q_{n-1} + m_n p_{n-1} \ (3).$$

$$\lambda_n \leq 2\, m_1 \ (4),$$

Defiition of λ_{n+1}. λ_{n+1} is the absolute value of the Pell number ${q_n}^2 - N{p_n}^2$ of the dyad (q_n, p_n).

Claim 1 $N - {m_n}^2 = \lambda_n\lambda_{n+1}$.[*Proof.* (q_{n-1}, p_{n-1})*$(m_n/\lambda_n, 1/\lambda_n) = (q_n, p_n)$.
By the multiplicative property of Pell numbers we have
$\lambda_n \cdot (N - {m_n}^2)/{\lambda_n}^2 = \lambda_{n+1}$, hence $N - {m_n}^2 = \lambda_n\lambda_{n+1}$].

Definition of m_{n+1}. We define m_{n+1} *the unique number* such that

$$m_{n+1} + m_n = T_{n+1}\lambda_{n+1}, \text{and}$$

$$m_{n+1} \leq m_1 < m_{n+1} + \lambda_{n+1}.$$

[With this definition of m_{n+1}, *kuttaka* is avoided, and according to Proposition of Sect. 10.4 this definition coincides with the dyad q_{n+1}, p_{n+1} in our reconstructed Theaetetus' proof of periodicity (Sect. 5)]

Definition of q_{n+1}, p_{n+1}. Define q_{n+1}, p_{n+1} by

$$\lambda_{n+1}q_{n+1} = m_{n+1}q_n + Np_n, \text{and}$$

$$\lambda_{n+1}p_{n+1} = q_n + m_{n+1}p_n.$$

[It is clear that this definition, on the one hand coincides with the Brahmagupta's *bhavana*

$$(q_n, p_n)*(m_{n+1}/\lambda_{n+1}, 1/\lambda_{n+1}) = \left(q_{n+1}, p_{n+1}\right), \text{and}$$

on the other hand coincides with the reconstructed Theaetetus' proof of periodicity (Sect. 5).

$$\gamma_n\varphi_{n+1} = (q_n b - p_n a)((a - m_{n+1}b)/\lambda_{n+1}) =$$

$$[((q_n + m_{n+1}\, p_n)a - (m_{n+1}\, q_n + Np_n)b)/\lambda_{n+1})b$$

$$= \left(\left(p_{n+1}a - q_{n+1}b\right)/\lambda_{n+1}\right)b = \gamma_{n+1}b].$$

Claim 2 q_{n+1}, p_{n+1} are natural numbers.

[Proof. $\lambda_{n+1}\, p_{n+1} = q_n + m_{n+1} p_n$

$$= (m_n q_{n-1} + N p_{n-1})/\lambda_n + m_{n+1}\, (q_{n-1} + m_n p_{n-1})/\lambda_n$$

$$= (m_n + m_{n+1}) q_{n-1}/\lambda_n + (N + m_{n+1}\, m_n) p_{n-1}/\lambda_n =$$

$$= T_{n+1}\, \lambda_{n+1}.q_{n-1}/\lambda_n + \left(N - {m_n}^2 + {m_n}^2 + m_{n+1}\, m_n\right) p_{n-1}/\lambda_n =$$

$$= T_{n+1}\, \lambda_{n+1}.q_{n-1}/\lambda_n + (\lambda_n\, \lambda_{n+1} + m_n\, (T_{n+1}\, \lambda_{n+1}))\, p_{\mathrm{n}-1}/\lambda_n$$

$$= \lambda_{n+1}\, (T_{n+1}\, q_{n-1}/\lambda_n + (\lambda_n + m_n\, T_{n+1})\, p_{n-1}/\lambda_n).$$

Hence, $p_{n+1} = T_{n+1}\, q_{n-1}/\lambda_n + (\lambda_n + m_n\, T_{n+1})\, p_{n-1}/\lambda_n$,
a natural number].
Similarly for q_{n+1}.

Claim 3 $\lambda_2 \leq 2\, m_1$.

[*Proof.* $N < (m_1 + 1)^2 = {m_1}^2 + 2\, m_1 + 1$,
hence $\lambda_2 = N - {m_1}^2 < 2\, m_1 + 1$,
hence $\lambda_2 \leq 2\, m_1$.]

Claim 4 $N - {m_n}^2 \leq 2\lambda_n m_1$.

[Proof. $N - {m_n}^2 < (m_1 + 1)^2 - {m_n}^2 = {m_1}^2 - {m_n}^2 + 2\, m_1 + 1$,

hence

$$N - {m_n}^2 \leq (m_1 + 1)^2 - {m_n}^2 - 1 = {m_1}^2 - {m_n}^2 + 2\, m_1.$$

By (2) $m_1 < m_n + \lambda_n$, hence $m_1 - m_n < \lambda_n$, hence $m_1 - m_n \leq \lambda_n - 1$.
Thus, by Claim 3,

$$N - {m_n}^2 \leq {m_1}^2 - {m_n}^2 + 2\, m_1 = (m_1 - m_n).(m_1 + m_n) + 2m_1 \leq$$

$$\leq (\lambda_n - 1)\, 2m_1 + 2m_1 = 2\lambda_n m_1].$$

Claim 5 $\lambda_{n+1} \leq 2m_1$.

[Proof. By Claims 1, 4, $\lambda_n\, \lambda_{n+1} = N - {m_n}^2 \leq 2\lambda_n\, m_1$,
hence $\lambda_n\, \lambda_{n+1} \leq 2\, \lambda_{n+1}\, m_1$, hence $\lambda_{n+1} \leq 2m_1$].

10.8 On the Cyclical Meaning of the Term *Cakravala*

The *cakravala* of Jayadeva and Bhaskara II is a general method for solving the *varga-prakrti*, the Pell problem. Why is it called by such a term implying *circularity,*

periodicity? Jayadeva or Bhaskara II do not give a reason for this terminology. Modern scholars have suggested that the method is called *cakravala* because of its *iterative* character:

> **"***Bhāskara II* provided *chakravāla* or *cyclic* process for the solution. *Chakravāla* means '*circle*'. It is observed that the same set of operations being applied again and again in a *continuous round* so the method is called *chakravāla*." (Datta and Singh 1938, p 162).

But this cannot be accepted as a satisfactory justification. For example, infinite anthyphairesis has an iterative character, but will never be described as cyclical unless it is periodic, a condition far stronger than infinite. In the same manner, the approximation of the area of a circle by say inscribed canonical polygons with number of vertices that doubles at every stage is clearly an iterative procedure, repeating itself ad infinitum, that would never be described as circular.

Weil suggests that the shortcuts employed by the Hindus, such as Brahmagupta's *bhavana*/composition rules (or those minimizing the absolute value of the distance of m^2 to N, we might add), ingenuous as they certainly are, tend to obscure the underlying periodicity; and that the periodicity to which *cakravala* refers is the anthyphairetic/continued fraction periodicity:

> Finally, we are told to iterate the process only till we find an "additive"
> m with one of the values $\pm 1, \pm 2, \pm 4$, and then to make use of the *bhavana*,
> i.e., of *Brahmagupta*'s procedure for that case.
> Actually this is no more than a shortcut,
> since it can be shown that the *cakravala*, applied in a straightforward manner,
> would inevitably lead to a triple (***p, q; 1***) as desired;
> while this shortcut is quite effective from the point of view of the numerical solution,
> it destroys the *"cyclic"* character of the method, which otherwise would appear from the fact
> that the "additives" ***m***, ***m', m''*** . . . would repeat themselves *periodically*,
> corresponding to the *periodicity* of the continued fraction of $\sqrt{N}$. (Weil 1984, p 24)

10.9 The *Cakravala* Method Appears to Have Similarities with the Reconstructed Theaetetus' Conjugacy Method

Weil has succeeded in presenting the *cakravala* method in pure form, stripped of the mathematically unnecessary *kuttaka* (Sect. 10.7.3), and unencumbered by ingenious shortcuts (that are in fact consequences of the reconstructed Theaetetus' conjugation method) that nevertheless obscure the periodic nature of *cakravala*.

The *cakravala* method in its pure form reveals close similarities with the reconstructed Theaetetus' conjugacy method presented in Sect. 5. Indeed, for a non-square number N, the *cakravala* process:

(a) Invariably starts with an arithmetized version ($q_1 = m_1, p_1 = 1$), [or else if N is nearer to $(m_1 + 1)^2$, with ($q_2 = m_1 + 1, p_2 = 1$)],

of the initial anthyphairetic remainder(s) $\gamma_1 = a - m_1 b$, or $\gamma_2 = (m_1 + 1)b - a$, respectively,

of the anthyphairesis of a to b, for $a^2 = Nb^2$ (Sect. 10.5).

(b) Proceeds inductively from the given dyad (q_n, p_n), whose Pell number ${q_n}^2 - N{p_n}^2$ has absolute value equal to λ_{n+1}, to the next dyad (q_{n+1}, p_{n+1}), with a Brahmagupta-type composition of the given dyad with the dyad $(m_{n+1}/\lambda_{n+1}, 1/\lambda_{n+1})$:

$$(q_n, p_n) * (m_{n+1}/\lambda_{n+1}, 1/\lambda_{n+1}) = \left(q_{n+1}, p_{n+1}\right),$$

where m_{n+1} is the greatest number satisfying the rule $m_n + m_{n+1} =$ multiple of λ_{n+1} and $m_{n+1} \leq m_1$.

This procedure appears close to the arithmetized version of the geometric rule $\gamma_n \varphi_{n+1} = \gamma_{n+1} b$, namely

$$(-1)^n (q_n b - p_n a)\, \varphi_{n+1} = (-1)^{n+1} \left(q_{n+1} b - p_{n+1} a\right) b,$$

where φ_{n+1} has been determined by the conjugation relation $\varphi_n (I_n b + \varphi_{n+1}) = b^2$, into which the anthyphairetic relation $\gamma_{n-1} = I_n \gamma_n + \gamma_{n+1}$ has been transformed and thus shown by conjugacy to have the form $\varphi_{n+1} = (a - m_{n+1}\, b)/\lambda_{n+1}$, where m_{n+1} is determined by the rule $m_n + m_{n+1} = I_n \lambda_{n+1}$, according to the Theaetetus' conjugacy method, as reconstructed in Sect. 5 (cf. Sects. 10.6, 10.7).

(c) Employs the term "*cakravala*," unmistakably indicating periodicity (and not just iteration).

This makes sense only as referring to the periodicity of the geometric anthyphairesis of a to b, for $a^2 = Nb^2$, and non-square number N. (Sect. 10.8)

11 Summary

In Sect. 2, the concept of geometric anthyphairesis is introduced as well as the consequent pre-Eudoxian, in fact Theaetetean, anthyphairetic definition of equality of ratios for magnitudes. The Logos Criterion for the periodicity of an anthyphairesis is also presented, as a natural consequence of the pre-mentioned definition.

In Sect. 3, the anthyphairetic base of the platonic philosophy is highlighted. Through the examples of the One of the second hypothesis in the *Parmenides* and the Divisions of the Angler and Sophist in the *Sophistes*, the following argument is supported: the basic entity in Plato's philosophical system, the intelligible Being, is a philosophical imitation of a dyad in periodic anthyphairesis.

In Sect. 4, the influence of Theaetetus' mathematical discoveries on Plato's later work is outlined. This influence is pointed out by Plato himself ("*πειρῶ μιμούμενος*"), in the *Theaetetus* 148d4-5. The philosophical statements closely

imitating Theaetetus' periodicity theorem are detected in both the *Politicus* 283a-287b (according to which intelligible Beings are generated by a double measurement: finite measurement and infinite generation) and the *Philebus* 16c, 23a-25e (where an intelligible Being is presented as the mixture of an anthyphairetic infinite/incommensurable and a finite/commensurable). It should be noted that Theaetetus' periodicity theorem, which is traced in semi-philosophical language in the *Theaetetus* 147c-148d, is the following: "if a, b are lines commensurable in power only, in particular if $a^2 = Nb^2$ for some non-square number N, then the anthyphairesis of a to b is periodic."

In Sect. 5, the reconstruction of the proof of Theaetetus' periodicity theorem takes place, by using tools found in the Theaetetean Book X of the *Elements* and more specifically the apotomae (Propositions X.73–107), the binomial lines (Propositions X.36–72), and their conjugation (Propositions X.112–114). The reconstruction is based on the realization that the anthyphairetic remainders γ_n of the anthyphairesis of a to b, for $a^2 = Nb^2$, N non-square, are simple apotome lines of the form $\gamma_n = (-1)^n(q_n b - p_n a)$, and that the anthyphairetic relations of the remainders ($\gamma_{n-1} = I_n\gamma_n + \gamma_{n+1}$) can be transformed into relations of conjugation $\varphi_n(I_n b + \varphi_{n+1}) = b^2$, with the necessary introduction of the reduced remainders φ_n defined by $\gamma_n\, \varphi_{n+1} = \gamma_{n+1}\, b$. It is noted that for the reconstruction, the *pigeonhole principle* is additionally needed, which is provided by Plato in the *Politicus* (272d-e).

In Sect. 6, it is shown that the definition of the Statesman in the dialogue *Politicus* is an imitation not only of periodic but *palindromically periodic* anthyphairesis. It is argued that Theaetetus had established the stronger palindromic periodicity theorem: "if $a^2 = Nb^2$, N non-square, then the anthyphairesis of a to b is palindromically periodic," a result that in modern era was gradually discovered by the work of Fermat, Brouncker, Lagrange, and Euler (seventeenth and eighteenth century).

In Sect. 7, the reconstruction of Theaetetus' strong theorem of palindromic periodicity takes place, using tools from the Book X of the *Elements* and the pigeonhole argument implied in the *Politicus*.

In Sect. 8, it is argued that Theaetetus had a strong motivation to deal with the Pell problem, since the case $N = 2$ was a venerated Pythagorean discovery not unnoticed by Plato (*Politeia* 546c), closely connected with their discovery of the periodicity of the anthyphairesis $a^2 = 2b^2$.

In Sect. 9, it is noted that the Pell property is actually included in the reconstructed proof of Theaetetus' palindromic periodicity theorem and thus the proof of the general Pell property (namely for every N non-square) is within easy reach. Hence it is assumed that Theaetetus had a proof for the general Pell property as well. It is argued that Theaetetus' preoccupation with the Pell problem can be indeed traced in Book X since the classification of apotome lines in six classes in X.84/85 makes sense only in relation to the Pell property. Specifically:

(i) The first and the fourth apotome are apotomae with *positive Pell number* (corresponding to the *even remainders* of the anthyphairesis of a to b for $a^2 = Nb^2$, N non-square, and b the assumed "one foot" line) while the second and the fifth with *negative Pell number* (corresponding to the *odd remainders*).

(ii) The first and the second apotome lines are those with *square Pell number*, the fourth and the fifth with *non-square Pell number*.

Moreover, Proposition X.97 produces an apotome with square Pell number from any apotome line.

Based on the above, a multidimensional but clear purpose can now be attributed to the Book X of the *Elements*: firstly, to provide the tools needed for the statement and proof of the palindromic periodicity of the anthyphairesis of quadratic irrationals and the closely related proof of the general Pell problem and secondly, to try to extend the anthyphairetic periodicity to space and the canonical solids.

In Sect. 10, the contribution of the Hindus to the Pell problem is presented and similarities between their methods and the reconstructed proof of Theaetetus are denoted. Specifically:

(i) The *bhavana* method of composition of Brahmagupta bears similarities to the arithmetized version of the geometric rule of "multiplying" simple apotomae or binomial lines, which is closely related to the Proposition X.97 of the *Elements*. *Bhavana* is applied as a rule of reductions for the solution of the Pell problem in cases where the Pell number is equal to -4 or $+4$, or -2 or $+2$, or -1.
(ii) The *cakravala*/circle method, mostly advanced by Bhaskara II, when studied stripped from the interfering (and mathematically unnecessary, as Weil remarks) method of *kuttaka* and from other (ingenious but disorientating) shortcuts, appears as the arithmetized version of the reconstructed conjugacy method used in the proof of the Theaetetus' palindromic periodicity theorem presented in Sect. 5.

Acknowledgments We express our warm thanks to Athanase Papadopoulos for encouragement and extensive and essential discussions that have led to a substantial improvement of our work.

Bibliography

Acerbi F (2000) Plato: Parmenides 149a7-c3. A proof by complete induction? Arch Hist Exact Sci 55:57–76

AVGI (2000) A mathematician throws new light on Plato-professor Stelios Negrepontis reads Plato anew, 30 July, 6 and 13 August 2000

Bachet de Méziriac CG (1612) Problèmes plaisants et délectables qui se font par les nombres, avec leur demonstration, 2nd ed (1624). Pierre Rigaud, Lyon

Bassiakou A (2004) Plato's statesman and the palindromic periodicity of the Anthyphairesis of quadratic irrationals. MSc Thesis, University of Athens [in Greek]

Bhanu Murthy IS (1992) A modern introduction to ancient Indian mathematics. Wiley Eastern, New Delhi

Bhattacharyya RK (2009) Brahmagupta: the ancient Indian mathematician. In: Yadav BS, Mohan M (eds) Ancient Indian leaps into mathematics. Birkhäuser, Boston, MA. https://doi.org/10.1007/978-0-8176-4695-0_12

Bombelli (1572) L'Algebra, Opera di Rafael Bombelli da Bologna, divisa in tre libri. Giovanni Rossi, Bologna
Brokou M (2014) Pell's equation in Greek and Hindu mathematics. MSc Thesis, University of Athens [in Greek]
Burnyeat MF (1978) The philosophical sense in Theaetetus' mathematics. Isis 69:489–513
Colebrooke HT translated (1817) Arithmetic and mensuration, from the Sanscrit of Brahmegupta and Bhascara, John Murray, Albemarle Street, London
Datta B, Singh AN (1938) History of Hindu mathematics, a source book, part II, Algebra. Motilal Banarsidass, Lahore
de Fermat P (1679) Varia opera Mathematica D.Petri de Fermat Senatoris Tolosani. Apud Joannem Pech, juxta collegium PP. Societatis Jesu, Tolosae
Dutta AK (2010) Kuṭṭaka, Bhāvanā and Cakravāla In: Seshadri CS (ed) Studies in the history of Indian mathematics. Culture and history of mathematics, vol 5. Hindustan book agency (India), New Delhi, pp 145–199
Emch GG, Srinivas MD, Sridharan R (eds) (2005) Contributions to the history of Indian mathematics. Culture and history of mathematics, vol 3. Hindustan Book Agency (India), New Delhi
Euler L (1767) De usu novi algorithmi in problemate Pelliano solvendo. Novi Commentarii academiae scientiarum Petropolitanæ 11:28–66
Fowler D (1987) The mathematics of Plato's academy, a new reconstruction, 2nd edn (1999). Clarendon Press, Oxford
Fowler D (1994) Could the Greeks have used mathematical induction? Did they use it? Physis 31: 253–265
Freudenthal H (1953) Zur Geschichte der Vollstandingen Induction. Archives Internationals d' Histoire des Sciences 6:17–37
Hankel H (1874) Zur Geschichte der Mathematik in Alterthum und Mittelalter. Teubner, Leipzig
Hardy GH, Wright EM (1938) Introduction to the theory of numbers. Clarendon Press, Oxford
Heath TL (1908) The thirteen books of Euclid's elements, three volumes, 2nd edn (1926). Cambridge University Press, Cambridge
Heiberg IL, Stamatis ES (eds) (1977) Euclidis elementa Vol. V Pars 2: Scholia in libros VI-XIII cum appendicibus, 2nd edn. Vieweg+Teubner Verlag [in German]
Iamblichus (1894) In Nichomachi arithmeticam introductionem liber. Pistelli H (ed). Teubner BG, Leipzig
Kahane JP (1985) La Théorie de Théodore des corps quadratiques réels. L'Enseignement Mathématique 31:85–92
Knorr WR (1975) The evolution of Euclidean elements: a study of the theory of incommensurable magnitudes and its significance for early Greek geometry. Reidel, Dordrecht
Knorr WR (1983) "La Croix des Mathématiciens": the Euclidean theory of irrational lines. Bull Am Math Soc 9:41–69
Krishnaswamy Ayyangar AA (1929–30), New light on Bhaskara's Chakravala or cyclic method of solving indeterminate equations of the second degree in two variables. J Indian Math Soc18: 225–248
Lagrange JL (1769a) Solution d'un problème d'arithmétique, Miscellanea Taurinensia, vol 4. Reprinted in: Serret JA (ed) (1867) Oeuvres de Lagrange, vol 1. Gauthier-Villars, Paris, pp 671–731
Lagrange JL (1769b) Sur la solution des problèmes indeterminés du second degré, Mem. Acad. Roy. Berlin 23. Reprinted in: Serret JA (ed) (1869) Gauthier-VillarsOeuvres de Lagrange, vol 2, Paris, pp 377–535
Lagrange JL (1770) Additions au mémoire sur la résolution des équations numériques. Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin 14:581–652
Morrow GR (1970) Proclus: a commentary on the first book of Euclid's elements, translated, introduction, notes. Princeton University Press, Princeton
Mueller I (1981) Philosophy of mathematics and deductive structure in Euclid's "elements". MIT Press, Cambridge, MA

Negrepontis S (1999) Plato's dialectic under the anthyphairetic scrutiny. In: Syros V, Kouris A, Kalokairinou E (eds) Lectures, academic year 1996–97, pp 15–58. Philosophy Group of the University of Cyprus, Nicosia

Negrepontis S (Notes by Kleftaki V) (2003) A reconstruction of the proof of the palindromic periodicity of the Anthyphairesis of commensurable in power only lines, employing only tools from book X. of Euclid's elements. Manuscript (22 pages)

Negrepontis S (2005) The Anthyphairetic nature of the platonic principles of infinite and finite. In: Proceedings of the 4th Mediterranean Conference on Mathematics Education, 28–30 January 2005, Palermo, pp 3–27

Negrepontis S (2012) Plato's theory of knowledge of forms by division and collection in the Sophistes is a philosophic analogue of periodic anthyphairesis (and modern continued fractions), arXiv:1207.2950, 12 Jul 2012

Negrepontis S (2018) The Anthyphairetic revolutions of the platonic ideas. In: Sialaros M (ed) Revolutions and continuity in Greek mathematics. Science, technology, and medicine in ancient cultures, vol 8. Walter De Gruyter, Berlin/Boston, pp 335–381

Negrepontis S (2019) Plato on geometry and the geometers. In: Dani SG, Papadopoulos A (eds) Geometry in history. Springer Nature, Cham, pp 1–88

Negrepontis S (n.d.) From Musical Intervals to Hapseis and Logoi: the Periodic Anthyphairetic Nature of the One of the Second Hypothesis in Plato's Parmenides and Consequences (Indivisible Line, TMA, Zeno Anthyphairetic). In: Hascher X, Papadopoulos A (eds) Proceedings of the Conference Mathématiques et musique: des Grecs à Euler, Strasbourg, 10–11 September 2015. Hermann, Paris, to appear

Negrepontis S, Farmaki V (2019) The history of ancient Greek mathematics, vol 1. Ekkremes, Athens. [in Greek]

Negrepontis S, Farmaki V (2021) The paradoxical nature of mathematics. In: Papadopoulos A (ed) Topology and geometry: a collection of papers dedicated to Vladimir G. Turaev. European Mathematical Society Publishing House, Berlin

Negrepontis S, Farmaki V, Kalisperi D (2021). Pre-Eudoxian Geometric Algebra. Ganita Bharati: Bulletin of the Indian Society for the History of Mathematics, vol 43

Plato (1921) Theaetetus, Sophistes, Statesman, In: Plato in Twelve Volumes, vol 12 (trans: Fowler HN). Cambridge, MA, Harvard University Press; London, William Heinemann Ltd.

Plato (1925) Parmenides, Phaedrus, Philebus. In: Plato in Twelve Volumes, vol 9 (trans: Fowler HN). Cambridge, MA, Harvard University Press; London, William Heinemann Ltd.

Plato (1967 & 1968) Nomoi. In: Plato in Twelve Volumes, vols 10, 11 (trans: Bury RG). Cambridge, MA, Harvard University Press; London, William Heinemann Ltd.

Plato (1969) Politeia. In: Plato in Twelve Volumes, vols 5, 6 (trans: Shorey P). Cambridge, MA, Harvard University Press; London, William Heinemann Ltd.

Procli Diadochi (1873) In Primum Euclidis Elementorum Librum Commentarii. Friedlein G (ed), Teubner BG, Leipzig

Procli Diadochi (1899–1901) In Platonis rem publicam commentarii. In: Kroll W (ed) Bibliotheca scriptorum Graecorum et Romanorum Teubneriana, 2 vols. Teubner BG, Leipzig

Raghavan S (1997) Cakravala method. In: Selin H (ed) Encyclopaedia of the history of science, technology, and medicine in non-Western cultures. Springer, Berlin-New York, pp 1007–1010

Selenius CO (1975) Rationale of the Chakravāla process of Jayadeva and Bhāskara II. Hist Math 2(2):167–184

Shukla KS (1954) Acarya Jayadeva, the mathematician. Ganita 5:1–20

Singh P (1983) Varga-prakṛti – the cakravāla method of its solution and the regular continued fractions. Indian J Hist Sci 19(1):1–17

Srinivasiengar CN (1988) The history of ancient Indian mathematics. The World Press Private Ltd., Calcutta

Taisbak CM (1982) Coloured Quadrangles. A Guide to the Tenth Book of Euclid's Elements. Museum Tusculanum Press, Copenhagen

Tannery P (1882) Sur la mesure du cercle d'Archimède. Mém Soc Sci Phys Nat Bordeaux 4: 313–337

Tarrant H (1985) Scepticism or Platonism?: the philosophy of the fourth academy. Cambridge University Press, Cambridge

Theon Smyrnaeus (1878) Expositio rerum ad legendum Platonem utilium. Hiller E (ed). Teubner BG, Leipzig

Unguru S (1991) Greek mathematics and mathematical induction. Physis 28:273–289

van der Waerden BL (1950) Ontwakende Wetenschap. Noordhoff, Groningen. English edition: van der Waerden BL (1954) Science Awakening (trans: Dresden A). Noordhoff, Groningen

van der Waerden BL (1976) Уравнение Пелля в математике греков и индийцев. Uspeki Math. Nauk 31(5):57–70. English edition: van der Waerden BL (1976) Pell's Equation in Greek and Hindu Mathematics. Russ Math Surv 31:210–225

Wedberg A (1955) Plato's philosophy of mathematics. Almqvist & Wiksell, Stockholm

Weil A (1984) Number theory, An approach through history from Hammurapi to Legendre. Birkhäuser, Boston

Ancient Indian Mathematics: Sulbasutras – A Mathematical Review

S. G. Dani

Contents

1 Introduction 1025
2 A General Introduction to Śulbasūtras 1027
2.1 Their Period and the Available Corpus on Them 1027
2.2 The Format 1029
2.3 The Role of Śulbasūtras 1030
2.4 Units of Measurement in the Śulbasūtras 1031
2.5 *Vedi*s and *Citi*s 1033
2.6 Mathematical Highlights 1035
3 Constructions for Basic Rectilinear Figures 1041
3.1 Construction of a Square by the "Compass" Method 1041
3.2 Producing Perpendiculars by *Nyañchana* Method 1043
3.3 Perpendiculars by the "Arrow" Method 1044
3.4 Toward Conceptualization of Perpendicularity 1045
3.5 Āpastamba's Varied Applications of *Nyañchana* 1047
3.6 Other Rectilinear Figures 1048
3.7 Dividing Line Segments into Parts 1050
3.8 Transformation of Figures by Elementary Methods 1051
4 The Pythagoras Principle 1054
4.1 The Formulation in Baudhāyana Śulbasūtra 1054
4.2 The Statement in Āpastamba and Kātyāyana Śulbasūtras 1055
4.3 The Formulation in the Mānava Śulbasūtra 1056
5 Applications of the Pythagoras Principle 1057
5.1 A Square Whose Area Is the Sum of Two Given Squares 1057
5.2 Enhancing the Size of Figures by Scaling the Unit 1057
5.3 Kātyāyana's Method for Putting Together Multiple Squares 1059
5.4 A Square Whose Area Is the Difference of Two Given Squares 1060
5.5 A Square with the Same Area as a Given Rectangle 1061
6 Miscellaneous Topics Around the Pythagoras Principle 1062
6.1 Square Roots 1062
6.2 Semicircle with the Same Area as a Circle 1063

S. G. Dani (✉)
UM-DAE Centre for Excellence in Basic Sciences, Vidyanagari Campus of the University of Mumbai, Mumbai, India

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_114

6.3 Pythagorean Triples 1063
6.4 The *Mahāvedi* and a Thesis of Seidenberg 1064
7 The Circle 1066
7.1 Circumference of a Circle 1067
7.2 Construction of a Circle with a Given Area 1068
7.3 Circles Appearing in Construction of *Citi*s 1070
7.4 Construction of a Circle with a Given Area: The Mānava Version 1072
7.5 Value of π? 1074
7.6 Squaring the Circle 1075
7.7 Alternative Factor for Squaring the Circle 1077
7.8 Mānava's Conversion of Circle to Square 1078
7.9 Understanding Baudhāyana's Formula 1078
8 Square Roots in the Śulbasūtras 1080
8.1 $\sqrt{2}$ in Mānava Śulbasūtra 1080
8.2 Explanations of Baudhāyana's Formula 1081
8.3 Datta's Explanation 1082
8.4 Irrationality of $\sqrt{2}$? 1083
8.5 Other Square Roots 1084
9 Configurations of Shapes and Numbers 1086
9.1 Dārśapaurṇamāsika Vedi 1086
9.2 Tiling of *Gārhapatya* 1087
9.3 Tilings of *Kāmya Citi*s 1088
9.4 Configuration of *Rathacakraciti* 1090
10 In Place of Conclusion 1092
Appendix 1093
Concordance of sūtras between Śulbasūtras 1093
References 1093

Abstract

Śulbasūtras, which are compositions from the first millennium BCE, related to the activity of construction of *vedis* (altars) and *citis* (fireplaces) for the performance of *yajñas* (fire sacrifices), constitute a unique historical resource from India from the ancient times, with explicit mathematical content. They manifest familiarity with various geometric constructions and also geometrical principles, including what is now commonly known as Pythagoras theorem. The converse of the Pythagoras theorem played an important role in many of their practical constructions. The task of interrelating the square and the circle in terms of areas was also addressed by them, though in a limited practical context, and the pursuit also seems to have led them to quite an accurate approximation of $\sqrt{2}$.

Study of the Śulbasūtras in the modern framework began in the second half of the nineteenth century, with the work of George Thibaut. Since then four Śulbasūtras, Baudhāyana Śulbasūtra, Āpastamba Śulbasūtra, Mānava Śulbasūtra, and Kātyāyana Śulbasūtra, have been studied by various scholars, to varying degrees of detail. While on the one hand there have been several illuminating findings, on the other hand a variety of poorly substantiated claims and theories have also emerged in historical writings on them.

This chapter aims at presenting a reasonably comprehensive account of the mathematical contents of the Śulbasūtras, together with a critical review of

various inferences drawn in the literature, and bringing out the issues calling for further studies and clarifications. We highlight also the differences between the different Śulbasūtras and comment on their significance in various respects, an aspect which has not received much attention so far.

Keywords

Śulbasūtras · Baudhāyana · Āpastamba · Mānava · Kātyāyana · Pythagoras principle · Vedis and citis · Geometrical constructions · Circling the square · Squaring the circle

1 Introduction

Around the beginning of the first millennium BCE performance of *yajñas* (ritual fire sacrifices) seems to have acquired tremendous importance among the Vedic people in India. The *nitya* (literally meaning "perpetual" or "regular") *yajñas* performed by normal householders (from the priestly class) would have had a longer tradition, presumably going back to the Ṛgvedic period. Following the rise of the ritualism associated with the *yajurveda* (cf. Bloomfield 1908, p. 31), eventually the *kāmya yajñas* became one of the salient features of the Vedic culture. Here "*kāmya*" refers to being undertaken in pursuit of "desires," and the desires involved could relate to both spiritual objectives as well as material gains, and the latter could in turn be in terms of health and physical strength, wealth, political power, etc. The schemes for *kāmya yajñas* evolved as ambitious large-scale projects, with the involvement of the kings and the rich, and brought a great deal of material as well as intellectual resources into the activity. This led, on the one hand, to elaborate formulation of intricate operational details for the performance of *kāmya yajñas*, both in terms of the associated littany, ritual elements, as well as physical structures and arrangements to be adopted. On the other hand, it gave a fillip to the development of know-how in various respects, as an enabling factor in the projects contemplated. The *vedis* (altars) for the *yajñas* were raised areas, in specified shapes, of sizes often over 50 feet in length and breadth. There were designated spots marked on the vedis for performance of various ritual events involved. The *citis* (fireplaces) to be erected for the main sacrificial event were often in intricate shapes, of various birds, turtle, etc., with sizes of the order of 20 feet in length and 30 feet in width. The preoccupation with drawing figures on the ground on these scales seems to have brought in an understanding of various principles in planar geometry, as also some miscellaneous mathematical ideas. The Śulbasūtras form our main, and almost exclusive, source of this mathematical content; in respect of geometry, they constitute the earliest corpus from India (in absence of any significant inputs relating to the Indus valley civilization in this respect).

The Śulbasūtras form a section of the *Śrautasūtras*, which in turn are parts of *Kalpasūtras* belonging to the *Vedāṅga* (appendages of Vedas) literature. The

Śrautasūtras deal with various aspects of the performance of the *yajñas* including, in particular, the *mantras* to be chanted in the course of various rituals, while the Śulbasūtras are concerned primarily with various aspects of laying of the *vedis* and *citis* for the purpose. Apart from the details with regard to plotting and erecting the *vedis* and *citis*, the Śulbasūtras also include enunciation of the geometric principles and constructions involved. Also, aside from the explicitly articulated knowledge, there are various ideas discernible in other parts, which also form an integral part of the study of the mathematical contents of the Śulbasūtras.

The Vedic people were a heterogeneous community with numerous branches (*śākhās*), and while they shared a common body of knowledge, the individual branches had their respective compositions that they followed. Thus, it is quite plausible that there may have been numerous Śulbasūtras.[1] There are at present four Ślbasūtras that have been edited, translated into western languages, and have been studied to varying degrees: Baudhāyana Śulbasūtra, Mānava Śulbasūtra, Āpastamba Śulbasūtra, and Kātyāyana Śulbasūtra; three more have been named in Datta (1932), as being extant, which are by Maitrāyaṇa, Varāha, and Vādhula, and two others *Maśaka śulba* and *Hiraṇyakeśi śulba*, among those that are lost; and one other, *Satyāṣaḍha śulba,* is named in Kulkarni (n.d., p. III).[2]

The studies and writings on the mathematical contents of the Śulbasūtras have largely focused on the body of knowledge contained in them as a whole, a typical question addressed being whether one or other mathematical idea is found in (any one of) them. While some statements involving comparisons do come up in the literature, and also there are some tabular concordances produced (the latter notably in Kulkarni n.d.), a serious comparative study of them as independent compositions, and whether and what it may signify with regard to the activity as a whole, seems lacking so far. This aspect seems especially important on account of the fact that they belong to widely differing time periods, and the differences between them may relate, in one or more ways, to the overall developments and changes over the time period involved. In this context, our objective here will be twofold: One is to give a comprehensive introduction to the mathematical content of the Śulbasūtras, and the second is to provide a comparison between the four Śulbasūtras, with comments on the similarities and differences.

We begin in Sect. 2 by recalling some general details concerning the Śulbasūtras. The later sections are devoted to thematic description on various topics covered in the Śulbasūtras, paying attention also to the distinctive features of the individual Śulbasūtras.

[1] An estimate of over a thousand is cited in Datta (1932, p. 1) as the number of "schools," each one "with its own Śrautasūtra and hence probably its own Śulba" – it is however unclear how many of them would have been substantially distinct from each other.

[2] It is also mentioned in Kulkarni (n.d., p. III) that while they come from different *Kalpasūtra*s, thereby inheriting their names, the Satyāṣāḍha śulba is the same as the Āpastamba śulba and similarly Maitrāyaṇīya śulba is the same as Mānava śulba (though the larger works may have differences).

2 A General Introduction to Śulbasūtras

The four Śulbasūtras to be discussed, Baudhāyana Śulbasūtra, Mānava Śulbasūtra, Āpastamba Śulbasūtra, and Kātyāyana Śulbasūtra, we shall refer, for brevity, as BSl, MSl, ASl, and KSl, respectively. In this section, we recall various general details which will set the context for the discussion in the later sections.

2.1 Their Period and the Available Corpus on Them

Determination of the period of the Śulbasūtras has been a problematic issue, with no intrinsic clues available other than their style of language and grammar,[3] and the estimates of various scholars differ considerably. According to Plofker (2009), their approximate chronological order corresponds to what we have adopted in the listing above; this is also consistent with the discussion on the issue by Sen and Bag (1983). In the wider range of scholarly works involved, there is no consensus even about the century of individual compositions and also on the chronological precedence between MSl and ASl. A broad date-range is introduced in Kulkarni (n.d.), citing C. G. Kashikar (1966), as Baudhāyana (800–500 BCE), Āpastamba and Mānava (650–300 BCE), and Kātyāyana (300 BCE–400 CE), which it seems worthwhile to keep in mind, except that in the case of Kātyāyana, from the extended range suggested, the early part seems more appropriate, in the context of estimates from various other works (cf. Sen and Bag 1983, page 3, in particular).[4] These differences in the estimates with regard to their period will not be of consequence for our discussion in the following sections.

There have been several commentaries on these compositions. Their dates, especially of the notable ones, are uncertain however, for want of adequate inputs. Based on the works they cite, or allude to, most of them are inferred to be from the

[3] The reference to "style of language and grammar" here is in terms of the contents of the respective *Śrautasūtras* containing the respective śulbas, as a whole, and not just the śulbas. While this provides an advantage in terms of comparability, it brings in an issue whether the inferences drawn would apply equally to the corresponding śulbas. In the overall context, the śulbas which were typically the last sections of the Śrautasūtras, comparable to Appendices in modern works, and played largely a facilitating role, which may not have been considered as sacrosanct as the core body, and could have been subject to flexible updating. It may be noted that there has been at least one suggestion, by a commentator, Rāma of the *Kātyāyana Śrautasūtra*, believed to be from the fifteenth century, that the śulba in that work was perhaps composed by a different person (cf. Datta 1932, p. 12, where however the possibility has been discounted).

[4] It may however be recalled here that according to many early scholarly works, from the nineteenth and early twentieth century, the Śulbasūtras are from a somewhat later period than those indicated above; see the account on the question in Seidenberg (1975, Sect. 10), where the author recalls the estimates of various scholars. It may also be mentioned that the Chronology Committee of senior historians of science, formed in the 1950s following a symposium on History of Sciences in South Asia organized by the Indian National Science Academy (INSA) at Delhi, under the Chairmanship of R.C. Majumdar (Other members were A. S. Altekar, P. C. Bagchi, S. L. Hora, D. S. Kothari, and A. N. Singh) which considered dating of various historical stages in India, recommended 500 BCE and later as the period of the Śulbasūtras (cf. Gupta 1990, 2001).

second millennium CE, with a few possibly from the later part of the first millennium. The commentaries have played a very important role in the preparation of the modern editions of the works, it being a difficult task to make out the meaning and significance of various parts of the original texts, without the benefit of the interpretations in the commentaries; the reader is referred to Datta (1932) for a detailed discussion on the topic; see also (Delire 2003, 2009).[5] On the other hand, it needs to be borne in mind that the level of mathematics witnessed in the commentaries does not necessarily correspond to that of the originals, given especially the time gap. The commentators were not taking on the role of a historian but were acting as interpreters of the contents of the Śulbasūtras for the community engaged in pursuing the ritual activity, and their interpretations would naturally incorporate some of the mathematical knowledge accumulated by their time. The mathematical practices and trends had also seen major changes over the span of the millennium separating the originals and the commentators.[6] We shall however not go, in this review, into the details on these points, except for occasional references.

The first English edition of BSl was produced by George Thibaut, together with his commentary, in 1875 (cf. Thibaut n.d.-a). Thibaut also edited and translated into English the first two chapters of KSl (cf. Thibaut n.d.-b). The next to attract the attention of the global scholarship was ASl, of which a German edition, with commentary, was brought out by Albert Bürk in 1901 (cf. Bürk 1901). The first modern language edition of the Mānava Śulbasūtra arrived only in 1963, with Jeanette M. van Gelder's translation of (the larger work) *Mānava Śrautasūtras* (van Gelder 1963). English editions of BSl (including the commentary of Dvārakānatha Yajvan) and of ASl (with commentaries of Kapardisvāmin, Karavinda, and Sundararāja) were brought out in 1968 by Satya Prakash and Ram Swarup Sharma, in (1968a, b), respectively. A comprehensive edition on the four Śulbasūtras together, accompanied by translations and the authors' commentaries, was brought out by S.N. Sen and A.K. Bag in 1983 (Sen and Bag 1983).[7] A comprehensive edition of the four Śulbasūtras together with several useful tables and indices, and a concordance, was brought out by Raghunath Purushottam Kulkarni, in Marathi, which has subsequently been translated also in Hindi (Kulkarni n.d.). A critical edition with translation and commentary in French, of the Baudhāyana Śulbasūtra, together with the commentary

[5] In Delire (2009), comparing some parts of the text of *Śulbadīpikā*, Dvārakanātha's commentary on BSl, with Sundararāja's commentary on ASl the author infers that Dvārakānātha copied from Sundararāja (negating the view held by B. Datta that the latter copied from the former, and confirming R.C. Gupta's suggestion made earlier). Based on the conclusion the author revises David Pingree's estimate of *terminus post quem* for the former to the second half of the fifteenth century CE; we may also recall here that 1609 is argued as a *terminus ad quem*.

[6] It may be worthwhile to quote here the following from Thibaut (1875, p. 272): "Trustworthy guides as they are in the greater number of cases, their tendency of sacrificing geometrical construction to numerical calculation, their excessive fondness, as it might be styled, of doing sums renders them entirely misleading."

[7] It may be noted here, however, that while the commentaries on BSl and ASl are fairly detailed, those on the other two, especially MSl, have tended to be rather cursory, warranting a better treatment.

Śulbadīpikā of Dvārakānātha Yajvan, was brought out by Jean-Michel Delire in 2016 (Delire 2016).[8]

There have, of course, also been numerous descriptive accounts in various books (we may mention in particular (Srinivasiengar 1967; Amma 1979; Bag 1979; Kulkarni 1983; Plofker 2009; Puttaswamy 2012; Joseph 2016; Divakaran 2018), the summary accounts (Gupta 2001, 2006) on the topic) and several papers on themes related to the topic. The works that we refer to will be cited accordingly.

2.2 The Format

Like other *Vedāṅga* literature, the Śulbasūtras are composed mostly in the "sūtra" (aphorism) style prose, except for some parts in MSl, which are in verse form. "Sūtra" means thread, and the style involved endeavors to present the text through concise arrangements facilitating memorization and recitation, rather than adhering to a regular complete-sentence format for a narrative. It may be noted that the originals do not consist of individually numbered sūtras. The text of the respective originals has been divided in the modern expositions in convenient individual segments which have been enumerated as sūtras, and grouped into chapters; some of the original works also have divisions like chapters, but their scheme has not always been adhered to in the modern editions. Unfortunately there is considerable diversity in the schemes adpoted in the different expository works with regard to segmenting and numbeing of the sūtras, (Sen and Bag 1983; Prakash et al. 1968a; Kulkarni n.d.). For instance, Thibaut in (n.d.-a) adopted a division of BSl into three chapters, with 113, 83, and 323 numbered sūtras, respectively, adding to 519 sūtras, while on the other hand in (Sen and Bag 1983), which now serves as a standard reference, the material appears in 21 chapters adding to 285 sūtras; the authors however include with each sūtra the reference number from Thibaut, in parenthesis. To provide the reader an overall sense of the size of the works, let me also mention further that as presented in Sen and Bag (1983) MSl has 16 chapters adding to 228 sūtras, ASl has 21 chapters adding to 202 sūtras, and KSl has 6 chapters adding to 67 sūtras; except in a few cases, each sūtra is in one or two lines.

Here we shall primarily follow the numbering as in Kulkarni (n.d.), which from the author's point of view has the advantage that it provides the text of the Śulbasūtras in the Devanagari script. For BSl these numberings are according to Thibaut's, as may be referred from Sen and Bag (1983), and for MSl they correspond to van Gelder's numbering. The author hopes to publish in the near future a concordance between the numbering of the sutras in various sources of mathematical Śulbasūtras.

[8] Recently an edition of *Maitrāyaṇīya Śulbasūtra* was brought out by R.P. Kularia (Kularia 2016); the author however has not had an opportunity to access it, and as such its contents will not be reflected here.

2.3 The Role of Śulbasūtras

Equipping the people involved with constructions of *vedi*s and *citi*s, with the necessary skills, would have been the task before the composers of the Śulbasūtras. The śulbas however are not quite in the format of simply describing, sequentially and minimally, the steps to be taken toward the desired end output, in the style of manuals for assembling equipment etc., even though they have often been referred to as manuals, and though their basic purpose could have been served in this fashion as well. They are seen to take on a larger role; in the case of BSl and ASl, the presentation evidently manifests a certain purposefulness, with a pedagogical objective of disseminating also geometric understanding around the procedures described; and in the case of MSl, though one does not find any disciplined pursuit of a pedagogical objective, a good deal of geometrical knowledge is disseminated anyway, along with the operational details meant to be conveyed, in what seems like a chatty style. In KSl, one finds somewhat of a mixture of the two approaches. It may be noted that on the whole many of the geometric principles and constructions that have been articulated are not directly involved in following the procedures for laying out the *vedi*s and *citi*s. Thus, a more comprehensive approach is seen to be taken in the context, toward disseminating the accumulated knowledge, along with explication of the core operational part involved in the activity.

In BSl, an exposition of geometry is taken up systematically at the outset. The first two chapters, as set out in Sen and Bag (1983), with 25 sūtras (the number comes to 62 in Thibaut's edition), contain an exposition on geometry[9] (together with an introduction to units of measurement). This part could have, for instance, been adopted in the form of a textbook, by other sections of the society at the time, not concerned with the activity of laying out the structures involved in the rituals; whether this happened on any significant scale is something we may never get to know.

Much in the same way as BSl, ASl also includes a course on geometry, at the beginning, and while there are some significant differences between the two (which we will have occasion to discuss in the later sections here), the divergence is similar to what one would find between two textbooks on the same topic. The same however cannot be said of MSl and KSl. Indeed, a good fraction of the propositions appearing in BSl and ASl, and also some that are not part of them (as will be seen later), are found in MSl and KSl, but they turn up strewn across the narrative; significance of the spots where they appear would itself be an interesting issue for exploration. Many of these statements, though not all, also seem to be more brusque than the corresponding part in BSl and ASl. One may wonder whether this suggests that a

[9] A quick account of this part with translations of sūtras and a brief commentary may be found in Plofker (2007); I may add a small clarification that *nyañchana* refers to a mark made for the purpose of plotting points on the ground with respect to it, and is not a word for "mark," as may be suggested by the parenthetical reference there, on p. 388; also, it perhaps ought to have been highlighted that sutra BSl I.48 (in Thibaut's numbering as adopted here) corresponds to what is called the Pythagoras theorem.

degree of geometric understanding is viewed as a prerequisite, and the statements are being recalled briefly as recapitulation; the suggestion is not quite borne out consistently however, if one considers the details that the sutras go into in certain cases. We shall discuss this further when we get to the details.

2.4 Units of Measurement in the Śulbasūtras

Before proceeding further, it would be convenient to give an introduction to the units of measurements involved, to build up a basic familiarity with them. Interestingly, Baudhāyana Śulbasūtra begins, rather uniquely among the Śulbasūtras, with a detailed account of various units, which serves as a good reference for our context here; there are some variations involved in the usage of units in other Śulbasūtras, which also we will discuss briefly.

Among the principal units (in terms of their usage in the exposition as in BSl) are *aṅgula, prādeśa, pada, aratni*, and *puruṣa*. These are manifestly based on human anatomy; the terms, respectively, mean "finger," "span" (between the thumb and forefinger), "foot," "elbow," and "man"; the interrelation between them is set out in BSl as 12 *aṅgulas* to a *prādeśa*, 15 *aṅgulas* to a *pada*, two *pradeśas* to an *aratni*, and five *aratnis* to a *puruṣa*. A *puruṣa* would thus be 120 *aṅgulas*. Two other measures occurring frequently are *prakrama* (which literally means a step or stride), defined to be two padas, and *vyāyāma*, defined to be four *aratni*s.

While the interrelations described above would fix the relative sizes of the units mentioned, further inputs would be needed to make them more definitive, as the anatomical sizes would vary to an extent from person to person. This would seem to have been achieved to an extent, though it is unclear whether this was indeed intended (see below for some related comments), by Baudhāyana by describing *aṅgula* as being equal to 14 *aṇus* and also alternatively as 34 *tilas*; *aṇu* here stands for a certain seed – and not an atom (which is the now common meaning associated with the term) – and *tila* means sesame seed, and in either case what is meant is that the length measure *aṅgula* equals that many of the seeds placed sideways in a row, widthwise, next to each other. On the other hand, Baudhāyana also states that modifications may be applied in respect of measurement of certain units, which include *pada, prakrama,* and *aratni* from the abovementioned ones, "when desired" (BSl I.18). Baudhāyana's text does not specify what circumstance would call for such a modification, but as explained by commentator Dvārakanātha (see Prakash et al. 1968a, p. 43 – in Sanskrit) this is for making an allowance for traditions, and in particular a provision was called for to conform to the practice of taking the *puruṣa* measure dependent on the *yajamāna* (the person performing or sponsoring the *yajña*), as the distance from the tip of the middle finger to the ground when he stands with upstretched arms; this convention was commonly used for fixing the sizes of the bricks for tiling the *citi*s (cf. Kulkarni n.d., p. XXV). In such a case, the *aṅgula* measure is taken to be 120th of the *puruṣa* measure thus determined.

Taking the height of the *yajamāna* from toe to upstretched arms to be $7\frac{1}{2}$ feet, viz, 90 inches, the *aṅgula* measure would be $\frac{3}{4}$th of an inch, and *pada* would be $11\frac{1}{4}$ inches, or close to a foot.

As indicated, unlike BSl the other Śulbasūtras do not present a list of units, but various units appear along the texts. Of the abovementioned units, *aṅgula, aratnī*, and *puruṣa* are shared by all. *Prādeśa* takes a new but equivalent name *vitasti*. The unit *pada* appears later only in ASl and KSl, but in KSl it stands for 12 *aṅgulas*.

In the account on units of measure in BSl mentioned above, there are also various other units described, which do not bear a simple relationship with others; among these, I may especially mention *īṣā* which is 188 *aṅgulas*, *akṣa* equal to 104 *aṅgulas* and *yuga* which is 86 *aṅgulas*; all the three terms are related to chariots, meaning the shaft, axle, and yoke, respectively, with primarily a contextual significance and, interestingly, are shared by all the four Śulbasūtras. Though these have been listed by Baudhāyana along with various units that are involved in describing various lengths in the course of the constructions, they are in fact sizes associated with a specific *vedi*, styled as a chariot in the shape of an isosceles trapezium, with *īṣā*, *akṣa,* and *yuga* as its altitude, base, and face, respectively. While apparently the sizes were expected to be known all around, and accordingly disseminated in all the Śulbasūtras, they had no manifest role in measuring other lengths, in multiples or parts.

The relation one *aṅgula* being 34 *tilas* would also seem curious. It turns out that the choice is likely to have to do with an approximate numerical expression for $\sqrt{2}$ (see Sect. 8) in which 34 is involved in the denominator; this was noted by Thibaut in Thibaut (n.d.-a),[10] observing also that the subdivision may have been chosen to simplify calculations in that context – with the sesame seed serving as a matching choice; whether it matched the actual size of the seed to a substantial degree of accuracy or not may not have mattered. On account of this background, the conversion holds wider interest from a historical point of view. Both of these finer units are specific to BSl and absent from the other Śulbasūtras. It may also be noted that in BSl the division into 34 parts is related to the theoretical issues of squaring the circle (see Sect. 7), which have generally been skipped in the other Śulbasūtras. On the other hand, the finer units being dropped from the later Śulbasūtras also suggests that they had not been of practical consequence.[11]

[10] It is argued in Delire (2005) that similarly the unit *aṇu*, defined to be the 14th part of *aṅgula*, is motivated by another approximation of $\sqrt{2}$, viz, as $\frac{7}{5}$. Though there is no reference to this approximation in BSl, it is found in MSl and is likely to have a longer history in the tradition. According to Delire (2005), the *aṇu* being defined as the 14th part of *aṅgula* may be considered as "the fossilized remnant of it."

[11] It may be worth recalling here that the Mānava Śulbasūtra mentions, in MSl 1.4.2–4, a different set of fine measures. These are introduced in order to build up the *aṅgula* and other larger measures, in a context when the *yajamāna*'s height is stunted, due to genetic reasons or a disease, and hence inadmissible in setting the units. Mānava takes an *aṅgula* to be six *yavas* (barley seeds), each *yava* to be six *sarṣapa*s (flax seeds), which are further subdivided, with the latter described as six hairs of a "3-year-old calf," and the last one further equated to six of lotus pollen. The scheme however is very likely a notional make-believe formulation, rather than anything of metrological significance, especially there being nothing to indicate their having the practical means to make comparisons at that level. At any rate, it is doubtful if the finer measures alluded to played any role at all, as the seed units would have sufficed for the purpose at hand; more descriptions would only lead to more ambiguity and confusion in matching them. It is interesting however that *aṅgula* in this formulation equals 36 *sarṣapa*s, quite close to Baudhāyana's subdivision of *aṅgula* as 34 *tilas*. It is also interesting that the subdivisions are repeatedly into six parts of the previous one.

Finally, it may be mentioned here that in measuring areas, in the Śulbasūtras the same name as any linear measure is used for the area of the square whose side is one unit in that measure; thus, one *puruṣa area* is what we may refer to as one square *puruṣa*[12]; in the sequel, we shall adopt the same practice as in the originals, and only occasionally include "square" in parenthesis, for clarity.

2.5 *Vedis* and *Citis*

The performance of the *kāmya yajña*s involved elaborate arrangements, with multiple structures and designated spots for various functions that were part of the sacrificial event. The steps for these are prescribed in quite minute detail; they are however not all part of the Śulbasūtras; rather, they go back to the older *Brāhmaṇa* works,[13] and one of the cited sources in this respect is the *Śatapatha Brāhmaṇa*; to some extent, the details may also have been passed on from generation to generation orally, as a part of the tradition, without being part of any composition. The Śulbasūtras recall only some features that are immediately relevant to their narrative, and the other aspects, including notably the names of various components in the overall organization, are taken for granted as known. Let me recall here a few of the details which would be helpful in following the discussion in the forthcoming sections; the reader is referred, in particular, to Sen and Bag (1983, pages 169–173) for an exposition in greater detail.

2.5.1 The *Mahāvedi* and the Three *Agnis*

An important component of the layout for the *kāmya yajñas* is the *Mahāvedi* (literally "the great altar," also called *saumikī vedi*, the soma altar). It is in the shape of an isosceles trapezium, with its axis of symmetry along the east-west line (called *prācī*)[14]; the stretch along the axis is 36 *steps,* and the (mutually parallel)

[12] This can be a source of some confusion, since when two units are involved in a description, say two *puruṣas* and one *aratni*, it would not be clear whether the area refers to the sum of the two square measures, or the area of the square with the sum as the linear measure; see the discussion related to this in Seidenberg (1978, pp. 336–337); and the issue however will not concern us here.

[13] *Brāhmaṇas* are scriptures intermediate to the *Vedas* and the *Kalpasūtras* (which are parent bodies of the Śulbasūtras), and at several places Baudhāyana makes statements citing their authority, distinguishing along the way between teachings of the Brāhmaṇa to which he owes allegiance and those of the others.

[14] Baudhāyana does not mention how the cardinal directions were determined. It has been suggested that it was determined by the shadow of the pole on the equinox day, and verified by the rising and setting points of the star Kṛttikā (see Sen and Bag 1983, p. 264). In KSl 1.2, a procedure is described to determine the east-west line as the line joining the two points on a circle drawn around the base of a pole, where the shadow of the tip of the pole falls on the circle in the course of a day. The north-south line is then obtained by drawing a perpendicular to the east-west line, the instruction for which (as given in KSl 1.3) corresponds to what we refer to as the arrow method, described in Sect. 3.3. Even so, KSl 7.35 describes determining East as the direction of rising of the stars Kṛttikā, Śravaṇa, or Puṣya, or as the midway of the directions of rising of Citrā and Swàtī. Presumably both procedures coexisted and were used for confirmation of each other.

sides on the western and eastern sides are 30 and 24 steps, respectively; while the description in *Śathapatha Brāhmaṇa* in this regard is in terms of the number of step strides (see Sect. 6.4 for more details on this), BSl and ASl assign for them that many units in size, the unit chosen being either a *pada* or a *prakrama*, which correspond to $11\frac{1}{4}$ inches and $22\frac{1}{2}$ inches, respectively (it may be recalled that *prakrama* by itself also means a stride). The area defines the region for various special ritual activities, of which we will not go into the details (see Sen and Bag 1983 for an exposition). I may mention however that there is a rectangular annexe to this, on the western side, symmetric about the axis of the *Mahāvedi*, called *pragvaṃśa*, where the priests, the members of the family, and guests assemble.

Every *yajña* mandatorily involved also three *agni*s (sacred fires), viz, *āhavanīya*, *gārhapatya,* and *dakṣiṇāgni*, and these are also situated in the *prāgvaṃśa*, on eastern, western, and southern side, respectively. Three different schemes for locating the position of *dakṣiṇāgni*, in terms of the other two agnis, which are along the east-west line, are described in BSl I.67–69. In Datta (1932), the author assumes that identical spot is being described in the three, and deduces from it the implied values of $\sqrt{5}$ and $\sqrt{2}$; however, the assumption is erroneous, as we contend in detail in Sect. 8.5. Datta's assumption is all the more surprising in the light of the fact that two of the three methods proposed in BSl, and also analogous descriptions in Āpastamba and Kātyāyana, give two options each (in ASL 4.6–4.7 and KSl 1.26–1.27, respectively), in which it is evident that they do not purport to describe the same spot.

2.5.2 *Kāmya Citis*

Depending on the motive of the *yajamāna* (one performing the *yajña*) in performing the *yajña*, the ensemble for it includes an *agniciti* of a special design, like in the shape of a bird, turtle, etc.; one desiring heaven would have the *citi* in the shape of one of the birds, *Śyena* (falcon) or *Kaṅka*, in a flying mode, with wings outstretched. *Droṇaciti*, in the shape of a trough (used for serving food), is when the desire is for plentiful of food; the trough designs, in square or circular shape, come equipped with a handle. *Chandściti* is the *citi* to be adpoted when cattle is desired.[15] *Rathacakraciti*, in the shape of a chariot wheel, was erected when annihilation of rivals was the desire involved.[16] The turtle-shaped *Kūramāciti*, described only in BSl, is for vanquishing *Brahmaloka*! (BSl III.270).

The *citi*s typically consisted of five layers of tilings of bricks, together forming a knee-high structure, with the cross-section, and the top view, in the desired shape, specific to the *yajña*. The area of each *kāmya citi* was to be $7\frac{1}{2}$*puruṣa*, while performing the *yajña* for the first time, and was enhanced in successive repetitions (the process of enhancements is discussed in Sect. 2.6.4 and in further detail in Sect. 5.2; when not

[15] Cattle was one of the major items of acquisition in the Vedic times, and evidence from the vedic literature shows that people maintained huge herds of cattle.

[16] A tabular list of *citis* together with the motives associated with them, as described in the *Taittirīya Saṃhitā* of the *Yajurveda*, is provided in Sen and Bag (1983, p. 6).

specified otherwise, in our dicussion we shall restrict to the former, the primary context).

2.5.3 The Brick Layouts

The brick-tiling plans for the various designs of the *kāmya citi*s involve intricate designs (even when the outline of the *citi* is simple, like a rhombus), and specific patterns are provided by the Śulbasūtras for the *citi*s they discuss.

Each layer was to have 200 bricks, so the total number of bricks would be 1000, and there are also some special instances involving larger number of bricks, but we need not go into it here. The first, third, and fifth layers follow a common design for brick-tiling, and the second and fourth share another common design, and the two designs are such that the edges of the bricks would not overlap between the successive layers, except at the outlying edges of the figure of the *citi*. Apart from these general restrictions, the designs are seen to incorporate various esthetic considerations. The tiles involved are in multiple shapes, though a majority of them are in square shape. We will have further occasion later to elaborate on the shapes involved (see Sect. 3.6).

The *āhavanīya* and *gārhapatya agni*s, which were used domestically, as well as mandatory parts of the elaborate sacrificial rituals, had an area one (square) *vyāyāma*. The *gārhapatya* was tiled with 21 bricks, adhering also to the stipulation of non-overlapping of the edges, as above; see Sect. 9.2 for more on their design patterns.

2.6 Mathematical Highlights

Construction of various figures, fulfilling a variety of stipulations, is the main theme of the mathematical content of the Śulbasūtras; other aspects of geometrical understanding, presented or alluded to, are supplementary. Though the *citi*s were indeed three-dimensional objects, it is the planar designs of the shapes involved that play the crucial part, and as such the main geometrical content of the Śulbasūtras is about geometry of planar figures.[17]

2.6.1 Rectilinear Constructions and Related Geometry

The typical figure that one encounters most often in the Śulbasūtras is that of isosceles trapezia, including the various special cases consisting of squares, rectangles, and isosceles triangles.[18] Indeed, the isosceles trapezia seem to have held a great appeal to the ancient Indian geometers, as witnessed already by the choice for

[17] There is an isolated instance of a *citi*, the *śmaśānaciti* (cemetery-shaped *citi*), described in BSl 8.1–17, whose shape is of a frustrum of a pyramid, and there is a discussion in Datta (1932, pp. 101–103) involving a volume computation of the shape; however, the formulation and inferences do not seem convincing enough to be recalled here.

[18] The shapes of some of the bricks used in the construction of *citis* are asymmetrical; see Sect. 3.6. However, there is no discussion of geometrical issues involved in respect to them.

the shape of the *Mahāvedi*, presumably inspired by esthetic considerations, or also metaphoric or spiritual aspects, associated with the activity.[19]

These figures, as drawn in the context of construction of the *vedi*s, had a bilateral symmetry, with the axis of symmetry along the east-west direction,[20] with the face (the shorter of the parallel sides) toward the east and the base (the longer parallel side) on the west, as seen in the case of the *Mahāvedi*. The description of the steps in the constructions is typically with reference to the cardinal directions; this applies not only to description of practical constructions of *vedi*s and *citi*s, but also in the enunciation of the general theory, even where it is taken up separately. This is seen to have served as a crucial reference framework in absence of notation for points and lines drawn in the course of the constructions.

As may be expected, construction of perpendicular is one of the key ingredients in drawing these figures. However, the notion of a perpendicular line is not explicitly formulated in any of the Śulbasūtras. Moreover, the constructions described are restricted to closed geometric figures needed in practical constructions. As such, tasks like drawing a perpendicular to a line at a point do not figure in the exposition by themselves, even though they are involved as part of various procedures that are described in the construction of the individual figures. In this respect, the works resemble manuals and reflect their limited aim. It is nevertheless clear that the theorists were aware of them as primary tools, though they may not quite have abstracted them as principles that could be communicated directly. In a critical appreciation of the results, it would be important to bear this aspect in mind. Some details of how the steps are communicated in the basic constructions will be described in Sect. 3.

2.6.2 The Pythagorean Principle and Pythagorean Triples

One of the main highlights of the Śulbasūtras geometry relates to the Pythagoras theorem, or rather the physical principle involved.[21] It is explicitly stated in all the four Śulbasūtras as above; the statements appear in terms of diagonals and sides of a rectangle, rather than a right-angled triangle as in Euclidean geometry. These are arguably the earliest extant explicit statements of the principle, anywhere in the world. We shall discuss the statements as presented in the individual Śulbasūtras, and their nuances, in some detail; see Sect. 4. Apart from the statements themselves, we find numerous theoretical geometric as well as practical applications of the principle in the Śulbasūtras. These will be discussed in the later part of Sect. 4.

While the Śulbasūtras geometers were aware of various methods of drawing perpendiculars, as we shall see in detail in Sect. 3, their favourite method was

[19] Engagement with isosceles trapezium in fact goes back to more ancient works. It was pointed out by A. Bürk that it occurs as a shape of a *vedi* described in *Taittirīya saṃhitā* (cf. Seidenberg 1975, p. 510).

[20] This was followed in respect of the larger plan as well, except in a few respects, such as the location of *Dakṣiṇāgni*, or choice of tiling designs of some of the *kāmya citis*.

[21] In the context of the Śulbasūtras, I prefer to use the term "principle" in this context, the term "theorem" being associated closely to framework of deductive systems. The Sanskrit word *nyāya* that is commonly used for statements from this category also corresponds to "principle," than an equivalent of "theorem."

based on the converse of the Pythagoras principle, that a triangle in which the square of the largest side equals the sum of squares of the other two sides is a right-angled triangle; this was used along with Pythagorean triples (triples of integers in which squares of two of them add up to the square of the third), typically (3, 4, 5) and (5, 12, 13) by a procedure called *nyañchana* (see Sect. 3.2 for details).[22]

Curiously, though it is used extensively in both theoretical and practical contexts involved in the works, the converse of the Pythagoras principle is not stated explicitly anywhere in the Śulbasūtras. It may also be noted that the usage is seen to appear in constructions even before the statement of the Pythagoras principle itself, in the text of the Śulbasūtras. One may wonder if this means that they viewed it as an independent basic principle by itself, which it was adequate to illustrate through construction of figures.

Usage of the converse of the Pythagorean principle, as indicated above, brings in the notion of Pythagorean triples, viz, integer triples (a, b, c) such that $a^2 + b^2 = c^2$. Their interest in the latter (for which no name was introduced) was anchored on its usage in *nyañchana*. For that, the small-sized triples as above are the most convenient, so one may wonder if they would have had any motivation for exploring the multitude of Pythagorean triples. However, various such triples, other than the two recalled above, do appear in BSl and ASl, as we shall see in Sect. 6.3; there is also a sūtra in KSl which has been linked in literature to Pythagorean triples, on which also I shall comment in Sect. 6.3.

2.6.3 Transformation of Figures

Areas of figures played an important role; in a way, along with the significance of the design, its area would have been a guage of the stature associated with it. While multiple designs were possible, depending on the context, the areas involved were chosen to be the same. It is generally supposed that the practice stems from a scriptural injunction, but there is no evidence to that effect. This seems to have been a matter of convention as it came to be practiced by different groups.[23] When looking for a new or different design for a product, for one or other reason, it is natural to maintain some of the vital parameters associated with the components to be the same. The choice of the area of the *citi*s being retained the same is likely to have been in the same spirit. Presumably it would have been meant for balancing the

[22] Use of the converse of the Pythagoras theorem for obtaining perpendiculars on an extensive scale is unique to the Śulbasūtras tradition. It was earlier believed that the Egyptians used triangles with sides 3, 4, and 5 (in some units) for construction of perpendiculars, but this has now been discounted (see Gillings 1972, Appendix 5).

[23] As Thibaut puts in succinctly in Thibaut (1875, p. 232) (see also Thibaut n.d.-a, p. 145) "... according to the opinion of some theologians, the *gārhapatya* had to be constructed in a square shape, according to the opinion of others as a circle, the difference of the opinions referred only to the shape, not to the size, and consequently there arose the want of a rule for turning a square into a circle."

level of importance associated with the *citi*s.[24] We shall discuss this in some further detail in Sect. 2.6.4.

Whatever the genesis of the convention, it motivated the issue of transforming figures into one another, retaining the same area, and the methods introduced in this respect stand out as a major theme found in the Śulbasūtras. While some of the procedures are simple cut and paste prescriptions, others involve application of the Pythagoras principle. Transformation of a rectangle into a square with the same area may be one of the notable instances of such a geometrical application of this form[25]; see Sect. 5.5 for details.

2.6.4 Enhancement of Figures

It was well understood that for the figures drawn, scaling up a figure by a factor enhances the area by the square of the factor, irrespective of the specific design of the figure, and this played an important role in their constructions for redrawing the figures in larger sizes. In certain contexts, it was needed to redraw an earlier figure with an area larger by a given factor (see Sect. 5.1 for details), and for this they introduced the technique of scaling up the unit by a factor equal to the square root of the given enhancement factor for the area, and then following the same steps as before, with the new unit; producing the enhanced unit is achieved as an application of the Pythagoras principle; see Sect. 5.2.

In particular, for doubling the area, the diagonal of a unit square would be used as the enhanced unit; on account of this, the diagonal (and its measure) was called *dvikaraṇī* (one which doubles). There seems to be no direct application of this involved in the Śulbasūtras, but presumably it had a pedagogical role as the simplest case of enhancement by a factor. On the other hand, it is used in defining *trikaraṇī* (one which triples), as the diagonal of a rectangle with one side the *dvikaraṇī* and other side the unit. In turn, the latter is involved in defining *tṛtīyakaraṇī* (one which modifies by a factor of one-third), as the third of the *trikaraṇī*. These terms are actually involved, in particular, in describing the area of *Sautrāmaṇī vedi* in relation with the *Mahāvedi* (cf. BSl I.85–86).

Baudhāyana also has a numerical formula for $\sqrt{2}$ which is the doubling factor, as $1+\frac{1}{3}+\frac{1}{3}\times\frac{1}{4}-\frac{1}{3}\times\frac{1}{4}\times\frac{1}{34}$. In decimal expansion, this corresponds to 1.1424156 . . ., which is accurate to five decimal places (six significant digits). It has been a major highlight on Śulbasūtras, and subject of much discussion in the literature (see Sect. 8 for further details on this). In its numerical context, the magnitude is also given a separate name as *saviśeṣa*. The twin names for the entity seem to have led to some confusion in interpreting a sūtra from KSl, in Thibaut (1875) and some later works, which we shall discuss in Sect. 3.5.

[24] It is important to keep this perspective in mind, as in application it means that there should be no perceptible difference between the two areas, as far as the people around can judge, and there would be no emphasis on exact equality by any absolute standards, or even a quest for achieving it.

[25] This corresponds to Proposition 14 of Book II in Euclid's *Elements*; the arguments involved differ somewhat, though both use the Pythagoras principle.

2.6.5 The Circle

There are various *citi*s which involve circles in their design. The *gārhapatya agni*, which is a regular component of the setting of sacrificial events, was often circular in shape. There are also various *kāmya citi*s, *rathacakraciti* (chariot-wheel shaped), *droṇaciti* (trough-shaped), and *parimaṇḍala kūrmaciti* (rounded turtle-shaped) which have circles in their design. Their sizes followed the convention of having areas equal to the other *citi*s and, as such, were defined in terms of area, and not by specification of the diameter. In this context, we find procedures for constructing a circle with area (meant to be) equal to a given square.

For the converse problem, of squaring the circle, BSl gives two numerical formulae for the ratio of the side of the square to the diameter of the given circle, rather than a geometrical construction as given for the other way. One of these is a crude one, while another one is refined and elaborate, even though the accuracy achieved does not happen to be particularly good (see Sect. 7.6). Only the crude one is found mentioned in the later Śulbasūtras. The converse issue has no application or connection with their practical context, which could have led to the detailed formula of Baudhāyana not finding a mention later.

The elaborate numerical expression is apparently obtained as the inverse of the ratio that would correspond to the ratio in the other way transformation. While there is consensus among scholars on this point, the question of how it was computed, and why it is given as a combination of terms in the way it is (see Sect. 7.6), has been a matter of much debate and can hardly be said to have been resolved satisfactorily. We shall discuss some of the points that have emerged and highlight the difficulties in coming to any conclusion.

The computation of the reciprocal of the ratio as alluded to above involves understanding $\sqrt{2}$ in terms of familiar numbers, which seems to have inspired the formula recalled in Sect. 2.6.4. With regard to this formula also, there has been much discussion on how it was arrived at, though one may say it has not been as perplexing as the other number mentioned above. We recall some of the suggestions with regard to it in Sect. 8.

Apart from geometry, there are also some arithmetical and combinatorial aspects that turn up in the course of the constructions. We shall examine these in Sect. 9. There are often claims made with regard to the topic, of solutions of indeterminate equation being treated/considered in the Śulbasūtras. It turns out however that they are not well-founded (see Sects. 9.2 and 9.3). On the other hand, the works show good prowess in dealing with numbers which is noteworthy, as we shall illustrate in Sect. 9.

2.6.6 Proofs?

One question that often arises is whether the Śulbasūtras have proofs. While it is reasonable to have curiosity about such a question, in the context of the overall development, the question seems to have got needlessly loaded by extraneous considerations than properly as a question in the subject of the history of mathematics.

A simple and straightforward answer to the question would be a no. Seidenberg in Seidenberg (1975, p. 519) (see also Seidenberg 1978, p. 332) takes to task the writers who say that there are no proofs there, adding that "We can only suppose that these

writers have not bothered to examine the work," and goes on to give a "counter-example." It comes from the sūtra ASl 5.7. After noting the area of the *Mahāvedi* to be 972 square *padas*, it is mentioned in the original that by transporting a triangular part from the south of the vedi to the other side, a rectangle can be produced. To this part, Seidenberg adds another sentence, included within what is presented as a quotation, saying "In this form one computes its area."; no equivalent of such a sentence is found in the standard editions as in Sen and Bag (1983), Prakash et al. (1968a), and Kulkarni (n.d.), so one wonders where our author got it from – so much for careful examination of the text – and then (euphorically) adds "This is a proof. And there are others."; this same example is cited in Datta et al. (1980, p. 123), with no claim about others. How far such an occasional extra comment, thrown in to help the listener grasp the point, or just in a chatty vein, can be taken to be in the spirit of giving a proof, in the natural context of the query, is highly questionable.[26] At any rate, this is an isolated instance and a comprehensive survey of the literature to look for more would not make any dent in the overall status in the matter. The crucial point is that there are no proofs found at places where one would really wonder how they arrived at the statement being made, such as say the Pythagorean principle, squaring the circle, etc., nor does the general format suggest, in spite of the pedagogical import of the material, any inclination to convey also the reasoning involved, together with the material being communicated. In fact at many places, one finds statements such as "This is known," "This is the knowledge of areas," "This has always been prescribed by ancient teachers," etc., with emphasis on authority and trust in traditional knowledge, rather than reasoning or proof.

And why should we expect, in the first place, to find proofs there? A historian's task is to understand how things were, and see them in context. We know from the context that these are documents meant for equipping the people concerned with the aspects of brick-laying for erecting the sacrificial structures, people roughly in the role of architects. Nowhere do the Śulbasūtras give the impression of aiming to be fountainheads of all geometric knowledge of the time – unlike many other parts of the Vedic literature, they had a limited aim (and generally formed a chapter at or near the end). Why would they be bothered to give, to the architects involved, proofs for mathematical statements, beyond an occasional explanation of how the area of a trapezium could be computed by relating it to a rectangle? We need to see the matter in perspective. Beyond their limited objective, there is also the point that while there are various mathematical statements involved, they are essentially in the nature of facts about the physical world. Once discovered, by some wise men, the trust in them

[26] In Datta et al. (1980, p. 123), the authors also recall "after describing the constructions necessary in a proposition the early Hindu geometers are found to have remarked *sa samādhiḥ*" and liken it to the expression Quod Erat Faciendum in Euclid's *Elements,* hinting at a proof-like structure for the narrative. The comparison is however not valid, since the clause mentioned (or its variants) occurs only occasionally in the Śulbasūtras, apparently highlighting some specific points, and not in the nature of announcing accomplishing any preplanned objective, unlike in the latter case. The mention as noted is significant from a pedagogical point of view but does not represent adoption of any formal discipline in the narrative.

would be based on the statements being consistent with physical observations and experiences and did not warrant any additional "proof" in terms of reasoning. It may be recalled here that even as late as the twelfth century Bhāskarācarya commented that "proofs" were for the dim-witted!

None of this means that they did not *have* proofs. Arriving at various results of course needed a thought process. The wise men who found them had to have a way of convincing *themselves* of various things being true, even if no one else. A statement like that of the Pythagorean principle cannot come out of nowhere. Though it seems to have been known in more ancient times elsewhere, like to the Babylonians, no credible route of possible transmission from outside has ever been suggested or conceived. How did it come up then? One possibility that is often suggested is that it was guessed or extrapolated from the examples of the Pythagorean triples that are found in the Śulbasūtras (see Sect. 6.3). This reasoning however seems flawed. Knowing a Pythagorean triple, viz, a triple (a, b, c) with a, b, c integers such that $a^2 + b^2 = c^2$, does not by itself point to that it corresponds geometrically to a right-angled triangle – making the connection, or looking for such triples, already involves knowing the Pythagorean principle. At any rate, there are several geometric ways they could have been convinced of the validity of the Pythagorean principle, based on the kind of constructions they were used to; while these would not be on the same footing as the Euclidean axiomatic proofs, they would be adequate for carrying conviction of their validity; and it would constitute a *proof* which is valid in a way that we are quite used to in our day-to-day practices, though not in a modern mathematical framework. We shall not go into the details of the possible proofs here, but refer the reader to Dani and Limaye (2021), Datta (1932), and Divakaran (2018); the last one presents a series of possibilities.

3 Constructions for Basic Rectilinear Figures

In this section, we recall the constructions of various basic rectilinear figures as described in the Śulbasūtras. Exposition of these procedures are contained in the preparatory mathematical sections in BSl and ASl.

3.1 Construction of a Square by the "Compass" Method

Baudhāyana's course on geometry begins, after an introduction to various length measures and their interrelations (see Sect. 2.4), with a systematic construction of a square, a basic task that everyone connected with the subject should be able to carry out. The method, introduced in sūtras BSl I.22–28, may be described as follows. The original description is in terms of a cord of length equal to the side of the desired square, and ties made at its ends. It would however be convenient here to recall it in terms of abstract geometric steps.

One starts with a point, say P (see Fig. 1), which is to be the midpoint of the desired square, and a line segment AB, set along the east-west line, with P as its midpoint; it

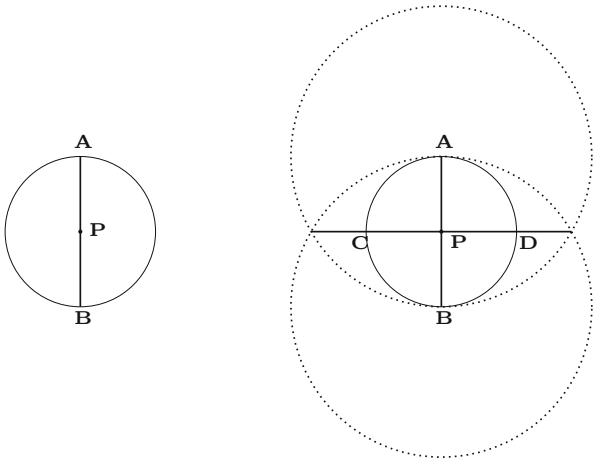


Fig. 1 First steps in Baudhāyana's construction of a square

may be borne in mind that A and B are to be the midpoints of the sides of the square on the east and west sides, respectively. A circle is drawn with P as the center, passing through A and B. The midpoints of the south and north sides are going to be identified on this circle. This is done by drawing two circles through A and B, passing through the other (B and A, respectively), and a line is drawn through their two points of intersection. This line is along the north-south direction, and the points C and D where it meets the first circle are the midpoints on the north and south sides. The vertices of the square are now identified as the points of intersection of circles drawn with adjacent midpoints (east and south, etc.) as centers, passing through P. It may be noted that the procedure essentially relies on the method for drawing a perpendicular to a line L at a given point P on it, by taking two points on L equidistant from P on either side, drawing circles through them with equal (sufficiently large) radius that would meet on either side of the line, and joining the two points of intersection. Here, however, specific choices for the circles are prescribed in terms of the length of the side, and the entire construction is presented as a package, a feature in which the exposition resembles a manual. In the light of the compass-like construction involved, we shall refer to this method in our discussion below as the compass method, though in their context it would be executed with poles and cords.[27]

[27] The original refers to "drawing circles," though the purpose involves only locating the relevant points by means of the intersecting segments of the circles. In literature, sometimes one finds quite adulatory comments on gratifying features of the figure that emerges from drawing all the full circles (e.g., Amma 1979, p. 29; Joseph 2016, p. 81); such comments however seem rather odd, as it would be pointless to draw the entire figure for getting the square, and one would be content identifying the relevant points of intersection! Surely for drawing nice figures they had other means and contexts!

3.2 Producing Perpendiculars by *Nyañchana* Method

The preceding construction is followed, in BSl I.30–35, by another method for construction of a square, which relies on the fact that in a triangle whose sides are in the proportion 3:4:5 the angle opposite to the longest side is a right angle. To construct a square of a given size, along the cardinal directions, one proceeds as follows: Taking a cord with twice the given length, say with endpoints P and Q as in Fig. 2, successively its midpoint R, the midpoint S of the half part RQ, and the midpoint X of RS are marked. The last one, which is called the *nyañchana*,[28] holds the key to the desired construction.

The *nyañchana* mark clearly divides the length of the cord in the proportion 3: 5. Tying the ends of the cord to poles at two points, say A and B, along the east-west line, at a distance equal to the desired side-length and stretching the cord on the ground, by holding it at the *nyañchana* mark, the corresponding point on the ground, C as in Fig. 2, is marked. This produces a right-angled triangle ABC, with the right angle at B. A similar one is produced with a right angle at A. Plotting the points on these perpendiculars at distance equal to half of AB, from the points A and B on both sides, produces the vertices of the desired square.

The reader would readily recognize that the fact used above about the proportion 3:4:5 is a particular case of the converse of the Pythagoras principle, and that (3, 4, 5) is a Pythagorean triple, viz, a triple (a, b, c), with a, b, c positive integers such that

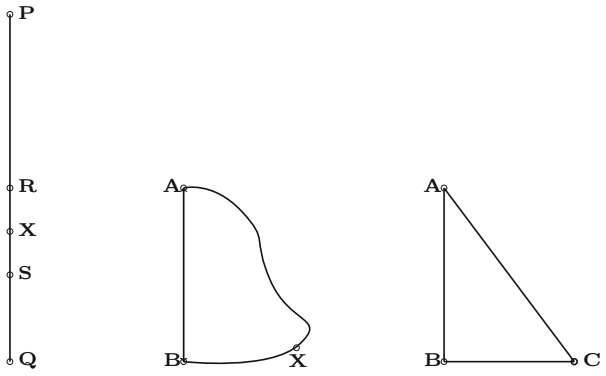


Fig. 2 The *nyañchana* method

[28] This term is reproduced here as in Thibaut's edition of BSl, which is also followed in Prakash et al. (1968) and Kulkarni (n.d.). In Sen and Bag (1983), the authors use *nyañcana*, which is a later-day Sanskrit equivalent of *nyañchana*. Etymologically the word is seen to have its root in the verb *nyañc*, which means "lying with face downwards." In MSl and KSl, it is called *nirañchana*. Curiously, in ASl though use of such a mark is involved in similar constructions, it is not given any name but is only referred to as "the mark" (*lakṣaṇaṃ*). For convenience, we shall uniformly use the term *nyañchana*.

$a^2 + b^2 = c^2$. An analogous method for drawing perpendicular using the Pythagorean triple (5, 12, 13) is adopted in a construction of rectangles that follows, in BSl I.42–44. In this instance, to draw perpendiculars to a segment PQ as above a cord with $1\frac{1}{2}$ times the given length is used, and on the additional part the *nyañchana* mark is made at $\frac{1}{6}$th from the separating mark, thus dividing the cord in the proportion 5:13. Tying the ends of the cord and stretching it by holding at the *nyañchana* mark produces the requisite perpendiculars; the construction is then completed by marking the vertices at the appropriate distances on them.

3.3 Perpendiculars by the "Arrow" Method

Interestingly, a method for producing perpendiculars which relies on the same geometric principles as the compass method, but is operationally different, is described in the course of the construction of rectangles and isosceles trapezia, in BSL I.36–40 and BSl I.41, respectively. Here, to draw a perpendicular to a line L at a point P, two points on L equidistant from P on the two sides are taken, say Q and R; see Fig. 3. A cord of length longer than that of QR is then tied to poles at P and Q, and points are plotted on the two sides of L by stretching the cord, holding it at its midpoint – joining the two points produces the desired perpendicular. The rectangle is drawn by plotting two points on the east-west line at the desired distance between the transversal sides, drawing perpendiculars to them by the method as above, and plotting points on them from the east-west line at half the size of the desidered width. For the isosceles trapezium, adopting the same procedure, the vertices are drawn at half the distance, from the east-west line, as the desired sizes of the base and face, respectively.

We note that while it shares its geometric basis with the compass method, it does not involve tracing the pairs of circular loci and taking points of intersection, but only on stretching a cord and marking a point on it on the ground. Equality of the distances of points as required is achieved through midpoint marks on the cords.

In the Āpastamba Śulbasūtra, the method is involved in the procedure described, in ASl 1.13–1.18, for construction of a square; here the method is further elaborately

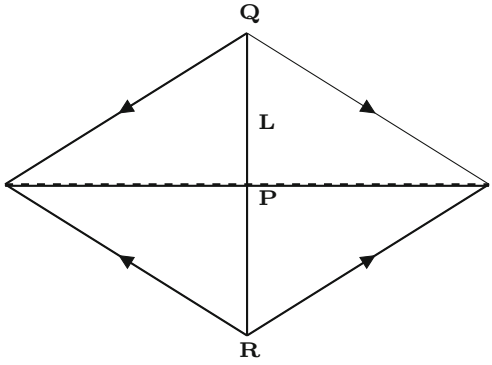


Fig. 3 Drawing perpendiculars by the "arrow" method

packaged, involving repeated application, to plot the vertices of the desired square. It may be worth recalling here for its interesting "packaging" feature.

Taking a cord of the given length a for the side of the square, markings are made at its midpoint, and the midpoints of the two halves. Poles are set along the eastwest direction at points A, B at distance a, at the midpoint M of AB, and at points P and Q which are midpoints of AM and BM, respectively; see Fig. 4. Tying the ends of the cord at P and Q, the midpoint on it is marked, after stretching on the ground, at R as in the figure. Next, tying both the ends of the cord to the pole at M and stretching the cord passing over R, the midpoint of the cord is marked, as C. In the same way, the point D is marked on the other side. Thus, C and D are the midpoints of the sides of the desired square, apart from A and B. Now, tying the ends of the cord at A and C and marking the midpoint, after stretching the cord on the ground (away from the midpoint M), yields the vertex in which the sides with A and C as midpoints meet; similarly the other vertices are determined.

Like in BSl, in ASl also an initial portion of the text is devoted to exposition of theory, and one finds interesting asides, as seen in the above construction. ASl however does not repeat the elaborate construction of the square by the compass method described in the beginning of BSl, as in Sect. 3.1. This does not amount to any loss in terms of ability of the intended learners to draw a square, for which an alternative, and relatively slick, method is provided, but with that the exposure to the compass method and the geometric understanding associated with it would no longer be a part of the knowledge of the learner.

3.4 Toward Conceptualization of Perpendicularity

It should be clear from the above discussion that in the underlying thought process, Baudhāyana implicitly had the idea of a perpendicular to a line at a given point,

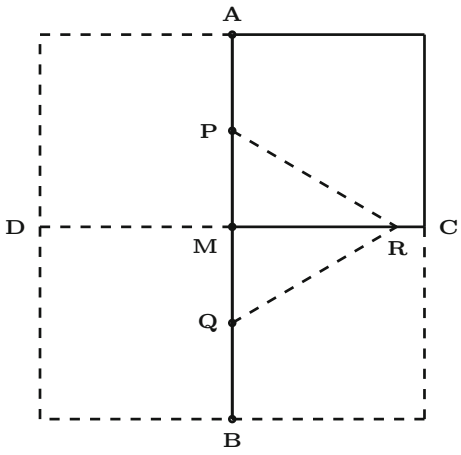


Fig. 4 A step in Āpastamba's construction of a square

together with various ways of constructing it. However, somehow the notion of perpendicular does not get formulated into a mathematical concept; or if it did in his mind, it was considered too abstract for the intended audience.

The idea of conceptualizing the perpendicular seems to have eluded the later Śulbasūtras as well, and nowhere does one find a direct instruction to draw a perpendicular (by first having a name for it) to a line at a given point; it comes only as a part of elaborate instructions for drawing the familiar *shapes* such as the squares, rectangles, and more generally, trapezia, typically with reference to the given dimensions of the figure individually. The closest they seem to have come in this respect is, as I pointed out in Dani (2010, p. 19), to use a common clause in certain contexts: The *vedi*s described in ASl 6.3–6.4 (*nirūḍhapaśubandha vedi*), ASl 6.6, ASl 6.7 (*paitṛkīvedi*), ASl 6.8 (*uttarāvedi*), and ASl 7.1 involve varied shapes (two trapezia of different dimensions, two squares, and a rectangle), but the construction of each of them refers to "Having stretched (the cord) by the mark at fifteen" (*pañcadaśikenaivāpāyamya*), and from the context it is clear that the clause refers to a process of drawing a perpendicular.

One may wonder what "stretching the cord by the mark at 15" has to with drawing a perpendicular. The answer is not far to seek. It may be recalled that one of the methods discussed, using the converse of the Pythagoras principle with the triple (5, 12, 13), consisted of using a (*nyañchana*) mark dividing the cord in the proportion 5:13. Actually it was more common to think in terms of its threefold multiple (15, 36, 39) because of its association with the *Mahāvedi*. Recall that the *Mahāvedi* is in the shape of an isosceles trapezium with a base of 30 units, height of 36 units, and face (side opposite to the base) of 24 units (see Sect. 2.5.1). Thus, the (right-angled) triangle formed by one of the vertices at the base together with the midpoints of the base and the face has sides 15, 36, and 39, and this is in fact used in one of the constructions of the figure; see Sect. 3.5. On account of this association with the *Mahāvedi,* the triple seems to have had an importance by way of tradition.[29] In this case, the cord involved would have the measure $54 = 15 + 39$ units, with the *nyañchana* mark at 15; thus, stretching by the mark at 15 is the key step in this way of drawing the perpendicular, and lent the name for the act of performing it, as mentioned earlier, in Sect. 3.2. Āpastamba does not have a term corresponding to *nyañchana*, so the name here directly refers to the mark at 15. This *operational* standardization led to associate "stretching by the mark of fifteen" with the process of drawing a perpendicular.[30]

[29] It may also be noted here that the triple is included in Baudhāyana's list, even though it is in a way redundant, being a multiple of another triple in the list, (5, 12, 13); see Sect. 4.1.

[30] It may be mentioned here that the altitude of an isosceles triangle (the segment joining the midpoint of the base to opposite vertex) had a name, *avalambaka*; mention of the term is found in MSl, including in particular in MSl 3.2.10. Etymologically its root is *avalamba*, which means "hanging down." The segment would readily have been known to be "perpendicular" to the base (say in terms of being like the sides of a rectangle), and this may well relate to the fact that in later Sanskrit literature the term *lambaka* stood for perpendicular, in a broader context. The connection however does not seem to have led to any insight toward the issue of conceptualizing the perpendicular, as discussed above.

3.5 Āpastamba's Varied Applications of *Nyañchana*

Āpastamba also takes forward the idea of using the converse of the Pythagoras principle, deploying other Pythagorean triples, as also through a triple like $(1, 1, \sqrt{2})$. Availing of the context of the *Mahāvedi,* he introduces various ways, in ASl 5.3–17, of constructing it, using the *nyañchana* method.

In terms of Fig. 5, starting with the east-west line segment of 36 units the trapezium ABCD is constructed, in various ways, as follows: The first construction which he notes to be an *ekarajjuvidhi* ("one-rope process") uses a cord of length 54 units, with a mark at 15 (as the *nyañchana* mark), to determine the vertices A and B by drawing the triangles EWA and EWB (see Fig. 5); the same marked cord is used to produce the direction of ED, and the vertex D itself is produced through a mark at 12 on the cord.

We note that the point M where the diagonals of the trapezium meet is at distance 16 units from E and 20 units from W. Availing of this, Āpastamba introduces a method for plotting A and B by drawing the triangles AMW and BMW by *nyañchana*, using the triple (15, 20, 25), and similarly C and D are plotted by drawinngs triangles ECM and EDM, using the triple (12, 16, 20).

The third construction is based on the triples (5, 12, 13) and (15, 36, 39); with P chosen as the point on the segment EW at distance five units from E, the vertices C and D are obtained by drawing the triangles ECP and EDP by *nyañchana*, with (5, 12, 13); the point P is of course of no significance to the main figure itself. The vertices A and B are obtained using (15, 36, 39) as in the first construction.

A fourth construction involves the triples (8, 15, 17) and (12, 35, 37); see (Dani 2010, pp. 16–17 for details). Taking a point Q on the segment EW at a distance 8 units from W, the vertices A and B are produced by drawing the triangles QWA and

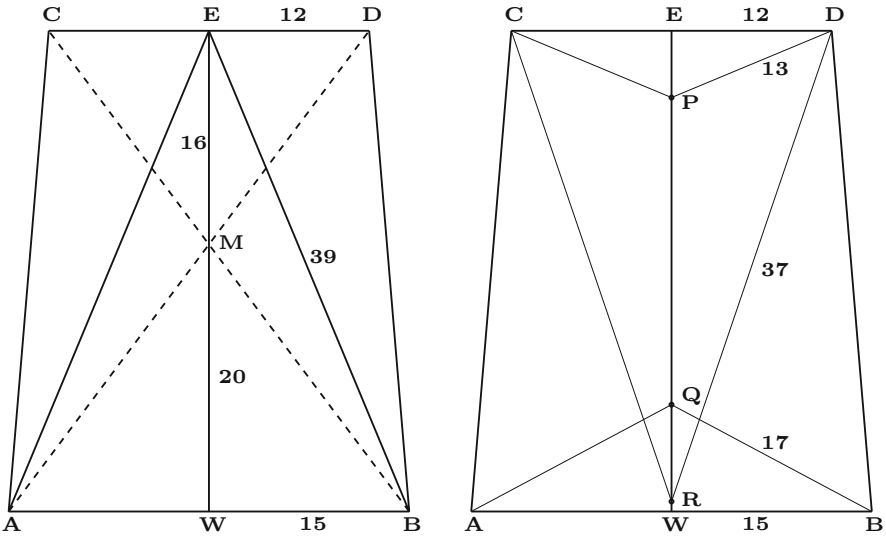


Fig. 5 The *Mahāvedi*

QWB with the triple (8, 15, 17), and with a point R on EW at a distance 35 units from E (equivalently, distance 1 from W), vertices C and D are obtained by drawing the triangles REC and RED using the triple (12, 35, 37).

It may be noted that having multiple ways of constructing the perpendicular had no practical relevance (nor has any been claimed in the original), but it is fun enough, and would serve as an exercise to enhance comprehension of the learner, of the basic idea involved.

There is also an application found in ASl 2.1–5, in a similar spirit, involving use of the nonintegral triple $\left(1, 1, \sqrt{2}\right)$. For this, a cord of length $1 + \sqrt{2}$ units would be used, with a *nyanchana* mark dividing it in the proportion $1 : \sqrt{2}$ (see Dani 2010, pp. 18–19); the mark for $\sqrt{2}$ on the cord would have to be preconstructed; thus, in effect one is replicating a previous construction of right angle, but this is very meaningful from a practical point of view. Unfortunately there seems to be a rather serious misunderstanding relating to this construction, going back all the way to Thibaut (1875), and persisting through some later works (cf. Heath 1926, p. 361, where the author refers it to papers of Bürk), and Seidenberg (1975, p. 511 – footnote 66). Quoting the sūtras ASl 2.1–5, Thibaut gives a translation as "Fix poles on both ends and the middle of the prishthya line, add to a cord of half the length (of the prishthya) its visesha, *i.e. its third plus fourth part of the third minus the thirty-fourth part of that fourth part*, . . . " (emphasis added). By recalling a meaning of visesha (*viśeṣa*) here, in the part noted with emphasis, the author means to be helpful, but it is misleading. First, it may be noted that the part does not correspond to anything in ASl 2.1–5, and the meaning is brought in from elsewhere, viz, ASl 1.10. While the latter gives a numerical value for *viśeṣa*, that term and *saviśeṣa* were also understood in terms of marks on a cord. In the context of the Śulbasūtras, where the thinking was largely geometrical, that would be the relevant meaning, and not the numerical value. In absence of any specific evidence to that effect, it would be incorrect to impute, and criticize them for it, that they were using the cumbersome way of adding indiviaually the components as described in the numerical value.[31]

3.6 Other Rectilinear Figures

One notable figure other than squares, rectangles, trapezia, and triangles appearing in the Śulbasūtras is of a rhombus. Interestingly, it figures also as the shape of a *citi*. In Baudhāyana's geometry such a figure, called *ubhayataḥ prauga* (literally "both ways triangle"), is constructed (cf. sūtra 1.57) starting with a square, by what we may call a cut and paste method, which we shall discuss in a Sect. 3.8. This method of

[31] As an illustration of the point being made, let me point out that if I tell my student to add e to an expression, even though by definition $e = 1 + \frac{1}{2} + \frac{1}{6} + \cdots$, I would not mean that she should add the terms $1, \frac{1}{2}, \frac{1}{6}, \cdots$, one by one; depending on the context, a readymade approximation may be used, or perhaps the expression absorbs adding of e in another way.

construction of a rhombus is also described in KSl (cf. sūtra 4.6). In either of the cases, the diagonals of the rhombus are in the proportion 2:1 (with the longer side along the east-west direction). Tiling of the *citi* in that shape in terms of rectangular and triangular tiles is described in BSl III.172–173; it is not indicated in the sūtra how the rhombus is constructed, but it is generally presumed that this is by the procedure taught in BSL I.57.[32] It is also described in ASl, but with a different tiling configuration, with most bricks also in the shape of a rhombus; the others are quarter portions, or halves in a few cases, of a rhombus, cut along the diagonals. MSl also mentions, in MSl 3.6.4–5, the construction as above (applicable to the outline) and aludes to the *citi,* but there does not seem to be any detailed description of the tiling design.

In KSl, the description in KSl 4.6 for transforming a square into a rhombus is followed by an analogous transformation into triangles, and also the converse aspect of transforming various figures to a square, which we shall take up in Sect. 3.8. There is no detail of what pentagons are meant, but the later commentators have postulated that it was in the shape of *haṃsamukhī* bricks (see the next paragraph).

Shapes of the bricks tiling the *citi*s would be another source of figures with which they were engaged. Actually in the construction of the *citi*s, most of the bricks are of square shape. However, they also used bricks whose shape is a triangle (arising as two of four parts of a square, cut along one or both diagonals); these are referred to as *ardhā* and *pādyā*. Note that in the former there are two sides equal to that of the original one while in the latter there is a unique one with the property. Triangular bricks arising from similar division of rectangles are also involved in the construction of certain *citi*s (specifically, in the tiling of *śyenaciti* described in BSl 119–124). Apart from these, they also used bricks of shapes which consist of adjoining with rectangular half-piece of the original square to one of these triangular shapes along a side equal to that of the original square. In the case of the *ardhā,* this yields an asymmetical trapezium, called *adhyardhā,* while with the *pādyā* it gives a pentagonal shape, called *haṃsamukhī* (see Sen and Bag 1983, pages 190–191 for various details in this respect – there however seems to be some confusion involved with regard to the sizes of the sides in Fig. 37 on page 191). The triangular and pentagonal bricks are used in producing the desired pointed parts in the outlines of the desired figure, such as for the tail or the beak of a falcon, as also for adjusting the total count of bricks in each layer, which was stipulated to be 200; we find many places in the layer designs where two or more triangular bricks are used in a layer covering a square, evidently for the purpose of enhancing the number of bricks laid in the layer. We will have an occasion to dwell on this point again in Sect. 9.3.

Interestingly the construction of *śyenaciti* in ASl involves bricks in the shape of a parallelogram, used in the wing part of the *citi* (see ASl, Chapters 15 – 20 in Sen and Bag 1983). There is however no description of how the parallelogram figure was constructed geometrically. The sizes involved being 24 *aṅgulas* (18 inches) and

[32] As pointed out in Seidenberg (1975, p. 330), use of the Pythagorean principle comes directly into play in this, unlike in the case of *citis* like *śyenaciti* or other complicated figures involving bricks of many shapes, in the construction of a square of area 15 (square) *puruṣa*.

20 *aṅgulas* (15 inches), the shape was presumably achieved by the masons through visual adjustments to fit the inclination desired for the designs involved in specific contexts. It may be noted that there is no notion of angle, or of a quantitative notion of slope in terms of ratios of sides etc., appearing in the Śulbasūtras. Achieving a given inclination would therefore be in terms of matching with given pair of lines set at an angle between them (say in the form of sticks clamped to each other suitably, or a metallic frame, etc.); the statement involved in item 9 in Datta et al. (1980, p. 130), listing construction of "a parallelogram having given sides and a given inclination" in the Śulbasūtras, needs to be read in this light.

3.7 Dividing Line Segments into Parts

Various geometric constructions described in the Śulbasūtras involve division of line segments, or figures such as rectangles, into a certain number of (equal) parts[33]; typically Śulbasutras mention only "parts" and the equality is implicit, and sometimes not mandatory. How a segment may be divided into two equal parts with the help of a cord is obvious; taking a cord (or a marked part on a cord) with the given length, one can simply fold it over and mark its midpoint. How this was done for larger number of parts is not clear. In Seidenberg (1978), Seidenberg mentions that at his request Frits Staal enquired with a priest performing a *yajña*,[34] and the reply was that they would fold the cord that many times over, adjusting it along the way. Seidenberg (1978, p. 336) comments that they might have solved the problem transforming the square with the given length as the side, into a rectangle whose one side is n times that, where n is the number by which division is sought; the other side will give the desired length-size of each of the equal parts. However, as he notes, there is no word on this, and it seems unlikely that they would have done it in that way.[35]

It is also conceivable that the length of the segment, taken in specific units, was numerically divided, converting into smaller units at each stage of remainders, down to the smallest unit, or one small enough for the context; that it would not be mathematically accurate is unlikely to have made a difference to them – the Śulbasūtras practices are seen to be quite open to approximations, making this a plausible explanation.

[33] Producing the *savišeṣa* (see Sect. 2.6.4) involves subdivision into 34 parts, and the procedure for the converse operation described involves dividing into 29 parts; these could however be theoretical or notional subdivisions. For more concrete examples, one may recall that in the construction of various *citis,* sides of squares or triangles are needed to be divided into a variety of number of parts like 7 (*Gārhapatya*, in BSl II.64), 12 or 13 (*praugaciti*, in ASl 12,7–11), etc.. Also, application of *nyañchana* with the triple (5, 12, 13) involves division of the additional part of the cord into six parts. It may also be recalled that the procedure for enhancement of citis of $7\frac{1}{2}$ *puruṣa citis* size into larger ones (cf. Sect. 5.2) involves division into 15 parts.

[34] Apparently a Vedic style *yajña* was performed during Staal's stay in India, and it was filmed by him.

[35] In Datta et al. (1980, p. 131, footnote 2), it is suggested that the division "was probably done by drawing parallels, as in Euclid"; the authors refer to sūtras concerning division of triangles as indicating that "surmise to be correct." However, the reasoning, which has not been elaborated, seems to be circular.

One may similarly ask about the division of a circle into equal radial parts. Division into two parts would result from plotting a diameter, and into four would follow from two diameters perpendicular to each other. They could also presumably bisect angles by considering the isosceles triangle formed by the two corresponding radii and taking the ray joining the midpoints of its base to the origin. In particular, it would be possible to divide the circle into 2^n parts for $n = 2, 3\ldots$. The design of the version of the *Rathacakraciti* with spokes involves radial division into 32 parts, to fit 16 spokes, with gaps of the same size as the spokes (see Sect. 9.4 for details). Whether the requisite division was indeed done by such a geometric process or by visual adjustments is not clear from available inputs. Also, division into six radial parts (and hence also three parts, and further into $3 \cdot 2^n$ parts for $n = 1, 2, \ldots$) of the circle would have been possible by drawing equilateral triangles; equilateral triangles, and the procedure for producing it, are found in the construction of the *Dārśapaurṇamāsikāvedi* (see Sect. 9.1). There are no instances of radial divisions of the circle outside these possibilities; it is unlikely they seriously considered the general issue.[36]

3.8 Transformation of Figures by Elementary Methods

One finds considerable preoccupation in the Śulbasūtras with transforming figures from one shape to another, retaining the same area, especially transforming a square into a figure with a desired shape and conversely transforming a given figure to a square shape with the same area. The methods introduced, though simple in many cases, have an intresting pedagogical dimension and have been discussed in Dani and Limaye (2021) in that light. In the case of rectilinear figures, the methods proposed are exact, while in the case of figures like circle or semicircle the methods

[36] In Joseph (2016, Sect. A.3.1.1, p. 79), there is a bulleted mention to the effect that Śulbasūtras have construction(s) to "Divide a circle into any number of equal areas by constructing diameters," with parenthetic reference to BSl II.73–74 and ASl 7.13–14; as is well-known, an accurate division of this kind is not possible geometrically, but even if one grants that approximate divisions in the overall spirit of the Śulbasūtras are meant, the above statement involves gross inaccuracies; the author has apparently relied, unfortunately, on some secondary source material aimed at lay public, not conforming to scholarly standards. First, the division in either of the instances cited is specifically into nine parts (and not "any number"), and second, and more importantly, the parts are not suggested or meant to be radial; the *Dhiṣṇya citis* which are the subject of BSl II.73–74 can be in the shape of a square or a circle; for the square version, the division is into nine square parts and is by a three-row grid, and the part in the center marked for being substituted by a stone in place of a brick; while the brief sūtra does not specify the division (in either case), in the case of the circle the context does not fit a radial division at all, and traditionally in the circular version the corresponding ninth part is taken to be a concentric circle with specified (eight *aṅgulas*) diameter, and the remaining annular part is divided into eight equal radial sectors (see Prakash et al. 1968, p. 101; Sen and Bag 1983, p. 184; Kulkarni n.d., p. 44); and the central part does not quite equal the rest. The sūtras ASl 7.13–14, in the other citation, are not about division of circle, at all, but of a square inscribed in a circle; the sutras concern tiling of the circular *Gārhapatya*, into 21 tiles, with 9 tiles in the inscribed square and 3 each in the four segments outside the inscribed square (see Prakash et al. 1968, p. 100; Sen and Bag 1983, pp. 183–184; Kulkarni n.d., pp. 43–44).

are approximate; these will be discussed in later sections. Some conversions involve the Pythagoras principle which we shall describe in Sect. 5.

In the descriptions below, we shall assume that the figures have their axis of symmetry along the east-west direction.

3.8.1 Transforming a Square into an Isosceles Trapezium

Given a square, an isosceles trapezium having the face of a desired size less than the side of the square is instructed, in BSl I.55, to be produced as follows: Mark out a rectangle at one end in the square along one of the sides, with the other side of length equal to the excess over the desired size of the face, cut the rectangle into two along the diagonal, take away one of the parts, and join it on the other side (the diagonal is chosen so that the shorter side that would be produced is toward the east).

3.8.2 Transforming a Square into an Isosceles Triangle

This could be done as a special case of the previous construction, but there is a different method described in BSl I.56, giving a diffrent triangle, with the base of the isosceles triangle equal to the diagonal of the square (and not twice the side of the square as one would get by the other method). A square is formed with the diagonal of the original square as the side. The isosceles triangle is formed by joining the vertices of one side to the midpoint of the opposite side; the latter would be situated toward the east.

3.8.3 Transforming a Square into a Rhombus

For this purpose, the sūtra BSl I.57 describes the following construction: Given a square, a copy of it is adjoined on the western side, forming a rectangle with the east-west sides twice the size of the north-south sides; the midpoints of the sides of the rectangle are joined cyclically to get the desired rhombus. (The way this is formulated, it is not just "cut and paste," on account of the introduction of a copy of the square to be adjoined, but it is in the same spirit and may still be considered a cut and paste method.)

3.8.4 Transforming a Square into a Rectangle: A Special Case

BSl I.54 is a sūtra for transforming a rectangle into a square with the same area, which involves usage of the Pythagorean principle, and will be discussed later in Sect. 5.5. Ahead of that, there are two sūtras on the issue of transforming a square into a rectangle, in the spirit of considering the converse issue. The resolution is based on elementary considerations in the same way as the above construction. Of the two sides of a rectangle with a given area, one could be chosen freely, so the converse issue as above needs an additional input of choice of the size of one side. BSl I.52 deals specifically with the case where one side of the rectangle is chosen to be the diagonal of the given square. The rectangle in this case is instructed to be produced by cutting the square along one of the diagonals and cutting one of the parts into two halves along the other diagonal; the three pieces can then be reassembled into a rectangle, with one side equal to the diagonal and the other

side half the diagonal. The same rule is also given KSl 3.4; it is not found in ASl and MSl.

Constructing the rectangle from a square having the diagonal as one of the sides would have been a novelty (and in fact may have motivated the *dvikaraṇī*) and hence explained systematically, while the practical case which comes next, with simple proportion of sides, may not have seemed to warrant an elaborate treatment.

3.8.5 Transforming a Square into a Rectangle: General Case

In the next sūtra BSl I.53, Baudhāyana moves on to the general question, but the response seems rather casual or informal. The instruction simply says "move the extra part to the side(s)." There have been many comments to the effect that the solution is unsatisfactory (see Prakash et al. 1968a, p. 56; unintelligible according to Seidenberg 1978 – p. 334; etc.). A slightly more elaborate instruction on the issue is found in ASl 3.1 (Kātyāyana does not discuss the general case, though the special case as above is included – MSl does not deal with the issue). Thibaut (1875, p. 246) comments "Āpastamba's rule helps us somewhat further"; this author however does not find any substantive difference. Rather than fully interpreting the sūtra in the light of the clarity claimed to have been attained, Thibaut then proceeds to discuss a special example, taking the side of the square to be 5 and the prescribed side of the rectangle to be 3, and describes (geometric) steps arriving at the other side to be $8\frac{1}{3}$ (as to be expected).[37]

Keeping the Śulbasūtras context in mind, where apart from imparting geometrical content at theoretical level, they aimed informing about steps to be taken in practical contexts, the following seems a plausible explanation, for both BSl I.53 and ASl 3.1, though the wording is not clear enough in either case; the explanation is for square with integer sides, which is the case they would have aimed to explain in these sūtras, for the practical contexts. In a way, the explanations of Thibaut and Bürk fit with this general version. Given a square with side m and the prescribed side for the rectangle being n (both integers – describing measures with respect to some unit), the procedure would be as follows: For simplicity, we shall assume $n < m$; the other case can be reduced to this by elementary considerations, which we shall not go into. We divide the quare into two rectangles one with sides m and n and the other with sides m and $m - n$. It now suffices to transform the second rectangle into one with one side n, so that it can be put alongside the first to get the desired rectangle. For this, we divide the side of size m into mn equal parts and the side of length $m - n$ into unit size parts and form a corresponding grid of rectangles. The resulting $mn(m - n)$ rectangles, with sides 1 and $\frac{1}{n}$, can be reassembled to get a rectangle with side n as desired, by stacking them into $m(m - n)$ columns of n each, alongside each other. More efficient plans involving fewer parts would be possible, but this is a

[37] As cited in Sen and Bag (1983, p. 157), Bürk also gives, in his 1902-paper (cf. Bürk 1901; the page number given in Sen and Bag 1983 seems incorrect, and this author has not accessed the original paper) a similar illustration with a square of side 6 and the prescribed side of the rectangle to be 4, leading to the other side being 9.

straightforward scheme which would be accessible to them, at least for small values of the integers that they would be concerned with. These kind of rectangular grids are witnessed, for instance, in the plan for tiling for the *gārhapatya agni* (cf. BSl II.64–65); similar grids are involved also in the tilings of *ubhayataḥ prauga citi*, the rhombus-shaped *citi* (cf. ASl 12.12–13).

4 The Pythagoras Principle

In Baudhāyana's scheme, the next topic to learn is the theorem now known after Pythagoras. It ought to be borne in mind that the notion of a theorem in the sense of Euclidean geometry, deduced from a set of axioms, etc., is not involved in the context of the Śulbasūtras. The statements made in the Śulbasūtras are in the nature of stating facts about the physical world, and no particular distinction was made between statements coming from empirical knowledge (those that could serve as "axioms" and those that may be deduced), though undoubtedly some of the statements would have emerged through a deductive process or cogitation, rather than through empirical observation.[38] It would therefore be more appropriate to discuss it as a principle known to them rather than as a "theorem," and we shall generally adopt the nomenclature accordingly. For convenience, we shall however continue to associate the name of Pythagoras, though the latter would be a later figure historically. In this section, we shall discuss the contents of the Śulbasūtras related to the principle.

The Pythagoras principle is stated in all the four Śulbasūtras, and it would be worthwhile to go over the formulations and the nuances involved in the different Śulbasūtras in some detail.

4.1 The Formulation in Baudhāyana Śulbasūtra

In BSl, the presentation on the topic begins with a sūtra, BSl I.45, which in close translation corresponds to "the diagonal of a square makes double the area"; it is implicit that the comparison is with the original square, and a linear size *making* a certain area means that the latter is the area of the square whose side is the given linear size. This is presumably meant as a warm-up toward the statement of the Pythagoras principle which is to follow, after one more intermediate observation. Based on the preceding observation, Baudhāyana refers to the size of the diagonal of the square as "*dvikaraṇī* (the doubler); the term is not introduced in the form of a definition, or nomenclature, but its usage in the next sūtra and linguistic considerations make this obvious.

The next statement is "when the transverse side is one unit and the stretch (*āyāma*) is its doubler, the diagonal is the tripler." While the first instance about

[38] An attempt is made in Datta (1932) to place the Śulbasūtras geometry on an axiomatic footing, but the effort does not seem to have led to any further development on the theme.

the diagonal of a square would be visually clear to a learner from a tiling, this second one is not so. Baudhāyana goes into one more digression, stating that one also gets from this a side-length which produces a third part of the area, by dividing the tripler (length) into three equal parts, and the division of the sides into three equal parts divides the square into nine equal parts. Following the preliminary observations, we come to the statement of the principle, in BSl I.48, which may be translated as:

> The diagonal of a rectangle makes the area that is made separately by the transverse and flank sides.

This is unmistakably a statement of the Pythagoras principle, though it is made in the framework of rectangles, and not right-angled triangles.[39] I may mention that the term used for the diagonal is *akṣṇayārajju*, which literally means "the cord that goes across."

The sūtra containing the statement of the principle is followed by one (BSl I.49) which states:

> Three-four, twelve-five, fifteen-eight, seven-tentyfour, twelve-thiryfive, fifteen-thirtysix, are pairs in which this is obtained.

The pairs of numbers described clearly correspond to the first two entries of the Pythagorean triples (3, 4, 5), (12, 5, 13), (15, 36, 39), (15, 8, 17), (7, 24, 25), (12, 35, 37), and (15, 36, 39). There is no clue as to how the triples were arrived at. They are apparently given as examples illustrating the principle, without involving surd magnitudes. The inclusion of the triple (15, 36, 39) here, though it is just a multiple, is very likely related, as commented in Sect. 3.5, to its association with the *Mahāvedi*.

4.2 The Statement in Āpastamba and Kātyāyana Śulbasūtras

While Baudhāyana comes to the statement of the Pythagoras principle in steps, ASl on the other hand goes straight to it, after introducing the methods for constructions of squares that used the converse of the principle, with the triples (3,4,5) and (5,12,13); the special case of the square is stated after that statement, and *dvikaraṇī* is introduced through it. The statement in ASl is the same except for a minor difference that the rectangle is referred simply as *dīrgha*, rather than *dīrghcaturasra* as in BSl.

Following the statement of the Pythagoras principle, ASl has a sūtra amounting to a remark that the constructions of squares described earlier are based on the knowledge of this principle. The clause involved is *tābhirjñeyābhiruktaṃ viharaṇaṃ*, and it translates as "the constructions noted above are known

[39] In certain books, one finds a statement that BSl gives only a restrictive statement of the theorem, and there is a complete one in ASl, which is not true; the statements in the two, as we shall see, are essentially equivalent.

(or knowable, or deducible) from this" (Kulkarni n.d.).[40] Such a remark is rather unique, with no analogous interconnections found elsewhere in the Śulbasūtras. It may be observed that strictly speaking the constructions are based on the *converse* of the Pythagoras principle. On the other hand, the converse statement, though used extensively in various constructions, is not stated explicitly in any of the Śulbasūtras. Apparently the converse statement was conflated with the statement of the principle itself, and the remark as above supports such an inference. The converse statement can of course be deduced from the principle through a simple argument, and visually it may have seemed to them a part of the statement itself.

KSl recalls the principle, as well as *dvikaraṇī*, *trikaraṇī, and tṛtīyakaraṇī* introduced in Sect. 2.6.4, in the second *kaṇḍikā* (Chapter) along with some aspects of construction of the *śamyā vedi*. The statement of the principle is same as in BSl. It is followed by the clause *iti kṣetrajñānaṃ*, "this is known about areas." One may wonder about the significance of such a clause. On the one hand, its being included especially in this context may signify some importance being associated to it being conveyed. On the other hand, it also suggests that the principle was being transmitted to the successive generations, providing no elaboration or reasoning toward its validity.

4.3 The Formulation in the Mānava Śulbasūtra

MSl, on the other hand, has a different take on the statement of the principle. The relevant sūtra there, MSl 3.1.10, translates as

> Multiply the length by the length, breadth by the breadth; it is known that adding them and taking the square root yields the diagonal.

It is noticeable that unlike in the other Śulbasūtras here the statement is not in terms of areas, but rather the lengths of the sides. Also, it is now seen to be presented more directly as a procedure for computing the size of the diagonal in terms of the sides. It is however not possible to ascertain whether these features reflect anything of significance, by way of change in mathematical practices, from a historical perspective. The concerned sūtra appears isolatedly, in a section with an assortment of others relating to bricks and brick-laying of the vedis, rather as one of the things that is especially selected for being remembered, and the scope for comparisons is limited on account of this.

[40] In Sen and Bag (1983), the authors recall (page 234) that according to Bürk the statement means that "the construction has been taught by means of these, of course, by means of such as are recognizable"; later authors have interpreted "recognizable" as referring to recognizable numbers (Sen and Bag 1983) or specifically to Pythagorean triples (Heath 1885, p. 146). Use of the verb "to know" in the clause relates the construction to the statement of the principle, and reference to "recognizable" numbers or triples is apparently based on a misunderstanding.

5 Applications of the Pythagoras Principle

The Pythagoras principle is applied in BSl and KSl for the construction of squares whose areas are prescribed in various ways, as we shall next discuss. In BSl, this part follows soon after the statement of the Pythagoras principle.

5.1 A Square Whose Area Is the Sum of Two Given Squares

The sūtra BSl I.50 describes how to construct a square whose area equals the sum of two given unequal squares; for equal squares, the issue is resolved by *dvikaraṇī* – the case could have still been included with the construction that is described, as given below, but is presumably excluded for the sake of maintaining clarity of some of the steps, where specializing could seem unnatural, or lead to some confusion.

The resolution, as may be expected, consists of plotting a rectangle with sides equal the sides of the given squares and taking the diagonal of the rectangle for the side of the desired square. In the sūtra, the rectangle is prescribed to be plotted within the larger square, based on one of its sides, marking the side length of the smaller one on the other side. See Fig. 6: PC is obtained as the side for the sum of squares whose sides are PA and PB.

The same procedure is described in ASl 2.10–12 and KSl 2.22. Mānava Śulbasūtra does not mention the proposition – it is however involved, and implicitly assumed to be known, in one of the applications of the proposition relating to enhancement of sizes of figures, which we discuss in the next subsection.

5.2 Enhancing the Size of Figures by Scaling the Unit

A rather unique practical application of the construction described in the last subsection, of combining two squares into one, is found in the Śulbasūtras; it is involved in MSl as well (in implementing the instruction as in MSl 3.3.6) though the

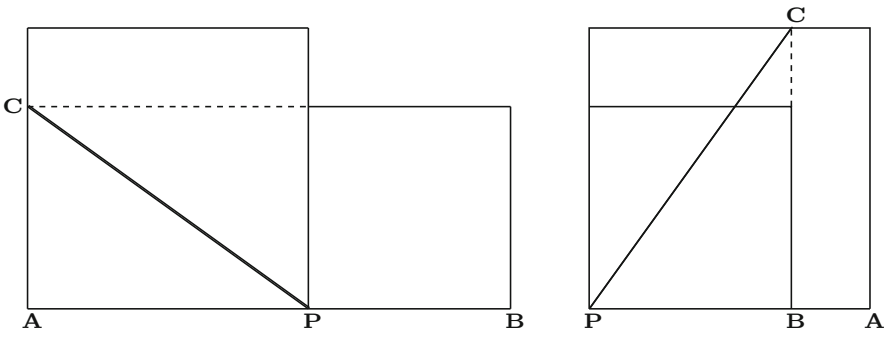


Fig. 6 The side for the sum of two squares

statement itself has not been recalled; apparently it is assumed as known; and the sūtra itself however pertains to a rather specific situation and is rather terse to be discussed here any further.

There was a certain practice or context to which this application has been linked. Certain *kāmya yajñas* were performed repeatedly, and there was a stipulation that on each successive occasion the *citi* should be in the same shape, but with area enhanced by one unit; the starting one would be $7\frac{1}{2}$ *puruṣa*, the second $8\frac{1}{2}$ *puruṣa*, the third $9\frac{1}{2}$ *puruṣa,* and so on (BSl II.1–5).[41] As the same figure was to be drawn each time, this was achieved by simply enhancing the basic measuring unit, as marked on the cords used, suitably, and following the same steps as before. Thus, to increase the area from $7\frac{1}{2}$ units to $8\frac{1}{2}$ units, the area needs to be enhanced by a factor of $\frac{2}{15}$, for which one would replace the original unit by the size of the side of the square obtained by augmenting the unit square by a square of size $\frac{2}{15}$; in this respect, BSl II.12 instructs that the unit area (*puruṣa*) be divided into 15 (equal) parts, 2 of which should be added with each unit area in the figure, to produce the enhanced figure of $8\frac{1}{2}$ *puruṣa*s.

Seidenberg (1975, p. 492) postulates that this stipulation may have been the motivation for the discovery of the above construction, and in turn of the discovery of the Pythagoras principle itself. While this may fit into the overall theme of "ritual origin" of ideas pursued there, there is no evidence to that effect. In fact it is quite in Indian tradition to incorporate modifications in the "requirements" for various rituals as newer possibilities arise through newly acquired knowledge. Thus, it is more likely that until some point in time the stipulation may have been only that on each successive occasion the area should be bigger, and the precise quantative version was introduced only after the principle, and the possible construction came to be discovered, in one or other general context. A few sūtras down the line from those mentioning the augmentation as recalled above, in BSl, there is a sūtra (BSl II.18) which says that according to some Brāhmaṇas, after performing a *yajña* with a *citi* of a certain size one should not perform it again with a *citi* of a smaller size; in terms of the overall requirement from a ritualistic perspective, this seems to be the key point, and the precise prescriptions would have been filled in when there were people around who had an idea of how to do it; it was accordingly followed by some Brāhmaṇas, which in this instance includes Baudhāyana's.

Curiously there is a variation in the application of the scaling scheme described above found in the Śulbasūtras, which seems to have gone without much notice in the literature. In the Āpastamba Śulbasūtra, in ASl 8.6 where the issue is dealt with, we see a variant as follows: It instructs to add $\frac{1}{7}$th (square) *puruṣa* per iteration to the unit square to get the new unit in the manner as above (the number of iterations is conveyed in the sūtra by words equivalent to "what is added to seven"). This in effect amounts to

[41] As described by Baudhāyana (BSl II.1–4) the sequence is supposed to go up to $101\frac{1}{2}$ *puruṣa*; one may wonder whether they would actually have continued that far, or it is simply a metaphor of some kind; anyway up to where it went does not make a difference to the mathematical content and will not concern us here.

adding $1\frac{1}{14}$ (square) *puruṣa* to the net area at each iteration. In the Kātyāyana Śulbasūtra, while some of the instructions on the issue of enhancements (specifically, those appearing in *kaṇḍikā* (Chapter) 5) are in conformity with Baudhāyana's formulation, in KSl 6.2 (in *kaṇḍikā* 6) this is seen not to be the case: For the *ekaśatavidha* (101-fold), *citi* the (original) *prakrama* unit is instructed to be replaced by the side of a square with area $14\frac{3}{7}$ (square) *prakramas*, thus implying that the area should be arranged to be $14\frac{3}{7}$ times that of the original *citi*. It can be seen that this does not match with the general scheme involving addition of one *puruṣa* to the area at each iteration, described earlier; the area works out to about 108.2 *puruṣa*s, considerably in excess of $101\frac{1}{2}$ *puruṣas* associated with the *citi*. On the other hand, it is seen to be in agreement with what Āpastamba has prescribed, as $94 \times \frac{15}{14} + 7\frac{1}{2} \approx 108.2$, to which the factor is adjusted by Kātyāyana in the sūtra. Apart from ASl 6.2, the subsequent sūtras ASl 6.3–5 also involve division by 7, and not $7\frac{1}{2}$. This seems to have been an independent practice, or tradition, than the scheme presented by Baudhāyana.

This alternative version has not received due attention in the literature. Notably, in particular, in the translation of the sūtra ASl 8.6 given in Sen and Bag (1983) the authors state "... should be divided into *seven and half equal parts* and each part added to each *puruṣa* ..." (emphasis added). This is however incorrect. There is no mention of "seven and half" in the original sūtra[42]; (the translation in Kulkarni n.d. (correctly) translates as division by 7). Having missed that instance, in the commentary in Sen and Bag (1983) on the sūtra KSl 6.2 (on page 270), where they notice the difference, the authors treat it, with some apparent puzzlement, as an exceptional instance, which "does not appear to be correct" or is a simplification.[43] It may be noted that in the alternative scheme the area added at each iteration is not one *puruṣa*, but $1\frac{1}{14}$ *puruṣa*. How much enhancement is to be made is after all a matter of convention, and variant practices may not be deemed incorrect.

Before concluding this discussion, a point may be made here, based on these observations, and also supported by other instances in the Śulbasūtras that we need not go into, that the prescriptions about the dimensions, etc. did not strictly follow any sacrosanct injunctions of permanent nature but were open to adjustments by individual teachers.

5.3 Kātyāyana's Method for Putting Together Multiple Squares

We recall next a method from KSl for constructing a square with area equal to an integer multiple of a given square. Note that the issue is tractable by repeated application of the method of augmentation of squares, described in Sect. 5.1; given a number, say n, of squares of a fixed size, to combine them, rather than joining one

[42] Perhaps the error goes back to Bürk's edition of ASl, but this author is not in a position to check this at this point.

[43] They also recall an explanation from Datta (1932, pp. 166–168) which seems contrived and misses the point.

square each time, n may be expressed as a sum of squares, as $n = \sum_{j=1}^{k} n_j^2$, and squares with sides $n_1, \ldots, n_k$ may be joined successively to get the desired square.

Kātyāyana describes another method, a more direct one, which also depends on the Pythagoras principle. For later reference, I include the details of the sūtra; see also Dani (2018).

The sūtra in question, KSl 6.7, translates as the following statement.

> To produce a square putting together a number of squares of a fixed size, construct an [isosceles] triangle with one less than the number as base and one more than it as the other two sides [together] and take the altitude of the triangle for the side of the desired square.

The isosceles triangle involved would be constructed by taking a cord of length equal to the two slanting slides together, with a mark at its midpoint, tying its ends to poles at the vertices at the base, of the desired size, and stretching the cord on the ground, holding at the midpoint mark. The prescription is of course based on the fact that a slanting side together with the corresponding half the base and altitude forms a right angled triangle, to which one can apply the Pythagoras principle, together with an identity which in symbolic terms may be expressed as

$$n = \left(\frac{n+1}{2}\right)^2 - \left(\frac{n-1}{2}\right)^2,$$

where n is the number of squares to be put together. There is no evidence in the Śulbasūtras of symbolic algebraic calculations, as involved here, being carried out. The identity as above would have been recognized by observing the square tilings, and realizing the difference of the squares matched along two perpendicular sides, in terms of the rows that are left on two sides.

5.4 A Square Whose Area Is the Difference of Two Given Squares

The procedure for this is explained in BSl I.51. For this, actually a right angled triangle (half of a rectangle along the diagonal) is constructed whose hypotenuse equals the side of the larger square and one of the sides is the side of the smaller square. The specific prescription is to take a rectangle, as in the preceding construction, and identify on the newly added side the point at the desired distance; see Figure 7.

In the case of longish rectangles, when multiple squares of the size of the shorter size can be taken out, Kātyāyana suggests that a square equivalent to those may be put together with one for the remaining rectangle.[44]

[44] The sūtra described in Sect. 5.3 appears later in the work, and whether it is meant to be used here is not apparent. While evidently the sūtra is intended as simpler or more elegant than the one suggested above for the general case, the precise details are not clear, and the explanations in Kulkarni (n.d.), following (van Gelder 1963), do not seem to make it clear enough.

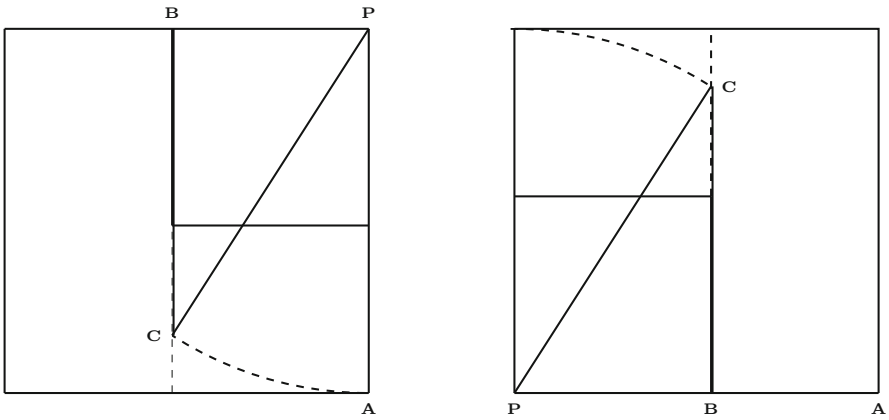


Fig. 7 The side for the difference of two squares

Unlike in the case of augmentation of squares, the result on producing a square equivalent to the difference of two given squares does not seem to have been applied directly in practice. We see, on the other hand, it being used in deriving a further consequence, a procedure for transforming a rectangle to a square, which was probably used in the constructions of some of the *citi*s (see below).

5.5 A Square with the Same Area as a Given Rectangle

Given a rectangle to be transformed into a square, it is first turned into a figure which is a difference of squares (a gnomon), by cutting of half a strip along the shorter side with width half the difference between the sides, and placed along the other side, aligned with the other shorter side (BSl I.54), as illustrated in Fig. 8. Having achieved that, Baudhāyana simply refers to the preceding construction, saying "resolution of this has been explained"; this may perhaps be the earliest case of back reference of a mathematical kind!

The process of transforming a rectangle into a square would have been needed in the construction of various *citi*s. Constructions of the *droṇaciti* (*citi* in the shape of a trough) and *parimaṇḍala kūrmaciti* (a rounded turtle shape) as described in BSl (sūtras BSl VII.2–3 and X.2–3, respectively) involve designating a certain part of the prescribed area to be turned into a circular region, and using the remaining area for other parts of the figure, a handle for the trough in the first instance and legs and head in the second; specifically, the prescribed area of $7\frac{1}{2}$ *puruṣa* is instructed to be divided into 120 square-brick-sizes (with sides 30 *aṅgulas* each), and the area of 119 and 115 of them, respectively, to be transformed into circular regions. It is not specified in the respective sūtras how this was meant to be accomplished, but in the overall context it seems reasonable to assume that first a rectangle would be constructed with the respective areas designated for the circles (say as 17×7 unit

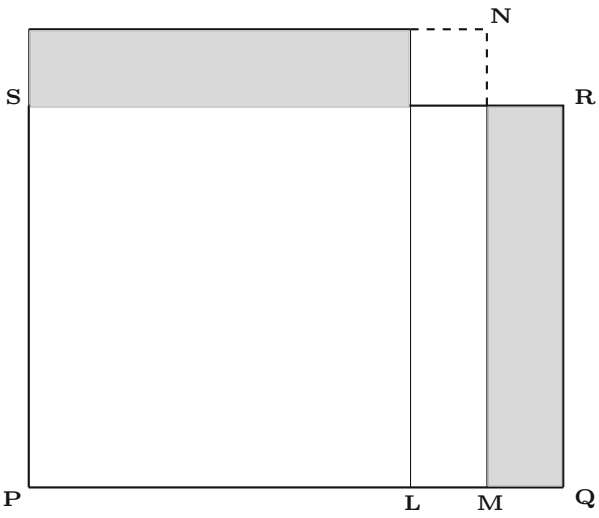


Fig. 8 Transforming a rectangle to a square

sizes in the former case and 23 × 5 unit sizes in the latter case), transfoming the rectangles into squares by the method described above, and then transforming the squares to a circles by the method which we describe in Sect. 7.2 below.

6 Miscellaneous Topics Around the Pythagoras Principle

This section is devoted to a variety of topics related to the theme of the Pythagoras principle, not covered in the earlier sections.

6.1 Square Roots

The methods described in Sects. 5.1 and 5.3 incorporate in particular means of finding square roots of integers, as geometric magnitudes. Rarely do we find numerical values associated with them; a rather unique case of $\sqrt{2}$ for which there is an impressive numerical expression, and some other plausible cases of numerical values associated with square roots, will be discussed in Sect. 8; we also point out that there are certain instances discussed in literature as computations of square roots in the Śulbasūtras are not quite so.

The major preoccupation in the Śulbasūtras relating to roots was about measuring out magnitures on ground which are specified by area, rather than their linear dimension. A simplest example of this, of determining a square whose area is two (square) units, was achieved by choosing for its side the diagonal of the unit square, viz, the *dvikaraṇī* as may be recalled. This is a special case of the Pythagoras

principle which has been specifically mentioned in BSl I.45, ASl 1.10–11, and KSl 2.12. The term *saviśeṣa* also appears in relation to $\sqrt{2}$; there seems to be a slight syntactical difference in the usage of the terms *dvikaraṇī* and *saviśeṣa*, with the former signifying a more geometrical entity and the latter a numerical one (scalable with the unit, nevertheless), but the distinction does not seem to be exclusive. The excess of *saviśeṣa* over the unit is referred to as *viśeṣa*, and the term appears in particular in ASl 2.2.

The side of a threefold square was also accorded a name, *trikaraṇī*, which is also mentioned in all Śulbasūtras (BSl I.46, ASl 2.6, and KSl 2.14). This is again constructed geometrically as the diagonal of a rectangle whose one side is the *dvikaraṇī* and the other side is one unit. Interestingly, there is also a name for the square root of $\frac{1}{3}$, *tṛtīyakaraṇī*; it was produced by dividing *trikaraṇī* into three parts, or by taking *trikaraṇī* of $\frac{1}{3}$ of the unit (BSl I.47, ASl 2.7–8, and KSl 2.16–18).

6.2 Semicircle with the Same Area as a Circle

We will be taking up aspects of mensuration of the circle in the next section. Here we only note an application of the idea of enhancement of areas as discussed in the last subsection to the circle, found in MSl 1.1.8. It may be recalled that the *dakṣiṇāgni* was semicircular; while *gārhapatya* was circular and *āhavanīya* a square, they were to have the same area. The abovementioned sūtra, which describes the constructios for the three *citi*s together, instructs the semicircular one to be obtained from the circlar one as follows: construct a square circumscribing the circle and draw a semicircle with its midpoint as the center, passing through its vertices; thus, for the new radius the area of the circle will be twice the original, so the semicircle will have the same area as the original circle, as desired.

6.3 Pythagorean Triples

We recalled earlier (in Sect. 4.1) that BSl mentions the Pythagorean triples (3, 4, 5), (12, 5, 13), (15, 8, 17), (7, 24, 25), (12, 35, 37), and (15, 36, 39) (the last entry of each triple being implicit). It is a specific list, and we do not know how it was arrived at. There is also no hint about more such triples being known to the author (such as through usage of an equivalent of "etc."). ASl does not present any list of such triples but illustrates the use in construction of the *Mahāvedi* (see Sect. 4.1), of the triples as above, other than (7, 24, 25).

The interest of the sulba geometers in the Pythagorean triples stemmed from their application in producing the perpendicular, through the *nyañchana* method; the reader is referred to Dani (2003) for a detailed discussion on this topic. Functionally, only one, the simplest one, viz, (3,4,5), would have sufficed their purpose. They also found the triple (12, 5, 13) and typically used it in the *nyañchana* procedure with appropriately changed markings. These triples are found used in all the four

Śulbasūtras. Driven by curiosity, a few more triples seem to have been found, which BSl notes as examples, following the statement of the Pythagoras principle, and ASl illustrates in terms of construction of the *Mahāvedi* in multiple ways. It is not clear how they might have arrived at them, and the small sample that is available tends to make any general hypothesis on this rather too speculative for our discussion here; see however (van der Waerden 1983) for some observations in this respect. It would have been noticed that the newer triples offer no particular advantage for usage with the *nyañchana*, and the procedure for requisite divisions, for the *nyañchana* markings, becomes more and more cumbersome as the sizes of the entries grow bigger. It is therefore unlikely that they would have had any further interest or motivation in finding more Pythagorean triples. And indeed no other triples, other than scalar multiples of the ones seen above, turn up in the Śulbasūtras explicitly.

It may be noticed however that the identity as in Sect. 5.4, noted by Kātyāyana, could be used to produce a sequence of Pythagorean triples, one for each odd square number n. This has been sometimes treated as a sequence of Pythagorean triples from KSl (cf. Datta and Singh 1962, II, pp. 204–205). It should be borne in mind however that Kātyāyana's sutra recalled in Sect. 5.4 has a different context, not involving Pythagorean triples, or solving indeterminate equations, and no connection is made in the original with the triples associated with the *nyañchana* procedure, nor is there any mention of some of the triples involved in the *nyañchana* procedure being particular cases of these. The restriction of choosing n to be the square of an odd integer, needed to get Pythagorean triples, is also not alluded to.

6.4 The *Mahāvedi* and a Thesis of Seidenberg

Let us recall that the *Mahāvedi* is in the shape of an isosceles trapezium whose altitude is 36 units, and the base and face are 30 and 24 units, respectively. The *Śatapatha Brāhmaṇa* and *Taittirıya Saṃhitā* give these dimensions explicitly, in terms of steps (strides) by the *yajamāna* as the unit; specifically, as described in Sen and Bag (1983, p. 170), following Eggeling's translation from *Śatapatha Brāhmaṇa* the description involves the following, with respect to Fig. 4: the *yajamāna* arriving at point W, the midpoint of the base, from a post to its west, taking 15 steps to his right where he drives a peg for the right hip, at the point B in the figure. Returning to W, he then takes 15 steps northward where he drives the peg for the left hip, at the point A. From W, he then strides 36 steps eastward and there drives the peg, at E, for the fore part. . . [the left and right shoulder points C and D are described similarly].

It may be noted that the triangle BWE has sizes (15, 36, 39) which is a Pythagorean triple, and that this fact is used in Āpastamba's construction of the vedi with what was called the *ekarajjuvidhi* (see Sect. 3.5). In Seidenberg (1978, p. 323), Seidenberg raises the question "did the priests at the time of *Śatapatha Brāhmaṇa* and the *Taittirıya Saṃhitā* know this Pythagorean triangle?" He goes on to note also that the figure "is loaded with" other Pythagorean triples (as have been used by Āpastamba in the other constructions), and arguing that this would be a major miracle if they were not to know the Pythagorean triples, and concludes that

"Aspect II was known at the time of *Taittirıya Saṃhitā*" which then forms a prop of his theory of a common predecessor[45] ("Aspect II" we may recall, without going into the details, concerns the Pythagorean triples).

Actually there does not seem to be much of a case in such an argument. First, though the author mentions all of Āpastamba's five Pythagorean triples (12, 16, 20), (15, 20, 25), (5, 12, 13), (8, 15, 17), and (12, 35, 37) in his reference to being "loaded with," in the last three the third point is actually manufactured, and not intrinsic to the description; Āpastamba was aware of this and makes a reference to a separate mark on the cord. The case for the "major miracle" is thus already overstated. More importantly the argument does not take into account the actual construction described in *Śatapatha Brāhmaṇa*, as recalled above. If the choice of the dimensions was motivated by knowledge of the Pythagorean triples, why would the procedure described be only in terms of taking steps left, right, or east? That would be a greater puzzle.

It seems to me that while we have here a very interesting choice of dimensions for an isosceles trapezium, it only depends on the familiarity with the triple (3, 4, 5), and the idea of scaling which was very prevalent. The trapezium is actually made up from vertices of two isosceles triangles, ABM and CDM as in the figure, one put over the other upside down. Both the triangles are scaled-up versions of an isosceles triangle with base 6 and altitude 4, and the halves of this by the altitude are (3,4,5) triangles. This seems a natural way of constructing isosceles trapezia, starting with an isosceles triangle. Scaling would be used to adjust to the desired overall shape. In this case, we see the triangles to be in the proportion 5:4, a fairly natural choice, and it yields the dimensions of the *Mahāvedi* as altitude 36 units, and the base and face 30 and 24 units. The fact that (15, 36, 39) is also a Pythagorean triple may not have been noticed until later, and even if it had been noticed, that would not have made much of a difference without being familiar with the Pythagorean principle.

This also explains why no geometry turns up in the construction described in *Śatapatha Brāhmaṇa.* While the figure was arrived at by "playing with" the (3,4,5) type triangles, the right angles themselves have gone into the background. It would then have seemed simpler, and also with a greater ritual appeal, to measure the requisite units than work through the two component triangles involved in the construction.

[45] The inference as above is used to deduce that the Indian knowledge from the Vedic period is not a derivative from the Greek tradition. In the current context however, such an argument would not be needed, as the period of the Śulbasūtras has got pushed back in the general discourse, and there is no serious contention in the field about Śulbasūtras knowledge being derived from the Greek tradition. The conclusion thus arrived at, that the Vedic knowledge of geometry is not a derivative of Greek geometry, is then used to make a case for a common predecessor. The present author has serious reservations to the theory of common predecessor as well; however, this will not be an appropriate place to go into detail about it.

7 The Circle

It may be recalled that drawing of circles was already involved in the construction of rectilinear figures; in those contexts, only certain segments were relevant, but the description is in terms of drawing circles, and the term *maṇḍala* for circle made an appearance there, right at the beginning of BSl's geometry course. The same term is found used in all the four Śulbasūtras throughout.

In the preliminary parts, the construction of a circle is described in terms of placing a pole at a point, to which a cord is to be tied, drawing the circle following a mark made on the (stretched) cord as specified in the individual contexts; the point at the pole would be the center, and the distance of the mark from it along the cord would be the radius of the circle; neither center nor the radius is referred to by any name. There was a term used commonly for the diameter, *viṣkambha*, which stands also for "width" in a wider context; in respect of the circle, the term was used as a reference measure of the size.

In the later practical contexts of construction of *citi*s however, the circles are mostly produced by transforming squares into circles (there are some notable exceptions – see below for details.). In this way, a circle would be constructed with the area desired for the context. The construction for this, which is explained in all the Śulbasūtras is described in Sect. 7.2 below. This aspect of the Śulbasūtras has been the subject of much discussion in the contemporary literature on the Śulbasūtras. It has sometimes been likened to the problem of doubling the cube, associated with the legend of the oracle of Delphi, in the Greek tradition (see Seidenberg 1975, 1978, 1981), as a motivating objective. In this context in Sect. 7.3, we shall go over the details of how the issue unfolds in the Śulbasūtras, and comment on the thesis as above, in the light of the observations.

The problem of squaring the circle, namely the converse of the process as above, of transforming a circle to a square, is also considered in Śulbasūtras. For this, Baudhāyana gives, rather than a geometric construction, an intricate numerical factor for the ratio of the side of the desired square to the diameter of the given circle. This will be discussed in Sect. 7.6.

There are two notable exceptions where circles, or rather arcs of circles, appear as part of the design of *vedi*s that are not constructed by transforming a square of the desired area into a circle. One is the *dakṣiṇāgni* mentioned earlier, which is in the shape of a semicircle (bounded by the diameter). As it is to have the same area as the *gārhapatya* which is circular, it is obtained simply by constructing a circle with twice the area as the latter (a scaled-up version in the same overall spirit as discussed in Sect. 5.2) and taking half it; the revised unit for this is produced by joining the end points of a quarter of the circle.

Another exception in this respect occurs in the construction of a *vedi* called the *Dārśapaurṇamāsika vedi*. The design involved is rather unique, in that it involves circular arc segments for its contours. It consists of first constructing an isosceles trapezium, with face, base, and altitude 48, 64, and 96 *aṅgulas,* respectively, and marking out circular arcs joining the pairs of adjacent vertices, lying within the

trapezium, the latter define the boundaries for the activities for the *yajña*. We shall discuss this *vedi* in some detail in Sect. 9.1.

7.1 Circumference of a Circle

The notion of circumference of a circle does not make any significant appearance in the Śulbasūtras, and in particular it does not feature in the geometrical expositions in BSl and ASl. There is however a passing reference to it in BSl, in the later part, in sūtras BSl I.112–113. This is in the context of describing the setting for the *Aśvamedha yajña*, which involves a row of circular pits. The pits are described to have 1 *pada* diameter (*viṣkambha*) and 3 *padas* circumference (*pariṇāha*). While that twofold part of the description is of no consequence there, and is simply made as an elaboration, it indicates that the ratio of the circumference to the diameter was taken to be 3, as was the case in many ancient cultures.[46]

There is however a sūtra in the Mānava Śulbasūtra, MSl 3.2.13, marking a departure from the belief. We recall the original sūtra here for reference to some of its parts below.

Viṣkambhaḥ pañcabhāgaṣca viṣkambhastriguṇaśca yaḥ |
sa maṇḍala parikṣepo na vālamatiricyate ||
A fifth of the diameter and three times the diameter; that is the circumference of the circle – not a hair-breadth remains.

Here *maṇḍala parikṣepa* is the term used for the circumference, it may be noted that it is different from the term *pariṇāha* that we came across in BSl, indicating that there was no uniform common term adopted for it yet. The last part *na vālamatiricyate* translates as "not a hair [breadth] remains." The ratio of the circumference to the diameter of a circle is thus assigned an upper estimate of $3\frac{1}{5}$, actually rather on the high side, and there is palpable jubilation over its not missing even a hairbreadth! The value is not seen recalled anywhere later however, whether in a practical context, or by way of dissemination.[47]

[46] In Dani (2022), I made a suggestion that this may have to do with the fact that fractions came into human thought process much later than circles – on account of wheels – and once an integer value was assigned to it, for various reasons the ratio may have come to be adhered to until much later, even after fractions became a part of the overall discourse.

[47] It may be recalled here that, on the other hand, the value $\sqrt{10}$ adopted for the same ratio by the Jaina scholars, from a somewhat later period from the same millennium BCE, had a long life of over a 1000 years and was used subsequently in theoretical and practical computations across many cultures; see Dani (2023) for details.

7.2 Construction of a Circle with a Given Area

Unlike the circumference, the area of a circle turned out to hold substantial interest for them. Indeed, there are many *citi*s, as we shall recall in some detail in the next section, involving construction of circles covering a given area. In implementing this, naturally a square would be an appropriate reference figure; we have seen earlier that conversion of rectilinear figures to a square was widely discussed. After obtaining the area in the shape of a square, it is to be converted to a circle using the procedure which we now describe; having described the procedure in his preparatory sections, Baudhāyana assumes that it will be applied accordingly wherever it is needed, and there is no specific recalling or back-reference mentioned at any later point.

In BSl, this is taken up immediately after the rectilinear geometry parts discussed in the earlier sections, in BSl I.58. Given a square, to get a circle equivalent to it Baudhāyana's prescription goes as follows: Take half the diagonal of the square and drop it from the center along the eastward direction, and draw the circle including a third portion of the part jutting out (see Fig. 9, with $\mathrm{QS} = \frac{1}{3}\mathrm{QR}$); it is implicit that the square is drawn along the cardinal directions and the half-diagonal is along the south-east or north-east directions.

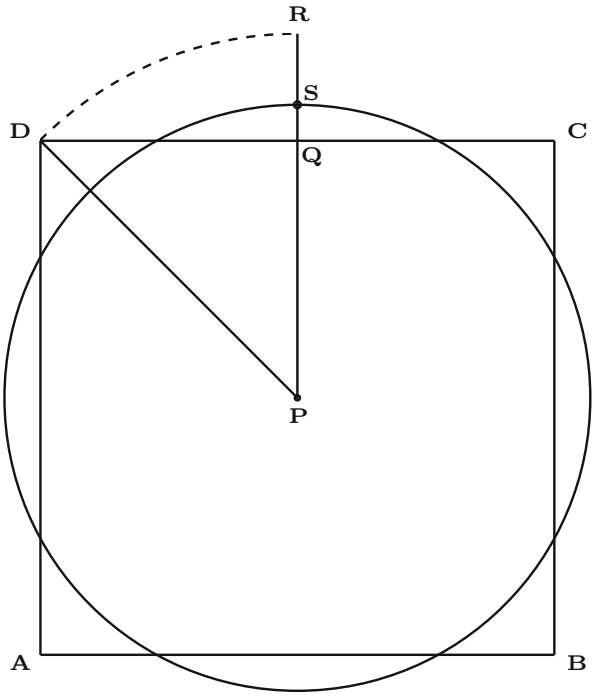


Fig. 9 Constructing a circle with the same area as a given square

For a square with side $2a$, the prescribed radius works out to be $a + \frac{1}{3} \times \left(\sqrt{2} - 1\right)a = \left(2 + \sqrt{2}\right)a/3$. For the unit square $\left(a = \frac{1}{2}\right)$, the area of this circle works out to be 1.01725…, a little more than that of the original square. If the circle is to be unit circle, then the initial square has to be correspondingly smaller and its area may be seen to be 3.0883…; the latter would thus "correspond" to the value in their context to the familiar constant π – it should be noted however they were not assigning any numerical value to the ratio involved, but only giving a construction of the desired circle in their context.

Realizing that the length of diameter of the circle would be between that of the side and the diagonal square, they would have looked for a suitable proportion of the extra length over the side to it. They could rule out half as unsuitable in this respect; one way to see this, at an intuitive or visual level, would be to see how much of the perimeter of the square lies inside. An easy calculation (with our present means) shows that for that choice over $\frac{3}{4}$th of the perimeter lies inside. This could have been easily noticed (even if not in precise terms), and considered far too much if the portions lying inside and outside are to be equal. For the next simple choice for the proportion to be $\frac{1}{3}$ however, the ratio of the segments lying inside and outside is significantly close to 1, the former being about 17% larger, but this may have seemed okay, keeping in mind that the corner points are farther from the circle than the midpoints of the sides.[48]

I may also recall here, as first pointed out in Dani (2010), that if they were to choose the proportion to be three on ten, a meaningful option in the context of their familiarity of the decimal system, they would have got much better results; however, they could not have anticipated it. Besides, although the decimal system was commonly used for integers it was seldom involved in fractions, or proportions.

Āpastamba gives the same construction for the task at hand (in ASl 3.2–3), though in somewhat different words. Interestingly it is then followed up (in ASl 3.4–5) with the comment "*sānityāmaṇḍalam | yāvaddhīyate tāvadāgantu |*." The second part means "as much is left out that much comes in." For the first part, there have been two alternative interpretations in the literature. One of the commentators Karavindasvāmī interprets the lines as "The circle is exactly as large as the square, for as much the circle falls short, so much comes in," while another commentator Kapardisvāmī breaks up *sānityā* as *sā anityā* and asserts that the circle construction is an approximate one (cf. Sen and Bag 1983, p. 161). Though linguistically both the interpretations are possible, it seems certain, as argued in Thibaut (1875, see

[48] Divakaran (2018) contains a suggestion on how the choice of $\frac{1}{3}$ as the factor may have been made. However, it seems far-fetched to believe that calculations as proposed would have been made at the time of Śulbasūtras. Second, the argument assumes implicitly that they knew that the ratio of circumference to the diameter, of a circle, and that of its area to the square of the radius to be the same, which has no justification. Besides, the argument fails to take note of the fact if the equality is assumed, and the ratio was taken to be 3, as has been recalled while discussing the details, they would not need Baudhāyana's construction as above to produce a circle with area as a given square – it would suffice to take the diameter of the circle to be $\frac{2}{\sqrt{3}}$ times the side of the given square, as seen to be implicitly involved in the Mānava Śulbasūtra (see Sect. 7.8).

pp. 252–253), that Āpastamba meant that the circle has the same size as the square – claiming no allowance as an approximation; without that, the subsequent part would be incongruous, as it does not quite make sense to say "this is an approximate construction; as much is left out so much comes in."[49]

The method has also been recalled in KSl 3.13; the sūtra seems a little less condensed than in BSl and uses some different wording, indicating an active retelling of the principle. MSl, on the other hand, has a different take on the issue. We shall defer the discussion on it to Sect. 7.4 and discuss next the application of the procedure as above in the construction of various *citi*s.

7.3 Circles Appearing in Construction of *Citi*s

It may be recalled that the three *agni*s, *āhavanīya*, *gārhapatya,* and *dakṣiṇāgni,* formed an integral part in the performance of all *kāmya yajña*s. There is detailed description of their relative locations in the overall plan, given in BSl I.63–69.

Of these, the *gārhapatya agni* could be either of square shape or circular, typically the latter (though not always, contrary to a prevalent belief in this respect), and the *dakṣiṇāgni* in the shape of a semicircle.[50] There are also other *citi*s, *rathacakraciti* (in the shape of a chariot wheel), *droṇaciti* (in the shape of a trough), and *parimaṇḍala kūrmaciti* (a round turtle shape), described in BSl, involving circles in their design; in the case of *gārhapatya,* the area is prescribed to be one (square) *vyāyāma*,[51] while in the other cases the area of the circles involved is prescribed, depending on the design so that the total area of the *citi* would agree with the prescribed area for the *yajñas* performed in similar contexts (typically $7\frac{1}{2}$ *puruṣa*). Baudhāyana describes two types for *rathacakraciti*, one corresponding to a wheel as a full flat disc, and another a wheel with spokes (connecting a nave to an outer rim), and for the latter type the construction involves three concentric circles, one corresponding to the nave of the wheel, and two others corresponding to the annular outer rim, the two parts being connected by a set of 16 spokes; the specific sizes are described in terms of distribution of the net area into the three parts (the

[49] It may also be noted here that the word "exact," especially in its formal usage, has a stronger connotation than the Sanskrit word *nitya*, as can be gauged from its alternative meaning as "regular," and Āpastamba would have more likely meant that the two areas are the same as far as normal judgment would go, and not by any intrinsic or universal standard, as may be suggested by the choice of the word "exact" in the English translations.

[50] The notion of the three *agnis* in combination appears in the Ṛgveda (5.11.2), without any names. The *Śatapatha Brāhmaṇa* includes names for two of the *agnis, gārhapatya* and *āhavanīya*, and the former is mentioned to be circular (see Datta 1932, p. 27, and Seidenberg 1978, p. 339, for related discussions). The mention of specific sizes for the *agnis* are known to appear for the first time in the Śulbasūtras; for all that one can say, the areas proposed for them *happen* to be equal, and there is no independent statement about their being required to be so; and quite possibly the choices were made only after there was better understanding of geometry, and means of achieving equality to an extent became part of it.

[51] For the square version that would be the length of the sides.

nave, rim, and spokes). The procedure for constructing this *citi* then consists of construction of three squares with a common midpoint, and transforming each of them into a circle by the procedure described in Sect. 6.2.[52]

It is generally believed that the areas of the three *agni*s *āhavanīya*, *gārhapatya,* and *dakṣiṇāgni* were stipulated to be the same, and this has been viewed in some works, especially (Seidenberg 1975, 1978, 1981), as the motivation for the invention of the procedure described in Sect. 6.2. It is true that the *citi*s are found to be constructed with equal areas (as far as the intent is concerned, though differences would arise due to inexactness of the procedure). This stands out prominently in MSl (sūtras 1.1.7 and 1.1.8) where all the three are presented together, along with prescribed constructions; Seidenberg (1978) cites the sūtra as the sole "explicit reference for the equality in the area" of the three *agni*s. There is however no explicit statement of the "stipulation" found anywhere in the Śulbasūtras. There is also no mention of anything like legend or influential statement or injunction, say attributed to the *Brāhmaṇa*s as one finds in various other contexts, that would correspond to the legend of the oracle of Delphi in the context of the problem of doubling the cube. The issue involved emerges as a practical matter and is seen to be handled as such, through routine practice, with no qualms about making approximate constructions. There is also no account, or indication, of a quest for finding a solution, or attempts to improve an ad hoc solution that was adopted, as we find in the case of the problem of doubling the cube. As seen in the other examples above, the task came down to constructing circles covering specific areas and in the case of *gārhapatya* the area having to match with *āhavanīya* is just one instance, and possibly a later one, since in BSl there is no emphasis on it. As such, the analogy drawn between the circling of a square as above with the developments around the problem of doubling the cube in the Greek context, as in Seidenberg (1975, 1978, 1981), seems misplaced.

It is also not apparent that the discovery of the method lends credence to the notion of "ritual origin," as made out in the abovementioned papers; it should be borne in mind that nowhere is there any statement in the Śulbasūtras suggesting that the method was discovered in the course of their own pursuits, and was not say part of folk knowledge from the time, arising from other pursuits; this is not to make a case for it, but to point out that it has not been ruled out; see also Joseph (2016, pp. 63–64), for analogous comments. It may also be recalled here that various similar geometrical facts were discovered by Jaina scholars of antiquity who did not have any ritual practices involving geometry; see Dani (2023), and other sources cited there.

[52] While the shapes of the bricks in the constructions of the *citis* are generally rectilinear, with the designs that are provided serving as rectilinear *approximations* to the desired figure (bird, turtle, etc.), in the instances involving circles in the design, the brick shapes are not always rectilinear and can have a rounded edge, and the design can match the desired figure exactly.

7.4 Construction of a Circle with a Given Area: The Mānava Version

The method described in Sect. 7.2 has also been described MSl. In fact, one finds it twice, once (in MSl 3.2.10)[53] in a general form (and accompanied by some emphasis with the preceding sūtra drawing special attention to it) and also in MSl 1.1.8, where it is recalled in the course of describing the construction of the *gārhapatya agni*. Still, there seems to be yet another sūtra, MSl 3.2.15, on the same issue, proposing another construction. There is no internal cross-referencing, and each appears as an independent statement. The sūtra MSl 3.2.15 reads as

caturasraṃ navadhā kuryāt dhanuḥ kotyāstridhatridha |
utsedhātpañcamaṃ lumpetpurīṣeṇeha tāvatsamaṃ ||

The first part readily translates as "divide the square into nine parts [by dividing] the transversal and vertical sides into three each." *utsedhātpañcamaṃ* means "a fifth of the jutting out part." Unfortunately the last part is problematic. Literally it just says something like "erase that much with rubble." And what is one supposed to get from that? The sequence of the sūtras around it also do not provide any clue about what is intended. The context seems to have prompted the authors of Sen and Bag (1983) to dismiss a set of sūtras including it with the comment "Possibly these are not problems of quadrature of the circle. Ordinary squares are drawn without any mathematical significance." In Kulkarni (n.d.), where the MSl part is based on van Gelder (1963), the sūtra is interpreted as describing a circle with the area of a given square, but the prescribed radius is taken, by way of interpretation of the second half of the sūtra as above, to be half the side plus $\frac{4}{5}$th of the jutting out part of the chords. The analysis following the interpretation, as described there, shows the result to be grossly erroneous; one way of highlighting how much it is in error would be to note that it corresponds to a value π which is about 2.63. It was pointed out in Dani (2010), and followed up in Dani (2022), that if one takes the circle passing through the points which are at $\frac{1}{5}$th along the chords from the corresponding sides of the square (see Fig. 10, with $\mathrm{QS} = \frac{1}{5}\mathrm{QR}$), then that indeed has area very close to the area of the square; skipping the details of the calculations, which may be found in Dani (2010), I may mention that for the unit square the area of the circle produced works out to be about

[53] In van Gelder (1963), and also in Sen and Bag (1983) and Kulkarni (n.d.) following the former, the sūtra as a whole is treated as describing the construction of a circle of equal area as a square by the same method as in Baudhāyana; it has however been noted in Hayashi (1990) that while the sūtra does contain in its second part a variant of Baudhāyana's method for circling a square, the first part is about squaring a circle; and the latter we will be discussing later. The miscomprehension involved in the earlier works seems to have been mainly due to the fact that in the original text as presented there a word, *caturasraṃ,* is missing in the previous sūtra MSl 3.2.9; the latter purports to provide an introduction to the issue being addressed, so missing the word affects comprehension of MSl 3.2.10, as has been clarified in Hayashi (1990).

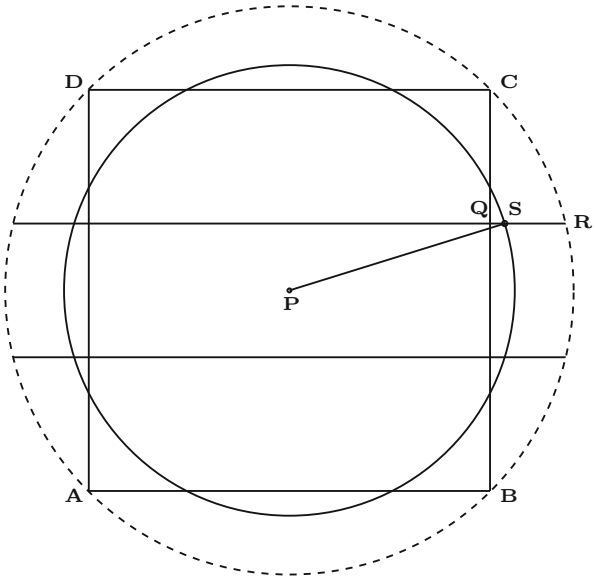


Fig. 10 Mānava's construction of a circle with the area as a given square

0.9947, involving an error of only about $\frac{1}{2}$%, considerably more accurate than Baudhāyana's prescription.

There are many pointers suggesting that this was in fact intended as a construction of the circle with the area of the given square. I have discussed in detail in Dani (2022) how the construction described here can be thought of as a refinement of Baudhāyana's construction, with the bisector replaced by the trisectors and the factor $\frac{1}{3}$ in Baudhāyana's construction replaced by $\frac{1}{5}$. That in particular motivates why the points on the jutting out parts at their $\frac{1}{5}$th from the side of the square are more appropriate, rather than from the other end at the circle, as chosen by van Gelder and followed in Kulkarni (n.d.).

There are a few disconcerting points however. First, that it never resurfaces again, either in the Śulbasūtras or in other Indian works of the period or later. Second, here the author shows no particular satisfaction over, or awareness of, the novelty and value of the construction, unlike in the case of his revision of the ratio of the circumference to the diameter of the circle (see Sect. 7.4), nor does he relate it to the other construction in any way. It should be borne in mind however that while we can now see it as a more accurate construction than the earlier one, in their time this may not have been so easily verifiable. In that sense, it was for them just another construction whose merits could not be ascertained.

I may mention finally here that there is also another interpretation of the verse, given by R.C. Gupta (1988), interpreting "utsedha" to mean "height," which in turn the author takes as the radius of the circumscribing circle. With those interpretations, the author deduces the area of the circle produced, starting with the unit square, to be

$8\pi/25$, so the procedure corresponds to a value of π as 25/8. I have however argued in detail in Dani (2022) that while it may lead to a good value, the interpretation is unsatisfactory in various ways, including especially that it does not involve the first part of the verse at all, and that interpretation of "utsedha" as radius of the circumscribing circle lacks justification in the overall context; we shall not go into further detail of this here.

7.5 Value of π?

At many places in the literature on the Śulbasūtras, one finds references like "the value of π known to the geometers of the Śulbasūtras," "Baudhāyana's value of π," etc. These can be quite misleading; Hayashi (1990, p. 77) has opted to say it seems meaningless. Such references implicitly suggest the following:

(i) That they knew the circumference of the circle is proportional to its diameter and/or the area is proportional to the square of the diameter
(ii) That they knew the ratio of the circumference to the diameter is the same as the ratio of the area to the square of radius (semidiameter)
(iii) That having determined such a number they used it to determine one from the other numerically

It is quite reasonable to assume that they know (i), in respect of both the circumference and the area, as the statements in the sūtras in respect of these are made independently of the size involved; though from a strict mathematical point of view it is a difficult issue, they would probably be convinced of its validity, in both cases, from the analogous facts about rectilinear figures. Also, they were quite used to the idea of scaling of figures, in the context of which these assertions would seem natural.

They just could not have been aware of (ii). In fact, if they were to believe that the two ratios were the same, then according to BSl I 112–113 they would consider it to be 3, and Baudhāyana would have given a solution in terms of the *trikaraṇī*, which would be exact, rather than the construction as in Sect. 7.2. Second, Mānava Śulbasūtra which takes the side of the square with the same area as the altitude of the inscribed equilateral triangle, which amounts having the (common) ratio to be 3, would not be bothered to suggest $3\frac{1}{5}$ as a bound for it as discussed in Sect. 7.1!

As to (iii), there is no evidence that the values determined for π were ever used as a numerical factor in another instance. The simple historical reality seems to be they never had an occasion to compare the two ratios, and with regard to areas they were content to start with squares and produce circles which they believed would have the same area. Their trust in that as the correct equivalent circle, and the geometric method as the unique way of addressing the task, is attested by the fact that Kātyayana describes the same method, several centuries after Baudhāyana.

7.6 Squaring the Circle

The converse problem of finding a square with the same area as a given circle is also considered, and a serious effort seems to have gone in to arriving at the resolution of it presented in BSl, as we shall see below.[54] This is especially notable, as the practical objectives, such as of arranging the areas of *āhavanīya* and *garhapatya* to be equal or fixing the size of *rathacakraciti* would be fulfilled by their construction(s) of a circle with the same area as a square (to the extent they could achieve it) discussed in Sect. 7.2. In practical terms, it was circles that needed to be constructed with a given area, which to begin with is described in square shape, as in the construction of *rathacakraciti*. On the whole also in the construction of the altars, one does not find any occasion to convert a circle into a square. Clearly therefore the task was undertaken out of curiosity. As commented in Delire (2003), it is a "manifestation of Baudhāyana's concern to complete his mathematical corpus."

Unlike the solution of the problem of finding a circle with the area of a given square, the resolution of this converse problem is not geometrical but is by describing a numerical factor as the ratio of the size of the desired square to the diameter of the given circle. One may wonder whether a geometric procedure was attempted, unsuccessfully,[55] or the process of trying to understand the numerical significance of the procedure for transforming a square to a circle led to consideration, and the resolution that is presented, of the converse problem. Given the overall context, it seems inconceivable that they could not have come up with a geometrical procedure, on a similar footing as the other way construction. As noted above, the exercise was driven by curiosity, and not a practical objective, and the curiosity could well be about better understanding of the numerical factor in the geometric construction in Sect. 7.2; once the factor is understood, in a certain way, inverting it would give a resolution of the converse problem.

[54] The problem considered here should however not be confused with its counterpart in the Greek tradition, which involved accomplishing the task only with a straight edge and compass; in the context of the Śulbasūtras, any procedure implementable by an artisan would be admissible. In this respect, discussions of the two on the same footing, as for instance in Seidenberg (1975), Sect. 4, seem quite out of place and unwarranted.

[55] It may be worthwhile to recall one of Seidenberg's theses, from (Seidenberg 1972, p. 175), in the author's own words: "This leads to our second main thesis: The first crisis occurred in mathematics because the ritualists could not reverse their (canonical) circulator of the square in a geometric way. The resulting efforts to find an arithmetic solution of the squaring of the circle gave rise to the algebraic tradition. Geometry was dislocated from its ritual base." While there is no doubt that the fact of looking for nongeometric solution is of enormous significance from an historical point of view, the characterization in this, as a crisis and its resolution, seems far-fetched. We do not really know whether geometric resolution was seriously attempted, or simply the effort to understand the surd magnitudes of a kind that were showing up for the first time led to the "algebraic" realization. As far as reading from the Śulbasūtras goes, the transition seems to have been seemless, and the matter even seems to have been forgotten quite soon, rather than leading to a new framework or insights.

The description of the factor by which to multiply the diameter of a given circle to get the side of the square with the same area, given in BSl 1.59, immediately following the construction described in Sect. 7.2 goes as follows: divide the diameter into 8 parts, further divide 1 of the parts into 29 parts and take away 28 of them, and also a sixth minus its eighth of the (remaining) part; in terms of numbers, the factor described literally is

$$1 - \frac{1}{8}\left(\frac{28}{29} + \frac{1}{29}\left(\frac{1}{6} - \frac{1}{6 \times 8}\right)\right). \tag{1}$$

This may equivalently be written as

$$\frac{7}{8} + \frac{1}{8 \times 29} - \frac{1}{8 \times 29 \times 6} + \frac{1}{8 \times 29 \times 6 \times 8}, \tag{2}$$

which is the form expressed in Thibaut (1875), and various other sources also have separated out the part $\frac{7}{8}$ from the expression above. The latter has the advantage of setting the magnitudes of the components in perspective. It would however be important to keep track of the specific wording used, especially as we do not yet know the precise significance of why the factor was described in terms of a combination of terms as above; the alternate presentations may convey the impression[56] that $\frac{7}{8}$ was first separated out (as a leading factor) and then the difference was analyzed in some way to get the remaining part of the expression. This may or may not actually be how calculations proceeded, though the choice made does seem to be borne out in some ways (see below).

We note that the fraction as in the expressions reduces on simplification to $\frac{9785}{11136}$ and for the circle with unit radius the expression yields the area 3.088..., a little more than 98.3% of the actual value. A notable point about this is that the error is of the same order as in the construction of the circle with the same area as a square, discussed in Sect. 7.2. We recall that the factor there was $\frac{2+\sqrt{2}}{3}$, and note that $\frac{2+\sqrt{2}}{3} \times \frac{9785}{11136} = 1.00000238\ldots$. The two numbers being such close reciprocals inescapably suggests that the latter was in fact computed as the reciprocal of the former, with the small difference attributable to an error introduced by the process of computation of the reciprocal. On the whole, there is a general consensus that the number would have been computed as the reciprocal of the factor $\frac{2+\sqrt{2}}{3}$ as in the convertion from a square to a circle. Beyond that, however, the formula has remained an enigma for various reasons. The key questions that remain are how was the reciprocal of $\frac{2+\sqrt{2}}{3}$ computed, and why is it presented as the combination as seen in the above expression. It has been the subject of numerous papers, but the issues associated with it still seem unresolved. We shall take up a discussion on this separately in the following subsection and continue here on other details on the issue.

[56] In fact, some of my comments in Dani (2010, on p. 29) are based on such an understanding, based on my familiarity at that time with only the alternative expressions as above.

7.7 Alternative Factor for Squaring the Circle

In BSl I.60, immediately following the sūtra with the formula as in the last subsection, Baudhāyana gives also another prescription. It is separated from the previous one by "or else" (*api vā*), and the sūtra seems to mention it to be "inexact"[57] indicating that it is proposed only as a simpler alternative. The prescription consists of dividing the diameter into 15 parts and dropping 2 of them, to get the side of the desired square; thus, the factor concerned would be $\frac{13}{15}$. This formula turns out to be very crude: For the circle of unit radius, it gives the square with side $\frac{26}{15}$, whose area would be $\left(\frac{26}{15}\right)^2 = 3.0044\ldots$, in place of the correct value 3.14159 ..., smaller by more than 4%. Even in terms of fractions with small denominators, there are better approximations possible.[58] The fraction $\frac{7}{8}$, which is recognizable as a substantial part in Baudhāyana's previous formula, would serve as a better approximate value, giving the area for the unit radius circle to be $\left(\frac{7}{4}\right)^2 = 3.0625$, than the suggested factor $\frac{13}{15}$. It is therefore not clear why Baudhāyana gave that as an alternative value, even as an approximation.

One explanation that has been put forth relies on the assumption that the area of the circle was taken to be $3r^2$, where r is the semidiameter (radius) of the circle, arrived at as the average of the areas of the inscribed and superscribed squares; see Datta (1932, p. 147), and Sen and Bag (1983, pp. 163–164). The factor for describing the side of the desired square in terms of the diameter would be $\sqrt{3}/2$, and $\frac{13}{15}$ was taken as approximation for it; see Sect. 8.5 for a discussion relating to this approximation. There is no specific evidence supporting such an explanation, but it seems plausible that the area of the unit circle was also taken to be 3, matching the ratio of circumference to diameter, especially in practical contexts.

A different explanation is found in Delire (2005), where it is argued that if we choose the approximation $\frac{7}{5}$ for $\sqrt{2}$ then the factor $\frac{1}{3}\left(2+\sqrt{2}\right)$ for obtaining the diameter of the circle from the side of the original square, as in Sect. 7.2, reduces to $1+\frac{2}{15}$, and the author comments that "they could well have considered $1-\frac{2}{15}$ as the inverse of $1+\frac{2}{15}$"; the latter part does not seem convincing, as the fraction involved is readily invertible, but perhaps may not quite be ruled out.

[57] Here again, there is an ambiguity comparable to the one seen near the end of Sect. 7.2: The clause "*eṣānityācaturasrakaraṇī,*" appearing at the end of the sūtra in this respect, can be interpreted in two ways, with *eṣānityā* split as *eṣā anityā* or *eṣā nityā*; Thibaut and those following him have adopted the former, with *anityā* interpreted as inexact, leading to what we noted above. On the other hand, the latter alternative possibility is adopted in Delire (2005, p. 68), with the word *nityā* taken to mean traditional – which is close to the currently common meaning as "perpetual." In a way, this explanation fits better with the fact that only this other factor appears in ASl and KSl (see below), as tradition is a more likely reason for adhering to it.

[58] The factor $\frac{8}{9}$ which was adopted by the ancient Egyptians gives for the unit circle the area to be 3.16049 ...; see Gillings (1972, Chapter 13).

Curiously, none of the other Śulbasūtras mention Baudhāyana's elaborate formula.[59] ASl and KSl recall only the approximate $\frac{13}{15}$ formula, right after recalling the procedure for producing a circle of equal area from a square, which they do essentially with the same wording as in BSl; it is as if the finer formula is forgotten by their time, and the approximate version follows the other way construction as a mandatory converse.

7.8 Mānava's Conversion of Circle to Square

The Mānava Śulbasūtra, which departs from BSl in many propositions, on the other hand, introduces another resolution for the converse operation: The first part of sutra MSl 3.2.10 prescribes, as explained by Hayashi in (1990) (departing from the earlier interpretations in van Gelder 1963; Kulkarni n.d.; Sen and Bag 1983), for a given circle, the altitude of the equilateral triangle with the diameter of the given circle as its base, as the side of the desired square. For a circle with unit radius, the altitude of the equilateral triangle will be $\sqrt{3}$ and hence the area of the prescribed square will be 3, in place of the desired value $\pi = 3.14159\ldots$ This is even cruder than the Baudhāyana's crude alternative value recalled by Āpastamba and Kātyāyana, as discussed earlier, as far as comparison with the true value goes. In the context of some of the interesting observations from Mānava Śulbasūtra noted earlier, this comes as an anticlimax! However, if the area of the circle was believed to be three times the square of the radius, a possibility mentioned earlier in Sect. 7.7, the construction gives exactly a circle with the desired area.

7.9 Understanding Baudhāyana's Formula

For much of the history in respect of analysis of the formula, it would be convenient to refer the reader to Kichenassamy (2006), which has been the latest notable contribution on the issue, and includes also a detailed account of the earlier explanations. Here we only summarize some of the crucial points that have emerged from the discourse.

It is well-known that Śulbasūtras contain a very good numerical approximation for $\sqrt{2}$, and we will be discussing it later in Sect. 8, given as

$$1 + \frac{1}{3} + \frac{1}{3 \times 4} - \frac{1}{3 \times 4 \times 34}. \tag{3}$$

[59] There is a suggestion in Seidenberg (1975, Sect. 14, p. 517) that the "squaring of the circle was interpolated"; presumably the reference is to the elaborate expression and not the approximate one. Its omission in the later Śulbasūtras gives some credence to the possibility. Nevertheless, the speculation does not seem worthy of serious consideration, since the issue involved is not known to have evoked an interest in the tradition at any later stage.

It should also be clear that the value would be of no consequence in their practical work; in practice, if they needed a length measuring $\sqrt{2}$ units they would naturally have used the *dvikaraṇī,* namely, the measure of the diagonal of the unit square. Evidently the value as above was determined in some theoretical context, and the context at hand presents the unique candidate in this respect. In other words, to determine the reciprocal of $\frac{2+\sqrt{2}}{3}$ they wanted to get an understanding $\sqrt{2}$ in terms of fractions.

In the expression for $\sqrt{2}$ as above, we find another number, apart from the ratio as in (1) written in a way that is quite unusual for us. This brings to mind the ancient Egyptian practice of writing each (rational) number as a sum of distinct unit fractions (viz, fractions with numerator 1). What is involved here is different however, and unfortunately we do not have a large enough sample of their writing numbers in this form, from which one could try to identify a pattern; in fact, the two fractions that we saw so far written in this manner are the only ones of the type in the entire body of Śulbasūtras; there are a few other fractions with simple descriptions.

It may be observed at this point that having determined $\sqrt{2}$ as above, if they had realized that $\frac{1}{2+\sqrt{2}} = \frac{2-\sqrt{2}}{2} = 1 - \frac{1}{2}\sqrt{2}$, the desired reciprocal could have been recognized as (approximately)

$$3 - \frac{3}{2}\left(1 + \frac{1}{3} + \frac{1}{3 \times 4} - \frac{1}{3 \times 4 \times 34}\right) = 1 - \frac{1}{8} + \frac{1}{8 \times 34}, \tag{4}$$

giving a much simpler expression for the desired ratio, so apparently that reduction of the surd was not recognized. Instead, the substitution for $\sqrt{2}$ may be considered to have been made in the denominator itself. If we substitute for $\sqrt{2}$ from (3), we get a fraction which on simplification turns out to be $\frac{1224}{1393}$. On the other hand, while the expression in (1) is close to that, its denominator turns out to be 8×1392. We may therefore further suppose after following whatever process they were adopting for the successive terms in the expansion some part was ignored, leading to the expression as above. Specifically it may be observed that $\frac{1224}{1393} = \frac{7}{8} + \frac{41}{1393}$, while the given expression as in (2) equals $\frac{7}{8} + \frac{41}{1392}$, which would support such an assumption.

A suggestion somewhat along these lines was made by Thibaut (1875), who seems to have assumed, based on workouts seen in the commentaries, that such a computation could have been made by the authors of the Śulbasūtras as well. The problem however is that such an assumption is untenable, as was pointed out by C. Müller in 1930 (cf. Kichenassamy 2006, p. 157, for details and reference), as there is no evidence to show that the authors of the Śulbasūtras had the computational skills as involved, and the explanations of the commentators, who are from more than a 1000 years later, cannot be relevant in this respect.

There have over the period been several attempts to explain the expression through various arithmetical stratagems (see Kichenassamy 2006 for details), but none of them seem, as noted in Kichenassamy (2006), to be satisfactory. A key point made in Kichenassamy (2006) is that though it is given as a numerical expression, the reasoning is likely to have involved geometric thinking, at least at a heuristic level, and especially

the author emphasizes likely use of the idea of scale, and scaling of figures depending on the context, which is very much involved in the Śulbasūtras. Kichenassamy provides an explanation based on these ideas. It is however not clear to the present author whether even the reasoning in Kichenassamy (2006) can be said to be conclusive. One substantial issue that comes to mind with regard to the reasoning is that the entire body of Śulbasūtra works, extending over centuries, is treated as a homogeneous knowledge system, and patterns in several much later calculations or applications have been brought in to explain how Baudhāyana may have reasoned. Second, on the whole it also has the feel of a "verification," with a bit of number-crunching, a criticism which has been leveled on the earlier proofs (see Kichenassamy 2006) though at a somewhat different level. Hopefully a clearer image would emerge along these lines, but we may have to wait a while for the last word to be said on the issue.

8 Square Roots in the Śulbasūtras

It was noted in Sect. 7.9, (3), that Baudhāyana presents an expression for $\sqrt{2}$, as

$$1+\frac{1}{3}+\frac{1}{3\times 4}-\frac{1}{3\times 4\times 34}$$

which is 1.4142156..., in place of 1.4142136..., accurate five decimal points (six significant decimal digits).[60] In this context, we review in this section that and other values for roots of numbers appearing in the Śulbasūtras.

8.1 $\sqrt{2}$ in Mānava Śulbasūtra

Curiously, the Mānava Śulbasūtra does not introduce the *dvikaraṇī*, and where the measure is involved in a construction, an approximation of $\sqrt{2}$ as $1+\frac{1}{3}+\frac{1}{12}$ is used in MSl 3.3.1–2, in a rather specific context. The latter value, equivalent to $\frac{17}{12}$, is evidently a cruder approximation compared to Baudhāyana's formula, which has not been mentioned. In the Mānava Śulbasūtra, $\sqrt{2}$ turns up also in another guise, and it has been quantified numerically. The context involved is to describe the size of the largest square contained in a circle (MSl 3.2.14). To quantify the side of the square in terms of the diameter of the given circle, the prescription is to divide the diameter into ten (equal) parts and take away three of them. This amounts to an approximation of $\sqrt{2}$ as $\frac{10}{7}$. Incidentally, the largest square contained in a circle is also involved in the designs for brick-tiling of some of the *citis* involving circular regions, in particular the *rathacakraciti* described in BSl V.5 and ASl 12.17; the tiling within that square is into a grid of 144 squares of uniform size, while that of outside

[60] This is comparable to the Babylonian value depicted on the cuneiform tablet YBC-7289, in sexagesimal system, which corresponds to $1+\frac{24}{60}+\frac{51}{60^2}+\frac{10}{60^3}=1.4142129\ldots$ (see Robson 2007).

segments is somewhat ad hoc and nonuniform, with 14 tiles in each of the four segments; the side of the square is not quantified, however, and is only mentioned as the largest possible square (*yāvat-sambhavet-samacaturasraṃ*-literally "as much possible square"); presumably it was meant to be determined by taking two diameters perpendicular to each other, but this is not specified.

8.2 Explanations of Baudhāyana's Formula

The elaborate formula (1) apparently had the context of the issue of transformation of a square into a circle and the other way, and it could specifically be for the purpose of computing the ratio of the side of the square to the diameter of the circle produced by the procedure as in Sect. 7.2. On the other hand, in ASl and KSl the formula is mentioned right after introducing *dvikarṇī*, apparently as something related to it; as noted earlier neither of them mentions Baudhāyana's refined formula (1) for the side of a square of area transformed from a circle.

How was Baudhāyana's formula (1) arrived at (by him or whoever may have done it prior to him)? Like the question about the expression for squaring a circle, this has also been the subject of much discussion, though it may be said that some of the explanations here seem much more plausible than was the situation with the previous question.

The approximation formula

$$\sqrt{a^2+b} \approx a + \frac{b}{2a} \tag{5}$$

for $a > 0$ and b smaller compared to a, of either sign, seems to have been known in ancient India. While there is no specific evidence in this respect, some calculations of square-roots going into several terms, especially in the Jaina literature (see Dani 2023), would be difficult to explain without recourse to such a formula. Second, the formula has a simple visual geometric explanation: When squares with area $a^2 + b$, $b > 0$, and a^2 are matched along a corner and two sides, the difference is readily seen to consist of two strips whose one side is a and the other is the excess over a that is desired, except for a small square which can be ignored; this readily illustrates that the desired excess should be $\frac{b}{2a}$; and a similar argument works for b negative. It is conceivable that (3) was derived from this approximation formula. If one is to start with simple-minded choice for a, the "first guess," say as 1 or $1\frac{1}{3}$ the formula, would have to be applied repeatedly to get (3). Actually all the known calculations, from those and even somewhat later times, seem to fit in the pattern of single application of the formula, so one may wonder if they may have applied the formula iteratedly in this instance. We note that the less accurate expression $1 + \frac{1}{3} + \frac{1}{12}$ used in MSl could be deduced in this way, choosing $a = 1 + \frac{1}{3}, b = \frac{2}{9}$. It could also be surmised that it was directly noticed that $17^2 = 289 = 2(144) + 1 = 2(12)^2 + 1$, which is quite plausible especially in the context of the square grids used in the tiling of *rathacakraciti*. If with that one chooses the approximation formula with $a = \frac{17}{12}$,

which is recognized to be close to $\sqrt{2}$, we have $b = -\frac{1}{144}$ and we get the result to be $\frac{17}{12} - \frac{1}{144} / \frac{34}{12}$ or $\frac{17}{12} - \frac{1}{12\times 34}$. This matches with the value of the expression in (3) but leaves open the question about expansion into components as appearing there. One possibility is that they understood the number as it is calculated here but expressed it in the components as a sort of natural representation that they were used to. We do not know enough about how they dealt with fractions, of which there are very few instances in the works, to be able to make a definitive judgment on this.

There have been seemingly arithmetic explanations of (3) by Thibaut (reproduced in detail in Datta 1932, pp. 189–192) and Rodet (see Rodet 1879, pp. 162–165, or Sen and Bag 1983, pp. 165–167). Thibaut's version, though somewhat belabored in its import, essentially matches with the arithmetic involved in the explanation as above, without the geometry. Rodet's explanation involves a more complicated approximation formula; there does not seem to be much justification in assuming that a formula of that kind would have been used in ancient times.[61]

8.3 Datta's Explanation

A more plausible explanation involving geometric reasoning which is consistent with the Śulbasūtras practices is due to Bibhutibhushan Datta (1932, pp. 192–194).[62] Consider two unit squares placed alongside as in Fig. 12, to be transformed into one square. We chop the second square into three equal strips vertically, and place the farthest one at the top of the first square, horizontally, matching the unit side. We now cut up the middle strip into three equal parts, each a square with side $\frac{1}{3}$, and shift the top one alongside the other strip places at the top of the first square. The figure now consists of a square with side $\frac{4}{3}$, and a strip of height $\frac{2}{3}$ and width $\frac{1}{3}$. We cut the strip vertically into four equal parts, the width of which will then be $\frac{1}{12}$. Two of these are placed horizontally at the top of the square with side $\frac{4}{3}$ formed earlier, and two on the side, vertically along the side. The figure produced so far is a difference of a square with side $1 + \frac{1}{3} + \frac{1}{3\times 4}$ and a small square of size $\frac{1}{12}$. At this stage, for futher adjustment Datta writes down a quadratic algebraic equation (together with an

[61] I may also mention here that in Dani (2010) I suggested that, keen on determining a value for $\sqrt{2}$, they may also have done it through direct measurements with cords. It was argued that such a determination would have been possible in their context since if one starts with a size of 35 feet for the side of the square and measure the diagonal in terms of pieces as suggested by the expansion (see Dani 2010 for details), then the last step would involve measuring out a piece of about 3 feet by one which is over an inch, and that this would indeed be possible. This argument has a lacuna however that it does not take into account that for a square drawn on the ground the measurements may not match closely enough with the respective mathematical values, and a small error can lead the result astray.

[62] Strictly speaking, Datta's original explanation is not entirely geometric, as he resorts, toward the end (see below), to algebraic reasoning, of a kind which actually is not evidenced anywhere in the Śulbasūtras. His basic idea however can be completed to arrive at Baudhāyana's formula, in the spirit of the Śulbasūtras constructions, as will be seen below.

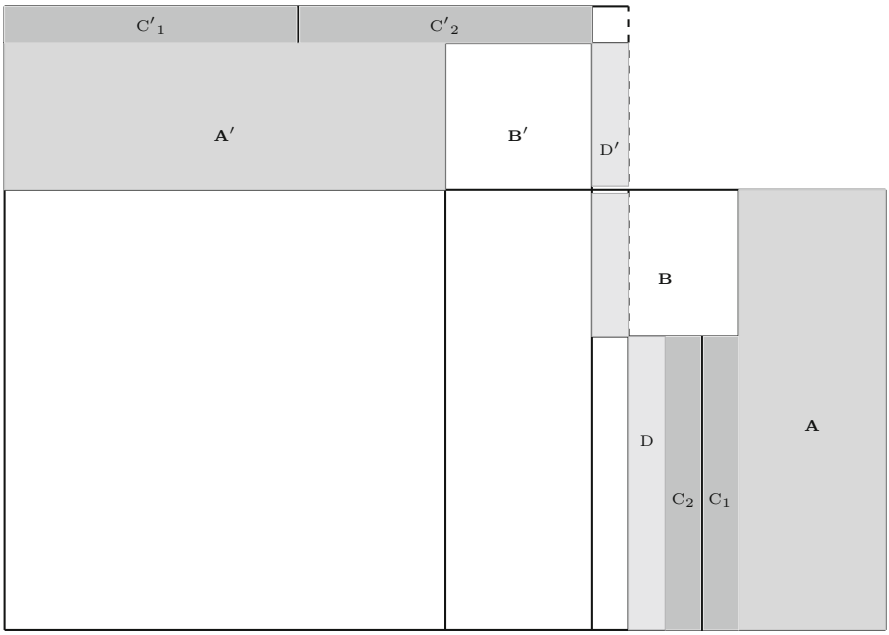


Fig. 11 Datta's construction of $\sqrt{2}$

algebraic approximation) to explain the last term. It can however be seen that we may continue along the lines of the argument as above and get the result; if strips of width $\frac{1}{12\times34}$ are cut along two perpendicular sides of the bigger square, they would make an area as much as the area of the smaller square, minus the area of a square of side $\frac{1}{12\times34}$ (on account of the overlap at the corner). Ignoring the latter tiny square, we see that the square with side $1+\frac{1}{3}+\frac{1}{3\times4}-\frac{1}{3\times4\times34}$ is (approximately equal to the original) 2 unit squares. This establishes (5) (Fig. 11).

This proof seems properly in the spirit of constructions in the Śulbasūtras and also incorporates a possible explanation of the number being expressed in the particular fashion; whether the latter is of significance, and valid, we cannot say, as it is not really understood what the pattern of numbers signifies and why it was adopted.

8.4 Irrationality of $\sqrt{2}$?

There are often claims made, imputing the knowledge of irrationality of $\sqrt{2}$ to the authors of the Śulbasūtras (see Datta 1932, p. 195; Sen and Bag 1983, p. 168). Sometimes the issue is linked, rather disingenuously, to knowing the value they assigned to be approximate. The Śulbasūtra geometers would indeed have known, at least at the time when the value was first determined, that it was approximate; arriving at such a value would have involved one or other computational process –

as the intricate value could obviously not have been the result of pure intuition – and no matter what the process was, whether one of the possibilities that we discussed or otherwise, in the course of the operation it would be recognized that the process has not come to an end, so the value is not exact. There is however no indication that it made any difference to them. Contrary to the general perception of thrust on precision or accuracy, the Śulbasūtras are seen to be open to adopting, for convenience, approximate values, even when more accurate values are available. Sometimes the word *saviśeṣa* for the adopted numerical value of $\sqrt{2}$ is brought in to argue that it was being explicitly declared as approximate (see Sen and Bag 1983, p. 168). This however involves a misinterpretation. *Viśeṣa* corresponds to extra, and in this instance the reference seems to be extra over the unit; this is suggested by the fact that description of *saviśeṣa* begins with the unit and prescribes adding parts to it (rather than by a direct representation), and this interpretation is also borne out by the use of *viśeṣa* in ASl 2.1, as has also been noted in Datta (1932, p. 197). In any case, the term carries no connotation of the difference from the implicit target being small, or eligible of being ignored, as with the word approximate.[63]

More importantly, even if one grants that they were quite conscious about the value being approximate and chose to put it on record, it still has little to do with being aware of the irrationality of $\sqrt{2}$ as a number. Asserting a number to be irrational involves that it can in no way be described as a ratio of two integers, and not just that the value one settled on, perhaps after a great deal of effort, does not quite represent it. Irrationality needs to be concluded by a proof and is not ensured by the fact that diligent efforts of even a great scholar did not produce a rational value for it. After all, rational numbers also can be approximated by other rational numbers and can arise as limits of sequences or series of rational numbers.

The Śulbasūtras were concerned with practical aspects of geometry of altar building. They are seen to have taken an enlightened view of the subject and identified various principles and constructions useful for the activity, but were nevertheless focused on practical matters. One does not find them getting involved with philosophical or arithmetical aspects of numbers, and it is wrong to indulge in wishful claims about it.

8.5 Other Square Roots

A numerical understanding of square roots, relating issues such as how they compare with other fractions, or their (approximate) value, did not hold much relevance for the Śulbasūtra geometers; the unique case of how they may have got interested in this respect in $\sqrt{2}$ we discussed above.[64] A couple of other instances have been

[63] The mention in Sen and Bag (1983, p. 169) of the use of *kiñcidviśeṣa*, by way of comparison, is not meaningful, since in the latter the part *kiñcid* makes all the difference.

[64] As pointed out in Thibaut (1875, p. 241), some commentators on the Śulbasūtras are seen to apply Baudhāyana's elaborate formula (3), for figures on the ground as well, but it is a useless exercise.

noted in this respect, in which the inference is indirect, and rather tentative; these are mentioned also in Seidenberg (1978). First, as mentioned earlier ratio $\frac{13}{15}$ proposed for the ratio of the side of a square and the diameter of a circle with the same area, as a simpler and common option, was perhaps meant to be for $\sqrt{3}/2$, based on the premise that the area of a circle is $\frac{3}{4}$th of the square of its diameter. Following the steps as in Datta's proof for the expression for $\sqrt{2}$, as described in Sect. 8.3 (excluding the last part of it involving subtraction), leads to the expression $\sqrt{3} \approx 1 + \frac{2}{3} + \frac{1}{15}$; first from the two squares to be added rectangles of width $\frac{2}{3}$ may be joined to the first square along two perpendicular sides, and a square with side $\frac{2}{3}$ may be filled in from the remaining rectangle with width $\frac{2}{3}$; thus, we get a square of side $1 + \frac{2}{3}$ and a balance rectangle of size $\frac{1}{3} \times \frac{2}{3}$, and the latter may cut up into ten strips of width $\frac{1}{15}$, and length $\frac{1}{3}$, and arranged along the edges of the square, making a square of side $1 + \frac{2}{3} + \frac{1}{15}$, except for an overlap of a square of side $\frac{1}{15}$ which is ignored; see Datta (1932, p. 194), if necessary, for details of arriving at this expression – the more refined version of the approximation obtained there is however an extrapolation which is hardly likely to have been involved in the Śulbasūtras.

Another instance relates to Baudhāyana's construction of *Droṇaciti*, shaped as a square trough with a rectangular handle, with net designated area of $7\frac{1}{2}$ (square) *puruṣa*. He chooses the square to have side length $3 - \frac{1}{3}$ *puruṣa* and the handle to have sides $\frac{2}{3}$ and $\frac{7}{12}$, with the areas of the two adding up to the designated size. It would seem that the broadly intended distribution of the areas would be 7 *puruṣa*s for the square trough and $\frac{1}{2}$ *puruṣa* for the handle. If so, $3 - \frac{1}{3}$ is meant to be an approximation for $\sqrt{7}$; of course it is in excess by $\frac{1}{9}$ *puruṣa,* and this has been adjusted in the area of the rectangle chosen. The surmise that $3 - \frac{1}{3}$ was chosen as an approximation for $\sqrt{7}$ fits with the fact that it conforms to the standard approximation formula $\sqrt{a^2 - b} \approx a - \frac{b}{2a}$ recalled in Sect. 8.2, with $a = 3$ and $b = 2$.

In Datta (1932, pp. 203–205), one finds also a reference to "a serious attempt" having been made in Baudhāyana Śulbasūtra to find an approximation to the value of $\sqrt{5}$, which has also been reproduced uncritically in Sen and Bag (1983, pp. 174–175). The discussion is based on Baudhāyana's description of three alternative ways for fixing the location of *dakṣiṇāgni* with respect to the *āhavanīya* and *gārhapatya agnis* (which are along the east-west line, with the former to the east), on the southern side from them. The locations described for the *dakṣiṇāgni* are clearly meant to be possible variants that may be adopted, conforming to an overall scheme, and not meant to describe the same spot.[65] Our author, on the other hand, treats it as if an identical spot is being described and deduces approximate values of $\sqrt{5}$, and

[65] Without going into the full details, I may mention that broadly speaking a spot is being marked on a line segment parallel to the one joining *āhavanīya* and *gārhapatya agnis* which is at a certain distance toward the south from the latter, and somewhat westward from the midpoint of the parallel segment; while in one method the point is described in terms of the level on the segment joining the *āhavanīya* and *gārhapatya agnis,* in the other two it is described in terms of the distances from the two of them.

$\sqrt{2}$; the author even alludes to it as "an attempt, though without much success, to find an approximation ..." The assumption is all the more surprising, as two of the three methods in BSl, as well as the corresponding descriptions in ASl and KSl, involve using the same pattern of construction with different numbers of parts being made, which would obviously not be intended to describe the same spot. Not surprisingly the approximations arrived at are no good, with values for $\sqrt{5}$ ranging from 2.16 to 2.4, and those for $\sqrt{2}$ varying from 1.2 to 1.44 (!). The source for the gross inaccuracy in the result lies in the author's unjustified assumption that the exact same spot was being described, and does not have to do with the capability of the sūtrakāras to determine a good approximation to $\sqrt{5}$; there is no evidence that such a problem or task occurred to them, in the first place. A purposeful "attempt" to describe $\sqrt{5}$ would surely have found a better expression than a description of the same ritual spot in multiple ways.

9 Configurations of Shapes and Numbers

There are some interesting configurations of shapes and numbers appearing in the construction of various *vedi*s and *citi*s, which it would be appropriate to recall here. On the other hand, there are several unwarranted associations and claims found in literature which ought to be clarified. This section is aimed at meeting this dual objective.

9.1 Dārśapaurṇamāsika Vedi

This is a *vedi* which as mentioned in Sect. 7 incorporates some circular arcs in the design, starting with an outline consisting of an isosceles trapezium, with face, base, and altitude 48, 64, and 96 *aṅgulas,* respectively. The arcs are drawn within the trapezium (see Fig. 12) joining the pairs of adjacent vertices. The *vedi* is described in BSl I.71–75, ASl 4.5, and MSl 1.4. It has a long tradition and is mentioned in *Taittirīya Saṃhitā* and *Taittiīya Brāhmaṇa* (see Sen and Bag 1983, p. 175).

Plotting of the arcs is described as follows: Taking a cord with double the size of the *pārśvamānī*, a mark is made at its midpoint. Tying the ends of the cord to poles at the two vertices on the southern side, the cord is stretched southward on the ground, and a pole is put at the location of the midpoint mark. Then tying both points of the cord to that pole, the southern arc of the design is drawn, by tracing the midpoint, turning the stretched cord around (note that it would be an arc passing through the two southern vertices). There is some ambiguity whether *pārśvamānī* stands for AD as in the figure or the altitude of the trapezium (more likely the former), but clearly the intended curve is the segment of the circle with that as the radius, with the point on the southern side which is equidistant from the two vertices. The northern arc is instructed to be drawn similarly, and for the arcs on the east and west the instruction proceeds analogously, with a cord of double the length of the side itself.

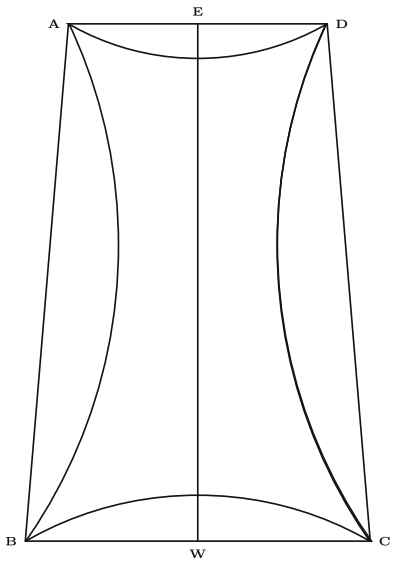


Fig. 12 *Dārśpaurṇamāsika vedi*

The arcs drawn are thus evidently a part of the design esthetics, coupled presumably with some metaphysical associations. One finds however a very fanciful interpretation of the arcs in Shylaja (2012), in terms of astronomy; it seems meaningless as there is no reference at all in the original to any external considerations, and the arcs are clearly meant to be segments of circles; while the name of the *vedi* only relates it to new moon and full moon, vague allusions are made in (Shylaja 2012) to eclipses and paths of planets, in this respect.

9.2 Tiling of *Gārhapatya*

For the *gārhapatya agni,* the layers were to be tiled with 21 bricks and for the five layers one needed two designs for the alternate layers, and they had to be such that the edges of the bricks in the successive layers do not overlap except at the boundaries of the whole figure. One such pair given is a simple one in which the square is divided into a grid of 21 congruent rectangles, with seven rows and three columns for one plan and three rows and seven columns for the other. Apart from this, there is another scheme described, consisting of division only into squares (of two different sizes); the tiling for the pair of layers is as in Fig. 13 (one presumes the arrangement was meant to be symmetrical as shown – the sūtras actually state only the number of tiles of the different sizes). For the first layer, from a square grid of four rows, the subdivision of the square formed by the middle four squares is replaced by one into a square grid with three rows, thereby enhancing the total number, by 5, to 21. For the second layer, starting with a square grid of three rows, four of the squares are subdivided into four squares each, bringing the total to 21;

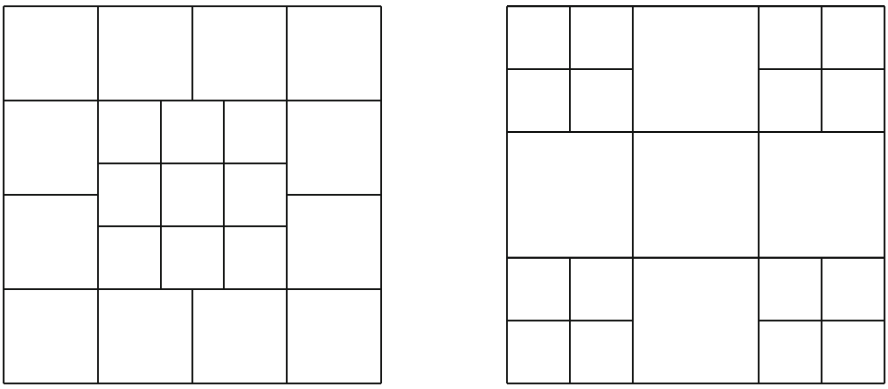


Fig. 13 Gārhapatya tiling

instead of the four corner squares as seen in the figure, choosing the four side squares would also produce a symmetrical design (there is no specification between these two possible choices) (Fig. 13).

In Datta (1932, pp. 180–182), the task of finding such patterns is projected as having been carried out via solutions of simultaneous indeterminate equations. It is quite far-fetched to suppose that Baudhāyana considered the simultaneous equations $\frac{x}{m^2} + \frac{y}{n^2} = 1$ and $x + y = 1$, as suggested in Datta (1932), and obtained solutions to the system (whether by a technique or by trial and error); there is no evidence to support such an assumption. On the other hand, interesting as the patterns produced, it can be seen that they are accessible to being carried out by intelligent boys and girls who have not yet been introduced to algebraic techniques. The geometers would in fact have been quite attuned to thinking of grids and making manipulations such as substituting a subgrid of two rows by one of three rows, etc.

Interestingly, for the circular version of *gārhapatya* a subdivision into 21 tiles is produced as follows: Starting with an inscribed square and a grid of squares in three rows, the chords are extended to the circle. This divides each of the four segments of the exterior of the inscribed square in the circle into 3 parts, thus adding 12 parts, taking the total number of parts to 21. The tiles at the boundary are of course no longer rectilinear but have rounded edges.

9.3 Tilings of *Kāmya Citis*

As noted earlier, the *kāmya citi*s consisted of five layers of brick-tilings, each layer of which had to have 200 bricks, and the edges of the bricks in the successive layers are not to overlap except at the edges of the *citi*. There is no restriction on the shapes of the bricks, and Baudhāyana uses four types of bricks. Recalling this background, Datta (1932) projects it as an issue of solving another system of simultaneous equations, which he introduces at the outset as "a much more difficult problem" (than the one involved in Sect. 9.2). Actually, while the tiling patterns introduced in

the Śulbasūtras for various shapes are esthetically appealing, and testify to noteworthy artistic skills, there is no serious mathematics involved in making the choices for the design; in particular, it does not involve solving any systems of linear equations. As has been commented in Seidenberg (1978, p. 338), one could first divide the designated area into suitably close to 200 parts, and follow it up by splitting up some of them to adjust for the deficit; similarly, if the parts go over 200 merging some of them could be considered, availing of the fact that multiple shapes were admissible for the tiles, providing flexibility in this respect. One can readily spot in the tilings where such adjustments have been made. As an example, let us consider Āpastamba's tilings for the simple design of *prauga citi* in ASl 12.7–11. It is a *citi* in the shape of an isosceles triangle (obtained by joining vertices at the base of a square to the midpoint of the opposite side). For the first layer, the scheme goes as follows. There is a primary subdivision of the triangle by dividing the three sides into 12 equal parts each, and joining the corresponding points (with the count starting from each vertex); this produces a division of the original triangle into 144 triangles. To increase the number to 200, a subdivision of a *small group of* triangles is undertaken. Some triangles are divided into four each (by joining the midpoints of its sides), while others are divided into two each (along the perpendicular bisector of the base). The former would increase the total number of parts by three with subdivision of each, while the latter will do so by one. With these provisions, one can then readily attain the target of 200 parts, as designated. As illustrated in Kulkarni (n.d.), 16 of the small triangles are divided into 4 each and 8 others into 2 each, contributing together to make up for the deficit of 56 – it could of course have been done in numerous other ways. For the second layer, a similar primary division is adopted with division of the sides into 13 parts each in place of 12, producing 169 triangles in the first instance, and dividing 7 of them into 4 each and 10 others into each 2 each making up the deficit of 31.[66] The designs described for various *citis* can be analyzed in a similar way in terms of a primary division and subdivisions for adjustments. It may also be noticed that sometimes the adjustments seem to have been done at the cost of losing bilaterial symmetry, which one may suppose they would normally have liked to maintain.

It may also be pointed out that the issue of nonoverlap of edges in successive layers would have been a routine issue in brick-laying, which is historically a much older activity, in the Indian context, if one keeps in mind the Indus valey civilization. Using lateral shifts or rotations of patterns would have been used, and one witnesses similar techniques applied in the Śulbasūtras, with some extra artistic still. There is however no serious mathematical issue at play here.

[66] One could also have first divided the triangle into 225 triangles in a similar way, merged nine groups of 4 from among them into double sized triangles, thereby reducing the number to 198, and then added 2 by splitting 2 triangles into 4, to bring the total to 200.

9.4 Configuration of *Rathacakraciti*

The *Rathacakraciti* came in two varieties, one a flat round wheel in the shape of a disc, and another one with spokes. The latter, described in BSl V.9–36, had three components, consisting of a (circular) nave, and outer annular rim and the middle part with the spokes (and the gaps between them), and the distribution of the designated total area for the *citi* ($7\frac{1}{2}$ *puruṣa*, when being performed the first time) is seen to involve an interesting play with numbers. For the purpose of the procedure of the construction, the designated area for the citi is realized as consisting of 225 square bricks of uniform size (of area $\frac{1}{30}$ *puruṣa* when the area is $7\frac{1}{2}$ *puruṣa*). Baudhāyana then adds to this the area of 64 such squares for the area of the gaps between the spokes in the middle region, thus getting 289 squares for the area of the circle corresponding to the outer edge of the felly (rim). The nave is assigned 16 of the squares, and the area of the circle corresponding to the inside edge of the felly is taken to be 144 squares. Thus, the area between the nave and the rim is 128 squares, half of which agrees with the 64 corresponding to the gaps. Now squares of area sizes 289, 144, and 16 of these squares, viz, 17, 12, and 4 times the side of the squares as above, are plotted with a common midpoint. These squares are then transformed into circles by the procedure described in Sect. 7.2, producing the concentric circles corresponding to the nave and the inner and outer edges of the felly. In the middle region, 16 spokes are plotted evenly, so that the gaps account for the area of 64, as prearranged by the above choices. Thus, the area of the *citi* works out to 225 of the squares as above, which equals the designated area for the *citi* as per the choice made. The tiling of the *citi* into 200 bricks is done independently. We shall not go into the details here, except to mention that it neatly uses splitting of 200 as $64 \times 3 + 8$; in one of the designs the 8 parts are assigned to the nave, 4 each to the 16 spokes and division of the felly into 2 rings of 64 parts.

One may spot in this procedure an interesting numerical relation $12^2 - 4^2 = 2(17^2 - 15^2)$; it may be noted that the task itself does not warrant the sizes involved to be in integers. As Seidenberg remarks, this shows noteworthy facility with numbers. It does not however indicate any systematic development of techniques with regard to indeterminate equations.

Before concluding this discussion, I would like to clarify a curious confusion found in the literature in respect of Baudhāyana's description of the procedure as above. The sūtras, after arriving at the collection of the 289 squares, BSl V.13–16 go on to say "The side of this square comprises 16 bricks. Thirty-three bricks still remain. These are to be placed on all sides round the borders (to make a square)" (cf. Prakash et al. 1968a, p. 165–166) (parenthetical explanations adjusted to the context here). In this respect, Thibaut, in his foundational work (Thibaut 1875, p. 261), comments "It is strange that we are not direted to construct the whole square at once, but are told to form at first a square out of 256 bricks and then to place the remaining 33 bricks around it." He then goes on to record a long-winded explanation of the commentator (Dvārakānātha) which he evidently finds unsatisfactory, and one would agree with him on it. Recalling Thibaut's question, in Seidenberg (1978,

p. 333) the author offers an explanation, which however seems to bring in irrelevant aspects and ultimately hardly connects meaningfully with the original text and Thibaut's query. In the translation in Sen and Bag (1983, p. 94), the authors parenthetically add "so as to obtain a square of side containing 17 bricks," throwing no light on why it was not taken up directly. The translations in Prakash et al. (1968a) and Kulkarni (n.d.) seem to be of no help on the issue either.

A simple explanation is as follows. Baudhāyana wanted to begin with a square whose size is made of an even number of the small squares, so that he can readily make sense of the *middle square* consisting of 16 of the original squares, for the nave (and also similarly a square of width 12 squares for the inner circle of the felly); it would not make sense to take middle 16 squares from a square-grid of 17 rows. Keeping that objective, Baudhāyana further wanted the remaining 33 squares to be placed around the 16×16 square without disturbing the centrality of the square for the nave; the commentator's interpretation to place them "on the north side and east side of the square," as included parenthetically in Prakash et al. (1968a) and other similar words in Sen and Bag (1983), is plain wrong. The commentators (ancient as well as modern) seem to have missed the point that Baudhāyana's instruction, in BSl 5.16, as "*tābhirantansarvataḥ paricinuyāt*" (These are to be placed on all sides round the borders) does not correspond to transporting the squares individually, but incorporating their area, in the same spirit as transformation of figures that we discussed in earlier sections. Thus, a natural choice for following the instruction would be to divide each square into two rectangular halves and to arrange the 66 rectangles along the border of the 16×16 square, turning it into a 17×17 square; see Fig. 14 (the figure in Sen and Bag

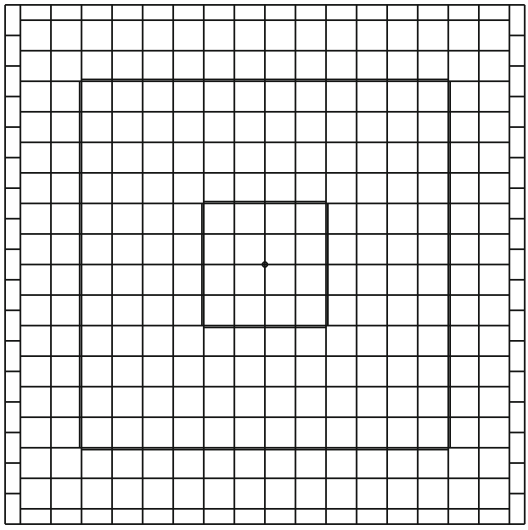


Fig. 14 Preparing for *Rathacakraciti*

1983, on p. 217, is misleading and does not correspond to Baudhāyana's instruction, though it presents a plan that Baudhāyana could have adopted, by constructing a 17×17 square straightaway, and describing the squares for the nave and the inner edge of the felly differently, to render them central).

10 In Place of Conclusion

The Śulbasūtras provide us a unique window to the mathematics in India in the millennium BCE; while there are indeed other mathematical developments traceable from that period, as also some of the earlier periods such as from the Indus valley civilization, they are based on sources that are either indirect or relatively scanty in comparison. The long time frame, extending over at least half a millennium if not much more, which they represent as a historical period with a relatively constant cultural milieu, brings an additional attraction, motivating their study.

In this review, I have highlighted various significant mathematical ideas evidenced in the Śulbasūtras, together with an overall profile of the mathematical edifice that got built up following the pursuit inspired by the projects for erecting elaborate and intricate altars and fireplaces for sacrificial worship. In doing so, I have endeavored to maintain a high degree of veracity and discussed the contours of various exaggerated assertions that have been in vogue and also tried to clarify some of the misunderstandings. Without any claims about the narrative being error-free, the author hopes that it will provide a clearer picture of the mathematical contents and their significance in respect of the historical development of mathematics.

The inquiry into the mathematics in the Śulbasūtras raises a variety of questions, or what we may even call puzzles. Some of these pertain to their mathematics, which we have noted in the course of this review. There are deeper puzzles from the point of view of history of mathematics. A principal one that comes to mind is why do the works seem so stagnant, or even declining in standard, with Baudhāyana's oldest Śulba standing out as the beacon for such a long tradition. While there are a few flourishes seen in the later Śulba works, that we have noted, a good deal of mathematically interesting parts are seen to be forgotten or ignored in the later woks. It may also be recalled, as noted in Datta et al. (1980, p. 124), geometry did not make much progress in the post-Vedic period in India, until much later, following the impact of Greek mathematics around the turn of the millennium. One may also wonder in this context about the mode of transmission, geographically across the country and from generation to generation, and how it may have brought about the change in content. Answers to these questions may hold the key to what lessons may be drawn from the study, for our times. It is hoped that this detailed review of the works will contribute in making progress toward understanding these issues.

Appendix

Concordance of sūtras between Śulbasūtras

Topic	Section	BSl	MSl	ASl	KSl
Constr. of perpendicular by "compass" method	3.1	I.22–28	–	–	–
Nyañchana method with (3, 4, 5)	3.2	I.30–35	1.3.6	4.11	1.17
Nyañchana method with (5, 12, 13)	3.2	I.42–44	1.1.11–12	1.3–6 1.7	1.12 1.14
Constr. of perpendicular by "arrow" method	3.3	I.36–40	–	1.13–18	–
Nyañchana method with other triples	3.5	–	–	5.13–17	–
Constr. of rhombus	3.6	I.57	3.6.4	12.12–15	4.6
Transf. of square to an isosceles trapezium	3.8.1	I.55	–	–	–
Transf. of square to a rectangle (special)	3.8.4	I.52	–	–	3.4
Transf. of square to a rectangle (general)	3.8.5	I.53	–	3.1	–
Pythagorean principle	4	I.48	3.1.10	1.8	2.11
Combining two squares	5.1	I.50	3.3.6	2.10–13	2.22
Enhancement of figures by scaling	5.2	II.1–5	–	8.16	5.x, 6.2–5
Combining multiple squares	5.3	–	–	–	6.7
Difference of two squares	5.4	I.51	–	2.14–17	3.1
Constr. of a square equivalent to a rectangle	5.5	I.54		2.21–23	3.2–3
Pythagorean triples	6.3	I.49	–	5.3–17	–
Circumference of a circle	7.1	I.112–113	3.2.13	–	–
Constr. of a circle equivalent to a square	7.2 7.4	1.58	1.1.8 3.2.15	3.2–5	3.13
Squaring the circle (refined)	7.6	1.59	–	–	–
Squaring the circle (approx.)	7.7	1.60	3.2.13	3.6–8	3.14
Evaluation of $\sqrt{2}$	8	1.61	3.3–1.2	1.10	2.13

References

Amma TAS (1979) Geometry in ancient and medieval India. Motilal Banarsidas, Delhi

Bag AK (1979) Mathematics in ancient and medieval India. Chaukhambha Orientalia, Varanasi/Delhi

Bloomfield M (1908) The religion of the Veda – the ancient religion of India. G.P. Putnam's Sons/The Knickerbocker Press, New York/London

Bürk A (1901) Das Āpastamba Śulbasūtra, herausgegeben, übersetzt und mit einer Einleitung versehen. Zeitschrift der Deutschen Morgenländischen Gesellschaft (in German) 55:543–591 and 56 (1902), 227–291
Dani SG (2003) On the Pythagorean triples in the Śulbasūtras. Curr Sci 85:219–224
Dani SG (2010) Geometry in the Śulbasūtras. In: Studies in the history of Indian mathematics, Culture and history of mathematics, vol 5. Hindustan Book Agency, New Delhi, pp 9–37
Dani SG (2018) Kātyāyana Śulbasūtra: some observations. Gaṇita Bhāratī 40(2):101–114
Dani SG (2022) Some constructions in Mānava śulvasūtra, Proceedings of Kanchipuram conference. In: Ram SS, Ramakalyani V (eds) History and development of mathematics in India. Published by National Mission for Manuscripts, New Delhi, copublished by D.K. Printworld (P) Ltd., New Delhi, pp 373–385
Dani SG (2023) Geometry in Ancient Jaina works – a review. In: Dani SG, Jain SK (eds) Mathematics in Ancient Jaina literature (Ohio University, USA), with the cooperation of Shah PK (eds) (Jain Centre of Greater Boston, MA, USA). World Scientific Publishers, pp 57–86
Dani SG, Limaye M (2021) On some geometric constructions in the Sulvasutras, from a pedagogical perspective, Aziz Premji University at Right Angles, Parts I and II, March and July issues, pp 7–16 (in both the issues)
Datta B (1932) Ancient Hindu geometry: the science of the Śulba. Calcutta University Press; reprint: Cosmo Publications, New Delhi, 1993
Datta B, Singh AN (1962) History of Hindu mathematics; A source book, Parts I and II. Asia Publishing House; reprint: Motilal Banarasidas, Delhi, 2001
Datta B, Singh AN, revised by Shukla KS (1980) Hindu Geometry. Indian J Hist Math 15(2): 121–188
Delire J-M (2003) The Baudhāyana Śulbasūtra's commentaries: Edition, translation and explanation. In: Proceedings of the XIth World Sanskrit Conference, Turin, 2000, Indologica Taurinensia, Official Organ of the International Association of Sanskrit Studies, vol XXIX, pp 123–143
Delire J-M (2005) Quadratures, circulature and the approximation of $\sqrt{2}$ in the Indian Śulba-sūtras. Centaurus 47(1):60–71
Delire J-M (2009) Chronological inferences from a comparison between commentaries on different Śulbasūtras. In: Wujastyk D (ed) Papers of the 12th World Sanskrit Conference, Helsinki, 2003, Mathematics and medicine in Sanskrit, pp 37–62
Delire J-M (2016) Des mathematiques de l'autel védique – Le Baudhāyana Śulbasūtra et son commentaire Śulbadīpikā, Ecole Pratique des Hautes Études, Sciences Historiques et Philologiques – II. Librairie Droz S.A, Genève
Divakaran PP (2018) The mathematics of India – concepts, methods, connections, culture and history of mathematics 10. Hindustan Book Agency, New Delhi
Gillings RJ (1972) Mathematics in the time of pharaohs. Dover Publications, New York
Gupta RC (1988) New Indian values of π from the Mānava Sulvasūtra. Centaurus 31:114–125
Gupta RC (1990) The chronic problem of ancient Indian chronology. Gaṇita Bhāratī 12(1–2):17–26
Gupta RC (2001) The study of history of mathematical sciences in India. Gaṇita Bhāratī 23(1–4):1–11
Gupta RC (2006) Śulvasūtras: earliest studies and a newly published manual. Indian J Hist Sci 41: 317–320
Hayashi T (1990) A new Indian rule for the squaring of a circle: Mānava Śulbasūtra 329–10. Ganita Bharati 12:75–82
Heath T (1885) Diophantus of Alexandria, A study in the History of Greek Algebra. Cambridge University Press; reprint: Dover publications, New York, 1964
Heath T (1926) Translation of "The Thirteen Books of Euclid's Elements" with Commentary, vol I. Cambridge University Press; Second Edition, Dover Publications, 1956
Joseph GG (2016) Indian mathematics, engaging with the world from ancient to modern times. World Scientific, Singapore
Kashikar CG (1966) A summary of Śrautasūtras. J Univ Bombay, No. 35 (New series), Part II

Kichenassamy S (2006) Baudhāyana's rule for the quadrature of the circle. Hist Math 33(2): 149–183
Kularia RP 2016 Maitrayaniya Sulbasutram (with Sanskrit commentary of Sankara, Hindi commentary Sulbamitra, English Translation, Explanations and Geometrical Figures), Abhishek Publications
Kulkarni RP (1983) Geometry according to Śulba Sūtra. Vaidika Saṃśodhana Maṇḍala, Pune
Kulkarni RP (n.d.) Char Shulbasūtre (in Marathi), Maharashtra Rajya Sahitya Sanskriti Mandal, Mumbai, 1978; Hindi translation Char Shulbasūtra published by Maharshi Sandipani Rashtriya Vedavidya Pratishthana, Ujjain, 2000
Plofker K (2007) In: Katz V (ed) The mathematics of Egypt, Mesopotamia, China, India and Islam, A sourcebook. Princeton University Press, Princeton/Oxford
Plofker K (2009) Mathematics in India. Princeton University Press, Princeton. xiv +357 pp
Prakash S, Sharma RS, Śulbasūtram B, with Sanskrit commentary by Yajvan D and English translation and critical notes by Thibaut G (1968a) The Research Institute of Ancient Scientific Studies, New Delhi
Prakash S, Sharma RS, Śulbasūtra, Ā, with commentaries from Kapardswamin, Karvinda and Sundararaja and English translation by Prakash S (1968b) The Research Institute of Ancient Scientific Studies, New Delhi
Puttaswamy TK (2012) Mathematical achievements of pre-modern Indian mathematicians. With a foreword by Ubiratan D'Ambrosio. Elsevier, Amsterdam, p xxiv +743
Robson E (2007) Mesopotamian mathematics. In: Katz VJ (ed) The mathematics of Egypt, Mesopotamia, China, India, and Islam: a sourcebook. Princeton University Press, p 143
Rodet ML (1879) Sur les méthodes d'approximation chez les anciens. Bull Soc Math France 7: 159–167
Seidenberg A (1972) On the area of a semi-circle. Arch Hist Exact Sci 9:173–210
Seidenberg A (1975) The ritual origin of geometry. Archr Hist Exact Sci (Springer, Berlin) 1(5): 488–527
Seidenberg A (1978) The origin of mathematics. Archr Hist Exact Sci (Springer, Berlin) 18(4): 301–342
Seidenberg A (1981) The ritual origin of the circle and square. Archr Hist Exact Sci (Springer, Berlin) 25(4):269–327
Sen SN, Bag AK (1983) The Śulbasūtras. Indian Natl Sci Acad, New Delhi
Shylaja BS (2012) Asymmetrical Vedis in Śulbasūtras. Indian J Hist Sci 47(2):271–280
Srinivasiengar CN (1967) The history of ancient Indian mathematics. World Press, Calcutta
Thibaut G (1875) The Śulbasūtras. J Asiatic Soc Bengal, Part I:227–285. Printed by C.B. Lewis, Baptist Mission Press, Calcutta
Thibaut G (n.d.-a) Śulvasūtra of Baudhāyana. The Paṇḍiṭ, vol 9 (1874), vol 10 (1875), n.s. vol 1 (1876–77)
Thibaut G (n.d.-b) Kātyāyana's Śulbapariśiṣṭa. The Paṇḍiṭ, n.s. vol 4 (t882)
van der Waerden BL (1983) Geometry and algebra in ancient civilizations. Springer Verlag
van Gelder JM (1963) The Mānava Śrautasutras, English translation, International Academy of Indian Culture and Aditya Prakashan, New Delhi

An Ethnoarithmetic Excursion into the Javanese Calendar

Natanael Karjanto and François Beauducel

Contents

1 Introduction 1098
2 Ancient and Modern Calendars 1101
2.1 Pre-Gregorian Calendars 1101
2.2 Gregorian Calendar 1102
3 Javanese Calendar 1104
3.1 Where Is Java? 1104
3.2 Who Are Javanese People? 1105
3.3 A Background of the Javanese Calendar 1105
3.4 Some Characteristics of the Javanese Calendar 1105
3.5 Computer Implementation of the Javanese Calendar 1117
4 Discussion and Epilogue 1119
4.1 Javanese View of Time 1119
4.2 *Pranatamangsa* Revisited 1120
4.3 Conclusion 1121
References 1122

Abstract

A perpetual calendar, a calendar designed to find out the day of the week for a given date, employs a rich arithmetical calculation using congruence. Zeller's congruence is a well-known algorithm to calculate the day of the week for any

N. Karjanto (✉)
Department of Mathematics, University College, Sungkyunkwan University, Suwon, Republic of Korea
e-mail: natanael@skku.edu

F. Beauducel
Université de Paris, Institut de physique du globe de Paris, CNRS, Paris, France

Institut de recherche pour le développement, Research and Development Technology Center for Geological Disaster, Balai Penyelidikan dan Pengembangan Teknologi Kebencanaan Geologi (BPPTKG), Yogyakarta, Indonesia
e-mail: beauducel@ipgp.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_82

Julian or Gregorian calendar date. Another rather infamous perpetual calendar has been used for nearly four centuries among Javanese people in Indonesia. This Javanese calendar combines the *Saka* Hindu, lunar Islamic, and western Gregorian calendars. In addition to the regular 7-day, lunar month, and lunar year cycles, it also contains 5-day *pasaran*, 35-day *wetonan*, 210-day *pawukon*, octo-year *windu*, and 120-year *kurup* cycles. The Javanese calendar is used for cultural and spiritual purposes, including a decision to tie the knot among couples. In this chapter, we will explore the relationship between mathematics and the culture of Javanese people and how they use their calendar and the arithmetic aspect of it in their daily lives. We also propose an unprecedented congruence formula to compute the *pasaran* day. We hope that this excursion provides an insightful idea that can be adopted for teaching and learning of congruence in number theory.

Keywords

Javanese calendar · Ethnomathematics · *Pasaran* · *Wetonan* · *Pawukon* · *Windu* · *Kurup* · Congruence · Indonesia

1 Introduction

Arithmetic and number theory find applications in various cultures throughout the world. In addition to solving everyday problems using elementary arithmetic operations, our ancient ancestors also developed and invented perpetual calendars without any aid of modern electronic calculators and the computer. A perpetual calendar is a system dealing with periods of time that occur repeatedly. It is organized into days, weeks, months, and years. A date of a calendar structure refers to a particular day within that system. A calendar is used for various civil, administrative, commercial, social, and religious purposes.

The English word "calendar" is derived from the Latin word which refers to the first day of the month in the Roman calendar. It is related to the verb *calare* ("to announce solemnly, to call out"), which refers to the "calling" of the new moon when it was visible for the first time (Brown 1993). Another source mentions that the modern English "calendar" comes from the Middle English *calender*, which was adopted from the Old French *calendier*. It originated from the Latin word *calendarium*, which meant a "debt book, account book, register." In Ancient Rome, interests were tracked in such books, accounts were settled, and debts were collected on the first day (*calends*) of each month (Stakhov 2009).

Generally, the calculation of and periods in the calendrical system are synchronized with the cycle of the Sun or Moon. Hence, the names solar, lunar, and lunisolar calendars, where the latter combined both the Moon phase and tropical (solar) year. Our current Gregorian and the previously Julian calendars are solar-based calendars, as well as the ancient Egyptian calendar. An example of the lunar calendar is the Islamic calendar. Prominent examples of the lunisolar calendar are the Hebrew,

Chinese, Hindu, and Buddhist calendars (Richards 1998; Reingold and Dershowitz 2018).

Studying calendar systems from any culture cannot be separated from attempting to gain insight and understanding not only in the arithmetic behind them but also in the interconnectivity with mathematics, the history of mathematics, astronomy, mathematics education, and sociocultural aspects. Hence, the study of perpetual calendars is encompassed by the fields of ethnomathematics and ethnoastronomy, the research areas where diverse cultural groups embrace, practice, and develop mathematics and astronomy in their daily lives.

Figure 1 summarizes a theoretical framework used for the discussion on the Javanese calendar considered in this chapter. Although the title of this chapter contains the term "ethnoarithmetic," indeed that covering a topic on the Javanese calendar embodies other aspects beyond the cultural and mathematical, in this case, arithmetical aspect per se. As we observe at the third level of the Venn diagram presented in Fig. 1, an excursion into the Javanese calendar intersects and includes four distinct but related disciplines: mathematics, education, culture, and astronomy.

Fundamental work in the area of ethnomathematics has been documented by Ascher (2002). The author elaborated on several mathematical ideas and their cultural embedding of people in traditional or small-scale cultures, with some emphasis on time structure and the logic of divination. In particular, the book also covered an explanation of the Balinese calendar. For discussion of the Balinese

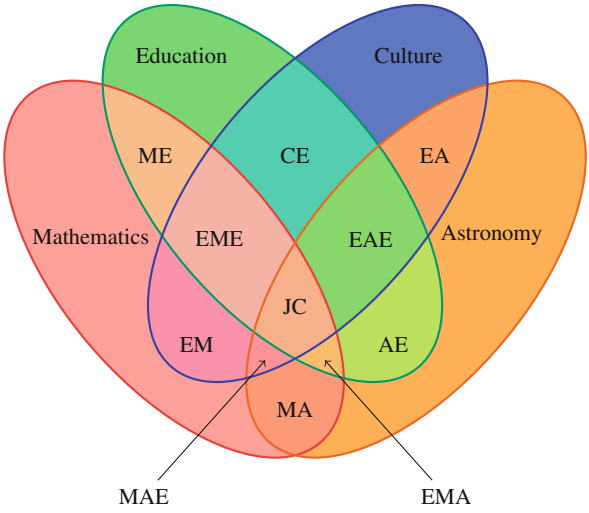


Fig. 1 A theoretical framework for an ethnoarithmetic study of the Javanese calendar. At the first level of the intersection, ME denotes mathematics education, EM denotes ethnomathematics, MA denotes mathematics astronomy, AE denotes astronomy education, EA denotes ethnoastronomy, and CE denotes cultural education. At the second level of the intersection, EME denotes ethnomathematics education, MAE denotes mathematics astronomy education, EMA denotes ethnomathematics astronomy, and EAE denotes ethnoastronomy education. The third level of the intersection is the heart of discussion; we have JC as the Javanese calendar

calendrical system and its multiple cycles, consult Vickers (1990); Darling (2004); Proudfoot (2007); Ginaya (2018); and Gislén and Eade (2019c).

To the best of our knowledge, the only monograph that discusses an intersection of three components of the theoretical framework is a volume edited by Rosa et al. (2017). The book covers diverse approaches and perspectives from ethnomathematics to mathematics education, particularly in several non-Western cultures. Since these topics stimulate debates not only on the nature of mathematical knowledge and the knowledge of a specific cultural group but also on the pedagogy of mathematics classroom, the field of ethnomathematics certainly offers a possibility for improving mathematics education across cultures.

A collection of essays dealing with the mathematical knowledge and beliefs of cultures outside the Western world is compiled by Selin and D'Ambrosio (2000). The essays address the connections between mathematics and culture, relate mathematical practices in various cultures, and discuss how mathematical knowledge is transferred from East to West. A coverage of calendars in various cultures is also briefly touched.

Since many calendrical systems are invented based on the movement of celestial bodies, particularly the Sun and Moon, we cannot dismantle the role of astronomy in the study of calendars, including the Javanese one. Various civilizations incorporate the cyclical movements of both the Sun and Moon into their calendrical systems (Ruggles 2015). Some examples of this calendrical-astronomical relationship can be observed among others in the Jewish (Cohn 2007), Indian (Dershowitz and Reingold 2009), Chinese (Martzloff 2000, 2016), mainland Southeast Asian (Eade 1995), and Islamic calendars (Proudfoot 2006).

An explanation of the mathematical and astronomical details of how many calendars function has been covered extensively by Reingold and Dershowitz (2018). A related series with non-Western ethnomathematics is a book on non-Western ethnoastronomy that has been edited by Selin and Sun (2000). In particular, the book also dedicated one chapter on an astronomical feature of the Javanese calendar, where a season keeper (*pranotomongso* or *pranata mangsa*) guides agricultural activities among rural peasants in Java (Daldjoeni 1984; Ammarell 1988; Hidayat 2000). The readers might be interested to compare this with the cultural production of Indonesian skylore across three ethnic groups: Banjar Muslim, Meratus Dayak, and Javanese peoples (Ammarell and Tsing 2015).

Integrating the Javanese calendar into elementary school education as an ethnomathematics study has been attempted by Utami et al. (2020). Another attempt is to embed some topics related to calendrical systems into the first-year seminar on the mathematics of the pre-Columbian Americas (Catepillán 2016). A study from Taipei, Taiwan, on the movements of the Moon at the primary level, introduced pupils to both the Gregorian and Chinese calendars in the context of physical classroom environments (Hubber and Ramseger 2017).

Although we propose an intersection of only four disciplines in our theoretical framework, the list is by no means exhaustive. Another possibility is to include psychology or the interaction between mathematics and psychology. For instance, a theoretical foundation on the calendrical system and the psychology of time has been

modeled by Rudolph (2006). In particular, the author proposed that Balinese (also equally applied to Javanese) time might be neither "circular" nor "linear," but *profinite*. Since the Javanese calendrical systems involve 5-day *pasaran*, 7-day *saptawara/dinapitu*, and 30 *wuku* cycles, it could be modeled by the group of *profinite integers* $\widehat{\mathbb{Z}}$. This group bundles together all different sets of the ring of p-adic integers $\mathbb{Z}_p$ and various finite sets $\mathbb{Z}/(p)$ of integers modulo p, where p denotes prime integers (Milne 2020).

A discussion on calendar timekeeping scheme in Southeast Asia is covered by Gislén (2018). An overview of the calendars in Southeast Asia is given by Gislén and Eade (2019a). In their subsequent papers, they and Lân also discussed calendars from Burma, Thailand, Laos, and Cambodia (Gislé and Eade 2019b), Vietnam (Lân 2019), and Malaysia and Indonesia (Gislén and Eade 2019c), eclipse calculation (Gislén and Eade 2019d), and chronicle inscriptions (Gislén and Eade 2019e). See also Eade (1995), Ôhashi (2009), and (Golzio 2012) for basic facts and further explanations of the calendrical systems in India and (Mainland) Southeast Asia.

This chapter is organized as follows. After this introduction, the following section briefly covers the calendars from ancient and contemporary times, which includes the pre-Gregorian and Gregorian calendars. After that, a discussion on the Javanese calendar in the context of ethnoarithmetic will be dedicated exclusively in one section. The final section presents discussion and epilogue.

2 Ancient and Modern Calendars

This section briefly covers several calendars in the pre-Gregorian era and Zeller's congruence algorithm for the Gregorian calendar.

2.1 Pre-Gregorian Calendars

Before the Gregorian calendar that we are using today was adopted, numerous calendars have been used in various parts of the world. The Egyptian calendar was among the first solar calendars with its history dated back to the fourth millennium BCE.

A Mesoamerican civilization of the Maya peoples also developed a calendar system where the year was divided into 18 months of 20 days (Ascher 2002; Stakhov 2009). The Mayan calendar system contains three separate calendars, and although they are not related mathematically, all of them are linked in a single calendar system. The first one is called the "long count," used by the Mayans to measure date chronology for history recording. The second one is the *Haab* calendar, a nonchronological civil calendar. The third one is the Mayan religious calendar, also nonchronological, called the *Tzolkin* (Ascher 2002; Cohn 2007).

Related to the Maya civilization is the Inca Empire, where the latter also developed its calendar system, which was based on several different astronomical cycles,

including the solar year, synodic and sidereal lunar cycles, and local zenith period (Urton and Llanos 2010). A variety of mathematical developments among the native Americans from the prehistoric to present has been compiled by Closs (1996). In particular, one chapter of the book discusses the calendrical system of the Nuu-chah-nulth (formerly referred to as the Nootka), one of the Indigenous peoples of the Pacific Northwest Coast in Canada (Folan 1986).

Some calendars are based on lunisolar, and one of them is the Jewish calendar. Also called the Hebrew calendar, it is still used until today, mainly for Jewish religious observance. Although the Jewish calendar was developed in its current format during the Talmud period in the fifth century CE by Rabbi Hillel II, the up-to-date counting has accumulated more than 5000 years (Ascher 2002; Cohn 2007). For our information, AM 5781 began at sunset on Friday, 18 September 2020, and will end at sunset on Monday, 6 September 2021 CE. Here, AM means *Anno Mundi*, the Latin phrase for "in the year of the world."

Although the traditional Chinese calendar is also a lunisolar type, the most fundamental component is the sexagenary cycle, i.e., a cycle of 60 terms marked by coordination between 12 celestial stems and 10 terrestrial branches, which is also known as the "Stems-and-Branches" or *gānzhī* (干支) (Martzloff 2000). A discussion on the mathematical aspect of the Chinese calendar has been covered extensively by Aslaksen (2001, 2002, 2006). A historical aspect of the Chinese calendar is discussed by Sun (2015). Astronomical aspects and the mathematical structures underlying the calculation techniques of the Chinese calendar are highlighted by Martzloff (2016).

The *Hijri*, or Islamic calendar, is a lunar calendar consisting of 12 lunar (synodic) months in a year of 354 or 355 days. It is still widely used in predominantly Muslim countries alongside the Gregorian calendar, primarily for religious purposes. The current counting started from 622 CE, commemorating the emigration of Prophet Muhammad and his companions from Mecca to Medina (*hijra*) (Hassan 2017). In the Gregorian calendar reckoning, the current Islamic year is 1422 AH, which approximately runs from 20 August 2020 until 9 August 2021. AH means *Anno Hegirae*, the Latin phrase for "in the year of the Hijra." In particular, the history of Muslim calendars in Southeast Asia in the historical and cultural context has been successfully traced by Proudfoot (2006).

2.2 Gregorian Calendar

The language of congruences is a basic building block in arithmetic and number theory. It allows us to operate with divisibility relationships in a similar way as we deal with equalities. Congruences have many applications, and one of them is to determine the day of the week for any date for a given perpetual calendar. In particular, the procedure for finding the day of the week for a given date in the Gregorian calendar has been discussed by Carroll (1887); Conway (1973); Burton (2011); and Rosen (2011). See also Gardner (1996) and Cohen (2000).

The Gregorian calendar was promulgated by Pope Gregory XIII and is a replacement of the Julian calendar, proposed by Julius Caesar in 46 BC. Starting from 1582, the Catholic states were among the first countries to adopt the Gregorian calendar by skipping 10 days in October: Thursday, 4 October 1582, was followed by Friday, 15 October 1582. Greece and Turkey are among the last countries that adopted the Gregorian calendar when they changed on 1 March 1923 and 1 January 1926, respectively.

Let W denote the day of the week from Saturday = 0 to Friday = 6, k denote the day of the month, m denote month, and N denote year. For the month, the convention is March = 3, April = 4,. . ., but January = 13, and February = 14. For the year, N is the current year unless the month is January or February, for which N is the previous year. The relationship between year and century is given by $N = 100\ C + Y$, where C denotes zero-based century and Y denotes a two-digit year. Note that the purpose of adopting this expression is not to confuse C with the standard century number, which is $C + 1$ for the first 99 years. The formula for finding the day of the week W of day k of month m of year N is given by Zeller's congruence algorithm (Zeller 1882, 1883, 1885, 1887):

$$W \equiv \left(k + \left\lfloor \frac{13(m+1)}{5} \right\rfloor + Y + \left\lfloor \frac{Y}{4} \right\rfloor + \left\lfloor \frac{C}{4} \right\rfloor - 2C\right) (\text{mod } 7). \tag{1}$$

Here, $\lfloor x \rfloor$ denotes the floor function or greatest integer function and mod is the modulo operation or remainder after division.

As an example, we are interested in finding the days of the week when a Javanese heroine and educator Raden Adjeng Kartini was born and passed away. She was born in Jepara, a town on the north coast of Java, located around 80 km to the northeast of Semarang, the present capital of Central Java province in Indonesia. Kartini was a pioneer for girls' education and women's emancipation rights in Indonesia, at the time when Indonesia was still a part of the Dutch East Indies colonial empire. She could be correlated with her European counterparts, including an English advocate of women's right Mary Wollstonecraft (27 April 1759 – 10 September 1797) or a Finnish social activist Minna Canth (19 March 1844 – 12 May 1897) (cf. Kartini 1911, 1920; Wollstonecraft 1792; Canth 1885).

Kartini was born on 21 April 1879, so we have $k = 21$, $m = 4$, $N = 1879$, $C = 18$, and $Y = 79$. She passed away on 17 September 1904, at the age of 25, and we identify that $k = 17$, $m = 9$, $N = 1904$, $C = 19$, and $Y = 4$. Using Zeller's congruence formula (1), the days when Kartini was born W_b and passed away W_d can be calculated as follows, respectively:

$$W_b \equiv \left(21 + \left[\frac{13(5)}{5}\right] + 79 + \left[\frac{79}{4}\right] + \left[\frac{18}{4}\right] - 36\right) (\text{mod } 7)$$

$$\equiv (21 + 13 + 79 + 19 + 4 - 36) (\text{mod } 7) \equiv 100 (\text{mod } 7) \equiv 2 (\text{mod } 7)$$

$$W_d \equiv \left(17 + \left[\frac{13(10)}{5}\right] + 4 + \left\lfloor\frac{4}{4}\right\rfloor + \left\lfloor\frac{19}{4}\right\rfloor - 38\right) \pmod 7$$
$$\equiv (17 + 26 + 4 + 1 + 4 - 38) \pmod 7 \equiv 14 \pmod 7 \equiv 0 \pmod 7.$$

Hence, Kartini was born on Monday, 21 April 1879 and passed away on Saturday, 17 September 1904. She was buried at Bulu Village, Rembang, Central Java, around 100 km east of Jepara.

3 Javanese Calendar

This section features the main characteristics of the Javanese calendar. We start with the geographical location of the island of Java, the Javanese people, and a brief historical background of the calendar. After providing a detailed discussion, we close the section by computer implementation of the Javanese calendar.

3.1 Where Is Java?

Java is an island in Indonesia, not a programming language, although the latter was renamed after some Javanese coffee by its founders. It is located in the southern hemisphere, around 800 km from the Equator. It extends from latitude 6° to 8° south, and longitude 105° to 114° east. With an area of 150,000 square kilometers, it is about 1000 km long from west to east and around 200 km wide from north to south. The island lies between Sumatra to the west and Bali to the east. It is bordered by the Java Sea on the north and the Indian Ocean on the south (see Fig. 2). It is the world's 13th largest island (Cribb 2000).

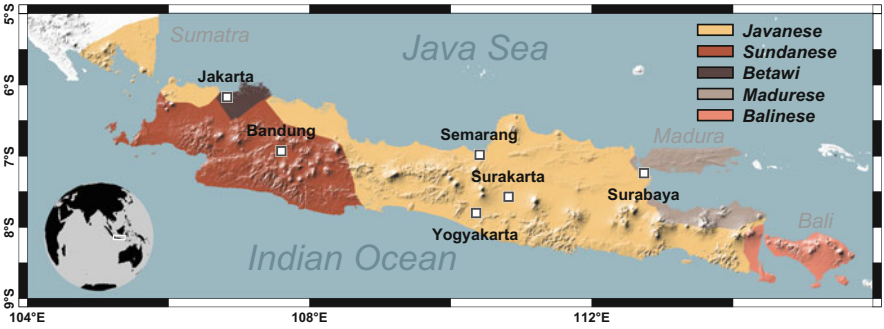


Fig. 2 The situation of the Java island, main cities, and present spoken languages: Javanese (Central and East Java, and a small enclave in North-West Java), Sundanese (West Java), Betawi (in and around Jakarta metropolitan area), Madurese (Madura Island and a part of North-Eastern Java), and Balinese (Bali Island and a small part of Eastern Java). Basemap uses ETOPO5 and SRTM3 topographic data and shaded relief mapping code (Beauducel 2020a). ETOPO5 is a 5 arc-minute resolution relief model for the Earth's surface that integrates land topography and ocean bathymetry dataset. SRTM3, the Shuttle Radar Topography Mission, is a 3 arc-second resolution digital topographic database of land elevation limited to latitudes from 60° south to 60° north

3.2 Who Are Javanese People?

The Javanese people are a native ethnic group to the island of Java. They form the largest ethnic group in Indonesia, with more than 95 million people living in Indonesia and approximately 5 million people living abroad. Although they predominantly reside in the central and eastern parts of the island, they are also scattered in various parts of the country (Taylor 2003; Ananta et al. 2015). See Fig. 2. Javanese people have their own language which is distinct from Indonesian. The Javanese language is a member of the Austronesian family of languages written in Javanese *hanacaraka* or *dentawyanjana*. Thanks to its long history and legacy of Hinduism and Buddhism in Java, the language adopted a large number of Sanskrit words (Marr and Milner 1986; Errington 1998).

3.3 A Background of the Javanese Calendar

Javanese people use the Javanese calendar simultaneously with two other perpetual calendars, the Gregorian and Islamic calendars. The former is the official calendar of the Republic of Indonesia, and the latter is used mainly for religious purposes. Prior to the adoption of the Javanese calendar in 1633 CE, Javanese people used a calendrical system based on the lunisolar Hindu *Saka* calendar (Ricklefs 1993).

The Javanese calendar was inaugurated by Sultan Agung Adi Prabu Hanyakrakusuma (1593–1645 CE), or simply Sultan Agung, the third Sultan of Mataram who ruled Central Java from 1613 CE to 1645 CE. Although the counting of the year follows the *Saka* calendar, the Javanese calendar employs a similar lunar year as the Islamic *Hijri* calendar instead of the solar year system like the former (Gislén and Eade 2019a). The Javanese calendar is sometimes referred to as AJ (*Anno Javanico*), the Latin phrase for Javanese Year. Since 2008, the difference between the Gregorian and Javanese calendars is about 67 years, where the current year 2020 CE corresponds to 1953 AJ (Oey 2001; Raffles 1817).

3.4 Some Characteristics of the Javanese Calendar

Different from many other calendars that employ a 7-day week cycle, the Javanese calendar adopts a 5-day week cycle, known as *pancawara*. Amalgamating with the 7-day week cycle of the Gregorian and Islamic calendars, namely the *saptawara* cycle, one obtains the 35-day cycle, known as *wetonan* (Darling 2004). This foundation cycle interferes with additional cycles:

- A 210-day cycle of 30 weeks, named as the *pawukon*.
- A more complex combination of the lunar month *wulan*, which has 29 or 30 days, the lunar year *taun*, which is 12 lunar months, *windu*, which is 8 lunar years, and finally *kurup* of 15 *windu* or 120 lunar years minus 1 day, which matches exactly the Islamic calendar cycle (Proudfoot 2007).

Table 1 Main cycles of the Javanese calendar

Cycle name	Length	Unit	Comment
Pancawara	5	Day	Javanese week 5 *pasaran* names (see Table 2)
Wuku	7	Day	Gregorian/Islamic week 7 *dinapitu* names (see Table 3) 30 *wuku* names
Wetonan	35	Day	35-day names as *Dinapitu* and *Pasaran*
Wulan	29 or 30	Day	Day number in a *wulan* is *dina* 12 *wulan* names (see Table 5)
Pawukon	30	*Wuku*	30 weeks $\equiv$ 210 days
Taun	12	*Wulan*	354 or 355 days *Taun* number starts on 1555 AJ 8 *taun* names (see Table 6)
Windu	8 4	*Taun* *Windu*	96 *wulan* $\equiv$ 81 *Wetonan* $\equiv$ 2,835 days 32 *taun*, 4 *windu* names
Lambang	2	*Windu*	16 *taun*, 2 *lambang* names
Kurup	15	*Windu*	120 *Taun* -1 day $\equiv$ 42,524 days 4 *kurup* names until today (see Table 8)

There are also additional cycles, but they are no longer used in the Javanese tradition. To be exhaustive, we list them here (Richmond 1956; Zerubavel 1989):

- A 6-day cycle called the *Paringkelan*: "Tungle," "Aryang," "Wurukung," "Paningron," "Uwas," and "Mawulu."
- An 8-day cycle called the *Padewan*: "Sri," "Indra," "Guru," "Yama," "Rudra," "Brama," "Kala," and "Uma."
- A 9-day cycle called the *Padangon*: "Dangu," "Jagur," "Gigis," "Kerangan," "Nohan," "Wogan," "Tulus," "Wurung," and "Dadi."

The list of main cycles and characteristics is summarized in Table 1 and is detailed in the following subsubsections. Figure 3 shows an original table of the main cycles' names written in the Javanese script *hanacaraka*.

3.4.1 *Pancawara* and *Pasaran*

The Javanese 5-day cycle is named *pancawara* and made of 5 days known as *pasaran*: "Lêgi," "Pahing" (or "Paing"), "Pon," "Wage," and "Kliwon." The word comes from "pasar" which means market. Historically, the market was held and operated on a 5-day cycle based on a *pasaran* day, e.g., "Pasar Kliwon" or "Pasar Lêgi." Until today, most of the markets in Java still have a *pasaran* name like the Kliwon Market in Kudus, Central Java, although they usually operate every day (Oey 2001).

Each *pasaran* is associated with some symbols, in particular, the five classical elements of Aristotle, colors, cardinal directions (note that the Javanese culture has

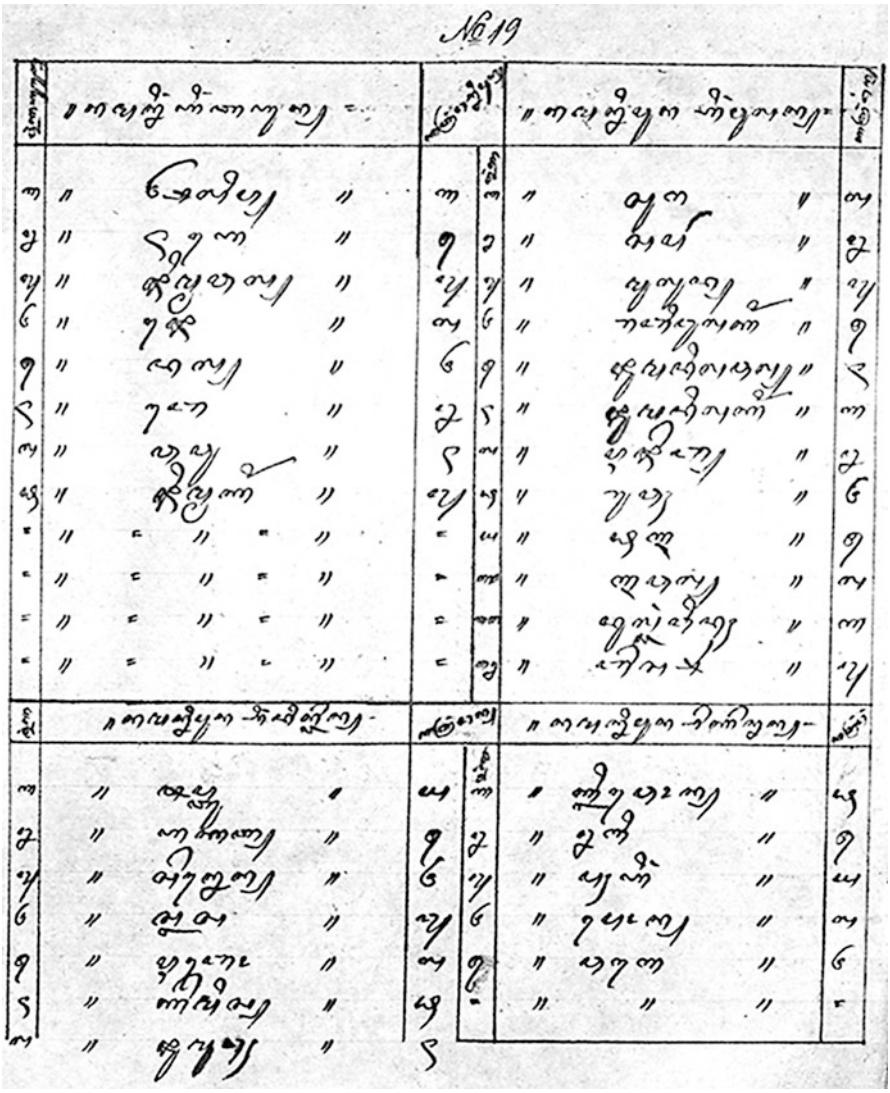


Fig. 3 Table list of 8 *taun*, 12 *wulan*, 7 *dinapitu*, and 5 *pasaran* names in Javanese Sanskrit, after Warsapradongga (1892)

five of them including a center), and human posture (Pigeaud 1977) (cf. Brinton 1893). See Table 2.

The computation of the *pasaran* day from the Gregorian or Islamic calendar date can be performed using a similar strategy as Zeller's congruence algorithm. Indeed, the formula is based on the observation that the day of the week progresses predictably based upon each subpart of the date, i.e., day, month, and year. In this

Table 2 Names of the five *pasaran* days in the Javanese week and commonly associated symbols

Ngoko	Krama	Meaning	Element	Color	Direction	Posture
Pon	Petak	–	Water	Yellow	West	Sleep
Wage	Cemeng	Dark	Earth	Black	North	Sit down
Kliwon	Asih	Affection	Spirit	Mixed color	Focus/center	Stand up
Lêgi	Manis	Sweet	Air	White	East	Turn back
Pahing	Pahit	Bitter	Fire	Red	South	To face

case, we will consider the 5-day Javanese week *pasaran*. Each term within the formula is used to calculate the offset needed to obtain the correct day.

Let P denote the *pasaran* day, in the order of Pon $=0$, Wage $=1$, Kliwon $=2$, Lêgi $=3$, and Pahing $=4$. We now propose the following *pasaran* day congruence formula:

$$P \equiv \left(k + \left\lfloor \frac{3(m+1)}{5} \right\rfloor + \left\lfloor \frac{Y}{4} \right\rfloor + 4C - 4\left\lfloor \frac{C}{4} \right\rfloor \right) \pmod 5. \tag{2}$$

where k, m, C, and Y denote the same variables as the ones defined for Zeller's congruence formula (see expression (1) in the previous section). Each term in (2) can be analyzed as follows:

- The variable k represents the progression of the day of the week based on the day of the month since each successive day results in an additional offset of one.
- The term $+\left\lfloor \frac{3(m+1)}{5} \right\rfloor$ adjusts for the variation in the days of the month. Indeed, starting from March to February, the days in a month are {31, 30, 31, 30, 31, 31, 30, 31, 30, 31, 31, 28/29}. The last element, February's 28/29 days, is not a problem since the formula had rolled it to the end. The number of days for the 11 first elements of this sequence modulo 5 (still starting with March) would be {1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1} which basically alternates the subsequence {1, 0, 1, 0, 1} every 5 months and gives the number of days that should be added to the next month. The fraction $\frac{3}{5} \equiv 0.6$ applied to $m+1$ and the floor function have that effect and will add the proper amount of days.
- Since there are 365 days in a nonleap year, and $365 \pmod 5 \equiv 0 \pmod 5$, there is no need to add an offset for the normal year.
- For the leap years, $366 \pmod 5 \equiv 1 \pmod 5$, so 1 day must be added to the offset value by the term $+\lfloor \frac{Y}{4} \rfloor$.
- There are 36,524 days in a normal century and 36,525 days in each century divisible by 400; the term $4C$ adds $36{,}524 \pmod 5 \equiv 4 \pmod 5$, 4 days for any century, and $-4\lfloor \frac{C}{4} \rfloor$ removes these 4 days for a century divisible by 400.
- The overall function, (mod 5), normalizes the result to reside in the range from 0 to 4, which yields the index of the correct day for the date being analyzed.

For example, to find the *pasaran* of the very first day in the Javanese calendar (1 Sura 1555 AJ Alip) which corresponds to 8 July 1633 CE, we have $k = 8$, $m = 7$, $C = 16$, and $Y = 33$. Using (2), we obtain $P \equiv 8 + 4 + 8 + 64 - 16 \equiv 68 \pmod 5 \equiv 3 \pmod 5$. Hence, 8 July 1633 CE was a "Lêgi." From the previously considered example, the *pasaran* days when Kartini was born P_b and passed away P_d can be calculated as follows:

$$P_b \equiv (21 + 3 + 19 + 72 - 16) \pmod 5 \equiv 99 \pmod 5 \equiv 4 \pmod 5$$

$$P_d \equiv (17 + 6 + 1 + 76 - 16) \pmod 5 \equiv 84 \pmod 5 \equiv 4 \pmod 5.$$

Hence, both 21 April 1879 and 17 September 1904 fall on a "Pahing."

3.4.2 *Dinapitu*, *Wuku*, and *Pawukon*

Dinapitu literally means "day 7" in Javanese and corresponds to the day names in the Gregorian/Islamic calendar week with exact equivalence, i.e., from Monday to Sunday: "Sênèn," "Selasa" (or "Slasa"), "Rêbo," "Kêmis," "Jemuwah" (or "Jumungah"), "Sêtu," and "Ngahad" (or "Ahad"). In recent literature, we often find the name of the days in the Indonesian language (from Monday to Sunday: "Senin," "Selasa," "Rabu," "Kamis," "Jumat," "Sabtu," and "Minggu"). They are also associated with particular symbolic meanings. See Table 3. The names of the days of the week are similar in both languages as they are absorbed from Arabic except for Sunday in Indonesian which was assimilated from the Portuguese "Domingo."

The 7-day cycle is called *saptawara* or *padinan,* and a week is named a *wuku*. The period of 30 *wuku* makes the *pawukon* cycle, i.e., 210 days (Proudfoot 2007). There are 30 different names of *wuku* (Soebardi 1965; Headley 2004) (see Table 4). Figure 4 shows an example of the Javanese calendar with *padinan*, *pasaran*, *wuku*, and *paringkelan* information, among others.

3.4.3 *Wetonan*

The *wetonan* cycle combines the 5-day *pancawara* cycle with the 7-day *wuku* week cycle. Each *wetonan* cycle lasts for $7 \times 5 = 35$ days, with 35 distinct combinations of the couple "*dinapitu pasaran*" which is called the *weton*. Figure 5 displays the 35-day *wetonan* cycle of the dual *dinapitu pasaran*. The 7-day *wuku* cycle is arranged clockwise from Monday to Sunday and the 5-day *pasaran* day progresses

Table 3 List of names of the 7 days in a *wuku* Gregorian week (cf. Rizzo 2020)

Dinapitu	Padinan	Week day	Symbol
Ngahad	Dite	Sunday	Silent
Sênèn	Soma	Monday	Forward
Selasa	Anggara	Tuesday	Backward
Rêbo	Buda	Wednesday	Turn left
Kêmis	Respati	Thursday	Turn right
Jemuwah	Sukra	Friday	Up
Sêtu	Tumpak	Saturday	Down

Table 4 List of names of the 30 different *wuku* in a *pawukon*

1	Sinta	7	Warigalit	13	Langkir	19	Tambir	25	Bala
2	Landep	8	Warigagung	14	Mandasiya	20	Medangkungan	26	Wugu
3	Wukir	9	Julungwangi	15	Julungpujut	21	Maktal	27	Wayang
4	Kurantil	10	Sungsang	16	Pahang	22	Wuye	28	Kulawu
5	Tolu	11	Galungan	17	Kuruwelut	23	Manahil	29	Dukut
6	Gumbreg	12	Kuningan	18	Marakeh	24	Prangbakat	30	Watugunung

from the center of the disk outwardly from Lêgi to Kliwon. It shows a spiral pattern that repeats seven times and sweeping five sectors each, indicated by solid black, solid blue, dashed-black, dashed-blue, dashed-dotted black, dashed-dotted blue, and dotted black spirals, respectively.

Although the *weton* can be calculated independently using either Eq. (1) or (2), it is also possible to propose a single congruence formula:

$$\begin{aligned} w &= k + \left\lfloor \frac{153(m+1)}{5} \right\rfloor + 15Y + \left\lfloor \frac{Y}{4} \right\rfloor + 19C + \left\lfloor \frac{C}{4} \right\rfloor + 5 \\ W &\equiv w \pmod 7 \\ P &\equiv w \pmod 5 \end{aligned} \tag{3}$$

where k, m, C, and Y are the same variables as defined for previous congruence formulas, w is a 35-day offset congruence, and W or P corresponds to the index of *dinapitu* or *pasaran* day, respectively. Each term of the congruence can be analyzed as follows:

- The variable k represents the progression of the day of the week based on the day of the month, since each successive day results in an additional offset of one.
- The operation $+\left\lfloor \frac{153(m+1)}{5} \right\rfloor$ adds the proper amount of days for each month, considering January and February as the 13th and 14th months of the previous year, respectively.
- The operation $+15Y$ adds an offset of $365 \pmod{35} \equiv 15 \pmod{35}$ days for the common, nonleap years.
- The operation $+\left\lfloor \frac{Y}{4} \right\rfloor$ adds 1 more day for the leap years since $366 \pmod{35} \equiv 16 \pmod{35}$.
- The operation $+19C$ adds $36{,}524 \pmod{35} \equiv 19 \pmod{35}$, 19 days for any regular century (a century with nonleap year).
- The operation $+\left\lfloor \frac{C}{4} \right\rfloor$ adds 1 more day for a century leap year that is divisible by 400 since $36{,}525 \pmod{35} \equiv 20 \pmod{35}$.
- The operation +5 adjusts the final offset to fit the indexes defined for *dinapitu* W and *pasaran* P after (mod 7) and (mod 5), respectively.

Let us take the same example of 8 July 1633 CE with $k = 8$, $m = 7$, $C = 16$, and $Y = 33$. Using (3), we have $w \equiv 8 + 244 + 8 + 304 + 4 + 5 \equiv 573$, $W \equiv 573 \pmod 7 \equiv 6 \pmod 7$, and $P \equiv 573 \pmod 5 \equiv 3 \pmod 5$. Hence, the 8 July 1633 CE was a

KARATON DALEM
NGAYOGYAKARTA HADININGRAT

1954 DESEMBER 2020

BAKDAMULUD
Windu Sangara
Warna Yaksa neng Gapura Nata

JIMAKIR

JUMADILAWAL
Lambang Langkir
Tan Rinengga Luhuring Paksa

NGAHAD (DITE)		6 Legi 20	13 Pon 27	20 Kliwon 5	27 Pahing 12
SENEN (SOMA)		7 Pahing 21	14 Wage 28	21 Legi 6	28 Pon 13
SELASA (ANGGARA)	1 Legi 15	8 Khoul SHB IV Pon 22	15 Tingalan Dalem Tahunan (SKS HB X) Khoul SHB VII Kliwon 29	22 Pahing 7	29 Wage 14
REBO (BUDA)	2 Pahing 16	9 Wage 23	16 Legi 1	23 Pon 8	30 Kliwon 15
KEMIS (RESPATI)	3 Pon 17	10 Kliwon 24	17 Pahing 2	24 Wage 9	31 Legi 16
JEMUWAH (SUKRA)	4 Wage 18	11 Legi 25	18 Pon 3	Natal 25 Kliwon 10	
SETU (TUMPAK)	5 Kliwon 19	12 Pahing 26	19 Wage 4	26 Legi 11	
Wuku	Wuye	Manahil	Prangbakat	Bala	Wugu
Paringkelan	Paningron	Uwas	Mawulu	Tungle	Aryang

Pranata Mangsa : Kanem, 43 dinten, 9 Nopember - 21 Desember 2020.

Candranipun : Rasa mulya kasucian.
Wit woh - wohan sami mawoh, Kathah bun upas mejahi taneman.
Peksi kuntul sami neba, Nyebar winih yen lintang luku sampun katingal hing wetan.

Fig. 4 An example of the Javanese calendar for December 2020 CE issued by the Kraton palace in Yogyakarta. It contains information on *Jimakir taun*, *Bakdamulud* and *Jumadilawal wulans*, *pasaran* day, *padinan* weekday, *wuku*, and *paringkelan,* among others. Each column in the wuku row lists *wuye*, *manahil*, *prangbakat*, *bala*, and *wugu*, respectively. Each column in the *paringkelan* row lists *paningron*, *uwas*, *mawulu*, *tungle*, and *aryang*, respectively. The information on *pranatamangsa* reads as "*kanem* (the sixth), 43 *dinten* (days), 9 November–21 December 2020." The text that follows *candranipun* (interpretation) says "*rasa mulya kasucian; wit woh–wohan sami mawoh, kathah bun upas mejahi taneman; peksi kuntul sami neba, nyebar winih yen lintang luku sampun katingal hing wetan*." It can be translated as follows "a noble sense of holiness; fruit trees–the same fruit grows, a lot of poisonous dew drops kills plants; heron birds are identical, spread the seed if the plowing star has been observed in the east."

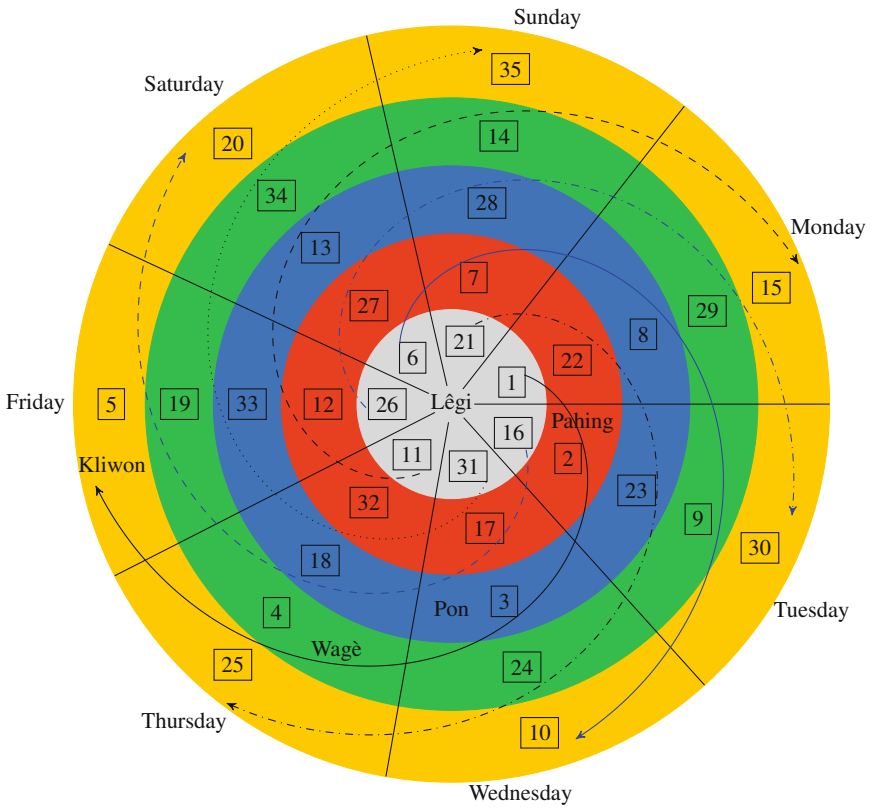


Fig. 5 A 35-day cycle of *wetonan* in the Javanese calendar. The regular Gregorian 7-day cycle is arranged clockwise from Monday to Sunday. The 5-day *pasaran* cycle is arranged from inside to outside. The white disk, red, blue, green, and yellow rings correspond to "Lêgi," "Pahing," "Pon," "Wagè," and "Kliwon," respectively

"Jemuwah Lêgi." Notice that we are able to replace both the congruence formulas (1) and (2) using our original single congruence relationship (3). Obtaining both *dinapitu* and *pasaran* days using a tabular method has been attempted by Arciniega (2020), and a *Java* application based on the Indian calendar for calculating the Javanese calendar has been developed by Gislén and Eade (2019e).

The *wetonan* cycle is especially important for divinatory systems, celebrations, and rites of passage as birth or death. Commemorations and events are held on days considered to be auspicious. In particular, the *weton* of birth is considered playing an important role in any individual personality, in a similar way as a zodiac sign does in Western astrology. The two *weton* days of future spouses are supposed to determine their background nature compatibility and are used to compute the best date of marriage using a strict arithmetic formula (Utami et al. 2019). It also figures in the timing of many ceremonies of *slametan* ritual meal and many other traditional divinatory systems (Utami et al. 2020). The eve of "Jemuwah Kliwon" is considered

particularly popular and auspicious for magical and spiritual matters (Darling 2004; Arciniega 2020).

The anniversary of Javanese birthday occurs every 35 days, so about a thousand times in a century (exactly 1043 times). For a newborn baby, the first occurrence of *weton*, i.e., aged 35 days, is named "selapanan" where the parents will cut hairs and nails of their child for the first time. For adults, the *weton* of birth is considered as an eminent day, not festive but expressing humility and blissfulness. For example, a person may fast (sometimes also the day before and the day after), stop his commercial activity, avoid taking any major decision, or simply be more generous to surrounding people through philanthropic actions such as share a blessed meal or some "jajan pasar" (early morning fresh sweet snacks from the market).

The *weton* for the birth and death of Sultan Agung is "Jemuwah Lêgi," which is also the first day of the Javanese calendar he created. This *weton* is therefore one of the recurrent noble days (see sub-subsection *Dina Mulya*) and considered as an important night for pilgrimage. Indeed, the *weton* of every King's birth is a special day; the present Sultan Hamengkubuwana X was born on 2 April 1946 CE, a "Selasa Wage." This *weton* has been chosen for his ascension to the throne, on 7 March 1989 CE. Every "Selasa Wage," the animated touristic center of Yogyakarta, Jalan Malioboro, is closed to motor vehicles, and the covered sidewalks "kaki lima," where commercial activity usually abound, are entirely cleaned.

As another especially prominent example, the present palace of Kraton Yogyakarta has been inhabited by Sultan Hamengkubuwana I and his regal suite on the "13 Sura 1682 AJ" (7 October 1756 CE), a "Kemis Pahing." In 2015, the Governor decided that every "Kemis Pahing," schoolchildren, public servants, and in particular those working in territorial services of Yogyakarta, must wear the traditional costume all day long, as a reminder of their regional culture. Merchants from traditional markets are encouraged to do the same.

3.4.4 *Wulan*

The lunar month is named a *wulan* and lasts for 29 or 30 days. There are 12 different names: "Sura," "Sapar," "Mulud," "Bakdamulud," "Jumadilawal," "Jumadilakir," "Rejeb," "Ruwah," "Puasa," "Sawal," "Dulkangidah" (or "Sela"), and "Besar." The length of each *wulan* is attributed as follows (see Table 5 and the next sub-subsections):

- "Sura," "Rejeb," "Puasa," and "Dulkangidah" are always 30 days.
- "Bakdamulud," "Jumadilakir," "Ruwah," and "Sawal" are always 29 days.
- "Sapar," "Mulud," "Jumadilawal," and "Besar" have lengths depending on the *taun* and *kurup*.

3.4.5 *Taun*

A *taun* is a cycle on 12 *wulan* and corresponds to the Javanese lunar year. There are eight different *taun* names: "Alip," "Ehé," "Jimawal," "Jé," "Dal," "Bé," "Wawu," and "Jimakir," formed by different *wulan* length sequences (see Table 5). For seven of the *taun*, the sequence alternates monotonically between 30- and 29-day lengths.

For the final *wulan* "Besar," it can be either 29 or 30 days, depending not only on the *taun* but also on the *kurup*, the 120 lunar year cycle, which has different day length sequences for the fifth *taun* "Dal." As a result, the total day length of a *taun* varies from 354 (short or normal year, named "Taun Wastu") to 355 (long or leap year, named "Taun Wuntu"), as described in Table 6:

- "Alip," "Jimawal," and "Bé" are always normal years, with a *wulan* "Besar" of 29 days.
- "Jimawal" and "Wawu" are always leap years, with a *wulan* "Besar" of 30 days.
- "Jé" and "Dal" are normal or leap depending on the *kurup.*
- "Jimakir" is a leap year for the 14 first *windu* but becomes a normal year for the final *windu* of a *kurup* cycle.

Each *taun* is assigned by a monotonic increasing number, based on the Indian calendar *Saka*. The reason was Sultan Agung decided to continue the counting from the Shalivahana era, which was 1555 at the time when inaugurating the Javanese calendar (cf. Nuraeni and Azizah 2017). Thus, the Javanese calendar began on "1 Sura Alip 1555 AJ," which corresponds to 8 July 1633 CE.

3.4.6 *Windu* and *Lambang*

Eight *taun* make a *windu* (Proudfoot 2006). Despite the variability of each *taun* length, the total length of a normal *windu* is constant since it always contains both five short and three long *taun*, which is a total of 2,835 days (about 7 years 9 months in the Gregorian/Islamic calendar). This corresponds to exactly 81 *wetonan*. This means that each "New Windu" day, dated as "1 Sura Alip," falls on the same *weton*.

Table 5 List of names of the 12 *wulan* Javanese lunar months and the associated number of days, depending on the *taun* and *kurup* (see Table 6)

		Wulan length (days)				
		Taun 1–4, 6–8	Taun 5			
No	Wulan Name	All Kurup	Kurup 1	Kurup 2	Kurup 3	Kurup 4
1	Sura	30	30	30	30	30
2	Sapar	29	29	30	30	29
3	Mulud	30	30	29	29	30
4	Bakdamulud	29	29	29	29	29
5	Jumadilawal	30	30	30	29	30
6	Jumadilakir	29	29	29	29	29
7	Rejeb	30	30	30	30	30
8	Ruwah	29	29	29	29	29
9	Pasa	30	30	30	30	30
10	Sawal	29	29	29	29	29
11	Dulkangidah	30	30	30	30	30
12	Besar	29/30	30	30	30	29
	Total (days)	254/355	355	355	354	354

Table 6 List of names of the 8 *taun* Javanese lunar years forming a *windu*, and the associated number of days, an alternate between 354 (short or normal year) and 355 (long or leap year, gray-shaded cells) days, depending on the *kurup*

No	Name	Krama	Meaning	Taun Length (days)			
				Kurup 1	Kurup 2	Kurup 3	Kurup 4
1	Alip	Purwana	intention	354	354	354	354
2	Ehé	Karyana	action	355	355	355	355
3	Jimawal	Anama	work	354	354	354	354
4	Jé	Lalana	destiny	354	354	355	355
5	Dal	Ngawanga	life	355	355	354	354
6	Bé	Pawaka	back and forth	354	354	354	354
7	Wawu	Wasana	orientation	354	354	354	354
8	Jimakir	Swasana	empty	355	355	355	355

Table 7 List of names of the four *windu* and two *lambang*

Windu name	Lambang name
Adi	Langkir
Kuntara	Kulawu
Sêngara	Langkir
Sancaya	Kulawu

There is an exception to that rule: The final *windu* of a *kurup* cycle (see the next sub-subsection) is always shortened by 1 day, with a 29 days *wulan* "Besar" during the final *taun* "Jimakir." This induces a shift in the *wetonan* cycle at each *kurup*.

Furthermore, there are four different names of *windu*: "Adi," "Kuntara," "Sengara," and "Sancaya" that compose a 32-*taun* cycle. Another cycle of 16 *taun* is combined using two different names of *windu*, called *lambang*: "Kulawu" and "Langkir." These two cycles are summarized in Table 7.

3.4.7 *Kurup*

The longest cycle in the Javanese calendar is called a *kurup*, formed by 15 *windu*, which is equivalent to 120 *taun* or 1440 *wulan* (Gislén and Eade 2019c). But the very last *wulan* of the cycle, i.e., the twelfth *wulan* "Besar" of the eighth *taun* "Jimakir" of the fifteenth *windu*, has only 29 days, so that the total length of a *kurup* is $2{,}835 \times 15 - 1 = 42{,}524$ days (about 116 years and 6 months in the Gregorian/Islamic calendar) (Rosalina 2013). The period has the same number of days as in 120 lunar years of the Tabular Islamic calendar, a rule-based variation of the Islamic *Hijri* calendar. Although the number of years and months are identical, the months are determined by arithmetical rules instead of observation or astronomical calculations.

Moreover, each *kurup* determines the following, as given in Tables 5 and 6:

- Which of the *taun* "Jé" or "Dal" has a long *wulan* "Besar."
- The sequence of *wulan* lengths in the *taun* "Dal."

Hence, the full date sequences in the calendar vary between *kurup*.

As the *weton* of the first day of a *kurup* repeats at each first day of the *windu*, a *kurup* is named using the corresponding *weton* falling on "1 Sura Alip." Table 8 lists the first five *kurup*.

Meanwhile, the Sultanate of Mataram was divided under the Treaty of Giyanti between the Dutch and Prince Mangkubumi in 1755 CE. The agreement divided ostensible territorial control over Central Java between Yogyakarta and Surakarta Sultanates. The former was ruled by Prince Mangkubumi, also known as Raden Mas Sujana or Hamengkubuwono I (1717–1792 CE), and the latter was administrated by Sinuhun Paliyan Negari, who was known as Pakubuwana III (1732–1788 CE) (Ricklefs 1974; Soekmono 1981; Frederick and Worden 1993; Brown 2004).

During the second *kurup* (1749–1821 CE), some experts realized that the Javanese calendar was still 1 day behind compared to the Islamic *Hijri* calendar. Hence, the King of Surakarta, Susuhunan Pakubuwana V (1784–1823 CE), decided to end the *Kurup* "Amiswon" in the year 1748 AJ, even though it had only been running for nine *windu* and two *taun*. So, the *taun* "Ehé" 1748 AJ, which was supposed to be a leap year, was made only 354 days, and the third *kurup* "Aboge" started on the *taun* "Jimawal" 1749 AJ. But some noticed that it would be more appropriate if the incrementation of *kurup* should have been carried out 2 lunar years before, namely on the *taun* "Alip" 1747 AJ. As a consequence of this delay, the third *kurup* "Aboge" is only 118 *taun* long. However, the Sultanate of Yogyakarta did not make a similar decision and pursued the second *kurup* normally, so that the calendar in the two concurrent regions was different during 46 years (see Table 9). On *taun* "Jimakir" 1794 AJ, the Sultan of Yogyakarta, Hamengkubuwana VI (1821–1877 CE), finally agreed and decided that the third *kurup* "Aboge" will also end with *taun* "Jimakir" 1866 AJ, reconciling the two calendars.

Table 8 List of names of the five first *kurup* Javanese 120 lunar year cycles, their short names (a contraction of the *weton* at each new *windu*, i.e., on "1 Sura Alip"), the first and last *taun*, the total amount of *taun*, and the starting dates in the Gregorian calendar

No	Kurup name	Short name	First Taun (AJ)	Last Taun (AJ)	Taun	Start date (CE)
1	Jamingiyah	A'ahgi	Alip 1555	Jimakir 1674	120	8 July 1633
2	Kamsiyah	Amiswon	Alip 1675	Ehé 1748	74	11 December 1749
3	Arbangiyah	Aboge	Jimawal 1749	Jimakir 1866	118	28 September 1821
4	Salasiyah	Asapon	Alip 1867	Jimakir 1986	120	24 March 1936
5	Isneniyah	Anenhing	Alip 1987	Jimakir 2106	120	26 August 2052

Table 9 List of dates of the second and third *kurup* in the Sultanate of Yogyakarta

No	*Kurup name*	*Short name*	*First Taun (AJ)*	*Last Taun (AJ)*	*Taun*	*Start date (CE)*
2	Kamsiyah	Amiswon	Alip 1675	Jimakir 1794	120	11 December 1749
3	Arbangiyah	Aboge	Alip 1795	Jimakir 1866	72	16 May 1866

The fourth and present *kurup* "Asapon" is planned to last a normal length of 120 lunar years and will end on "29 Besar Jimakir 1986 AJ," which is 25 August 2052 CE. The following day will start the fifth *kurup* "Anenhing" ("Alip Senen Pahing") on "1 Sura 1987 AJ Alip," but the sequence of leap years has not yet been decided, so it is formally impossible to calculate the exact *dina*, *wulan*, and *taun* for such a distant future date. Nevertheless, there is no obstacle with other strictly monotonic cycles such as *weton*, *wuku*, and *windu*.

3.4.8 *Dina Mulya*

Dina mulya are the noble days in the Javanese calendar. Except for the "Siji Sura" which is the new lunar year and falls on the first day of the first *wulan* every *taun*, others are associated with a specific *weton* and specific *taun* or *wuku* (see Table 10).

3.5 Computer Implementation of the Javanese Calendar

The computation of the full Javanese calendar has been implemented using *GNU Octave*, a Matlab®-compatible scientific language with a single function "weton.m" (Beauducel 2020c). When using a computer, however, the determination of the weekday or *pasaran* day from a date in the Gregorian calendar does not require the congruence formula (3). In fact, most computer languages are able to calculate the exact number of days that last from any reference date, correctly taking into account leap years. In the case of the Javanese calendar, the linear timeline will be the number of days counted from 8 July 1633 CE, which falls on "Jemuwah Lêgi 1 Sura Alip 1555 AJ Jamingiyah," the first day of the first *kurup*. Using that facility, calculation of the 7-day week cycle can be made by a simple modulo 7, the *pasaran*

Table 10 List of the noble days *dina mulya*

Dina Mulya	Weton	Wuku	Dina	Wulan	Taun	Occurrences
Siji Sura	–	–	1	Sura	–	1 every 354/355 days
						(new lunar year)
Aboge	Rêbo Wage	–	–	–	Alip	10 during the *Taun*
						Every 7/8 years
Daltugi	Sêtu Lêgi	–	–	–	Dal	10 during the *Taun*
						Every 7/8 years
Kuningan	Sêtu Kliwon	Kuningan	–	–	–	Every 210 days
Hanggara Asih	Selasa Kliwon	Dukut	–	–	–	Every 210 days
Dina Mulya	Jemuwah Kliwon	Watugunung	–	–	–	Every 210 days
Dina Purnama	Jemuwah Lêgi	–	–	–	–	Every 35 days

cycle by a modulo 5, the *wetonan* cycle by a modulo 35, and the *pawukon* cycle by a modulo 210 functions. Moreover, since any modern digital calendar is able to set a periodic event, a specific *weton* date repeated every 35 days will smoothly give the corresponding *wetonan* cycle over the whole calendar.

On the other hand, the computation of *dina*, *wulan*, *taun*, *windu*, and *kurup* is more complicated since the exact sequences vary throughout history following human decisions, making these cycles neither exactly cyclic nor monotonic. Hence, the proposed computing strategy is to construct, for each *kurup*, a *windu* table as a 8×12 matrix of rows *taun* versus columns *wulan*, containing the day length of that specific *wulan*. This matrix is repeated 15 times to form a complete *kurup*, or less for the second and third *kurup*. Then, all the matrices are concatenated. The cumulative sum of this table elements, in the row order, gives the total number of days that last from the origin at the beginning of each lunar month, and can be compared to the linear timeline described above. Thus, a simple "table lookup" function will give the corresponding *kurup*, *windu*, *taun*, and *wulan* indexes, and the *dina* will be given by the remainder.

We have discussed earlier that the Javanese months *wulan* adopt the lunar month, and this is closely related but does not necessarily correspond exactly with both the Islamic *Hijri* and Tabular Islamic calendars. For the former, the length of the months, whether 29 or 30 days, depends on the moon's visibility or weather conditions. The latter, however, adopted a convention of 30-day months for odd-numbered months (including the 12th month in a leap year) and 29-day months for even-numbered ones. On the other hand, the number of days in a Javanese *wulan* depends on a complex combination of *taun* and *kurup*, as we observed in Table 5. As a consequence, discrepancies are inevitable between the Javanese and Islamic calendars. The two calendars correspond exactly only during the first *kurup*, i.e., the first 120 lunar year cycles. Table 11 displays the corresponding Gregorian dates of the Javanese and (Tabular) Islamic New Year, 1 Sura and 1 Muharram, respectively, for the past half century. While more than 85% of the dates match, we observe that, in general, a 1- or 2-day gap may occur between the two calendars.

The Gregorian dates in Table 11 can be obtained using the function "weton.m" by typing in

```
>> weton('all',' 1 Sura')
>> weton(datenum(1633,7,8):now,' 1 Sura')
```

The two syntaxes are strictly equivalent. The output spits out 400 items from 1555 AJ to 1954 AJ. Note that the space before "1" is important, else, 11 and 21 Sura will be spitted out as well. Additionally, it is also essential to introduce the time window from 8 July 1633 CE onward as the dates without this will be considered invalid in the Javanese calendar. Here are other examples (recall Kartini's birthdate):

```
>> weton(1:now,'28 Bakdamulud 1808')
Error using weton (line 247)
Some dates are invalid (before 08-Jul-1633)

>> weton('all','28 Bakdamulud 1808')
```

Table 11 Examples of the Javanese and (Tabular) Islamic New Year for the past 52 years and their corresponding Gregorian dates. The bracketed (6) December 2010 was the observed date

Javanese year	Gregorian date (CE)	Islamic year	Gregorian date (CE)
1900 AJ	30 March 1968	1388 AH	31 March 1968
1906 AJ	24 January 1974	1394 AH	25 January 1974
1914 AJ	29 October 1981	1402 AH	30 October 1981
1922 AJ	3 August 1989	1410 AH	4 August 1989
1927 AJ	11 June 1994	1415 AH	10 June 1994
1930 AJ	8 May 1997	1418 AH	9 May 1997
1935 AJ	16 March 2002	1423 AH	15 March 2002
1944 AJ	8 December 2010	1432 AH	(6) 7/8 December 2010
1951 AJ	23 September 2017	1439 AH	21/22 September 2017
1954 AJ	20 August 2020	1442 AH	19/20 August 2020

```
Senen Pahing Langkir 28 Bakdamulud 1808 AJ
Be Adi Langkir Arbangiyah (Aboge), 21 April 1879 CE

>> weton(1879,4,21)
Senen Pahing Langkir 28 Bakdamulud 1808 AJ
Be Adi Langkir Arbangiyah (Aboge), 21 April 1879 CE
>>
```

Observe that the last two distinct inputs produce identical outputs. In the Islamic *Hijri* calendar, 21 April 1879 CE corresponds to 28 Rabi'ath-Thani 1296 AH. More details of the Islamic calendar and its relationship with the Gregorian calendar are beyond the scope of this chapter. Interested readers may consult Proudfoot (2006); Cohn (2007); Reingold and Dershowitz (2018), and the references therein.

4 Discussion and Epilogue

4.1 Javanese View of Time

Different from other calendars where the calculation is often based mainly on the movement of celestial bodies, such as the Sun, the Moon, and the stars, the elements of the Javanese calendar are composed of several cycles, including the solar-based Gregorian and lunar-based Islamic calendars. The Javanese people have a different perspective on time, for which some of these cycles are not all related to the movement of the celestial bodies or seasonal changes, such as the 5-day cycle *pasaran*, 35-day cycle *wetonan*, and 210-day cycle *pawukon*. This could be compared with the 260-day cycle *Tzolkin* in the Maya calendar (Jenkins 1994; Rice 2007; Cohn 2007; Ascher 2002). On the other hand, *dina*, *wulan*, *taun*, *windu*, and *kurup* correspond to the familiar cycles we encounter in other ancient and modern calendars, i.e., daily (lunar), monthly, yearly, 8-yearly, and 120-yearly cycles, respectively.

Thanks to its geographical location, where the island of Java is only less than 1000 km from the Equator, the change of seasons should be interpreted differently from the one in the subtropical regions. The climate tends to be relatively even year-round. There is only a slight variation between daylight hours, sunrise, and sunset throughout the year. For example, the longest and shortest daylights during the southern hemisphere summer (December) and winter (June) solstices in Yogyakarta (7.78° S, 110.37° E) are 12 hours and 38 minutes and 11 hours and 43 minutes, respectively (Beauducel 2020b). Although one could hardly notice this 1-hour difference over the 6-month period, Javanese residents do notice the difference between the time of sunlight and sunset. The morning civil twilight, when the geometric center of the Sun is 6° below the horizon, begins at 05:30 when the Sun reaches the Tropic of Cancer in June, but this morning twilight can be observed before 05:00 when the Sun touches the Tropic of Capricorn in December.

There are neither significant events during these both solstices nor during both equinoxes among the Javanese community. For the people who live in subtropical regions, these astronomical events can be considerably important and mark some significant events, both chronologically as well as culturally. For example, the traditional East Asian lunisolar calendars divide a year into 24 solar terms *jiéqì* (節氣). Traditionally, they celebrate the *Dōngzhì* (冬至) festival during the winter (December) solstice on one of the days between 21 and 23 December, where people make and eat the soup ball *tāngyuán* (湯圓), made from glutinous rice flour and served in a hot broth or syrup.

Nonetheless, we would not interpret that the changes of season in Java are completely absent. The following subsection briefly discusses another "calendrical" system uniquely designed for the Javanese climate.

4.2 *Pranatamangsa* Revisited

Pranatamangsa means "the regulation or arrangement of the season." It is a "calendrical" system used not only for agricultural activities among rural peasants in Java with two millennia of history but also for fish catching and capturing among Javanese fishermen. Thanks to the classical and regular seasonal rhythms, the farmers and fishermen were able to conduct and organize their activities productively using this "season keeper" (Daldjoeni 1984; Hidayat 2000).

The *pranatamangsa* calendar is arranged to fit with the solar calendar of 365- or 366-day annual period, but the length of each period (*mangsa*) is unequal, which ranges from 23 to 43 days. The number of days exhibits a symmetric pattern between the first and second semesters. Coincidentally, the 1st, 2nd, and 6th *mangsa*, and hence the 12th, 11th, and 7th *mangsa*, have prime numbers 41, 23, and 43, respectively as their lengths of days. The cycle starts at the beginning of the dry season on 23 June, around the winter solstice in the southern hemisphere (Van Den Bosch 1980; Daldjoeni 1984). Table 12 lists the names of *mangsa*, each of its period in *pranatamangsa*, and the four main seasons.

Table 12 *Pranatamangsa* at a glance

No.	Starting day	Name of *mangsa*	Length (days)	Season
1	23 June	Kaso	41	*Kemarau* or *Ketiga* (Dry season)
2	3 August	Karo	23	
3	26 August	Katelu	24	
4	19 September	Kapat	25	*Labuh* (Transition; beginning of rainy season)
5	14 October	Kalima	27	
6	11 November	Kanem	43	
7	23 December	Kapitu	43	*Penghujan* or *Rendheng* (Wet season)
8	4/5 February	Kawolu	26/27	
9	2 March	Kasanga	25	
10	27 March	Kasadasa	24	*Maréng* (Transition; ending of rainy season)
11	20 April	Désta	23	
12	13 May	Saddha	41	

When the length of the 8-month *Kawolu* is 26 days, the year is a common year or *wuntu*. When a year becomes a leap year or *wastu*, *Kawolu* will have 27 days. According to the most popular climate classification system, the Köppen-Geiger scheme, the island of Java features both tropical rainforest and monsoon climates, i.e., Af and Am, respectively (Beck et al. 2018, 2020). Thus, dry and wet seasons occur intermittently throughout the year. However, as we observed in Table 12, there are two seasons of *Labuh* and *Maréng* that serve as a transition from one distinct season to another (cf. Kristoko et al. 2012; Zaki et al. 2020).

There are several of future implications of the *pranatamangsa* system. First, due to climate change and global warming, seasonal rhythms tend to be irregular than what they used to be in the past. Second, with the advent of technology, the historically twice harvest-time annually may be extended to more than two times. Third, with the regional and global industrialization, many farmers abandoned or sold out their lands, worked in factories, and relocated to urban areas.

The topic of *pranatamangsa*, these aforementioned implications, and other possible ramifications are separate topics in themselves. An in-depth discussion is beyond the scope of this chapter and should be covered elsewhere. We wrap up this section and chapter with the following conclusion.

4.3 Conclusion

In this chapter, we have discussed the cultural, historical, and arithmetic aspects of the Javanese calendar. Along with the internationally acknowledged Gregorian and the majority-embraced religiously Islamic *Hijri* calendars, the Javanese calendar has its unique place in the hearts of many Javanese people in Indonesia as well as Javanese diaspora overseas. Although many Javanese people have adopted modern

lifestyle, the Javanese calendar is still utilized in various daily affairs, including to choose the best possible time for arranging a wedding day.

While determining the day of the week for any given date can be computed using Zeller's congruence algorithm of modulo 7, the *pasaran* day of the Javanese calendar can be calculated using a new congruence formula of modulo 5. Additionally, we have also proposed a unique and combined congruence formula for calculating both the day of the week and *pasaran* day for any given date in the Gregorian calendar. Furthermore, using a computer program *GNU Octave* "weton.m" (Beauducel 2020c), all the cycles of the Javanese calendar, i.e., *wetonan*, *wuku*, *wulan*, *windu*, *lambang*, and *kurup* can also be determined straightforwardly.

Dedication NK would like to dedicate this chapter to his late father Zakaria Karjanto (Khouw Kim Soey, 許金瑞) who introduced and taught him the alphabet, numbers, and the calendar in his early childhood. Karjanto senior was born in Tasikmalaya, West Java, Japanese-occupied Dutch East Indies on 1 January 1944 (Saturday Pahing) and died in Bandung, West Java, Indonesia on 18 April 2021 (Sunday Wage).

Acknowledgment The authors would like to thank Matthew Arciniega (Vortx, Inc., Ashland, Oregon, United States of America) for sharing the contents of his old website and Roberto Rizzo (University of Milano–Bicocca, Italy) for pointing to the article written by Proudfoot (2007). FB warmly thanks Alix Aimée Triyanti for her cultural influence and inspiration, and all his Javanese friends for their enthusiastic support.

References

Ammarell G (1988) Sky calendars of the Indo-Malay archipelago: Regional diversity/local knowledge. Indonesia 45:85–104

Ammarell G, Tsing AL (2015) Cultural production of skylore in Indonesia. In: Ruggles CLN (ed) Handbook of Archaeoastronomy and Ethnoastronomy. Springer, New York, pp 2207–2214

Ananta A, Arifin EN, Hasbullah MS, Handayani NB, Pramono A (2015) Demography of Indonesia's Ethnicity. Institute of Southeast Asian Studies, Pasir Panjang, Singapore

Arciniega M (2020) Personal communication. Matthew Arciniega has written various articles on the Javanese calendar on his defunct website www.xentana.com from the 1990s until around 2004. Although the website is no longer accessible, the readers may still find his articles in the Internet Archive *Wayback Machine* web.archive.org

Ascher M (2002) Mathematics Elsewhere: An Exploration of Ideas Across Cultures. Princeton University Press, Princeton

Aslaksen H (2001) Fake leap months in the Chinese calendar: From the Jesuits to 2033. In: Chan AKL, Clancey GK, Loy HC (eds) Historical Perspectives on East Asian Science, Technology, and Medicine. World Scientific, Singapore, pp 387–393

Aslaksen H (2002) When is Chinese New Year? Griffith Observer 66(2):1–11, National University of Singapore preprint, last updated 13 March 2009

Aslaksen H (2006) The mathematics of the Chinese calendar. National University of Singapore pp 1–52, (Preprint), last updated 17 July 2010

Beauducel F (2020a) Dem: Shaded relief image plot (digital elevation model). https://github.com/IPGP/mapping-matlab. Retrieved 31 Dec 2020

Beauducel F (2020b) Sunrise: sunrise and sunset times. MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/64692-sunrise-sunrise-and-sunset-times. Retrieved 31 Dec 2020
Beauducel F (2020c) Weton: Javanese calendar. https://github.com/beaudu/weton. Retrieved 31 Dec 2020
Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data 5:180214
Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2020) Publisher correction: Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data 7(1):1–2
Brinton DG (1893) The native calendar of Central America and Mexico. A study in linguistics and symbolism. Proc Am Philos Soc 31(142):258–314
Brown C (2004) A Short History of Indonesia: The Unlikely Nation? Allen & Unwin, Crows Nest
Brown LE (1993) The New Shorter Oxford English Dictionary. Oxford University Press, Oxford
Burton DM (2011) Elementary Number Theory, 7th edn. McGraw-Hill Education, New York
Canth M (1885) Työmiehen vaimo (The worker's wife). CreateSpace Independent Publishing Platform, (in Finnish)
Carroll L (1887) To find the day of the week for any given date. Nature 517
Catepillán X (2016) An ethnomathematics course and a first-year seminar on the mathematics of the pre-Columbian Americas. Mathematics Education, Springer, pp 273–290
Closs MP (1996) Native American Mathematics. University of Texas Press, Austin
Cohen EL (2000) What day of the week is it? Cubo Matemática Educacional. Math J 2(1):01–11
Cohn M (2007) The mathematics of the calendar. Lulu Press, Morrisville. www.lulu.com
Conway JH (1973) Tomorrow is the day after doomsday. Eureka 36:28–31
Cribb R (2000) Historical Atlas of Indonesia. Routledge Curzon Press/University of Hawaii Press, Surrey/Honolulu
Daldjoeni N (1984) Pranatamangsa, the Javanese agricultural calendar–Its bioclimatological and sociocultural function in developing rural life. Environmentalist 4(7):15–18
Darling D (2004) Marking time. Latitudes 36:upaginated
Dershowitz N, Reingold EM (2009) Indian calendrical calculations. In: Yadav BS, Mohan M (eds) Ancient Indian Leaps into Mathematics. Springer, New York, pp 1–31
Eade CJ (1995) The Calendrical Systems of Mainland South-East Asia. Brill, Leiden
Errington JJ (1998) Shifting Languages: Interaction and Identity in Javanese Indonesia. Cambridge University Press, Cambridge
Folan WJ (1986) Calendrical and numerical systems of the Nootka. In: Closs MP (ed) Native American Mathematics. University of Texas Press, Austin, pp 93–108
Frederick WH, Worden RL (1993) Indonesia: A Country Study. United States Government Publishing Office for the Library of Congress, Washington, DC
Gardner M (1996) The Universe in a Handkerchief: Lewis Carroll's Mathematical Recreations, Games, Puzzles, and Word Plays. Copernicus, Springer, New York
Ginaya G (2018) The Balinese calendar system: From its epistemological perspective to axiological practices. Int J Linguistics Lit Cult 4(3):24–37
Gislén L (2018) On lunisolar calendars and intercalation schemes in Southeast Asia. J Astron Hist Herit 21(1):2–6
Gislén L, Eade JC (2019a) The calendars of Southeast Asia. 1: Introduction. J Astron Hist Herit 22(3):407–416
Gislén L, Eade JC (2019b) The calendars of Southeast Asia. 2: Burma, Thailand, Laos and Cambodia. J Astron Hist Herit 22(3):417–430
Gislén L, Eade JC (2019c) The calendars of Southeast Asia. 4: Malaysia and Indonesia. J Astron Hist Herit 22(3):447–457
Gislén L, Eade JC (2019d) The calendars of Southeast Asia. 5: Eclipse calculations, and the longitudes of the Sun, Moon and planets in Burmese and Thai astronomy. J Astron Hist Herit 22(3):458–478

Gislén L, Eade JC (2019e) The calendars of Southeast Asia. 6: Calendrical records. J Astron Hist Herit 22(3):479–491
Golzio KH (2012) The calendar systems of ancient India and their spread to Southeast Asia. In: Boschung D, Wessels-Mevissen C (eds) Figurations of Time in Asia. Wilhelm Fink Verlag, München, pp 204–225
Hassan A (2017) Muslim calendar. In: Çakmak C (ed) Islam–A Worldwide Encyclopedia. ABC-CLIO, Santa Barbara/Denver, pp 1129–1130
Headley SC (2004) The Javanese *wuku* weeks: Icons of good and bad time. In: Le Roux P, Sellato B, Ivanoff J (eds) Poids et measures en Asie du Sud-Est (Weights and Measures in Southeast Asia–Metrological Systems and Societies), vol 1, Institut de Recherche sur le Sud-Est Asiatique (Southeast Asian Research Institute), Centre national de la recherche scientifique (CNRS) et Université de Provence (The French National Center for Scientific Research and University of Provence Aix-Marseille I), Marseille, France, pp 211–236
Hidayat B (2000) Indo-Malay astronomy. In: Astronomy Across Cultures: The History of Non-Western Astronomy. Springer, Dordrecht, pp 371–384
Hubber P, Ramseger J (2017) Physical learning environments for science education: An ethnographic field study of primary classrooms in Australia, Germany and Taiwan. In: Hackling M, Ramseger J, Chen HL (eds) Quality Teaching in Primary Science Education. Springer, Cham, pp 51–77
Jenkins JM (1994) Tzolkin: Visionary Perspectives and Calendar Studies. Borderland Sciences Research Foundation, Eureka
Kartini RA (1911) Door duisternis tot licht: Gedachten over en voor het Javaansche volk (Through darkness to light: Thoughts about and for the Javanese people). G.C.T. van Dorp, Semarang and Soerabaja, Dutch East Indies (Indonesia) and 's-Gravenhage (The Hague), the Netherlands, (in Dutch, with a foreword by J. H. Abendanon)
Kartini RA (1920) Letters of a Javanese princess. Wentworth Press and University Press of America, Sydney, Australia and Landham, Maryland, United States of America, (translated from Dutch by A. L. Symmers, with a preface by E. Roosevelt, edited and with an Introduction by H. Geertz)
Kristoko H, Eko S, Sri Y, Bistok S (2012) Updated pranata mangsa: Recombination of local knowledge and agro meteorology using fuzzy logic for determining planting pattern. Int J Comput Sci Issues (IJCSI) 9(6):367–372
Lân LT (2019) The calendars of Southeast Asia. 3: Vietnam. J Astron Hist Herit 22(3):431–446
Marr DG, Milner AC (1986) Southeast Asia in the 9th to 14th Centuries. Institute of Southeast Asian Studies, Singapore
Martzloff JC (2000) Chinese mathematical astronomy. In: Selin H, D'Ambrosio U (eds) Mathematics Across Cultures: The History of Non-Western Mathematics. Springer, Dordrecht, pp 373–407
Martzloff JC (2016) Astronomy and Calendars – The Other Chinese Mathematics 104 BC–AD 1644. Springer, Berlin/Heildelberg
Milne J (2020) Class field theory (v4.03). www.jmilne.org/math/
Nuraeni Z, Azizah N (2017) Application of number theory in the calculation of Java calendar. In: Proceedings of Ahmad Dahlan International Conference on Mathematics and Mathematics Education, Yogyakarta, pp 177–181
Oey E (2001) Java, 3rd edn. Periplus Editions Publishing Group/Tuttle Publishing, Jakarta/North Clarendon
Ôhashi Y (2009) Mainland Southeast Asia as a crossroads of Chinese astronomy and Indian astronomy. In: Yadav B, Mohan M (eds) Ancient Indian Leaps into Mathematics. Birkhäuser, Boston, pp 193–200
Pigeaud TGT (1977) Javanese divination and classification. In: de Josselin de Jong PE (ed) Structural Anthropology in the Netherlands: A Reader, Koninklijk Instituut voor Taal-, Land- en Volkenkunde (Royal Institute for Linguistics, Geography, and Ethnology)–Translation Series, Foris, Dordrecht, pp 64–82

Proudfoot I (2006) Old Muslim Calendars of Southeast Asia. Brill, Leiden
Proudfoot I (2007) In search of lost time: Javanese and Balinese understandings of the Indic calendar. Bijdragen tot de Taal-, Land-en Volkenkunde (Journal of the Humanities and Social Sciences of Southeast Asia) 163(1):86–122
Raffles TS (1817) The History of Java. Black, Parbury, and Allen, London, United Kingdom, (two volumes)
Reingold EM, Dershowitz N (2018) Calendrical Calculations, The ultimate (Fourth) Edition. Cambridge University Press, Cambridge
Rice PM (2007) Maya Calendar Origins: Monuments, Mythistory, and the Materialization of Time. University of Texas Press, Austin
Richards EG (1998) Mapping Time: The Calendar and Its History. Oxford University Press, Oxford
Richmond B (1956) Time measurement and calendar construction. Brill Archive
Ricklefs MC (1974) Jogjakarta under Sultan Mangkubumi, 1749-1792: A history of the division of Java, London Oriental Series, vol 30. Oxford University Press, Oxford
Ricklefs MC (1993) A History of Modern Indonesia since *c*. 1300, 2nd edn. Macmillan Press, Houndmills/Basingstoke/Hampshire/London
Rizzo R (2020) What's in a name? How Indonesian Buddhism gained its second 'D'. Globe–Lines of Thought Across Southeast Asia (Published 14 October 2020. The Wednesday *Buda* can sometimes also spelled as *Budha*. Some government officials and civil servants in Indonesia seem to be ignorant by using this term when registering the religious data for the Buddhist community in the country. The correct Indonesian word for Buddhist is *Buddha*, with double 'd'. The error has been rectified in a letter issued on 31 December 2018 by the Indonesian Ministry of Religion No. B.3727/DJ.VII/Dt.VII.I.2/BA.00/12/2018)
Rosa M, Shirley L, Gavarrete ME, Alangui WV (2017) Ethnomathematics and Its Diverse Approaches for Mathematics Education. Springer, Cham
Rosalina I (2013) Aplikasi kalender Islam Jawa dalam penentuan bulan Qomariyah: Penyesuaian kalender Saka dengan kalender Hijriyah (Application of Javanese Islamic calendar in determining Qamariyyah month: Adjustment of Saka and Hijri calendars). Undergraduate thesis, Universitas Islam Negeri Maulana Malik Ibrahim, Malang, East Java, Indonesia. http://etheses.uin-malang.ac.id/87/ (in Indonesian)
Rosen KH (2011) Elementary Number Theory and Its Applications, 6th edn. Pearson, London
Rudolph L (2006) The fullness of time. Culture & Psychology 12(2):169–204
Ruggles CLN (2015) Calendars and astronomy. In: Ruggles CLN (ed) Handbook of Archaeoastronomy and Ethnoastronomy. Springer, New York, pp 15–30
Selin H, D'Ambrosio U (2000) Mathematics Across Cultures: The History Of Non-Western Mathematics, Science Across Culture: The History of Non-Western Science Series, vol 2. Springer, Dordrecht
Selin H, Sun X (2000) Astronomy Across Cultures: The History of Non-Western Astronomy, Science Across Culture: The History of Non-Western Science Series, vol 1. Springer, Dordrecht
Soebardi (1965) Calendrical traditions in Indonesia. Madjalah Ilmu-Ilmu Sastra Indonesia (Indonesian Literature Magazine) 3:49–61
Soekmono R (1981) Pengantar Sejarah Kebudayaan Indonesia (An Introduction to Cultural History of Indonesia). Kanisius, Sleman, Special Administrative Region of Yogyakarta, Indonesia, (in Indonesian)
Stakhov A (2009) The Mathematics of Harmony: From Euclid to Contemporary Mathematics and Computer Science, Series on Knots and Everything, vol 22. World Scientific, Singapore, (assisted by Olsen, S)
Sun X (2015) Chinese calendar and mathematical astronomy. In: Ruggles CLN (ed) Handbook of Archaeoastronomy and Ethnoastronomy. Springer, New York, pp 2059–2068
Taylor JG (2003) Indonesia: Peoples and Histories. Yale University Press, New Haven
Urton G, Llanos PN (2010) The Social Life of Numbers: A Quechua Ontology of Numbers and Philosophy of Arithmetic. University of Texas Press, Austin

Utami NW, Sayuti SA, Jailani J (2019) Math and mate in Javanese primbon: Ethnomathematics study. J Math Educ 10(3):341–356

Utami NW, Sayuti SA, Jailani J (2020) An ethnomathematics study of the days on the Javanese calendar for learning mathematics in elementary school. Elem Educ Online 19(3):1295–1305

Van Den Bosch F (1980) Der Javanische Mangsakalender (The Javanese *mangsa* calendar). Bijdragen tot de Taal-, Land-en Volkenkunde (Journal of the Humanities and Social Sciences of Southeast Asia) 136(2/3de Afl):248–282, (in German)

Vickers A (1990) Balinese texts and historiography. Hist Theory 29(2):158–178

Warsapradongga MD (1892) Koleksi Warsadiningrat (MDW1892a). Yayasan Sastra Lestari, Surakarta

Wollstonecraft M (1792) A Vindication of the Rights of Woman: With Strictures on Political and Moral Subjects. Cambridge University Press, Cambridge

Zaki MK, Noda K, Ito K, Komariah K, Sumani S, Senge M (2020) Adaptation to extreme hydrological events by Javanese society through local knowledge. Sustainability 12(24):10,373

Zeller JCJ (1882) Die Grundaufgaben der Kalenderrechnung auf neue und vereinfachte Weise gelöst (The basic tasks of the calendar calculation solved in a new and simplified manner). Württembergische Vierteljahrshefte für Landesgeschichte (Württemberg Quarterly Books for Regional History), W Kohlhammer, Stuttgart, Germany 5:313–314, (in German)

Zeller JCJ (1883) Problema duplex Calendarii fundamentale (The fundamental of the calendar double problem). Bull Soc Math France (Bull French Math Soc) 11:59–61, (in Latin)

Zeller JCJ (1885) Kalender-Formeln (Calendar formulas). Math-naturwissenschaftliche Mitteilungen des mathematisch-naturwissenschaftlichen Vereins Württemberg (Math Sci Rep Math Sci Assoc Württemberg) 1(1):54–58, (in German)

Zeller JCJ (1887) Kalender-Formeln (Calendar formulas). Acta Math 9(1):131–136, (in German)

Zerubavel E (1989) The Seven Day Circle: The History and Meaning of the Week. University of Chicago Press, Chicago

Historiography of Mathematics from the Mathematician's Point of View

Viktor Blåsjö

Contents

1 History of the Relation Between History and Mathematics 1128
2 The Role of Present Mathematics in Historiography 1135
3 Reconstruction as Historical Method 1141
4 Conclusion 1145
5 Cross-References 1145
References 1146

Abstract

Mathematicians used to be highly invested in the study of the history of their own field, but their voice in historiographical discussions has diminished in influence in the past century. One prominent narrative paints this as a justified fall from grace: mathematicians wedded to present mathematical values looked at the past with prejudiced eyes, whereas a new generation of historians were better able to appreciate the past proper, in its own terms. But the best internalist mathematical historiography of old needed no such external corrective. It was already committed to avoiding presentism and anachronism, for reasons that were not in opposition to mathematical values but rather derived directly from a positive vision of the role that history could play in the mathematical community. In this vision, a historical understanding of how a field developed is a proxy for first-hand research experience in that field. It follows that it is essential for historical accounts to thoroughly convey the scope and limitations of alternative conceptions and approaches, including dead-end developments, since this is precisely what sets the critical knowledge gained by first-hand research experience apart from the doctrinal knowledge gained merely from a textbook. Hence, from this point of view, presentist historiography is not the natural outlook of the mathematician, but rather a direct antithesis of the mathematician's most fundamental

V. Blåsjö (✉)
Universiteit Utrecht, Utrecht, The Netherlands
e-mail: V.N.E.Blasjo@uu.nl

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_67

reason for studying history. To study past mathematics precisely as it appeared to active researchers at the time is not foreign to the mathematician, but a direct corollary of the mathematician's core conviction that only those with first-hand research experience in a field of mathematics truly understand it. A sophisticated internalist historiography derived from these ideals was articulated to a greater extent in the past than is commonly recognized today. By going back to its roots, the mathematician's historiography could revive some of the virtues that have been neglected in recent years.

Keywords

Historiography · Internalism · Presentism · History of the history of mathematics

1 History of the Relation Between History and Mathematics

The "right" way to study history depends on what one wants historical inquiry to achieve. Historiographic practices – such as using modern mathematical concepts when analyzing past thought – are not inherently good or bad, but can only be judged relative to the intended goals of historical scholarship. To be able to say *how* we should do history, we must therefore first understand *why* we do history at all. So to understand the mathematician's relation to history, we must first ask ourselves: What purposes does history serve to the mathematical community? To answer this question, it will be instructive to look at some snapshots of how the relationship between mathematics and history has evolved over time.

In a way, mathematics can be regarded as inherently ahistorical. "Mathematics is . . . under the spell of the belief in its objective nature. . . . Motives [for undertaking particular investigations] are often more hidden than revealed" since "it is not the style of mathematics to speak about subjective things." But experience shows that the Platonic dream of achieving disembodied perfection is alas "a chimera." By contrast, Aristotle "laid out with the greatest care how thinkers before him thought about all the matters he was dealing with, and with this kind of comparative study saw how he might make a better start towards obtaining general norms." In mathematics, as much as in any other field, this suggests the value of "didactical diagnosis and therapy, on the basis of historical analyses, where these analyses serve only to turn one's attention in the right direction." History "can steer one's attention to differences and conceptual distinctions which would never come to mind by way of direct, abstract considerations." (Toeplitz 1926/2015, 308–309).

The "Platonic" perspective is thus absolutist. It aims for a state of finality and timeless perfection. History is antithetical to its purposes: it aims precisely to transcend that element of human fallibility with which past thought is contaminated. The "Aristotelian" perspective, by contrast, does not aspire to deny or erase the situatedness of knowledge. Instead, it sees past and present thought as inherently intertwined. This conflict remains at the heart of history of mathematics. Mathematicians dream of "the perfect embodiment of immutable laws," while historians are

all but defined by a "moral commitment to recording and remembering the earthly, perishable detail"; truly, "our muses … haunt us with their divergent demands" when we try to do justice to both of these perspectives (Guicciardini 2018, 148–150).

Aristotle's disciple Eudemus wrote a history of geometry – the first known treatise on the history of mathematics. The work itself is no longer extant but its overall outlook can be surmised from surviving fragments and references. Eudemus seems to have envisioned his history of geometry being used as a starting point for methodological reflection, similar to how Aristotle engaged with the views of earlier philosophers to articulate his own systems. In part, Eudemus seems to have aspired to present a kind of historical raw data to serve as the basis for such reflection, but he was evidently not shy to add interpretative analysis as well. Thus, his approach is in part comparative rather than purely chronological; he passes judgement on past mathematics, and crafts a narrative of cognitive evolution toward higher states of knowledge, culminating in the mathematics of his own time (Zhmud 2017).

The Aristotelian perspective, then, emphasizes history but treats it instrumentally. History is not expounded for its own sake, but as part of an articulation and analysis of present thought in the field. Mathematical authors of all ages have found it useful to deploy history in such a role. Lagrange and Bourbaki, for instance, incorporated extensive historical notes in treatises where they advanced systematizations of previous knowledge and reconceptualizations of mathematical fields.

This way of utilizing history to inform a systemic survey of a body of knowledge was perhaps especially natural in classical antiquity, when mathematics and philosophy advanced rapidly and were fundamentally transformed. Authors further removed from creative frontiers may take a more antiquarian approach to historical material. One may think, for instance, of how Eutocius, in the sixth century, meticulously catalogued a dozen different cube duplications by various past mathematicians (Netz 2004, 273–306) while showing few, if any, aspirations beyond simply amassing the material.

In the seventeenth century – another vastly creative and transformative era – it was again à la mode to situate research contributions in a potted historical narrative. Major works such as Descartes's *Géométrie* (1637), Wallis's *Algebra* (1685), and Newton's *Principia* (1687) all open by immediately characterizing their projects relative to a particular interpretation of what "the ancients" had done in the subject.

The seventeenth century had no time to produce any full-scale histories of mathematics. The eighteenth century took up this task with enthusiasm, producing several sprawling histories packed with information. After the creative explosions of the seventeenth century, the direction of mathematical research was now on a relatively stable and monotonous trajectory. Once again history was no longer needed to help envision major reconceptualizations of the field, but more to keep track of a growing body of accumulated work. This also fit with the fashion of the day in the sciences and intellectual work generally, which embraced massive encyclopedic surveys, cabinets of curiosity overloaded with assorted knick-knacks, and Baconian accumulation of empirical raw data (Goldstein 2008). Indeed, Montucla's *Histoire des Mathématiques* (1758) repeatedly references Francis Bacon in its preface.

As biology focused on gathering any and all specimens and sorting them into descriptive taxonomies, so history of mathematics was likewise preoccupied with assembling masses of material (Bullynck 2016, 10). "Lacroix's history of eighteenth-century analysis can be seen as the culmination of the eighteenth-century tradition of narrative mathematics. Coverage, rather than coherence, was his major goal." (Richards 2006, 707). Similarly, "Kaestner's four-volume *Geschichte der Mathematik* (1796–1800) is essentially a description of the author's large mathematical library, so that a malicious reviewer could dispatch Kaestner with the words that he seems to have written his work on the book ladder." (Stäckel 1900, 252). Historical works of this time largely assume that direct acquaintance with primary sources is a feasible next step following a standard education from current textbooks (Dauben and Scriba 2002, 88), so that the task of the historian is to survey and summarize rather than to interpret or explain.

As the quantity of mathematical publications grew rapidly, history answered "a need amongst mathematicians for information about who produced which piece of mathematical knowledge when"; in particular, "it would be one task of the historian of mathematics to attribute mathematical credit to the right person." (Schneider 1992, 620–621). "The preoccupation at this early stage in the history of the history of mathematics was an almost exclusive interest in 'priorities' – in mathematical 'firsts'." (Dauben 1999, 4).

Such "firsts"-oriented historiographies effectively assume that the development of mathematics is cumulative rather than conceptually transformative. Only then is a one-to-one matching of each tenet of present mathematics with some past work possible. Because of the conceptual stability of eighteenth-century mathematics, this view was sustainable for a while. But in the nineteenth century, it was quickly becoming obvious that mathematics was no longer progressing purely by adding knowledge, but by systematically reorganizing entire theories in ways that made much earlier work obsolete. Thus, the enterprise of encyclopedic "firsts"-oriented history lost its purpose. "In the current state of analysis we may regard these discussions [of past mathematics] as useless, for they concern forgotten methods, which have given way to others more simple and more general," Lagrange observed in 1806. Lacroix made a similar point in his *Essais sur l'enseignement* (1805): "to educate students in the ways of [historical] authors, even the most celebrated of them, must cease because since they wrote, the science has made progress which has completely changed the connections of the propositions and often their language." (Richards 2006, 709–710).

In response, the nineteenth century saw a revival of fortunes of the Platonic ideal of mathematics as historically disembodied perfect reason. Mathematical exposition should follow "purely rational principles," Lacroix observed as a contrast to his rejection of the utility of historical perspectives. Cauchy, setting the tone for decades to come,

> reconceptualized and restructured mathematics to be rigorous rather than narrative. . . . Cauchy's definitional precision effectively separated mathematics from the rest of human

> thought. . . . Cauchy clearly recognized this. The separation of mathematical argument from human reason was essential to his rigorous program. (Richards 2006, 700, 712)

This vision of mathematical rigor left little or no place for the relevance of history to mathematical understanding.

The end of the nineteenth century saw of resurgence of interest in the history of mathematics. The new generation noted how they had to break free from the anti-history sentiments widespread in the mathematical community:

> When I began to study mathematical and historical research about 30 years ago, most of the mathematicians I came into contact with thought that thorough research in this field was a waste of time, so that such work should not be considered meritorious but rather reprehensible. Now, however, such a view may be regarded as a curiosity, and most of my colleagues in the field agree that that the history of mathematics is not only relevant in itself, but also has great educational value for the study of mathematics, and that the understanding of it can often be useful even for the advancement of science. (Eneström 1900, 4)

This increased status of history can be attributed to internal mathematical developments (Stäckel 1900, 251): foundational mathematical research in the late nineteenth century undermined the simplistic notion of an absolute (and ahistorical) standard of rigor and truth, thereby reversing the rationale for anti-historicism of rigorization movements such as that of Cauchy.

History also gained new appeal as mathematics fragmented into more specialized subfields. Specialized research mathematics can often only be understood by one's closest colleagues. History, by contrast, provides a rich store of mathematical material that is accessible and interesting to all mathematicians. Therefore "one of the advantages in the study of the history of mathematics is to bring colleagues together" (Struik 1980, 24) around shared experiences and interests. Thus it became a selling point of the field that "only historical thinking can form a counterweight against the much-lamented specialisation" that was becoming more and more marked in contemporary research (Neugebauer 1927, 44).

Increased specialization also meant that that mathematicians more frequently found themselves teaching or engaging with mathematical theories with which they had little direct experience in their own research and personal mathematical practice. In such cases, there would often be an abundance of solid sources for picking up the necessary technical content, but these could often not convey the same sense of the lay of the land that an insider with hands-on experience in the field would have. History can be a substitute for this kind of first-hand experience. Retracing the thought of the pioneers of the field is a digest surrogate for personal experience. Without this perspective of what a field looks like to the active researcher, it is easy to misjudge the purpose and importance of ideas and the connections between them. "The historical investigation of the development of a science is most needful, lest the principles treasured up in it become a system of half-understood prescripts, or worse, a system of prejudices." (Mach 1883/1919, 255).

It was commonplace to argue that students of mathematics "must naturally follow the same course of development on a smaller scale, that the science itself has taken

on a larger" (Klein 1896, 248). In other words, mathematics education should follow the biological principle that key stages in the evolutionary history of a species are often recapitulated in the fetal development of each individual organism (Cajori 1896, v; Poincaré 1899, 159; Klein 1932, 268). Another sign of the unity of research and educational work envisioned at the time is that, in the nineteenth century, "in a remarkable number of cases research work in mathematics [was] stimulated by a perceived educational need: the new theory appeared more in textbooks and lecture notes than in ordinary papers" (Grattan-Guinness 1990, 155). Lectures by, for instance, Monge, Cauchy, and Weierstrass spun off entire research programs. History and teaching had a natural place in mathematics: the educator, the historian, and the research strategist shared a common goal in understanding the gist of existing work in a field, the scope and limitations of its methods, and the connections and motivations of its key ideas.

The early twentieth century was a golden age – "the most distinguished period" (Grattan-Guinness 1990, 150) – for the status of history in the mathematical community. At its peak as a world-leading research center, the Mathematical Institute at Göttingen offered a rich range of advanced courses on history of mathematics taught by leading mathematicians (Toeplitz 1926/2015, 303; Tobies 2019, 412–413). In Frankfurt, top mathematicians and students came together for in-depth seminars on a range of specialized topics in history of mathematics every semester from 1922 to 1935 (Siegel and Lenzen 1979, 224–226). Affordable editions of important primary sources in the history of mathematics aimed at the learned public sold well, as a historical perspective on mathematics was seen as befitting "every educated reader" (Dauben and Scriba 2002, 127, 130–131). History was pursued as "an honorable thing to do," and not for dogmatic ideological reasons such as an "attempt to legitimize the new mathematics by identifying it with the core element of the tradition" (Gray 2008, 372).

In this context, a new generation of mathematical historians emerged. But it was a rude awakening for the mathematical community to realize the ramifications of its neglect of history in the preceding decades. When mathematicians had turned their back on the history of mathematics in the nineteenth century, philosophers and cultural theorists had picked up the slack. The old eighteenth-century encyclopedic histories of mathematics, "understood only by professional mathematicians" (Dauben and Scriba 2002, 13), did not meet the needs of the nineteenth century. Mathematicians had abandoned history just as the transformative developments of the early nineteenth century created a need for higher-order narratives offering prominent interpretative syntheses (Bullynck 2016). Authors from humanistic fields had been happy to fill this void with works that integrated history with grand epistemological or sociocultural theories. But excessive enthusiasm for such macro perspectives had led to factual errors as well as neglect and ignorance of sources (Brentjes 2016, 50; Høyrup 2016, 65; Dauben and Scriba 2002, 19).

The new generation of mathematical historians were keen to position themselves in opposition to the excesses of the humanistic histories of the nineteenth century. An instinctive aversion to any attempts at grander "synthesis" was a natural part of this reaction (Neugebauer 1951/1969, viii; Rowe 2016, 139), as was a greatly increased

emphasis on primary mathematical sources (Dauben and Scriba 2002, 135). Propagation of errors and lack of scholarly rigor was also seen as a major threat to the status of historical scholarship (Eneström 1912/13). The fear was often expressed that mathematicians in particular might see history as an unserious field.

> If mathematicians who are not historians see that even highly praised mathematical-historical works contain a large number of unreliable data and uncritical statements, they easily get the impression that the history of mathematics does not yet exist as a real science, and that no special knowledge or training is needed to write a mathematical-historical work. If they then have any reason to want historical articles on certain mathematical subjects, it may easily happen that they turn to persons for this purpose who lack certain very important prerequisites for truly scientific research in the field of mathematics and history. (Eneström 1908/09, 9)

Moritz Cantor's multivolume *Geschichte der Mathematik* (1877–1901) became a focal point for these concerns, and stood as a symbol for the style of history that had resulted when mathematicians neglected the field (Lützen and Purkert 1994, 1–7).

> Its total lack of mathematic competence as well as its moralizing and anecdotal attitude seriously discredited the history of mathematics in the eyes of mathematicians, for whom, after all, the history of mathematics has to be written. . . . Had [Cantor] studied the texts themselves, he would have avoided countless misinterpretations and inaccuracies which have become commonplace. (Neubegauer 1956)

Attacks on Cantor's history such as this were reiterated again and again in the early twentieth century, as the new generation of mathematical historians saw this as a crucial contrast defining their distinctive historiography.

The efforts to rehabilitate history of mathematics in the eyes of mathematicians and revive its status as an integral part of mathematics itself in the early twentieth century were successful. But this was soon disrupted by war and fascism. When the continental model of university and PhD education was exported to other parts of the world, great emphasis was placed on the importance of leading students to the frontiers of original research (Turner 2020, 58). This had indeed been one element that set the PhD system and academic culture at the world-leading continental universities apart from the wider geographic range of academic institutions that now quickly rose to greater research prominence. However, broader intellectual elements of the tradition were largely discarded in this process. The outlook that sophisticated historical and philosophical awareness is essential even for a well-rounded understanding of mathematics itself had won considerable traction at leading continental institutions, but emerging academic institutions building up their rapidly expanding research infrastructure did not continue this trend.

Hence a new generation of more narrowly trained research specialist mathematicians emerged. Insofar as they addressed history, they did so with less humility, less historical and philosophical awareness, and greater overconfidence in the superiority of their own style of mathematics. In this way, they exposed themselves to critiques from historians. It became easy for historians coming from humanities disciplines to

dismiss mathematicians dabbling in history as naively writing "without consideration of basic questions of historiography" (Grattan-Guinness 1990, 157). A standard narrative emerged, alleging that the second half of the twentieth century brought "a new and more sophisticated historiography," whereas older work on the history of mathematics was commonly marked by rampant presentism and simple-minded priority questions. In older work, allegedly,

> the effect was to "explain" the subject in anachronistic concepts and terminology, to select topics for study only insofar as they had a connection to more recent developments, and to praise or damn these efforts on the basis of whether they anticipated (took a step toward) the current state of mathematical knowledge.

In the "new historiography," by contrast,

> the focus is on understanding the thought of the period, independent of whether it is right or wrong by today's account. The historiography is more philosophically sensitive in its understanding of the nature of mathematical truth and rigor, and it recognizes that these concepts have not remained invariant over time. (Kitcher and Aspray 1988, 24–25)

This narrative was very agreeable to a new generation of historians of science who "wanted independence from the sciences" and sought "to supply their own agenda for their discipline rather than accepting that of the scientist" (Lightman 2016, 2). In this way, they could characterize their own historiography as an objective improvement of standards, and the historical outlook of mathematicians as inherently naive. This basic framing remains fundamental in historiographical discussions to this day (even though, to be sure, actual professional affiliations are more fluid and complex than a neat terminological division into "mathematicians" and "historians" might suggest; Goldstein 1995, 7).

Historical circumstances facilitated over-universalization of this narrative. The single-minded focus on specialized research in the twentieth-century growth spurt of academic institutes of mathematics, often in places that had little tradition of intellectual history, produced a neglect of historical appreciation among mathematicians. What could better be regarded as a legitimate critique of this contingent circumstance has instead all too often been perceived as proof of an intrinsic fault inherent in any internalist-mathematical approach to history.

In the early twentieth century, history was a notable presence in elite mathematical culture. Leading mathematical historians were deeply anchored in this culture rather than coming from separate academic departments. This state of affairs was driven by a deeply felt conviction in the mathematical community of the rich and multifaceted ways in which history can be of integral value to the mathematical enterprise. This era set out the beginnings of what may be called a critical internalist historiography of mathematics.

History was not as central to subsequent mathematical generations. Having lost touch with the intellectual roots that informed the turn-of-the-century vision of history and its purpose within mathematics, many mathematicians reverted back to

the more simplistic views of history that had been commonplace in earlier centuries. This may be called a presentist internalist historiography of mathematics.

Critical internalism emphasized the parallels between historical understanding and creative mathematical research. The same skill of grasping the forces that have shaped a field is equally essential in historical scholarship and in the context of formulating an informed and forward-looking mathematical research program of one's own. To learn a new field is to reenact the course of its evolution in one's own mind; to digest the argument of a past work of mathematics is to feel precisely that tension between research goals and the tools available that the working mathematician of that time struggled with: this is an analog in the domain of historical scholarship of the view that the research-active mathematician of today has an authoritative claim to understanding of their field qualitatively different from any dilettante outsider account, no matter how descriptively and factually accurate. In this way, history is key to develop understanding of new fields, and to form critical judgements about it based on independent thought.

Presentist internalism, on the other hand, does not accord history any such parity of authority with present research. Critical and presentist internalisms are both committed to the ideal of active research experience as the sine qua non of mathematical understanding. But critical internalism interprets this ideal inclusively, and sees history as a way to effectively, in distilled form, obtain more of that precious form of understanding than we could have time for through conventional personal research. Presentist internalism demotes history from this perch. By conceiving the ideal of the primacy of active research understanding narrowly, to include current mathematical research only, it cuts off the critical internalist rationale for seeing history as an integral part of mathematical thought itself.

It is unfortunate that these two very different approaches are so often conflated. Much historiographical discussion in the second half of the twentieth century was framed in terms of a fundamental dichotomy between internalist-mathematical approaches on the one hand and properly historical approaches on the other. Critiques of the internalist-mathematical approach have naturally focused on the faults of presentist internalist historiography, under the mistaken impression or insinuation that this undermines internalist approaches altogether. The above contextualization suggests instead that, rather than attributing historiographic naiveté to inherent characteristics of the mathematical psyche, a better way to understand the mathematician's relation to history is in terms of the purposes history can serve to the mathematical community (Barany 2020, 1077).

2 The Role of Present Mathematics in Historiography

The mathematician's way of doing history is often portrayed as historiographically flawed in that it distorts past thought through the use of anachronistic perspectives and presentist interpretative biases. Indeed, it is often suggested that it was not until historians took over the field from mathematicians in the latter half of the twentieth

century that this rampant problem was recognized and cured (Unguru 1975; Rowe 1996, 10).

However, this diagnosis applies only to presentist internalism. Far from being inherent in any internalist approach, anachronism and presentism were strongly condemned by the critical internalists, long before they became a fault line of twentieth-century historiographic debates. Nesselmann, who studied mathematics under Jacobi before turning to history, put it clearly long ago:

> Nothing is more usual and natural when reading ancient works than our substituting the standpoint of the ancient authors with our own, … [so that] it is one of the most difficult tasks of the historian to understand and convey the conceptual background not only of each author, but also of each period of time in its individual character, and at the same time to keep the modern viewpoint in mind, without lifting the ancient author from his peculiar sphere of thinking into our present one. (Nesselmann 1842, 37–38; Smadja 2015, 229)

The same point was repeated over and over again in the critical internalist context.

> There is nothing more common and natural, especially when reading older works, than to substitute one's own point of view for that of the author, and to take for granted that the knowledge and insights we are familiar with are also taken for granted by that author, and to 'read into' him things he never thought of. It is therefore absolutely necessary to distinguish clearly the conclusions drawn from the words of an author from his own expressions. (Stäckel 1900, 255)

This strong rejection of anachronism and presentism was no mere lip service, but a fundamental recognition that these historiographic vices were in direct conflict with several of the core purposes the critical internalists intended history to serve in the broader landscape of mathematical thought.

The critical internalist ideal was that of "truly historical thinking united with personal research activity [in mathematics]." "Somewhat like history of philosophy," the history of mathematics must combine the genuinely historical with the "spirit" of research in the field itself. If it does, "even the purely professional mathematician would stand to gain from engagement with it"; if not, it will be relegated to "an entirely meaningless separate existence." (Neugebauer 1927, 39, 44). It is precisely because critical internalist history of mathematics intends to make a genuine contribution in the domain of actual mathematical thought that it must not be subservient to modern mathematics. For if historical scholarship merely passively mirrors the values of modern mathematics, it would be redundant to mathematicians and hence doomed to a "meaningless separate existence" in their eyes. Genuinely historical scholarship, on the other hand, could make relevant contributions to the mathematician by putting existing conceptions in an unexpected light, and to uncover and clarify illuminating connections between them, which is exactly what new research mathematics does.

The value of history for obtaining a form of understanding somewhat analogous to that gained through active research became all the more relevant as specialization increased. The more mathematics grew, the more mathematicians found themselves

having to engage with areas of mathematics in which they had no personal research experience. This drove home to them the insight that "it is impossible to grasp even *one* mathematical concept without having assimilated all the concepts which led up to its creation, and their connections" (Klein 1928, 1). To address this, one of the goals of historical scholarship was to recreate a distilled equivalent of first-hand research experience at a certain time and with certain tools.

> [History requires] acquainting oneself so well with the notions and procedures in use at that time that one is able to judge what could be achieved within the area known at the time and under the given conditions, and moreover by examining what these tools have in fact been used for. A thorough understanding will not be obtained until one finds a complete connection between the usability of the tools at their disposal and the range of the results they gained with them. (Zeuthen 1903, 555; Lützen and Purkert 1994, 16)

This describes not only sound historical methodology but also, and equally, the mindset of a mathematician immersing themselves in a foreign research field, or conscientiously preparing to teach a course in an area outside their expertise.

It is evident already from these considerations that an approach to history that is only interested in identifying precursors of one's preexisting notions is profoundly antithetical to critical internalist historiography – indeed, it immediately voids the core motivations for studying history at all. This applies to many other critiques of presentist internalist historiography as well.

It is common to contrast the mathematician's selective interest only in work that turned out to be important in hindsight with the historian's equal attention to dead ends and ideas that did not go anywhere (Grabiner 1975, 439–440; Grattan-Guinness 2011, 4). Again, nothing could be further from the goals of critical internalist historiography. The core aim of history to provide a critical perspective and sense of research-problem dynamic for parts of mathematics outside one's direct expertise depends very essentially on an understanding of the limitations of alternative approaches.

> Not only a knowledge of the ideas that have been accepted and cultivated by subsequent teachers is necessary for the historical understanding of a science, but also that the rejected and transient thoughts of the inquirers, nay even apparently erroneous notions, may be very important and very instructive. (Mach 1883/1919, 254–255)
>
> The historical study of the sciences must not only be concerned with tracing the progress of the human mind in the knowledge of truth, but also with recalling its errors; it is precisely the healthy appreciation of these errors which alone can make the true importance of the sciences clear. (Tannery 1930, 3)

In this respect, it is precisely the study of approaches that do *not* map neatly onto modern ones that is mathematically enlightening, and hence the particular interest of mathematicians who want history to be relevant to their concerns.

> It is certainly of interest to know when an idea or method that is part of present mathematics reached a form in which we immediately recognise it, or when a technique whose utility we now know was first used. But it is of greater interest to pay attention to those ideas and

> methods that diverge from present ones. … The modern method will generally have far more extensive uses, and it will be stated in rules that can be applied purely mechanically; but precisely because the related older methods were not so designed, but rather had to be adapted to each case, they could led to a deeper penetration and a more versatile investigation than the corresponding modern treatment. The mere fact that the less developed form required a greater effort of thought led to observations that would easily escape one who now reach the same main result with railroad speed. (Zeuthen 1903, 554–555)

The critical internalist approach sees the study of the history of mathematics as in many ways analogous to original mathematical research. "Whoever wants to find something new must often free themselves from finished rules, and this can be learnt from those who have not yet organised their procedures into finished rules." (Zeuthen 1903, 554–555). Just as mathematical innovation often consists in getting a firm grasp of ideas that are "in the air," so also historical understanding depends on that same talent to perceive the essence in a bustling body of burgeoning work (Weil 1978, 438). On a more big-picture scale, forming a clear overview of the development of ideas in a field up to a certain point is a task common to both the historian and the mathematician setting out an original research program, suggesting that mathematicians (especially research strategists) and historians of mathematics share skillsets and interests (Weil 1978, 436–437). These analogies work only if past mathematics is understood in its own terms, including its challenges and limitations as they appeared at the time. Only then is historical research analogous to mathematics research. All of these analogies break down completely if past mathematics is merely mined for precursors of modern ideas or is presumptuously judged by their degree of agreement with modern theories. This again shows why such presentist perspectives are anathema to critical internalism.

Supporting these goals often involves interpreting and clarifying past mathematics in modernized form. The historian's "most important function as far as the study of mathematics itself is concerned" is that of "studying and analyzing the works of the great mathematicians throughout history in order to make those works accessible to modern readers and researchers" (Edwards 2020, 66). Mathematics should be useful "above all to the mathematicians who are its most reliable consumers and the creators of its raw material" (May 1975, 450). Of course, "when a professional mathematician wants to cultivate the history of his science, he must, obviously, first of all submit to the rules that apply to everyone who wants to know the historical truth." But the goals of the historian and the mathematician are not incompatible. On the contrary, the mathematician can "contribute something of his own" to historical analysis by bringing current mathematical tools to bear on illuminating past thought.

> In very different expressions one can recognize the same truths. … In this way, the mathematically educated cultivator of the history of mathematics has the opportunity to interpret texts which otherwise seem to be incomprehensible, or which have been misunderstood, to find connections among historical statements which otherwise may seem to be concerned with different things, to detect the preparation of a discovery which otherwise seems to have been due to the unique gift of prophecy of one single ingenious man, and above all to find and understand the coherence in the research and knowledge of a given period and thereby its connection to the points of view of the earlier and later periods from or

> to which impulses are given. Not only will historical knowledge thereby increase and be consolidated but one will gain exactly that type of knowledge of old mathematics that will give the greatest benefit for mathematicians and pedagogues. (Zeuthen 1917, 201–202; Lützen and Purkert 1994, 17)

Such analyses must of course be careful to avoid the traps of anachronism. But the mere potential fallibility of the mathematician's method of analysis is no reason to abandon it; other historical methods have equal pitfalls of their own.

> It is not uncommon for past mathematicians to present their work in a way that differs substantially from that of modern science, and this can lead one to pay too little attention to the real train of thought when studying a text, so that one reads the modern view into the presentation. . . . By going back to the sources [and remaining close to the original form of expression], a mathematical-historical author will of course be able to avoid certain mistakes [of anachronism], but it is very possible that he will be guilty of other and perhaps even greater mistakes [such as deficiencies of mathematical understanding]. (Eneström 1911/12, 6–8)

The only foolproof way to avoid misrepresentation is to reproduce the original expressions verbatim, without commentary or analysis, but this would serve only to make the historian superfluous and capitulate on a major goal of historical scholarship.

Critical internalism has sound tools to avoid the dangers of anachronism. The structural relations between ideas, rather than their superficial surface form, is the truest indicator of their nature and essence.

> The value of a method used in the past does not depend on the greater or lesser superficial resemblance to those whose utility we now know, but on what they could be – and actually were – used for in their own time. Likewise, reaching a particular result has significance in and of itself; but the discovery of such a result may be more or less coincidental, and the value one should attribute to that discovery at a given time depends on whether the significance of the result, and its uses, were understood at the time. The correct assessment of the various mathematical advances of the past can therefore only be obtained by studying the mathematics of these times as a whole, that is to say by immersing oneself in the then customary ideas and methods. (Zeuthen 1903, 555)

The viability of interpretative hypotheses and the faithfulness of modernizing paraphrases should be judged at this deeper level of conceptual analysis rather than by excessive attention to surface form.

This point is much misunderstood. Instead of judging fidelity to historical thought in these sophisticated terms, historians often fall back on adherence to old notation and language as a proxy for non-anachronism – a conceptually simplistic and impoverished notion of historical fidelity (Blåsjö and Hogendijk 2018) that conflates critical with presentist internalism. Thus, historians congratulate themselves that "only comparatively recently has the practice of translating [historical] works into modern notation become unacceptable," praising the "new and sophisticated literature" that accords with these "historically sensitive" standards (Bertoloni Meli 2006, 9; Blåsjö 2014, 114). It is even claimed that analyzing past thought in modern

mathematical terms "implies ... a strong affirmation of the superiority of modern ideas over those of the past" (Rowe 1996, 10). But this is exactly contrary to the goals of critical internalism, to which providing a critical perspective on current mathematics is one of the main purposes of historical scholarship in the first place. Critical internalists were always very clear about this:

> The goal of my mathematical historiography ... requires ... a study that does not assume that our time has become wiser than [past mathematicians] were, but that seeks to learn something from them, and particularly to grasp that [aspect] of their way of thinking which may have been lost during the later generalising development. (Zeuthen 1903, 555–556)

Modern historians have been equally critical of the "Aristotelian" mode of engaging with history that uses an analysis of past thought as a basis for formulating new and forward-looking ideas. Cherry-picking historical evidence for self-serving ends is not history but "history-satire" (Grattan-Guinness 1973, 449), or "ransacking history" (Dauben 1994, 7). History proper asks "what happened in the past?" whereas these presentist narratives instead ask "how did we get here?" This is not "history" but "heritage": a self-centered delineation of a "royal road to me" (Grattan-Guinness 2004, 164–165).

But this modern dichotomy is too uncharitable a view of the Aristotelian mode of engagement with history. The point of the Aristotelian approach is not to use the present to judge the past, or to use the past to flatter the present. Rather, it is to use the past to conceive and understand the present. The question "how did we get here?" presupposes that the "here" is already fixed before historical inquiry begins, rather than being shaped through historical investigation. The defining characteristic of Aristotelian accounts of history is perhaps not presentism but historicism: instead of condemning such histories as abusing and distorting the past for instrumental ends, we should perhaps praise them for making history an integral part of any creative and critical thought.

The confluence of historical analysis and current thought can be seen as a source of strength and purpose rather than as a threat to historical purity and autonomy. "The historical notes in Bourbaki's *Eléments* were intended to show the antecedents of his ideas – an important goal, since a reorganization of established theories should be a natural outcome of previous mathematical development and not a rootless caprice." (May 1975, 449). Lagrange did the same "in many of his publications," basing his own systematization of the field and forward-looking research strategy on "excellent examples of serious historical research" (Dauben and Scriba 2002, 72). Indeed, it is natural for the interests of the mathematician and the historian to be united in such enterprises, since an incorrect understanding of past work will lead to a wrong-headed synthesis and hence a research program that fails to learn from the work that has already been done.

Instead of the dichotomy between "good" history pursued with a supposedly pure mind, and "bad" history contaminated by the author's own ideas, we should perhaps question the assumption that historical inquiry could and should be cleanly detached

from other dimensions of human thought, as such a critique presupposes. Indeed, one could argue that this assumption is a demonstrable logical impossibility.

> The idea of forgetting the later past of an historical episode, often put forward as desirable historiography, is impossible to achieve, since the historian has to know which notions *are* later, and this requires the historical task already to have been accomplished. . . . We have some historical text before us; say Euclid's *Elements*. How can we read it in a historical spirit? A popular answer, put forward for all kinds of history, goes as follows. When reading Euclid's work, forget all theories in the field involved that have been developed since; step into his shoes . . . and read his work with his eyes. Unfortunately, . . . this method suffers from a difficulty; namely, it is completely useless. For in order to ignore all knowledge produced since Euclid one needs to know what that knowledge is in the first place. But in order to know that we must be able distinguish it from the knowledge produced before and during Euclid's time. But in order to know that we need to know the history of Euclid's work – before studying the history of Euclid's work! (Grattan-Guinness 2004, 169, 179)

Altogether, the war on anachronism in modern historiography has sometimes been more bark than bite. It is a powerful image that translating past mathematics into modern notation and terminology is an act of unsympathetic erasure and dogmatic self-centeredness. As history of mathematics moved outside mathematics departments, this evocative motif was a neat rallying call for the self-identity of the new generation. Still, it may be fair to say that the number of actual scalps won by this method has been a bit underwhelming compared to the passion with which it has been advocated. Concrete attempts to prove its superiority over the critical internalism of old in specific cases have perhaps on the whole been less resounding than the rhetoric surrounding them (Blåsjö 2016). But if smoking-gun evidence for the culpability of old internalism is in limited supply here, then modern historiographers have had a second major prong of attack, to which we now turn.

3 Reconstruction as Historical Method

Mathematicians can be impatient readers of history. When they approach a text, they tend to ask themselves "what's really going on here?" (Rowe 1996, 9). They may prioritize "a sort of rationally reconstructed or 'distilled' history," dismissing minor aberrations with the cavalier attitude that "actual history is frequently a caricature of its rational reconstructions" (Lakatos 1976, 5, 21; Lakatos 1970).

Modern historians take a diametrically opposite view. Trying to explain why mathematical ideas developed as they did is seen as inherently dubious.

> The student of history should not ask such childish queries. His purpose cannot be to give a completely logical account of the past, for such account is obviously out of the question. It is only here and there that a few logical knots can be tied, for the rest, we must be satisfied with a faithful description of the possibilities which materialized. . . . The capriciousness of mathematical development cannot be emphasized too much. (Sarton 1957, 18)

In keeping with this conviction,

> whereas the history of science previously stressed a big picture focusing on the theoretical progress made by great scientific heroes . . ., the field is now dominated by scholars offering rich, thickly descriptive, local studies. (Lightman 2016, 1)

The mathematician is most interested in higher-order connections between ideas, while the historian finds authenticity precisely in local and idiosyncratic particulars. One may use a biological metaphor to illustrate these contrasting approaches.

> [Historical sources can be likened to] the fossils that tell us about life on earth in distant epochs, and just as the palaeontologist dares to reconstruct the fauna and flora of that time on the basis of these remains, one could demand that the historian of mathematics should try to establish the connection of mathematical ideas on the basis of the facts established by exact research and thus to bring that mathematical thought back to life. (Stäckel 1900, 260)

But there are two levels to this enterprise: the global and the local. One concerns an "evolutionary" view of history – trying to relate a given historical occurrence to its precursors and antecedents – while the other is "phenomenological" – striving for the most direct, raw experience of a historical moment in and of itself (Dijksterhuis 1953).

Both are essential for critical internalist history. The phenomenological perspective corresponds to the goal of recreating an authentic picture of the genuine research experience as it appeared to mathematicians of the past.

> The historian must re-enact the past in his own mind. . . . He must see what the [underlying] problem was, of which his author is here stating his solution. He must think that problem out for himself, see what possible solutions of it might be offered, and see why this particular [author] chose that solution instead of another. (Collingwood 1993, 282–283)

The mathematician's tendency to draw parallels to known modern problem contexts in order to create "useful pasts" (Goldstein 2010) can more charitably be seen as a way of getting at this form of authenticity rather than as a dogmatic translation into a preconceived framework.

The evolutionary perspective is equally needed, since traditional internalism intends history of mathematics to serve a supporting function to mathematical research and understanding. Purposes such as drawing up and critically assessing research programs and educational curricula demand a big-picture view of overarching developments – a need that history is well suited to address. Thus, to the internalist, "truly, the main purpose of the history of the development of mathematics is to provide information about the connection between mathematical ideas." (Eneström 1910/11, 6)

Based on confidence in a fundamental unity of mathematical through across all ages and cultures, mathematicians tend to be optimistic about the possibility of successful reconstructions on both of these levels.

> The paleontologist undertakes to reconstruct the whole organism from scarce remains, assuming that the laws of biology observed in beings living at present have always held. Who would deny that this assumption has proved to be true, that, even if some errors were made, the results of this research are more than an uncontrollable game of the imagination? It seems to be precisely because of the unity of the human spirit that the laws of productive mathematical creation which are currently in force remain applicable to earlier times. (Stäckel 1900, 262)

Today's historians have much less confidence in this assumption of unity of past and present thought, and the ability of a modern mathematician to intuitively empathize with historical thought in a way reliable enough for historical reconstruction. It is felt that the old attempts based on these hubristic assumptions underestimated their own subjectivity, and overidealized and over-linearized the irregular and contingent nature of historical development (Saito 1998, 134; Sidoli 2014, 35). To avoid these dangers, today it is almost axiomatic – "a kind of null-hypothesis" – "that there is a discontinuity between mathematical thought of the past and that of the present" (Fried 2018, 16). In terms of the biological metaphor, this corresponds to restricting of paleontology to the descriptive, factual level, and treating with great suspicion the possibility of genuinely grasping the lived experience of other organisms, as well as viewing the overall path of evolution as governed more by chance circumstance than explicable rational principles.

These problems were well known to the old generation of critical internalists.

> Doesn't one leave the field of science with the demand that the historian should not only represent the content but also the connections of mathematical ideas? Are not the historical facts then replaced by arbitrary and imaginative structures, so that such reconstructions have only subjective value? Certainly, no one will deny that taking this path brings with it great dangers, and it has rightly been warned that philosophical constructions of history made without sufficient expertise have never truly advanced our understanding. Nevertheless, I am convinced that this demand is not only justified, but that it can also be carried out. . . . Far from speaking out in favour of unbridled subjectivism, I consider it the right, the duty of the historian to go beyond the 'bare facts', and I am convinced that it is possible to fulfil this demand without having to give up the achievements of the critical school. . . . If such conjectures, which of course must always be called conjectures, explain all the facts of a complex of phenomena in a simple way, then one will be able to regard them, just as in the natural sciences, as a stage in the journey from error to truth. (Stäckel 1900, 261–263)

The divergent intuitions of the mathematician and the historian cast the viability of the method of hypotheses in the history of mathematics in very different lights. Historians, inclined to assume major discontinuities between past and present, give great credibility to documents of the time and little to later interpretations. Mathematicians, by contrast, place great faith in the brotherhood of their kind, and feel that the work of a mathematician is better understood by a mathematician of another age than by an outsider who happened to be closer merely in time and space. Already in the days of Zeuthen and Cantor, this conflict was articulated and debated at some length (Lützen and Purkert 1994, 28–33). Similarly, historians are predisposed to seeing idiosyncrasies and the particular, whereas mathematicians have a sense that their field develops almost quasi-deterministically, which implies good prospects for empathetic

understanding and reconstruction (Paumier and Aubin 2016, 209; Bourbaki 1994, 173). "Beethoven and Shakespeare did unique work, and we would not have *Fidelio* or *King Lear* if they have not lived. However, symmetric functions would have been discovered by others if Newton had not lived." (Askey 1988, 202).

Comparing the merits of these incommensurable outlooks objectively is all but impossible, and so is tallying their track records of actual outcomes. One could argue that internalism enjoyed some remarkable predictive successes: the discoveries of Archimedes's lost work *The Method* and the decoding of Babylonian mathematics, for example, can be regarded as spectacular vindications of hypotheses that had been previously stipulated on purely internalist grounds (Lützen and Purkert 1994, 25; Rowe 1996, 7). On many smaller points as well the critical internalists found that "whether I have been on the right track with [my historical interpretations], my work itself has often given me an opportunity to test" (Zeuthen 1903, 556), seeing themselves vindicated by the results. The necessity of using mathematical understanding to go beyond explicit sources can also be argued from the nature of present mathematics. "For example, Weierstrass's new foundation of the theory of functions was largely spread orally by his students. If a future historian of the nineteenth century would neglect this source of transmission, he would come to completely wrong conclusions." (Stäckel 1900, 261).

But the method of hypotheses has also had its failures. The early twentieth-century generation is often faulted for being too much under the spell of the mathematics of its own time rather than that of the historical era they were supposed to be studying (Høyrup 2016, 71; Chemla 2012, 21–26; Gray 2008, 373). A notorious example is the alleged crisis in early Greek mathematics resulting from the discovery of incommensurability, or the irrationality of $\sqrt{2}$ – a prominent topic of speculation to the critical internalists generation. Today's historians are inclined to view this as dubious history contaminated by the concerns of the mathematical community at the time – a mere projection of the *Grundlagenstreit* with Hilbert and Brouwer as the main antagonists (Knorr 2001; Saito 1998, 133).

Indeed, it is probably indisputable in retrospect that the old internalists sometimes let an inflated sense of the mathematical preoccupations of their own time go to their heads and cloud their judgement on historical matters in some cases. But one may note that the critical internalist historiographical program contains within it the tools to cure this problem. Namely, in its oft-repeated principle that history should be a vehicle for a critical perspective on current mathematical practice.

In addition to anachronistic biases, subjectivity is another charge against the method of reconstruction: "the number of plausible, reconstructed histories is often equal to the number of scholars working along such lines" (Sidoli 2014, 35). This severely undermines the allegedly semi-deterministic and unified character of mathematical development that was invoked to justify the reconstructivist enterprise in the first place. But in this case too, a corrective principle can be derived from within sophisticate internalism itself. The critical internalists were under no delusion that their reconstructions were anything but fallible, but they were confident that lively clashes of bold and clear ideas would be a self-correcting and fruitful enterprise. Daring hypotheses are valuable "if only to provoke further study." "It is the fate of all

scholarship to be valid only for a few generations; . . . let us stir up as many ideas as we can. Our grandchildren will benefit from this." (Tannery 1889, 496; Pineau 2010, 377). Many modern historians are more likely to greet the proliferation of hypotheses with exasperation than combative enthusiasm. Internalist studies that would be useful in the sense intended by Tannery, assuming an ambient culture of debate, are bound to look reckless in the risk-averse climate of today's historiography. What is constructive in one context is folly in the other. A multitude of competing alternative reconstructions is consistent with a standard conception of scientific progress, and is not in itself proof of a pathological level of subjectivity in the field. Theoretical quantum mechanics was not abandoned in the early twentieth century just because empirical findings were subject to several viable but incompatible interpretations.

In sum, relatively free-wheeling reconstruction as a historical method is naturally associated with the goals and convictions of the mathematician. Although out of fashion historiographically, this may not be a straightforward case of good riddance. Reconstructions may have a checkered track record in some respects, but the core convictions that motivated this enterprise also entail self-correcting mechanisms for overcoming precisely the weaknesses nowadays commonly associated with such approaches.

4 Conclusion

Mathematicians lost the historiography wars of the twentieth century and they have only themselves to blame. Due to an increasingly exclusive focus on specialized research productivity, mathematicians lost touch with the ideals that had led to history being accorded a central place in mathematics around the beginning of the century. The critical internalist historiography articulated by that generation gave way to a more simplistic internalism. Historians with a vested interest in wresting the field from the mathematicians pounced on this weakness and portrayed it as irrevocable. Having weakened their commitment to the intellectual unity of mathematical research, education, and historical inquiry, mathematicians were left without the tools to defend themselves. The mathematicians lost the historiography wars not for the fallibility of their approach to history, but for neglect of the ideals on which this approach was founded. The historiography of mathematics may yet have much to gain by reviving these ideals.

5 Cross-References

- ▶ Algebraic Versus Geometric Thought and Expression in the Early Calculus
- ▶ Mathematical Practice as Philosophy, with Galois, Riemann, and Grothendieck
- ▶ What Happens, from a Historical Point of View, When We Read a Mathematical Text?

References

Askey R (1988) How can mathematicians and mathematical historians help each other? In: Kitcher P, Aspray W (eds) History and philosophy of modern mathematics, Minnesota studies in the philosophy of science XI. University of Minnesota Press, Minneapolis, pp 201–217

Barany MJ (2020) Histories of mathematical practice: reconstruction, genealogy, and the unruly pasts of ruly knowledge. ZDM 52:1075–1086

Bertoloni Meli D (2006) Thinking with objects: the transformation of mechanics in the seventeenth century. Johns Hopkins University Press, Baltimore

Blåsjö V (2014) A critique of the modern consensus in the historiography of mathematics. J Humanist Math 4(2):113–123

Blåsjö V (2016) In defence of geometrical algebra. Arch Hist Exact Sci 70(3):325–359

Blåsjö V, Hogendijk JP (2018) On translating mathematics. Isis 109(4):774–781

Bourbaki N (1994) Elements of the history of mathematics. Springer, Berlin

Brentjes S (2016) Practicing history of mathematics in Islamicate societies in 19th-century Germany and France. In: Remmert VR, Schneider M, Kragh Sørensen H (eds) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Basel, pp 25–52

Bullynck M (2016) The history of mathematics in the progress of mankind: modifying the narrative around 1800. In: Remmert VR, Schneider M, Kragh Sørensen H (eds) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Basel, pp 9–23

Cajori F (1896) A history of elementary mathematics with hints on methods of teaching. Macmillan, New York

Chemla K (ed) (2012) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge

Collingwood RG (1993) The idea of history. Oxford University Press, New York

Dauben JW (1994) Mathematics: an Historian's perspective. In: Sasaki C, Sugiura M, Dauben JW (eds) The intersection of history and mathematics, science networks historical studies XV. Birkhäuser, Basel

Dauben JW (1999) Historia mathematica: 25 years/context and content. Hist Math 26(1):1–28

Dauben JW, Scriba CJ (eds) (2002) Writing the history of mathematics: its historical development. Birkhäuser, Basel

Dijksterhuis EJ (1953) Doel en methode van de geschiedenis der exacte wetenschappen, inaugural address. Utrecht University, Utrecht

Edwards HM (2020) The role of history in the study of mathematics. Math Intell 42:66–69

Eneström G (1900) Ziele und Aufgaben eines Organs für mathematisch-historische Forschung und für aktuelle Fragen auf dem Gebiete der mathematischen Wissenschaften. Bibl Math 1(3):1–7

Eneström G (1908/09) Über kritische Behandlung der Geschichte der Mathematik. Bibl Math 9 (3): 1–14

Eneström G (1910/11) Über Probleme der mathematischen Geschichtsschreibung. Bibl Math 11(3): 1–10

Eneström G (1911/12) Über die Bedeutung von Quellenstudien bei mathematischer Geschichtsschreibung. Bibl Math 12(3):1–20

Eneström G (1912/13) Wie kann die weitere Verbreitung unzuverlässiger mathematisch-historischer Angaben verhindert werden? Bibl Math 13(3):1–13

Fried MN (2018) Ways of relating to the mathematics of the past. J Humanist Math 8(1):3–23

Goldstein C (1995) Un theoreme de Fermat et ses lecteurs. Presses Universitaires de Vincennes, Saint-Denis

Goldstein C (2008) How to generate mathematical experimentation and does it provide mathematical knowledge? In: Fest U et al (eds) Generating experimental knowledge. MPIWG, Berlin, pp 61–85

Goldstein C (2010) Des passés utiles: mathématiques, mathématiciens et histoires des mathématiques. Noesis 17:2010

Grabiner JV (1975) The mathematician, the historian, and the history of mathematics. Hist Math 2(4):439–447
Grattan-Guinness I (1973) Not from nowhere: history and philosophy behind mathematical education. Int J Math Educ Sci Technol 4(4):421–453
Grattan-Guinness I (1990) Does history of science treat of the history of science? The case of mathematics. Hist Sci 28(2):149–173
Grattan-Guinness I (2004) The mathematics of the past: distinguishing its history from our heritage. Hist Maths 31(2):163–185
Grattan-Guinness I (2011) History or heritage? An important distinction in mathematics and for mathematics education. Am Math Mon 111(1):1–12
Gray J (2008) Plato's ghost: the modernist transformation of mathematics. Princeton University Press, Princeton
Guicciardini N (2018) Un altro presente: on the historical interpretation of mathematical texts. BSHM Bull 33(3):148–165
Høyrup J (2016) Mesopotamian mathematics, seen "from the inside" (by Assyriologists) and "from the outside" (by historians of mathematics). In: Remmert VR, Schneider M, Kragh Sørensen, H (eds) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Basel, pp 53–78
Kitcher P, Aspray W (1988) An opinionated introduction. In: Kitcher P, Aspray W (eds) History and philosophy of modern mathematics, Minnesota studies in the philosophy of science XI. University of Minnesota Press, Minneapolis, pp 3–60
Klein F (1896) The arithmetizing of mathematics. Bull Am Math Soc 2(8):241–249
Klein F (1928/1979) Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert. Teil I., Springer-Verlag, Berlin. Quoted from the English translation, Development of Mathematics in the 19th Century, Math Sci Press, Brookline
Klein F (1932) Elementary mathematics from an advanced standpoint: arithmetic, algebra, analysis, Macmillan, New York, reprinted by Dover Publications
Knorr WR (2001) The impact of modern mathematics on ancient mathematics. Revue d'histoire des mathématiques 7(1):121–135
Lakatos I (1970/1971) History of science and its rational reconstructions. In Buck RC, Cohen RS (eds) PSA 1970: proceedings of the biennial meeting of the philosophy of science association, Boston Studies in the philosophy of science VIII, Springer, Dordrecht, pp 91–136
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge, UK
Lightman B (ed) (2016) A companion to the history of science. Wiley Blackwell, Chichester, UK
Lützen J, Purkert W (1994) Conflicting tendencies in the historiography of mathematics: M. Cantor and H. G. Zeuthen. In: Knobloch E, Rowe DE (eds) The history of modern mathematics, volume III: images, ideas, and communities. Academic Press, Boston, pp 1–42
Mach E (1883/1919) The science of mechanics: a critical and historical account of its development, Open Court Publishing, La Salle, IL, 1919. First published as Die Mechanik in ihrer Entwickelung historisch-kritisch dargestellt, Leipzig
May KO (1975) What is good history and who should do it? Hist Math 2(4):449–455
Nesselmann GHF (1842) Versuch einer kritischen Geschichte der Algebra I: Die Algebra der Griechen. Reimer, Berlin
Netz R (2004) The works of archimedes, vol 1. Cambridge University Press, Cambridge
Neubegauer O (1956) A notice of ingratitude. Isis 47(1):58
Neugebauer O (1927) Über Geschichte der Mathematik. Mitteilungen des Universitätsbundes Göttingen 9(1927):38–45
Neugebauer O (1951/1969) The exact sciences in antiquity, 2nd edn. Dover Publications, New York
Paumier A-S, Aubin D (2016) Polycephalic Euclid? Collective practices in Bourbaki's history of mathematics. In: Remmert VR, Schneider M, Kragh Sørensen H (eds) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Basel, pp 185–218
Pineau F (2010) Historiographie de Paul Tannery et réceptions de son œuvre: sur l'invention du métier d'historien des sciences, PhD dissertation, Université de Nantes

Poincaré H (1899) La logique et l'intuition dans la science mathématique et dans l'enseignement. L'Enseignement Mathématique 1:157–162
Richards JL (2006) Historical mathematics in the French eighteenth century. Isis 97(4):700–713
Rowe DE (1996) New trends and old images in the history of mathematics. In: Calinger R (ed) Vita mathematica. Mathematical Association of America, Washington, DC
Rowe DE (2016) Otto Neugebauer's vision for rewriting the history of ancient mathematics. In: Remmert VR, Schneider M, Kragh Sørensen H (eds) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Basel, pp 123–141
Saito K (1998) Mathematical reconstructions out, textual studies In: 30 years in the historiography of greek mathematics. Revue d'histoire des mathématiques 4(1):131–142
Sarton G (1957) The study of the history of mathematics, and the study of the history of science. Dover Publications, New York
Schneider I (1992) The history of mathematics: aims, results, and future prospects. In: Demidov SS et al (eds) Amphora: Festschrift für Hans Wussing zu seinem 65. Geburtstag. Birkhäuser, Basel, pp 619–629
Sidoli N (2014) Research on ancient Greek mathematical sciences, 1998–2012. In: Sidoli N, Van Brummelen G (eds) From Alexandria, through Baghdad: surveys and studies in the Ancient Greek and Medieval Islamic mathematical sciences in Honor of JL Berggren. Springer, Berlin, pp 25–50
Siegel CL, Lenzen KM (1979) On the history of the Frankfurt mathematics seminar. Math Intell 1(1979):223–230
Smadja I (2015) Sanskrit versus Greek 'Proofs': history of mathematics at the crossroads of philology and mathematics in nineteenth-century Germany. Revue d'histoire des mathématiques 21(1):217–349
Stäckel P (1900) Review of Cantor, M., Vorlesungen über Geschichte der Mathematik. Göttingische gelehrte Anzeigen 162(1):251–264
Struik DJ (1980) Why study the history of mathematics? Undergrad Math Appl J 1:3–28
Tannery P (1889) Review of Paul Tannery, Pour l'histoire de la science hellène. Archiv für Geschichte der Philosophie 2:492–498. Reprinted in Paul Tannery, Mémoires Scientifiques, XI, 360–366
Tannery P (1930) Mémoires Scientifiques, Volume X, (edited by Heiberg JL, Zeuthen HG). Gauthier-Villars, Paris
Tobies R (2019) Felix Klein: Visionen für Mathematik, Anwendungen und Unterricht. Springer, Berlin
Toeplitz O (1926/2015) The problem of University courses on infinitesimal calculus and their demarcation from infinitesimal calculus in high schools. Sci Context 28(2):297–310
Turner LE (2020) Cultivating a research imperative: mentoring mathematics at Stockholms Högskola 1882–1887. Hist Math 50(2020):50–83
Unguru S (1975) On the need to rewrite the history of Greek mathematics. Arch Hist Exact Sci 15(1):67–114
Weil A (1978/2009) History of mathematics: why and how, lecture at the international congress of mathematicians in Helsinki, 1978, reprinted in Oeuvres scientifiques – collected papers III: 1964–1978. Springer, New York
Zeuthen HG (1903) Ved forelæggelsen af "Mathematikkens Historie i 16. og 17. Aarhundrede," In: Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger. Copenhagen
Zeuthen HG (1917) Hvorledes Mathematiken i tiden fra Platon till Euklid blev rationel videnskab, Det Kongelige Danske videnskabernes selskabs skrifter, Naturvidenskab og Mathematik, Afd. 8, Raekke I.5. Copenhagen
Zhmud L (2017) Eudemus' history of mathematics. In: Bodnar, Fortenbaugh (eds) Eudemus of Rhodes, Rutgers University Studies in Classical Humanities XI. Routledge, New Brunswick, NJ, pp 263–306

A Tale of Three Cities: Thebes, Babylon, and Alexandria

Maurice Burke

Contents

1 Introduction 1150
2 Thebes 1151
3 Babylon 1155
4 Alexandria 1158
4.1 The DMSS System 1158
4.2 Rational Expression Arithmetic 1160
5 Interweaving the Three Tales with Current Strands 1162
5.1 Measurement Systems and Decimals 1162
5.2 Common Fractions and Algebraic Reasoning 1163
5.3 The Number Line, Measure Numbers, and Precision 1164
6 Collecting the Threads and Tying the Present to Its Recent History 1165
6.1 A Current State of Affairs 1165
6.2 Observations from Recent History 1169
6.3 Other Voices of Note 1171
7 Conclusion 1173
7.1 Suggestions for Consideration 1173
7.2 The Lesson of Thebes 1175
References 1175

Abstract

The interface of the "scientific tradition" of mathematics with the traditions of American elementary education has grown in the past 200 years with state curriculum standards currently a highpoint of interactions. Episodes from the history of measure number provide a perspective suggesting current standards regarding the teaching of common fraction in American schools might be a bit wrongheaded and in need of reconsideration. Deeply rooted traditions are stubborn things, but the historical perspective provides clues for new directions.

M. Burke (✉)
Montana State University, Bozeman, Montana, USA
e-mail: burke@math.montana.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_100

Keywords

Common fractions · Decimals · Standards · Common core · Measurement · History of number

1 Introduction

The philosopher, Philip Kitcher (1984), presents a theory of mathematical knowledge residing in and warranted by evolving mathematical practices that are primordially rooted in human interactions with the world. Kitcher's historical examples are drawn almost entirely from what Høyrup (1989) refers to as the "scientific tradition" of mathematical knowledge. This tradition is *theoretical* in contrast to the practical art of subscientific mathematical knowledge, employed in the trades, commerce, and elementary schooling well into the early modern world. While Høyrup explains the extinction of the subscientific mathematical tradition in Early Modern Europe of the late sixteenth century, DNA from that tradition is richly present in elementary mathematics education today in the USA, with its rather unique semiautonomy and guild-like practices of local control that stem from its germination in American soil in the early seventeenth century. In fact, this Americanized variant might respond to Høyrup, like Mark Twain, "The report of my death was an exaggeration" (New York Journal, June 2, 1897). Just as Kitcher's theory must not ignore the "irrational" transitions or retrograde motions Høyrup (1989, 76) describes in the subscientific tradition, epicycles should also be noticed within the American tradition. This chapter shines a light on one of these epicycles relating to practices now driven by debatable curriculum standards.

The interface of the mathematical science tradition with the mathematical practices of elementary education in the USA has radically changed in the last 200 years. One of the key developments is the relatively recent emergence of standards for school mathematics. The "standards movement" in the USA surged between 1980 and 2010, with state and national insistence on assessment-driven accountability in the outcomes of elementary mathematics schooling. Its most recent high point has been the *Common Core State Standards for Mathematics* (CCSSM). To borrow Høyrup's terminology once again, such standards and their attendant frameworks and assessments are "index fossils" marking the boundary of a critical transition in the evolution of mathematical practice at the elementary level, a transition surely to be measured by many decades *yet to come* and not just a few years.

The history of mathematics makes thoroughly clear how much we do not know and is riddled with undocumented gaps. Even so, it does offer contrasts with the present, and this chapter examines historical contrasts to the fraction schemas being taught and notoriously not well learned in USA schools in recent decades, if they ever were. Its brief survey of historical practices raises questions about the current attempts to demand of elementary school children, through what are effectively national standards, fluency and proficiency in common fraction representations, interpretations, notations, operations, and standardized algorithms. The CCSSM

stipulations for common fractions have been embedded, virtually word-for-word, in the standards of 44 states and imitated in the rest for over a decade (NGA Center and CCSSO 2010). From the perspective of history, this might be a bit wrongheaded and worthy of reconsideration. The chapter closes with suggestions for future standards and practices regarding common fractions.

When using examples from ancient history, it is sometimes challenging to find consistent meaning in the words used by translators of ancient texts. This is not a criticism but an insinuation of the nature of a discipline. To paraphrase Wittgenstein, it is possible that the nuanced denotations of ancient symbols and written terms, passing through generations and cultures from the Bronze Age to us, like many strands twisted and braided into a rope, convey no thread of *meaning* running the full length (Wittgenstein 1958, 32). Therefore, like a statistician sampling a diverse population, this chapter attempts to connect the present to the past by interweaving multiple strands of meaning from the mathematical stories of three ancient cities with a sufficiently diverse collection of threads running through current elementary mathematics practice and its recent history. With these disclaimers and foreshadowing, please hang on to the gnarled rope that is woven, and welcome to Thebes.

2 Thebes

Imagine Thebes, the exotically picturesque capital of ancient Egypt in 1550 BC. There worked Ahmosé, the scribe who created the Rhind Mathematical Papyrus, a reproduction of a mathematical text from the time of Egypt's Middle Kingdom. This ancient scribe wrote in hieratic script, a cursive related to hieroglyphics almost as a symbol-by symbol shorthand. Ahmosé's hieroglyphic for $249\frac{2}{7}$ would faintly resemble the script in Fig. 1.

Wide spaces are used to separate the hieroglyphic into *five* symbols, each symbol a unique arrangement like the dots on a dice. They are, from right to left: 200, 40, 9, $\frac{1}{4}$, and $\frac{1}{28}$. A special mark, ⌓, was used above 4 and 28 to distinguish them as sub-multiples of the unit, i.e., $\frac{1}{4}$ and $\frac{1}{28}$. Ahmosé's hieratic script for this number would contain five corresponding symbols in the same order.

There are several important things to note about Ahmosé's fractions:

1. The five symbols forming the word $249\frac{2}{7}$ belonged to a larger palette containing 38 grouping symbols representing integer numerical sizes 1, 2, 3, …, 9, 10, 20, 30, …, 90, 100, 200, 300, …, 900, 1000, 2000, …, 9000, 10,000, and

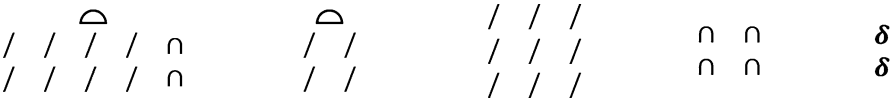


Fig. 1 Hieroglyphic for $249\frac{2}{7}$ read from right to left

100,000. The other grouping symbols on Ahmosé's numeral palette were $\frac{2}{3}$, and those some would call unit fractions, like $\frac{1}{4}$ or $\frac{1}{28}$, generated from their corresponding reciprocals with the use of the special sign.

2. The Egyptian $\frac{1}{7}$ is not the same as the unit fraction $\frac{1}{7}$, which is one of seven "equal quantities" or "aliquot parts" that compose a whole. Ahmosé seems to have thought of $\frac{1}{7}$ as a *grouping size* in the sense that a whole, as a measured quantity, has only *one* size that is its seventh *partsize*, while each of the seven equal aliquot quantities composing the whole shares that size. Thus, the partsize insinuates a ratio of an aliquot quantity to the whole quantity. Grouping sizes are not counted in Ahmosé's notation for sizing a quantity in terms of a single unit, though counting of unit quantities was certainly involved in the process of measuring and assigning grouping sizes. In contrast, the decimal 35.6, a multiplicative notation connoting $3 \times 10 + 5 \times 1 + 6 \times 0.1$, is a counting of grouping sizes: *3 tens 5 ones 6 tenths.* Indeed, what appears to be an exception, $\frac{2}{3}$, was very useful in Ahmosé's tools and methods, particularly in his doubling table and his methods for multiplication and division. Instead of being used as a count, it seems to be a special symbol for $\frac{1}{2}\ \frac{1}{6}$, perhaps preserving an earlier usage such as the size of *the-unit-less-its-third*.

   Unlike Ahmosé's fractions, a *common fraction*, as a measure number, is a count of a unit fraction. As a quantity, the common fraction $\frac{1}{7}$ can be considered a subunit, counted like a unit (hence the name *unit fraction*), and expressed as such in language and notation as, for example, *two-sevenths* or $\frac{2}{7}$. Thus, the word *partsize* is coined for this chapter to convey the Wittgenstein-like warning that Ancient Egyptian interpretations of fraction words and symbols are not common usages today.

3. A measurement in Ancient Egypt was done in the way typical of most ancient cultures: as a *simple measurement* using a count of a single unit or as a *compound measurement* using a decreasing sequence of counted amounts in a decreasing sequence of units and subunits, e.g., *4 cubits 6 hands 1 finger*. For computation purposes, given the algorithms of Egyptian practice, the latter measurement would be sized as a *single measure number in a single unit*: $4\frac{1}{2}\ \frac{1}{4}\ \frac{1}{7}$ cubit. In general, for Ahmosé to state the size of a quantity in terms of a single unit quantity meant to express the size as *a strictly decreasing sequence of the grouping sizes* from his numeral palette (Allen 2014, 123). Therefore, Ahmosé would not simply write $\frac{2}{7}$ or $2(\frac{1}{7})$ or $\frac{1}{7}\ \frac{1}{7}$, not even in his scratch work, to size the quantity we would represent as $\frac{2}{7}$. He would no more do this than one would use the numeral 32,020 to represent 340 just because $20 + 20 = 40$. In a very consistent practice, Ahmosé sized two of the seventh parts of a unit as $\frac{1}{4}\ \frac{1}{28}$. It was his way of sizing the ratio of a quantity to a single unit quantity.

The contours of Ahmosé's quantitative reasoning seem to move fluently from sizes expressed by his numerals to quantities having those sizes. In this regard, he was fully capable of reasoning with quantities in ways analogous to reasoning with

common fractions. For instance, when no measurement unit was given in a problem about an unknown quantity, he would identify the quantity with its size number. He could easily conceive of "5 hands, each the 7^{th} partsize of a cubit," a quantity he would size as $\frac{2}{3}\ \frac{1}{21}$ or as $\frac{1}{2}\ \frac{1}{7}\ \frac{1}{14}$ and would consider equivalent to the seventh aliquot part of a length that measured 5 cubits. (With common fractions, this corresponds to the insight that $\frac{5}{7} = 5 \div 7$, a theorem relating two separately defined numbers.) When performing arithmetic operations, Ahmosé could consider these five quantities as a "*heap*" of sevenths, an unsized quantity yet to be rendered as a single measure number of cubits. In fact, during a computation involving multiple heaps, Ahmosé sometimes jotted down lists of partsizes in no particular order and repeating a partsize several times.

Ahmosé would often reify partsizes into easy-to-work-with quantities by stipulating they referred to a quantity measured by a select number of units. For example, he could render $4\frac{2}{3}\ \frac{1}{5} + 1\frac{1}{2}\ \frac{1}{5}\ \frac{1}{15}$ as $5 + \frac{2}{3} + \frac{1}{5} + \frac{1}{2} + \frac{1}{5} + \frac{1}{15}$ and stipulate these sizes referred to wholes that each measured 30 units of the same kind. Thus, ignoring for the moment the 5 wholes, the fraction part of the sum amounted to putting together 20 + 6 + 15 + 6 + 2 into an unsized heap of 49 quantities each of the 30th partsize of one of the wholes. When sizing these 49 quantities by a single measure number, one of his many approaches would likely be to sort them into a heap of 30 quantities, which would be one whole, and then make heaps of size 15, 3, and 1, a *decreasing sequence* of quantities that were aliquot parts of 30. Accordingly, the size of all 49 quantities put together is $1\frac{1}{2}\ \frac{1}{10}\ \frac{1}{30}$, thereby giving $6\frac{1}{2}\ \frac{1}{10}\ \frac{1}{30}$ as his final sum. Ahmosé knew the sorting was not unique and tended not to use 1 as a "heap" size. So, instead of 15, 3, and 1, he could have created heaps of size 15, 2½, and 1½, thereby rendering the combined fractional quantities as $1\frac{1}{2}\ \frac{1}{12}\ \frac{1}{20}$ of the whole, which also might be preferred because of the larger partsizes.

Ahmosé's conceptions of fractions and sizing probably evolved from peculiarities of an oral tradition of measurement, perhaps the all-important system of grain measures. Grain measures were based on *hekat* units (≈ 5 quarts), and the naming sequence for most other grain units was developed by repeated halving and doubling of the *hekat* (Chace 1986, 18). Such a binary system would lend itself to reducing a measurement in multiple units (e.g.,1 gallon 2 quarts 3 cups), to a decreasing sequence of partsizes in one unit ($1\frac{1}{2}\ \frac{1}{8}\ \frac{1}{16}$ gallon). With relatively efficient computational algorithms, such a binary scale might have been the impetus for the evolution of the Egyptian system of measure numbers.

Unless disrupted by outside cultural influences, any oral tradition would develop computational algorithms symbiotically with its measurement systems. An arithmetic of halving and doubling seems "peculiarly" compatible (Chace 1986, 31) with the binary-scale system of grain-measure units. By the time of Ahmosé, Egyptians multiplied numbers using a doubling strategy, sometimes interleafed with multiplications by ten. If we think of multiplication originating as a scaling operation in the context of measurements, converting a grain measure from a larger unit to a smaller unit would involve a doubling strategy with the occasional multiplication by 10. (Ten entered with the *ro*, which was $\frac{1}{320}$ *hekat.*) Chace (1986, 31) suggests this

connection to grain measure as a possible source of the Egyptian multiplication technique.

To multiply 21 × 24, Ahmosé built the table shown in Fig. 2, with marks by the addends of 21:

Since $21 = 1 + 4 + 16$, Ahmosé knew $21 \times 24 = (1 + 4 + 16) \times 24 = 24 + 96 + 384 = 504$.

With fractions, such a multiplication technique demanded a way to double the partsizes on Ahmosé's numeral palette. Thus, the Rhind Papyrus opens by constructing a large table giving the double of $1/n$ for all odd integers n from 3 to 101. Ahmosé knew the table's values were the same as those when 2 is divided by n, i.e., the number that is multiplied by n to give 2. For example, the double of $\frac{1}{97}$ is given as $\frac{1}{56}\ \frac{1}{679}\ \frac{1}{776}$. Several methods were probably employed in computing the table. The simplest pattern seems to be the use of $\frac{2}{3}$, i.e., $\frac{1}{2}\ \frac{1}{6}$, as the template for entries of the form $\frac{2}{3n}$, which are all rendered in the form $\frac{1}{2n}\ \frac{1}{6n}$. The even values of n, such as 26, were not included in the table. Ahmosé knew that $2 \times 13 = 26$, and hence the *double of* $\frac{1}{26}$ is $\frac{1}{13}$. Indeed, he found quite useful and frequently invoked this *inverting rule*: $m \times p = n$ implies $m \times \frac{1}{n} = \frac{1}{p}$. Not surprisingly, Ahmosé's doubling table was a crucial tool in his computational algorithms.

These preliminaries help us understand Problem 33 in the Rhind Papyrus: *A quantity* [sometimes translated as *"heap"*], *its* $\frac{2}{3}$, *its* $\frac{1}{2}$, *and its* $\frac{1}{7}$, *to go in together, become 37. What is the quantity?* Today, $X + \frac{2}{3}X + \frac{1}{2}X + \frac{1}{7}X = 37$ and, consequently, $2\frac{13}{42}X = 37$ come to mind. Hence, $X = 37 \div 2\frac{13}{42} = 16\frac{2}{97}$. For Ahmosé, however, the problem required finding a number that multiplied $1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7}$ to get 37. He seemed satisfied to use $1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7}$ instead of reducing it to $2\frac{1}{6}\ \frac{1}{7}$ since it satisfied his sizing criteria and it conveniently expressed the problem. Doubling $1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7}$ four times, he gets $16 \times 1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7} = 36\ \frac{2}{3}\ \frac{1}{4}\ \frac{1}{28}$, which he found to be $\frac{1}{21}$ less than 37, possibly by using an abacus, and stipulating the fractions were partsizes of 84. But $\frac{1}{21}$ is the double of $\frac{1}{42}$, and he notes $1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7} \times 42 = 97$. Therefore, adapting his inverting rule, $1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7}$ times the double of $\frac{1}{97}$ is the double of $\frac{1}{42}$, or $\frac{1}{21}$. So, combining his products, $1\frac{2}{3}\ \frac{1}{2}\ \frac{1}{7} \times \left(16 + \textit{double of}\ \frac{1}{97}\right) = \left(36\ \frac{2}{3}\ \frac{1}{4}\ \frac{1}{28}\right) + \left(\frac{1}{21}\right) = 37$. Finally, getting the double of $\frac{1}{97}$ from his table, Ahmosé's answer is $16\frac{1}{56}\ \frac{1}{679}\ \frac{1}{776}$, agreeing with our $16\frac{2}{97}$.

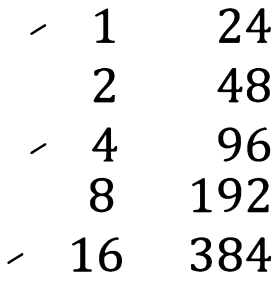


Fig. 2 Doubling table for multiplication 21 × 24

Such was Ahmosé's fraction arithmetic in Thebes, 1550 BC. He fluently calculated with "fractured numbers," but his arithmetic was morphologically different from common fraction arithmetic as a cognitive and mathematical practice.

3 Babylon

When Hammurabi (1792–1750 BCE) elevated Babylon into one of the preeminent city-states of the world, his empire had already been the crossroads of cultures for millennia. In this milieu, the Sumerian base-60 numeration system and the Sumerian-Akkadian sexagesimal system of calculation evolved, probably to cope with the diversity of number bases and metrologies used for centuries by the peoples mingling there.

Hammurabi's scribe had 5 different numeral palettes for expressing measurements, down from 12 palettes used a thousand years before. His five numeral systems pertained to length, area and volume and bricks, liquid capacity, weight, and the counting numbers for sets of discrete objects (Robson 2008, 15). This meant, in context, a numeral referred not only to an amount of a quantity, but it also signified a type of quantity.

In the discrete counting palette of Old Babylon, the numerals in Fig. 3 were used:

Therefore, 7390 would be written as shown in Fig. 4:

This represents 3600 + 3600 + 60 + 60 + 60 + 10. Like the Egyptian system, the discrete numeral system was not a place-value system. While having some base-10 grouping traits, it was the world's only known base-60 system.

The Babylonian scribe also possessed a *sexagesimal place-value system* (**SPVS**). Here, the number 7390 was represented as a *sexagesimal*, shown in Fig. 5, using only ones and tens:

In today's standard notation, this represents $2 \times 60^2 + 3 \times 60^1 + 10 \times 60^0$. However, the scribe had no numeral for zero, nor a sexagesimal point for distinguishing fraction from integer. These missing elements seem like deficits until we realize,

Fig. 3 Palette of numerals representing discrete counts

Fig. 4 Babylonian numeral for 7390

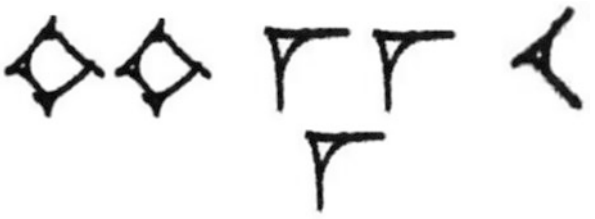


Fig. 5 Sexagesimal for 7390

"*The SPVS … was only a calculational device: it was never used to record measurements or counts*" (Robson 2008, 16). So, while the sexagesimal represented the number 7390, it could also represent $123\,\frac{1}{6}$ (i. e. , $2 \times 60^1 + 3 \times 60^0 + 10 \times 60^{-1}$) and countless other quantities.

SPVS evolved from calculations within the measurements systems of diverse Sumerian metrologies. By using smaller and smaller named and symbolized units in each metrology, they effectively dealt with fractional parts of a unit quantity without pursuing the path of Egyptian scribes. While Babylonians had a phrase for the equi-partitioning of a unit or the n^{th} *part of* a unit and indeed made tables showing the n^{th} parts of numbers like 60 for various *n's*, their scribes did not have a notation or arithmetic resembling our common fractions. Their palette of special symbols for fractional parts seemed to be limited to those most useful for their ordinary needs: $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\frac{1}{6}$, $\frac{1}{10}$, $\frac{2}{3}$, and $\frac{5}{6}$. These fractional parts were only used to estimate remainders or when rounding off a measurement. They were easily converted to sexagesimals when calculations were needed. For example, the *surveyor's rod* ($\approx$ 6 meters) equaled 12 cubits, and each cubit equaled 30 fingers. Thus, the compound measurement *3 rods 5 cubits 23 fingers* might be rounded to *3* $\frac{1}{2}$ *rods* or rounded to *3 rods* 5 $\frac{1}{2}\,\frac{1}{4}$ *cubits*.

By Hammurabi's time, the Babylonians had adjusted their major metrologies to form "*what might fairly be called the world's first metric system*" (Powell 1995, 1956). It was aligned to their base-60 numeration and used the rod-length as a common determinative element similar to the meter in our metric system. For example, the principal area unit, the *sar*, equaled one square rod. Fifty sar was 1 ubu, but, for calculation purposes using SPVS, sixty sar could be called a *big sar*, $\frac{1}{60}$ of a sar was called an *area shekel*, and $\frac{1}{60^2}$ (or $\frac{1}{3600}$) of a sar was called a *little shekel*. So, there are notable parallels to the metric system. Moreover, the use of *shekel* as the ubiquitous 60th of a measure unit, when no other name was standardized, allowed Babylonians to compute fractional quantities in terms of a "*generalized-compound-measurement" arithmetic*, the SPVS. Of common fraction arithmetic, this is not.

Babylonian calculations were performed on the side and not shown on the clay tablets (Robson, p. 78). However, error analysis of the results of sexagesimal calculations strongly suggests the scribes used a base-60 abacus, perhaps just lines on a table or grooves in the sand, along with tokens standing for tens and ones (Høyrup 2002, 193).

Imagine a scribe adding an area *1 ubu 13 sar 21* $\frac{1}{3}$ *shekels* to an area *47 sar 12* $\frac{5}{6}$ *shekels*. For simplicity, imagine the scribe placing Tens (T) tokens and Ones (|)

tokens on a contrived abacus to represent the two areas. See Fig. 6. With the first area, the $\frac{1}{3}$ shekel equals 20 little shekels and the ubu combines with 10 of the sar to make 1 big sar, leaving three Ones in the sar column. Similarly, *49 sar 12* $\frac{5}{6}$ *shekels* is represented in the second row: ($\frac{5}{6}$ *shekels* = *50 little shekels*).

At this stage, the specific metrology and its related numeral system, i.e., area measure, become irrelevant because the conversion ratio of 1 to 60, from column to column, is the same for all metrologies entered on this abacus. Thus, the scribe thinks of the columns simply as *generalized compound measurement units* during the calculation. The scribe combines the tokens in the two rows and simplifies to get the configuration shown in Fig. 7.

Having completed the calculation, the scribe jots on the clay tablet the sexagesimal measure number shown in Fig. 8.

The jottings exactly record the number and type of counting tokens in the abacus result. But the scribe is not finished. The answer should be reported as an absolute area measurement using the numeral palette for area. Therefore, since 50 sar combine to make an ubu, and 2 ubu equal *1* iku, the scribe records on the clay tablet *1iku 2 sar 34* $\frac{1}{6}$ *shekel,* using a roundoff fraction.

Thus, Babylon's sexagesimals indicate no particular absolute numbers, counts, or measurements except when being interpreted by a scribe within a particular context, usually after translation into the notation and metrology of the given problem (Robson 2008, 78). They were the inputs and outputs of formal calculations within a generalized compound-measurement arithmetic, peculiarly adapted to the

(Big Sar)	(Sar)	(Shekel)	(Little Shekel)
I	III	TT I	TT
	TTTT IIIIIIIII	T II	TTTTT

Fig. 6 Ledger suggesting the layout of tens and ones tokens on an abacus

I	TTTTT II	TTT IIII	T

Fig. 7 Ledger showing the result of the addition on the abacus

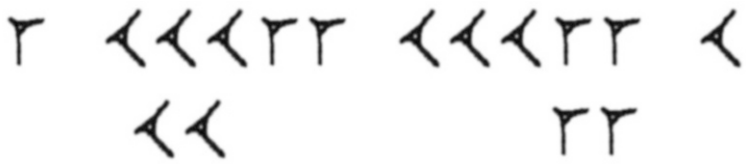


Fig. 8 Sexagesimal representing results on abacus

Babylonian "metric" system. In addition, the SPVS operations and algorithms on fractional parts of units like the sar were neither instances of nor justified by a broader arithmetic of fraction numbers. It was an arithmetic of 60s-progression numerals aided by an abacus. (This contrasts significantly with the decimal system today when it is treated as a "corollary" of the system of common fraction arithmetic due to the stipulation that a decimal is a common fraction.)

The Babylonians viewed sexagesimal division as multiplication by reciprocals and seemed to believe, on geometric grounds, that all sexagesimals had reciprocals, either exact or approximate, making no distinction between rational and irrational numbers. The product of the sexagesimals in a reciprocal pair was usually thought of as 60, which is easily viewed as an appropriately placed 1 in an abacus computation. For example, the reciprocal of 17; 46 40 is 3; 22 30, and the reciprocal of a whole number like 20 is 3, not a fraction $\frac{1}{20}$. From known reciprocals, the Babylonians used an area-based, geometric method to find new reciprocals (Robson 2008, 108). They were aware of rules like *the reciprocal of a product is the product of the reciprocals* since they frequently doubled one sexagesimal in a reciprocal pair while halving the other to produce a new reciprocal pair. The Babylonian SPVS, including its insights into the properties of sexagesimal reciprocals, had a significant indirect influence on the next city of our tale.

4 Alexandria

One thousand years after Ahmosé and Hammurabi, a Greek-centered Hellenistic culture emerged along the Mediterranean seaboard. Over several centuries, academies of learning and scholarship developed in prominent cities, with the greatest of these located in Alexandria, Egypt, at the Mouseion and Library created by Ptolemy 1 Soter (367–287 BCE). Two "fractional arithmetics" became prominent in this cosmopolitan city.

4.1 The DMSS System

The first arithmetic, the Degree-Minute-Second System (DMSS) or *astronomers' numbers*, found its home in the definitive work of the Alexandrian scientist Claudius Ptolemy (100–170 CE). Inspired by the Babylonian SPVS, Ptolemy and his immediate predecessors expressed measures of arcs and lengths in a sixties-based number system of degrees, minutes, seconds, thirds, etc. While we think of *degree* as a unit of angle measure indicating a known amount of rotation, Ptolemy's word for it was "section," and he applied it to circular arcs (i.e., 360 sections in the circumference of a circle) as well as diameters (120 sections in a diameter), both of which were considered "lines" by Greek mathematicians. In a given circle, a degree of arc length is *incommensurable* with a degree of diameter length, something Ptolemy might

have believed but could not have proven. His words for *minute* and *second* are literally translated "first sixty" and "second sixty."

Thus, DMSS is built upon the notion of a generalized unit of line measure. The *section* is a *generalized unit* in that it does not represent any absolute, known amount of length, like cubits. Also, DMSS was not a pure base-60 system since it expressed degrees in the base-10 integer scale. Unlike sexagesimals, it is built around measure numbers that refer to absolute numbers in a generalized unit, while retaining the sense of being compound measurements in known standard units (degrees, minutes, seconds, etc.) when taken on a particular circle or segment. One might call it a *compound-number arithmetic in generalized, denominatized units.* (A pint is a known amount in its own right but is "denominatized" when expressed as an eighth of a gallon.)

Even though Ahmosé's system of fractions was still dominant in Alexandria and had been adapted by the Greeks, Ptolemy and other Alexandrian astronomers preferred to use DMSS for their scientific calculations. He explains, "*In general, we shall use the sexagesimal system for the numerical calculations owing to the inconvenience of having fractional parts, especially in multiplications and divisions, and we shall aim at a continually closer approximation, in such a manner that the difference from the correct figure shall be inappreciable and imperceptible*." (Ivor Thomas 1993, 415). Ptolemy would use Egyptian fractions when rough estimates were sufficient and great precision not needed (Toomer 1984, 7).

As evidenced by Al Uqlidisi's tenth century *Arithmetic* (Saidan 1978), DMSS quickly adapted to Hindu-Arabic place-valued numerals for degrees, minutes, and seconds. Notably influenced by the Greek number theory found in Euclid, Al Uqlidisi viewed the numerical unit to be indivisible, which meant fractions, theoretically speaking, were not *numbers*, a position Ptolemy and all Greek mathematicians up to Diophantus seem to have taken (Gow 1968, 112). Al Uqlidisi wrote: "*All the preceding work* [referring to his own text] *on whole numbers and fractions and all that we have mentioned on parts of numbers refer in fact to this concept, namely, the degree. Number is indivisible, but the degree is divided into parts and parts of parts indefinitely without end*" (Saidan 1978, 84). Thus, the *compound-measurement* standpoint of the Babylonian system was carried forward into the *compound-number* arithmetic of the DMSS system, establishing it as a prototype of the positive real numbers. It finally bequeathed its position to the decimal system in the sixteenth and seventeenth centuries, the era when Egyptian fractions effectively went extinct and the use of common fractions in science started to recede.

It is worthwhile tracing into modern times the Wittgensteinian strand of decimals that entwines with the DMSS threads. Al Uqlidisi's *Arithmetic* is the earliest known Arab work that discusses decimal fractions and uses a notation similar to ours. Decimal arithmetic had evolved in Chinese and Indian cultures centuries before Al Uqlidisi and long before transmission to Europe. However, the efficiency and power of this arithmetic did not seem to be broadly appreciated in Europe until Simon Stevin published his influential pamphlet *De Thiende (The Tenth)* in 1585.

Significantly, Stevin, like Al Uqlidisi, modeled his development of decimals after the DMSS system. Stevin called each tenth part of unity a *Prime*, denoted with a ①,

each tenth part of a Prime a *Second* (denoted with a ②), each tenth part of a Second a *Third* (denoted with a ③), and so on. He would write 27.843 as 27⓪8①4②3③. Thus, in the words of Stevin, decimals consisted "*in characters of ciphers, whereby a certain number is described and by which also all accounts which happen in human affairs are dispatched by whole numbers, without fractions or broken numbers*" (Stevin 1608, First Definition).

Like Ptolemy's comparison of DMSS to Egyptian fraction arithmetic, Stevin lauded the simplicity of decimal computation over common fraction representations and computations. However, he recognized an advantage of common fractions. When reducing a compound measurement to a measure number with a single unit, such as *1 foot 7 inches* to *1* $\frac{7}{12}$ *feet* instead of *7.5833 feet*, common fractions had an advantage. But Stevin viewed this advantage as temporary and predicted sovereigns would eventually adopt measurement systems wherein the units formed progressions of powers of ten, paralleling decimal numeration. The French adopted the metric system a hundred years later.

It bears mentioning that Stevin, contra Euclid, believed number, and hence decimals flowed in a linear *continuum*. But Europe would have to wait for Descartes and Wallis in the seventeenth century before a real number line model was clearly articulated, particularly one that included negative and positive numbers. With his conveyance of the DMSS into the realm of decimal fractions, the symbols in a decimal suggest a conventional compound measurement in generalized, tens-based units analogous to degrees, minutes, and seconds. This conception of decimals is echoed in the research of Wearne and Hiebert (1988) where base-10 blocks, money, and metric measures provide the referents for decimal numerals. Interestingly, their research has shown such compound-measurement thinking benefits the semantic processing of decimal notations by fourth graders, leading to better abstraction of decimals as measure numbers.

4.2 Rational Expression Arithmetic

The second "arithmetic" that emerged in Alexandria, and practiced by Diophantus (200–284 CE), made use of "algebraic" reasoning in an arithmetic of numerical expressions containing one unknown. A close look at Diophantus' works reveals that many of the problems he investigated were likely drawn from subscientific traditions (Høyrup 1989) and shared by a diverse community of mathematicians for whom his work was definitive. Consider Problem 36 of Book IV (Heath 1910, 194; Christianitis 2004, 334–335): *To find three numbers such that the product of any two bears to the sum of those two a given ratio.* Diophantus knew that his trove of find-the-numbers puzzles typically had multiple solutions, but he would only offer one. He represented each problem in terms of one unknown number, which we might call "*x*" or "the unknown," and proceeded to reason algebraically, stipulating restrictions until he reduced the problem to a single equation.

In the above problem, Diophantus starts by stipulating that the product of the first and second numbers is three times their sum, the product of the second and third numbers is four times their sum, and the product of the first and third numbers is five times their sum. He designates the second number as his unknown, and therefore, the first number equals *three of the unknown in part of the unknown less three units*. In modern notation, where the first, second, and third numbers are denoted by *f*, *x,* and *t*, he is saying $f = \frac{3x}{x-3}$, and similarly, $t = \frac{4x}{x-4}$. Therefore, since the product of the first and third is five times their sum, $\frac{3x}{x-3} \times \frac{4x}{x-4} = 5\left(\frac{3x}{x-3} + \frac{4x}{x-4}\right)$.

For most of its history in Alexandria, "*2 in part of 7"* or "*of 2 the 7th part*" connoted not only Ahmosé's "*2 quantities each of partsize* $\frac{1}{7}$*,"* meaning an unsized "heaping," but also an *unclosed* operation of *2 divided by 7*. (Either way, in Alexandria, the closure was $\frac{1}{4}\ \frac{1}{28}$.) Because Diophantus was immersed in an arithmetic involving unknown numbers, closure was not an option. Therefore, he had rules for unclosed operations involving expressions in a single unknown number, analogous to *rational expression* arithmetic taught in high school algebra courses. Accordingly, he applies a rule for multiplying unclosed divisions to get $\frac{3x}{x-3} \times \frac{4x}{x-4} = \frac{12x^2}{x^2+12-7x}$. He then applies another rule to get $\frac{3x}{x-3} + \frac{4x}{x-4} = \frac{7x^2-24x}{x^2+12-7x}$. Thus, x must satisfy $12x^2 = 5(7x^2 - 24x)$, which he solves to get *120 in part of 23*. He is content with this answer, instead of closing it in the Alexandrian manner to get $5\frac{1}{6}\ \frac{1}{46}\ \frac{1}{69}\ \frac{1}{92}\ \frac{1}{276}$, a result so unsightly that it evokes an image of Diophantus concurring with Ptolemy's ghost on the inconvenience of Egyptian fraction divisions.

Was Diophantus participating in an established tradition of common fraction arithmetic? This is a disputed question (Christianitis 2004, 331–335). He clearly thinks the unit is divisible and that Egyptian fractions are numbers. He also thinks that unclosed divisions and unclosed heapings of addends determine unique numbers. However, there are several aspects of his work that increase prospects of the narrative stopping here, and not implying the embrace of a system of common fractions by Diophantus or by the mathematical literati of Alexandria.

First, in his solution to the above problem, Diophantus felt compelled to explain to his mathematical audience his rule for adding two unclosed divisions. Thus, after indicating $\frac{3x}{x-3} + \frac{4x}{x-4} = \frac{7x^2-24x}{x^2+12-7x}$, he explains: $\frac{3x}{x-3} + \frac{4x}{x-4} = \frac{3x(x-4)+4x(x-3)}{(x-3)(x-4)}$. This detour shows that Diophantus possessed a rule for adding rational expressions, and, ipso facto, a rule for adding common fractions. It also suggests that this rule was *not* very familiar to his mathematical audience, or he would not have mentioned it.

Second, Diophantus often expressed *m in part of n* with the shorthand notation m^n, a notation scheme for unclosed divisions that had been in use in Alexandria for several centuries by people who apparently were thinking in terms of Egyptian fractions (Fowler 2004, 370–371). Diophantus usage seems no different in this regard (Fowler 2004, 371fn). In his rare use of the mixed numeral form, he used Egyptian fractions. For example, he would express *6041 in part of 16* as $377\frac{1}{2}\ \frac{1}{16}$.

To summarize, it seems that Diophantus' arithmetic of rational expressions in one unknown was not a generalization of an undocumented practice of common-fraction arithmetic in the Hellenic world. Indeed, the historian David Fowler (2004, 368),

after an extensive review of primary sources, concluded: "*Nowhere do I find any convincing evidence for the proposal that 'the Greeks' used anything like our notations for common fractions and our ways of performing fractional arithmetic."* The earliest documentable prototype of common-fraction arithmetic is found in the Chinese bamboo text *Suàn shù shū* (Cullen 2007) circa 200 BCE. Diophantus evidently grasped the concepts underlying the operational rules of division that would eventually govern common fraction algorithms. However, these algebraic rules used in solving equations may well have been discovered through reasoning about unclosed divisions, invoking properties of reciprocals known and transmitted through subscientific traditions since Old Babylon. So, while Diophantus could operate with m^n expressions as if they were common fractions, it seems to be a pattern in the history of mathematics that concepts are often exploited long before their reference potentials are delimited with names and notations in mathematical definitions, and before they are constrained by systemic properties in rules or axioms (Kitcher 1984, 187–203). Hence, the most we can probably say about this second Alexandrian arithmetic, rational expression arithmetic in a single unknown it is, and common fraction arithmetic it is not.

5 Interweaving the Three Tales with Current Strands

5.1 Measurement Systems and Decimals

The Old Babylonian scribes were very proficient with the arithmetic of measure numbers, including highly precise calculations involving irrational numbers. They did this without the use of anything like common fractions or even Egyptian fraction arithmetic, a system they likely had encountered by Ahmosé's time. *The Babylonian experience suggests that decimal fluency does not depend on common fraction fluency, whether speaking mathematically, psychologically, or pedagogically.* It also suggests how the semantical underpinnings of decimals diverge from those of common fractions. They share cognitive roots in equi-partitioning and unitizing, leading to measure numbers in a single unit. But from there the semantics of decimals, like the sexagesimals of Old Babylon, is situated within that of *ideal measurement systems* and abstracts the sense of a conventional compound measurement as an enumeration according to a decreasing sequence of groupings in a single conversion ratio. Not so for common fractions.

When one thinks of a *system* of measurement like the apothecary pounds, ounces, drams, scruples, and grains, one appreciates that in this system the smaller units are divisors of all the larger units, but one frowns at the plethora of conversion ratios: 1: 12, 1:8, 1:3, and 1:20. More ideal are systems where the conversion ratios in the progression of units are a constant, like 1: 60 in the case of the DMSS system or 1:10 in the metric system. In the conceptualization of its groupings (tenths, hundredths, tens, hundreds, etc.), decimal semantics is thus rooted in an *ideal* measurement system whose powers-of-10 place values resemble units *used in actual metrologies*, where conversions of measurements to equivalent measurements are easily imagined

and comparisons of size easily established. Decimal operations and algorithms can thus be made very transparent as natural extensions of whole number operations and algorithms, but not without subtleties related to the uses of 0 and the decimal point.

Common fractions have a more abstract relation to actual measurement systems and processes. Smaller unit fractions are typically not divisors of larger unit fractions and typically have no correspondence with units used in actual measurement systems. Indeed, unit fractions seem like a counting system of general units where successive units differ by one in the denominator, obscuring not only the notion of equivalence, but also the associated conversion and comparison processes. This is reinforced when, in the effort to build conceptions of common fractions as a system of measure numbers, emphasis is placed on students creating equi-partitions of geometric figures and discrete sets to represent artificial measurement systems inhabitable by arbitrar*y* unit fractions. In these dissections, it is left to the student to "unitize" a component that they can count to create a measure number. To represent operations with their diagrams, students must then use highly variable, compound dissections with visual units merging to form other visual units. The sense of measurement by a decreasing sequence of units might easily be lost in all of this, particularly its approximate nature wherein the count of the final unit in the sequence is an estimate.

5.2 Common Fractions and Algebraic Reasoning

The best algebra of the Bronze Age, a notable algebra at that, was being done in Babylon with sexagesimals and geometrically based identities of area and length, without common fractions. More pertinently, Diophantus' story demonstrates that *algebraic reasoning about rational expressions does not depend on prior common fraction algorithmic competency, or even common fraction representations of fractional measures.* It also suggests that, once the connection is made between common fractions and the division operation in the elementary grades, proficiency in common fraction operations and algorithms can be developed much later as part of operating on real number fractions and rational expressions in a real variable, semantically anchored in the properties of the division operation. Indeed, after listing the usual "Why?" questions people have about the unique features of common fraction algorithms and their justifications, Lortie-Forgues et al. (2015), 208) explain: "*All of these questions have answers, of course, but the answers are not immediately apparent, and they often require understanding algebra, which is generally taught after fractions, so that students lack relevant knowledge at the time when they learn fractions and might never learn how algebra can be used to justify fraction algorithms after they gain the relevant knowledge of algebra*."

It is worth noting that the Dolciani high school algebra textbook series, perhaps the most popular textbook series in the USA between 1963 and 1986, derived the operational rules of rational expressions in real variables using carefully scaffolded lessons, while citing only token examples of common fractions to illustrate (and certainly not to validate) the rules. The derivations were simple and based on

assumptions about multiplicative inverses, i.e., reciprocals, and not common fractions, thereby supporting the notion that Diophantus might have derived his rules for unclosed divisions in a similar fashion (Dolciani et al. 1982).

5.3 The Number Line, Measure Numbers, and Precision

In one sense, as we survey the history of number, measure numbers were the result of successive abstractions from evolving measurement systems, often involving multiple numeral palettes depending on metrology. With traceable threads of meaning, this evolution connects our specie's earliest measurement systems to the very notion of "measure space" encountered in courses on real analysis. The notion of a measure space assumes two things, an underlying set of measurable quantities or magnitudes and a measure function assigning a size (a representation or description in terms of known quantities) to each measurable quantity.

For most historical measurement systems that evolved arithmetized fractional numbers, when a compound measurement was reduced to a single measure number in a single unit, the measure number *sized*, in a manner compatible with the particular historical system, the ratio of the quantity to the known quantity expressed by the unit. Therefore, the measure function connects to the idea of a *ratio scale*. Some measurement devices like rulers, spring scales, and graduated cylinders provide a visual sense of ratio scales that pair measurable quantities to measure numbers. All ratio scales have an origin, a unit, a built-in linear ordering of (quantity, size) pairs, and a gradation system that assumes the ratio of two quantities being measured is proportional to the ratio of their gradation-based numerical measures on the scale.

The historical development of the real number line can thus be characterized as an abstraction of an ideal ratio scale of measurement, *the ideal ruler* so to speak. It was not, at least in Western traditions, a generalization of the system of rational numbers. Indeed, the ancient Greeks discovered that the ratio of the diagonal of a square to the side of the square is inexpressible as a ratio of two counting numbers. This implied that counts of a fixed unit, and its n^{th} parts, were inadequate for the task of expressing the exact sizes of the magnitudes on a continuum with that unit. Simply put, ratios of integers, the precursors of rational numbers, fail to fully capture the essence of measure. Instead, the Eudoxian theory of proportions, commencing in Book V of Euclid, heralds the beginning of Greek measure theory and uses ratios of lines to represent the ratio of magnitudes.

A century after Simon Stevin, measure numbers on a ratio scale, like the DMSS scale, were being abstracted as real numbers in a linear continuum. Like the particle-wave duality of quantum objects, this abstraction supports the notion that a real number is a point on a graduated line and the notion that a real number is the address or "coordinate" of such a point. Perhaps the duality of perspectives can be bridged by simply saying a *real number* is a point on a number line, and an infinite decimal is an *idealized decimal measure number* associated with a point on a *decimal-scaled* number line. Both perspectives are important.

It is easy to get lost in arguments about what a real number is and forget the fundamental raison d'etre of real numbers, i.e., measurement. Long before the Greeks, ancient cultures were aware that physical measurements of continuous quantities were intrinsically approximate. Therefore, an important consideration in differentiating decimals and common fractions when taken as measure numbers is their pliancy in representing all real numbers, in computing-derived (indirect) measure numbers, and in determining the precision of derived measure numbers obtained from calculations.

Long ago, Ptolemy appreciated the infinite pliancy of the DMSS system for the indirect calculation of the chord sizes in his trigonometric table. Similarly, in modern school algebra, where the focus is on reasoning with continuous quantities, functions, and graphs, the Dolciani Modern Algebra Series chose to introduce real numbers using an infinite decimals approach, even stating a completeness axiom stipulating all real numbers have a decimal representation and all decimals represent real numbers. This was not a rational number approach, along the lines Richard Dedekind defined in the nineteenth century (Dedekind 1901) and was pedagogically speaking the better choice. While the notion of infinite decimals at first seems daunting, children encounter infinite decimals by necessity in the earliest grades when they attempt to complete simple long division problems of whole numbers or convert even the simplest common fractions to decimals. Furthermore, science instruction at the precollege level for the most part uses decimals for measure numbers, usually with metric units, and relatively simple "significant figure" rules for roundoff precision, testifying to the efficiency of decimals in deriving measure numbers and in determining the precision of such derived numbers.

In summary, history reveals that the real numbers grew from the needs of measurement, leading to the theories of geometric measure and to the algebra of operations needed for the indirect calculation of measures used in sciences like Ptolemy's astronomy, where exactness demanded better number tools. Although all physical measurements of continuous quantities are approximate and their associated measure numbers usually can be regarded as *rational*, the real number system was not abstracted from a defined system of rational numbers but was rooted in Euclidean measure theory and even more ancient traditions.

6 Collecting the Threads and Tying the Present to Its Recent History

6.1 A Current State of Affairs

In 1989, the National Council of Teachers of Mathematics (NCTM) lent its weight to the fledgling standards movement in the USA. Not afraid to challenge the status quo, the Council published its rather bold *Curriculum and Evaluation Standards for School Mathematics.* Unlike CCSSM, NCTM did not provide specific grade-level performance expectations. It set general goals for broad grade-level bands. The *Standards* did pronounce some controversial value statements in the form of topics

deserving increased or decreased attention. One topic designated for *decreased* attention in grades K-4 was "Paper-and-Pencil fraction computation." In grades 5–8, regarding common fraction operations, NCTM recommended mastery of a small number of basic facts, like $\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$, and that proficiency in operations be limited to fractions "*with simple denominators that can be visualized concretely or pictorially and are apt to occur in real-world settings*" (NCTM 1989, 96). Many state supervisors of school mathematics, who were members of NCTM, used the NCTM Standards as a template for developing the standards in their own states during the 1990s.

In many respects, the CCSSM movement was a reaction to NCTM's 1989 standards, as well as state officials grappling with *No Child Left Behind* calls for grade-specific expectations and assessments. Like the New Math movement, much of the leadership behind CCSSM came from well-intentioned university mathematicians (Høyrup's scientific tradition) concerned about the failure of the elementary curriculum (subscientific tradition) to prepare students for algebra. Indeed, these deficiencies were at the heart of the National Mathematics Advisory Panel (NMAP) Report, which stipulated that all children should successfully complete the equivalent of Algebra II in high school. CCSSM embodies the changes called for in the Panel's Benchmarks (NMAP 2008, 20) by making fluency with common fraction operations and algorithms the most urgent (NMAP 2008, 18) of the three pillars identified by the Panel in the *critical* foundation of algebra.

While recommendations can be inspiring, actual assessments of student achievement are far more sobering. The 2019 National Assessment for Educational Progress (NAEP) revealed that mathematics scores for fourth and eighth graders have essentially flatlined since 2009, the year before CCSSM emerged as the dominant standards in the USA. In NAEP 1978, only 24% of eighth graders tested chose 2 as the whole number closest to $\frac{12}{13}+\frac{7}{8}$ when given the options 1, 2, 19, 21, and "I don't know." Lortie-Forgues et al. (2015), 202) posed this same item to 48 eighth graders enrolled in an algebra course in an affluent, suburban middle school, which had based its curriculum on CCSSM since 2010:

> In 2014, 27% of the 8^{th} graders identified '2' as the best estimate of $\frac{12}{13}+\frac{7}{8}$. Thus, after more than three decades, numerous rounds of education reforms, hundreds if not thousands of research studies on mathematics teaching and learning, and billions of dollars spent to effect educational change, little improvement was evident in students' understanding of fraction arithmetic.

The difficulties children in the USA have in mastering common fractions has been well known for several hundred years. Young adults are faring no better. Kloosterman (2010, 50), reporting on trends in NAEP data for 17-year-olds between 1978 and 2004, notes the percent correct declined from 66.7% to 45.3% on the item stated as "$3 \times 2\frac{1}{3} =?$" The decline typified all the common fraction items he was given access to. Kloosterman concluded: *"This makes sense given that exposure to common fractions outside of school is decreasing, whereas exposure to decimals fractions outside of school is increasing."*

Kloosterman is not the first to notice the decline, if not outright extinction, of common fractions in the world around the child. (A similar phenomenon is being noted by NAEP with regard to certain reading skills in the age of video and cell phones.) With metric tools and digital everything, the objections of the trade guilds to this extinction claim are not compelling. STEM professionals also object, but their counterexamples generally exist in the technical milieu of *real number fractions*, which include both rational and irrational fractions. Such STEM examples are critically important, but translating them into a deferred gratification argument justifying the CCSSM approach has serious problems, one being the "Free" in "Land of the Free" and another being the L.P. Benezet refutation of deferred gratification arguments.

Consider first the "Free." NAEP data is unsettling to some because students in East Asian countries fare much better on common fraction literacy. This was certainly on the minds of the National Mathematics Advisory Panel when it made its recommendations in 2008. However, if history reveals anything, it is that culture, geography, and social traditions matter in the mathematical values and practices of a people. Local control is part and parcel of the history and social fabric of education in the USA. *One might, by analogy, mention the same phenomenon of local control when assessing the outcomes of this nation's response to a pandemic in comparison to East Asian countries.* Indeed, once the CCSSM standards became "national" in 2010, the political blowback was substantial. Ultimately, local control translates into an enlarged role for parents and students in curriculum decision-making, in contrast to highly centralized educational systems. Indeed, according to some experts, parents and students might play the most important role (Klein and M. 1991). NAEP suggests how American students and parents are casting their ballots vis-à-vis the learned curriculum. To most of them, except for the most basic examples like a half and a fourth, common fractions are *hidden numbers*, absolutely of no use, and never encountered outside the classroom.

Though well written and the source of many improvements (e.g., nicely coherent learning progressions and minimal redundancy from grade to grade), CCSSM has not helped the situation by making common fraction computational fluency perhaps *the central imperative* of the K-6 elementary curriculum. In the CCSSM K-6 standards, the words *whole number(s)* and *integer(s)* appear less than 100 times; the words *decimal(s)* and *decimal fraction(s)* appear about 33 times, while (common) *fraction(s)* appears over 180 times! Decimals are defined by common fractions. Consequently, initial decimal experiences of students, instead of being grounded in measurement, presumably involve much work relating decimals back to their fraction definitions, operations, and notations, thereby creating a dependency of decimal understanding on a student's common fraction understanding, which all too often is fragile.

In a logical/structural sense, decimals can be defined by common fractions, and decimals algorithms can be viewed as simple corollaries of common fraction algorithms within CCSSM's well-thought-out common fraction development. But if we believe Morris Kline's (Kline 1974, 100) critique of the New Math of the 1960s, an earlier attempt by the scientific tradition to rigorize elementary

arithmetic, such structure "*cannot be significant (to the child) at this stage*." If anything, the history of mathematics surveyed above suggests the claims of the National Mathematics Panel about the critical pillars of school algebra need qualification. Furthermore, research by Hiebert, Wearne, and others clearly articulates the deep complexity of developing decimal competency, regardless of any work on common fractions.

This leads to "The Story of an Experiment" (Benezet 1935), an article which Harvard mathematics professor Dr. Andrew Gleason found so interesting that he shared it with many, including this author. Gleason is mentioned because he was a mathematics editorial adviser of the very popular Dolciani Modern Algebra Series previously mentioned and because he directed the Cambridge Conference of research mathematicians who met in 1963 to produce a set of recommendations, which will be discussed later, titled *Goals for School Mathematics*.

In 1924, Benezet was the superintendent in Manchester, New Hampshire, facing a crisis wherein 20% of the first graders were not promoted to second grade because they had failed arithmetic. *This was particularly pronounced in minority schools with immigrant families who often spoke limited English.* After many informal assessments in the district's schools, in which he asked children quantitative reasoning questions often dealing with fractions and rates, the embarrassing results of traditional instruction convinced half of the schools in the district, particularly the minority schools, to eliminate formal arithmetic until sixth or seventh grade (no use of primers, algorithms, etc.). They were, instead, to focus on relevant quantitative reasoning tasks, mediated by *math talks,* in all subjects taught. Benezet (Benezet 1935, 2410) wrote: *"... we waste much time in the elementary schools, wrestling with stuff that ought to be omitted or postponed until the children are in need of studying it...The whole subject of arithmetic could be postponed until the seventh year of school and it could (then) be mastered in two years by any normal child."*

After pre- and post-testing the children from the four experimental schools that resumed the use of primers in the sixth grade, the results were compelling. Upon entering sixth grade, the experimental children, who had yet to learn standard operation algorithms for whole numbers, fractions, and decimals, were far behind the traditional classes. By the end of the year, on the arithmetic skills expected of sixth graders, which included common fraction proficiency, the experimental students performed just as well as the traditionally trained children and were more fluent and confident in their thinking about numbers and operations.

Perhaps Benezet was a bit underhanded using fraction and rate problems in his demonstrations that something was amiss in his Manchester School District's elementary mathematics curriculum. But he at least deserves credit for making one question the traditions of that curriculum regarding common fractions and the demand for early fluency in fraction algorithms, *particularly the recurring negative effects of this curriculum on certain populations within school districts.* For that reason, several cursory remarks must be made about the history of the American precollege curriculum.

6.2 Observations from Recent History

At the end of the nineteenth century, according to Snyder (1993), the typical school district in the USA had 144 days in its school year. The typical student attended class 99 of those days and nearly completed eighth grade, with great disparities due to race and gender. (The typical black male completed fifth grade while the typical white male completed eighth grade.) Between 1800 and 1900, these numbers were much smaller, including length of school year and attendance levels. In 1900, few Americans completed a course in algebra or geometry and only about 7% ever graduated from high school. Algebra was considered utterly abstract and intended as a mental discipline for the college-bound, which was very few students.

To say the least, during the nineteenth century and early twentieth century, arithmetic was the terminal mathematics course for the vast majority of students, and it was designed to be just that: "to prepare the young for the common avocations in life" (Greenleaf 1845, Preface). It is worthwhile exploring samples of arithmetic texts written in English from the seventeenth to the twentieth century to gain a better appreciation of this goal. (See *Resourceaholic* in References.) As their terminal mathematics course, children encountered substantial work with compound numbers, fractions, and decimals. The diversity of measurement units in common use required fractions when expressing a compound number in terms of its largest unit. *Importantly, common fraction fluency was not intended nor generally thought of as an essential foundation for algebra.*

In the twentieth century, high school populations gradually expanded, especially during the Great Depression. By the mid-twentieth century, the surge of enrollments in colleges, driven by the G.I Bill, and the calls for universal high school graduation led to the reevaluation of the high school curriculum as a preparation for college. Mathematicians of prestigious universities became involved in the goals and redesign of secondary school mathematics as a preparation for college. Such efforts as the *University of Illinois Committee on School Mathematics*, founded in 1951, and the 1958 AMS *School Mathematics Study Group* began to produce sample high school text materials, eventually morphing into a K-12 reformation known as the *New Math* movement, which barely lasted into the 1970s at the K-8 level before being driven "Back to Basics."

A report produced at the time, the Goals for School Mathematics (1963), was distributed among mathematicians and mathematics educators at universities, across the country, but received little acclaim in part because its stated purpose was to be futuristic, which it definitely was. However, the report is pertinent. Dr. William Martin of MIT and Dr. Andrew Gleason of Harvard cochaired the four-week conference. Most participants were well-known mathematics researchers from prestigious American institutions. The purpose of the conference was to produce a coherent and radical set of long-term goals for K-12 mathematics.

Section 5 of the report is devoted to the K-6 mathematics curriculum. The first sentence of the section captures the spirit of their recommendations. "*As we have indicated above, the objective for mathematics instruction in the elementary grades*

[K-6] *is familiarity with the real number system and the main ideas of geometry.*" The report goes on to emphasize the following:

- Early use of the real number line in the context of much measurement activity.
- Reasoning with inequalities and "considerable" experience with approximations, including the effects of "round off" and significant figures.
- Fractions with small denominators to name additional points on the number line.
- "explicit study of the decimal system of notation including comparisons with other bases"
- Study of decimals for rational and irrational numbers.
- Early focus on mathematical properties of real number system with algorithms coming "considerably later in the curriculum".
- Study of geometric transformations, symmetries, and similarity of figures.

What is very clear in the report is that the real numbers, the properties of their operations, the number line, measurement, decimals, and the geometry of transformations and similarity are dominant in the vision of the Cambridge Conference for the elementary curriculum. Common fractions can help to populate the number line but seem otherwise not to be emphasized. Fluency with common fraction operations and algorithms is clearly not a "critical" priority, and possibly not even a goal of these K-6 recommendations of professional research mathematicians.

The back-to-basics movement in the 1970s pushed common fraction fluency back into the forefront of the elementary curriculum, until the ascendancy of the NCTM 1989 Standards and various systemic reform efforts of the 1990s seemed to stall the trend. But the backlash to the NCTM Standards and the related curricula spawned by systemic reform efforts was severe. As already mentioned, the CCSSM doubled down and firmly embedded fluency with common fractions and their standard algorithms as perhaps the central imperative of the K-6 curriculum. Along with the National Mathematics Panel pronouncement about the critical foundation of algebra, CCSSM lent weight to the sense that the tradition of early common fraction fluency was a *mathematical imperative*, a status it had never previously had in American education. While lifting the burden of very large denominators, the CCSSM common fraction standards lock in, among several newer practices, the traditional computational practices of the decades before the "New Math." So, while many in the country claimed the CCSSM standards were enforcing a national curriculum and stepping on states' rights and local control, the CCSSM standards, except for nominal changes in branding, remain the standards in nearly all states. One reason, not surprisingly, may be that the CCSSM demands for fraction and decimal computational fluency using standard algorithms have strong appeal to many adults based on long-standing American traditions.

Nonetheless, states are hard-pressed to enforce standards when faced with the low performance of their students on recognized exams like NAEP and the lack of improvement, if not the outright decline in the performance of a large portion of their students on topics emphasized in their standards, like common fractions. Local control is exercised at many levels, and state intrusions into the affairs of school

districts have their limits. Even with state backing, CCSSM advocates are now, after a decade of implementation, facing growing concerns by parents and local educators. These CCSSM advocates are resorting to calls for patience (Goldstein 2019). Though painted with sweeping strokes, this story of epicyclic motion within the thinking about common fraction fluency in American elementary education is salient.

6.3 Other Voices of Note

In his essays on *Measure of Magnitudes* published in a series between 1931 and 1935, Henri Lebesgue (1966) introduced a construction of the real numbers as infinite-decimal measures of segments on a number line. He sought to illuminate a pathway for the precollege curriculum to develop the real numbers while passing over the study of common fractions.

> "But would we still speak of fractions at the primary level or in the first two grades of the secondary level? No, since this is not essential for the theory and serves no practical purpose. I believe that people will agree with me when I say that performing operations on twenty-seconds and thirty-sevenths is a martyrdom that we inflict on twelve-year-olds out of pure sadism without any justification based on usefulness as an extenuating circumstance. Yes, I know that by searching hard, one can find certain "applications" of fractions, that certain mechanics make calculations with their fractions in cutting threads on screws. . . . ."

> ". . .The reform would be effective if we agreed to have the children no longer study two kinds of numbering, namely, numbering in terms of n^{ths} for commensurable numbers and the decimal numbering system, that is if we permitted them to answer 0.428 where the answer is 3/7."

> *"True, a divided by b or a/b is still read "a b^{ths}" when a and b are integers, but this way of speaking no longer compels us to develop the entire theory of fractions any more than the word quarte-vingt-douze [French word for 92, literally "four-twenty (plus) twelve"] compels us to study the numbering system with base twenty."* (Lebesgue 1966, 31–32)

Lebesgue's views impacted the Cambridge Conference thinking. They also influenced Russian psychologist V. V. Davydov and his colleagues as they developed an elementary curriculum based on measurement and the real numbers. The Davydov curriculum provides an existence proof that the elementary curriculum can embody the core recommendations regarding real numbers and fractions found in the Cambridge Conference Report and build real number arithmetic from the fundamental notions of measurement. One of the overarching principles Davydov (1975, 113) drew from Lebesgue is stated as follows:

> *But what this means in terms of curriculum design is no less than an end to the arithmetic of fractions as it is interpreted in the school. The shift from whole numbers to real numbers is a shift from arithmetic to "algebra," to laying the foundation for analysis.*

The first half of Davydov's first-grade curriculum starts with comparing, composing, and decomposing continuous quantities of the same kind, without numerical

associations, while identifying measurable traits like volume and area and mass, and representing their actions on quantities symbolically and with lines. Children thereby develop algebraic relations of equality and inequality between composed and decomposed *prenumeric* quantities represented in symbols and on a linear scale, akin to the properties of general magnitudes investigated by Euclid (e.g., if a, b, and c are magnitudes of the same kind, $a > b$ implies $a + c > b + c$).

Only near the end of first grade are numbers introduced as multitudes of units by using tokens or tally marks to communicate the amount of a unit in a measurement, based on a 1–1 comparison. Counting numbers are then introduced as a sequence of names for those multitudes of units. As with the prenumeric quantities, the focus of work with numeric quantities is to develop algebraic relations of equality and inequality between composed and decomposed quantities, but this time akin to the properties of multitudes of general magnitudes investigated by Euclid. (E.g., if m is a number, i.e., "multitude of unity," then $m(a + b) = ma + mb$.) On this foundation in the first grade, Davydov proceeds to expand his arithmetic of measure numbers, as he said, in a shift toward algebra and analysis.

Davydov's curriculum is sometimes classified as "early algebra" because of its early use of algebraic symbolism to express general patterns encountered by students in attempting to solve problems of measurement and quantity. The teacher's role in this is not a case of handing students fully baked a priori abstractions, but more akin to providing words and symbols for patterns children have exploited in their problem-solving activities as well as guiding them in stipulating the general scientific principles governing those patterns. In this regard, Davydov's Vygotsky-based social constructivism seems to be following the historical roots of mathematical learning as described in Kitcher's constructivist epistemology.

Informative examinations of the philosophical/pedagogical details, the mathematical foundations, and the successful classroom implementations of Davydov's curriculum can be found, respectively, in Schmittau and Morris (2004), Bass (2019), and Dougherty (2008). Dougherty's words deserve mention here. Summarizing the *Measure Up* (MU) Project, a longitudinal experiment immersing elementary children, regardless of prior achievement or "ability level," in a modified Davydov curriculum, she concludes:

> *An approach to elementary mathematics that focuses on non-specified, generalized quantities is often thought to be too abstract and thus, not accessible by young children. MU preliminary results, however, support Davydov's claims that it offers young children a meaningful foundation on which to build sophisticated and complex mathematics. Understanding the structure and properties of mathematics creates a way for children to construct solid underpinnings that lead to substantive mathematics. . .It builds confidence so that even within non-routine or unfamiliar situations, children can reason through the relationships in the problem.* (p. 411)

Davydov's approach is but one possible development of elementary mathematics on the basis of measurement and the real numbers without passing first through the gauntlet of common fraction operational fluency. Significant approaches to the elementary mathematics curriculum centrally founded on measurement, and not on

common fraction fluency, have already been developed and studied in this country, such as *Developing Mathematical Processes* (Romberg 1977). It seems reasonable that similar approaches to real number numeracy will prove more than viable if given a chance for development.

7 Conclusion

7.1 Suggestions for Consideration

The historical strands surveyed above underline the significant cultural variations in the fundamental concept of *measure number*. Indeed, successive abstractions employed in the human effort of expanding number to include the "continuous number domain," i.e., the real numbers, led Richard Dedekind to philosophize: *Numbers are free creations of the human mind; they serve as a means of apprehending more easily and more sharply the difference of things.* (Dedekind 1901, Preface to First Edition). This general portrayal of number is quite appealing as a summary of number's social history.

Since the time of Plato and Euclid, the final stages of a mathematical creation in the scientific tradition have been those of stipulation and systematization, with definitions and axioms. In the context of education, Poincaré (1952, 116) cautions as follows:

What is a good definition? For the philosopher or the scientist, it is a definition which applies to all the objects to be defined, and applies only to them; it is that which satisfies the rules of logic. But in education, it is not that; it is one that can be understood by the pupils.

So, if we are to believe Poincare and Dedekind, a major goal of the elementary curriculum, per the Cambridge Conference Report, is to guide students, brief stipulation by brief stipulation rooted in experimental facts, to a good definition of the continuous number domain. The task of pedagogy is to assure this is done with understanding and in a way that allows students to recognize that this free creation truly helps them in "apprehending more easily and more sharply the difference of things" *in their world*. To the child first developing understanding of magnitude, number, and measurement relationships, the *things* are in their physical environment, not simply the "manipulatives" used in school, and in the world of quantities and comparisons invoked by their parents and by their social groups.

In Kitcher's philosophy, the mathematical reality of the "ideal human subject" is clear: It resides in the "ings" of mathematical practice, ontologically and epistemologically prior to the "ions"; e.g., abstract mathematical notions like *collection* are derived from idealizations of mathematical activities like *collecting* (Kitcher 1984, 110). Constructivist philosophies in education are consistent with this view: The protomathematical learning of children is located in reflective operating on their world. To engage reflectively, their actions must be purposeful. By way of an analogy whereby numbers are paints on a palette, painting a picture of an actual object or scene chosen by the child with *a palette of colors actually visible in the*

setting is usually more purposeful than a painting-by-the-numbers of an artificial, random scene handed to them and using colors that are never seen. Curriculum standards must pay attention to this.

For years, educators involved with the learning issues surrounding minority students have called for curriculum practices that reflect relevant community, cultural, and social purposes in the learning activities of these students (e.g., Educating American Indian/Alaska Native Elementary and Secondary Students 1995). But such practices are needed by all students if they are to apprehend more easily and more sharply the difference of things actually in their world. Hence, given the near extinction of common fractions in the environment of the child, there needs to be a compromise between the thinking of the 1989 NCTM Standards and the 2010 CCSSM Standards. Here are some suggestions that hopefully can provoke discussions and lead to a better interface, with fewer epicycles, between the scientific mathematical tradition and the mathematical practices of elementary education in the USA.

1. (*Echoes of Ptolemy*) The real numbers are an appropriate milieu for all of school mathematics, starting in kindergarten. The decimal system, with its cognitive roots in the denominate thinking of ideal measurement systems and in the geometry of a number line, provides the appropriate numerical idiom. Properties of decimal operations can be developed by means of measurement activities and geometric definitions vis-à-vis the number line. Decimal algorithms can be developed from the very concrete compound measurement operations from which they arose historically without appeal to common fraction algorithms.
2. (*Echoes of Babylon*) Elementary school, through Grade 8, is the appropriate place to develop fluency in decimal operations and algorithms. Today's children live in a world of decimal numbers. Simple common fraction literacy, *without* the goal of fluency in standard fraction algorithms, is still a relevant backdrop at this level. However, with the place-value system, calculators, the metric system, and other decimal referents, such as percentages, commonplace in their environment, the potential is considerable for providing children with semantical support of decimals as measure numbers on a real number line.
3. (*Echoes of Diophantus*) Algebraic thinking can be developed throughout elementary school, particularly with the operational properties of real numbers; but the algebra of grades 8 through 10 is a proper setting for developing fluency in common fraction algorithms as a corollary of the broader development of real numbers and simple rational expressions. All young adults need this foundation for a world in which so many career options depend on mathematical literacy. Moreover, based on the recommended certifications practices at this level, teachers would be prepared to help students navigate the algebraic reasoning required.
4. (*Echoes of Benezet*) State standards must not preempt the advantage *responsible local control* can bring to education in America. Year by year, benchmarked proficiencies can entrench a well-intentioned but not ideal curriculum for generations, and stifle the resourcefulness exhibited by Superintendent L.P. Benezet in thinking outside the box to address serious learning challenges in local

environments. Regarding fractions, educational research (Moss and Case 1999) has found compelling evidence that the most effective progression for developing understanding of fractional measures is percentages – decimals – and then common fractions. Standards that are effectively national, like CCSSM, should not stifle such longitudinal approaches to curriculum experimentation.

7.2 The Lesson of Thebes

One of the important lessons of the tale of Thebes is the rather notable fact that the Egyptian fraction system survived untarnished for *many thousands* of years in even non-Egyptian cultures, well into the age of common fractions. It is a warning that the common-fractions-based, elementary curriculum, with its deep roots in American tradition, can outlive its usefulness for centuries, long after CCSSM standards are forgotten and regardless of the toll taken on large segments of the USA population. Looked at historically, the standards movement in the USA seems superficial when faced with the challenge of perturbing the deeper equilibrium of traditional practices. Even so, it is never too late to rethink what is being done. As Benezet discovered in his school district, *unnecessary* hurdles, like the demands of early computational fluency with common fractions, can only exacerbate inequities in the elementary education system and ultimately lead to very poor educational outcomes overall. This appears to be the case today. An agenda that prioritizes decimal number fluency and early algebraic reasoning, while delaying common fraction fluency, is a compromise that seems possible in spite of the stubbornness of traditional practices.

References

Allen JP (2014) Middle Egyptian: an introduction to the language and culture of hieroglyphs, 3rd edn. Cambridge University Press, Cambridge UK

Bass H (2019) Is the real number line something to be built, or occupied? In: Weigand HG (ed) The legacy of Felix Klein, ICME-13 monographs, p 67–77. https://link.springer.com/chapter/10.1007%2F978-3-319-99386-7_5. Accessed 10 June 2021

Benezet LP (1935–1936) The story of an experiment (in three parts). J National Edu Assoc 24, no. 8: 241–243; 24, no. 9: 301–03; 25, no. 1: 241–43

Chace AB (1986) The Rhind mathematical papyrus. NCTM, Reston. The Rhind Papyrus Volume 1 (wikimedia.org). Accessed 10 June, 2021

Christianitis J (ed) (2004) Classics in the history of Greek mathematics. Kluwer, Dordrecht

Cullen C (2007) The Suàn shù shu, "writings on reckoning": rewriting the history of early Chinese mathematics in the light of an excavated manuscript. Hist Math 34:10–44

Davydov V (1975) The origin of concepts and its importance in structuring the school subject. In: Steffie L (ed) Soviet studies in the psychology of learning and teaching mathematics, vol 7. University of Chicago, Chicago, pp p110–p122

Dedekind R (1901) The nature and meaning of number. In: Beman WW (trans) Essays on the theory of numbers. Open Court, Chicago, p14–58

Dolciani M, Sorgenfrey R, Brown R, Kane R (1982) Algebra and trigonometry structure and method, book 2, 2nd edn. Houghton Mifflin, Boston

Dougherty B (2008) Measure up: a quantitative view of early algebra. In: Kaput J, Carraher D, Blanton M (eds) Algebra in the early grades. Lawrence Erlbaum, New York, pp 389–412
Educating American Indian/Alaskan native Elementary and Secondary Students (1995) American Indian Science and Engineering Society. Available via ERIC. https://files.eric.ed.gov/fulltext/ED385404.pdf. Accessed 10 June 2021
Fowler D (2004) Logistic and fractions in early Greek mathematics: a new interpretation. In: Christianitis J (ed) Classics in the history of Greek mathematics. Kluwer, Dordrecht, pp 367–380
Goals for School Mathematics: The Report of the Conference on School Mathematics (1963) Educational Services Inc. Cambridge. https://www.maa.org/sites/default/files/pdf/CUPM/first_40years/1963CambConf.pdf. Accessed 10 June, 2021
Goldstein D (2019) After 10 years of hopes and setbacks, what happened to the common core? In: New York Times Dec. 6 edition
Gow J (1968) A short history of Greek mathematics. Chelsea Publishing, New York
Greenleaf B (1845) Introduction to the National Arithmetic. Robert S. Davis Publisher, Boston. https://www.resourceaholic.com/p/digitised-antique-maths-textbooks.html. Accessed 10 June, 2021
Heath TL (1910) Diophantus of Alexandria: a study in the history of Greek algebra. University Press, Cambridge UK. https://archive.org/details/diophantusofalex00heatiala/page/130/mode/2up. Accessed 10 June, 2021
Høyrup J (1989) Sub-scientific mathematics: observations on a pre-modern phenomenon. Hist Sci 28:63–86
Høyrup J (2002) A note on old Babylonian computational techniques. Hist Math 29:193–198
Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press, Oxford
Klein F, M. (1991) The politics of curriculum decision-making. In: Klein FM (ed) A conceptual framework for curriculum decision-making. SUNY Press, Albany, pp 24–41
Kline M (1974) Why Johnny Can't add: the failure of the new math. Vintage Books, New York
Kloosterman P (2010) Mathematics skills of 17-year-olds in the United States: 1978 to 2004. J Res Math Educ 41:20–51
Lebesgue H (1966) Measure and the integral, May K (ed and trans). Holden-Day, San Francisco
Lortie-Forgues H, Tian JJ, Siegler RS (2015) Why is learning fraction and decimal arithmetic so difficult? Dev Rev 38:201–221
Moss J, Case R (1999) Developing children's understanding of the rational numbers: a new model and an experimental curriculum. J Res Math Educ 30:122–114
National Council of Teachers of Mathematics (NCTM) (1989) Curriculum and evaluation standards for school mathematics. NCTM, Reston
National Governors Association Center for Best Practices and Council of Chief State School Officers (2010) Common Core state standards for mathematics. Common Core state standards (college- and career-readiness standards and K–12 standards in English language arts and math). NGA Center and CCSSO, Washington DC. http://www.corestandards.org. Accessed 10 June, 2021
National Mathematics Advisory Panel (2008) The final report of the National Mathematics Advisory Panel. Department of Education, Washington DC
Poincaré H (1952) Science and method. Dover Publications, New York
Powell MA (1995) Metrology and mathematics in ancient Mesopotamia. In: Sasson JM (ed) Civilizations of the ancient near east, Vol. III. Charles Scribner's Sons, New York, pp 1941–1957
Resourceaholic website.: https://www.resourceaholic.com/p/digitised-antique-maths-textbooks.html. Accessed June 10
Robson E (2008) Mathematics in ancient Iraq: a social history. Princeton University Press, Princeton
Romberg TA (1977) Developing mathematical processes: the elementary mathematics program for individually guided education. In: Klausmeier HJ, Rossmiller RA, Saily M (eds) Individually guided elementary education: concepts and practices. Academic Press, New York
Saidan AS (1978) The arithmetic of Al-Uqlîdisî. D. Reidel Publishing, Dordrecht

Schmittau J, Morris A (2004) The development of algebra in the elementary mathematics curriculum of V. V. Davydov. Math Educ 8(1):60–87

Snyder TD (1993) 120 years of American education: a statistical portrait. National Center for Education Statistics, Washington DC

Stevin S (1608) The art of tenths, or decimal Arithmetike. Robert Norton, London. https://adcs.home.xs4all.nl/stevin/telconst/10ths.html. Accessed 10 June 2021

Thomas I (1993) Greek mathematical works, volume II. Harvard University Press, Cambridge

Toomer, GJ (1984) Ptolemy's almagest. Gerald Duckworth & Co., London. https://archive.org/details/PtolemysAlmagestPtolemyClaudiusToomerG.5114_201810 Accessed 10 June 2021

Wearne D, Hiebert J (1988) Constructing and using meaning for mathematical symbols: the Case of decimal fractions. In: Hiebert J, Behr M (eds) Number concepts and operations in the middle grade. NCTM, Reston, pp p220–p235

Wittgenstein L (1958) Philosophical investigations. Basil Blackwell Ltd, Oxford

Sir Isaac Newton and *Opticks*

Jonah Britton and Bharath Sriraman

Contents

1 Sir Isaac Newton 1180
2 *Opticks*: Book I 1181
3 *Opticks*: Books II and III 1185
4 Conclusion 1186
References 1187

Abstract

Isaac Newton (1643–1727) is a polarizing figure in the history of mathematics (and science) because of the inordinate attention given to priority disputes with Hooke and Leibniz in his numerous biographies. Nevertheless Newton's *Opticks* endures and is remarkable for its adherence to the axiomatic style of classical Greek texts like the Elements. In this chapter, we examine influences in Newton's ensuing mathematical practices, and present examples of Newton's work which illustrate his genius and his place at the helm of the pantheon of natural philosophers. In particular, we examine Newton's use of arithmetic, his work on color, and his theory of light.

Keywords

Isaac Newton · *Opticks* · Optics · Light · Color · Series · Infinite series · Ratios · Proportions · Physics · Particle light · Corpuscular light · Rainbows

J. Britton (✉)
University of Connecticut Storrs, Storrs, CT, USA
e-mail: jonah.britton@umconnect.umt.edu; jonah.britton@umontana.edu; jonah.britton@uconn.edu

B. Sriraman
University of Montana, Missoula, MT, USA
e-mail: sriramanb@mso.umt.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_145

1 Sir Isaac Newton

Isaac Newton needs no introduction to the community of science. His name is ubiquitous with the force of gravity itself. There are scholars who specialize in Newton studies akin to those that do Leibniz studies, and numerous biographies that examine Newton's life and career. His work in mathematics, astronomy, and physics have provided fodder for centuries of scientific investigation culminating in advances in modern physics, construction of astronomical devices like telescopes, and last but not least initiating the modern view of the calculus of variations. This merely scratches the surface of the iceberg representing his entire body of work. No chapter can do justice to the scope and brilliance of Newton. Our attempt here is to delve into some of Newton's mathematical practices that systematically furthered the works of the ancient Greeks and ensuing generations of natural philosophers.

Sir Isaac Newton was born in Woolsthorpe, England, on December 25, 1642. He grew up separate from his mother (his father dying before his birth) and instead grew up with his grandparents after his mother remarried Westfall (2023). Newton was a reserved child, prone to potentially violent outbursts. He spent his time reading and learning, rather than being outside or with friends. After his grandparents passed in 1653, his mother insisted Newton move back in with her and eventually take over the farm and the estate. However, Newton wanted nothing to do with farm affairs, often reading on the job rather than watching the animals. His uncle saw that Newton not only preferred academia, but was incredibly good at it and convinced his mother to let him be enrolled at Trinity college in 1661 when he was nineteen Davidson (2015), Westfall (2023). After graduating, Newton was offered a professorship at the college teaching optics. Yet, Newton would not publish anything on optics for several years. Instead, during this time he developed his "method of fluxions" which would then become calculus Davidson (2015). However, Newton's fear of criticism would stop him from publishing his findings until Gottfried Leibniz published his own work independent of Newton. He accused Leibniz of plagiarizing his work, claiming himself as the sole inventor of calculus Davidson (2015), Westfall (2023). Newton's first published work came in 1672 when he presented some of his work light and color to the Royal Society. This began to form the basis for his optical treatise, *Opticks* Davidson (2015), which he published in 1702. Newton's *Philosophiae Naturalis Principia Mathematica* (often shortened to just *Principia*) was published in 1687. Newton's methods of calculus was hinted at in *Principia*, however, it would not be officially published until it was appended to a verion of *Opticks* in 1704 Westfall (2023).

Newton's advancements in science and mathematics were not constructed out of thin air. Newton build his developments from the works of those who came before him. Notably, he used the theory of proportions from Euclid's *Elements* to prove many of his works. These theories are thought to originate from Eudoxus and presented as a part of the *Elements* Joyce (1996). Euclid explains and uses proportions in the *Elements* to compare quantities of the same type through ratios. Magnitudes can only be proportional with other magnitudes. Likewise, areas can only be in proportion to areas, volumes to volumes, etc. Euclid expands on these

concepts too by introducing continued proportions to compare multiple quantities at once through a long ratio. Newton uses these concepts as a basis for his discoveries in both his *Principia* Wawrzycki (2021) and *Opticks*.

Newton's mathematical background goes beyond that of the Greeks. Newton had explored and read the works of many scientists and mathematicians as he studied and performed his experiments. This includes works written by Arab scientists, such as Alhazen who was one of the first people to work on telescopes Marshall (1950). Newton's intimate knowledge of various works allowed him to build his own theories which he proved (or at least attempted to prove) in his works.

Newton's philosophical background also allowed him reject popular philosophies at the time, including Aristotelian doctrine Albert (1997). Newton went against the popular Cartesian doctrine which emphasized hypoetheses, and instead argued for the use of analysis in his works, opting to prove his theories through experiments and observations rather than thought experiments Albert (1997). Newton wanted to support his science through mathematical arguments and observations. His adaptation of the "way of analysis" coupled with his mathematical background from the likes of the Greeks and the Arabs culminated in *Opticks* which became a stepping stone away from traditional natural philosophy and toward more modern scientific and mathematical research methods Albert (1997), Evans (1955).

2 *Opticks*: Book I

Opticks was originally published in 1704 and was published four times, with the fourth edition being published posthumously Davidson (2015). In it, Newton explained color, reflection, refraction, and diffraction. He did all this with what was at the time a new theory of light: he treated light as corpuscular (light is made of particles) Kumar (2006). The particle nature of light led to heavy use of geometry in his work. In fact, *Opticks* mimicked the set up and style of Euclid's *Elements*. The work begins with a series of axioms, followed by a series of definitions, and finally the propositions. Furthermore, *Opticks* is broken into three books, each with their own themes. The first book sets up the building blocks of optics and explains how light behaves in reflection and refraction.

Opticks is full of many interesting and useful propositions. For the scope of this chapter, however, only a few propositions will be examined both for their mathematical content and for their historical and modern relevance. Proposition VI, Theorem V in Part I of Book I states: "The Sine of Incidence of every Ray considered apart, is to its Sine of Refraction in a given Ratio" Newton (1730). This is a precursor to Snell's Law which states that $n_i \sin(\theta_i) = n_f \sin(\theta_f)$ where n_i is the index of refraction of the initial medium, n_f is the index of refraction of the second medium, θ_i is the angle of incidence of the initial ray, and θ_f is the angle of incidence of the final ray. Newton does not use the index of refraction, a property of the medium itself, in *Opticks* and thus could not derive Snell's law in its modern form directly using this method. However, Newton lays the groundwork by observing that the sines of the two angles form the same ratio when diffracted through the same pairs of

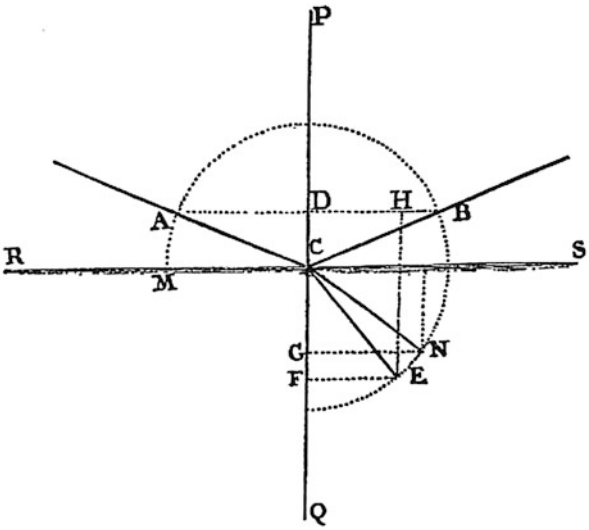


Fig. 1 The geometry behind Axiom V Newton (1730)

media. Newton uses his fifth axiom, "The Sine of Incidence is either accurately or very nearly in a given Ratio to the Sine of Refraction" Newton (1730) along with proportion and ratio properties from the *Elements* to prove this proposition.

Figure 1 is the diagram Newton uses to explain his fifth axiom and to prove this proposition. For this proposition, he lets AC and MC be two rays which are to be refracted through a medium along QP. Next he "let[s] MC, AD, be the Sines of Incidence of the two Rays, and NG, EF, their Sines of Refraction" Newton (1730). Then he breaks the "motion" or path of the ray up into a perpendicular part, AD, and a parallel part, DC. He breaks the motion of the refracted rays into parallel and perpendicular components as well: "whereof the perpendicular ones are MC/NG × CG and AD/EF × CF" Newton (1730). According to his previous proposition, since the perpendicular component of CN is $\frac{MC}{NG} \times CG$, then the perpendicular component of CE, $\frac{AD}{EF} \times CF = \sqrt{CDq + \left(\frac{MCq}{NGq} \times CGq\right)}$ for some q. Then, squaring both sides, adding ADq and $MCq - CDq$ to both sides, and dividing by $CFq + EFq$ and $CGq + NGq$ yields $\frac{MCq}{NGq} = \frac{ADq}{EFq}$. But AD is the sin of incidence and EF is the sin of refraction for the incident ray AC. So the sin of incidence is in a ratio with the sin of refraction. Newton uses this result in his next proposition: "The Perfection of Telescopes is impeded by the different Refrangibility of the Rays of Light" Newton (1730) where "refrangibility" is how much the light ray is refracted upon entering a new medium.

In part II of Book I, Newton begins to explore the physical differences of color. Newton explained that color, while relegated to light, was more of a mental construction Newton (1730). He observed the white light could be split into a rainbow, starting at red and moving through to violet. However, not all colors seemed to appear in this rainbow. Magenta was created by overlapping red light and violet light. It was not a "true" color, but rather one constructed in the mind as a result of

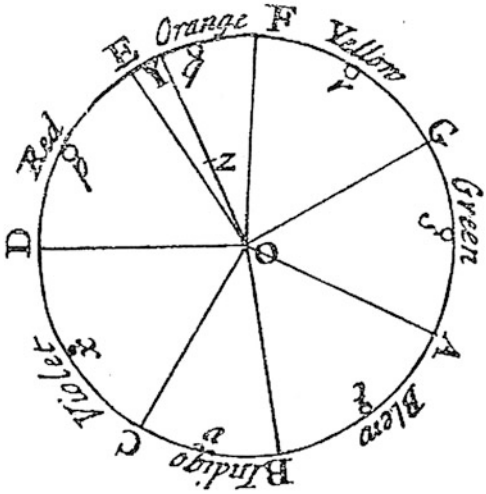


Fig. 2 Newton's color wheel (Newton 1730)

red and violet light overlapping. He links the seven "true" colors he observed – red, orange, yellow, green, blue, indigo, and violet – to seven song notes. He uses the differences in notes in harmonics to sort the colors sometimes in his experiments Newton (1730). Having defined each color, he begins to explore the physical differences in each, often using the musical notes analogy.

In proposition III, he finds that blue and violet light refract more than red light. Using this property, in proposition VI, he defines a color wheel which he uses as a tool to determine the color of light given its constituent parts. Figure 2 shows the color wheel Newton devised. First, Newton splits the color wheel into "seven Parts DE, EF, FG, GA, AB, BC, CD, proportional to the seven Musical Tones or Intervals of the eight Sounds, Sol, la, fa, sol, la, mi, fa, sol, contained in an eight, that is, proportional to the Number 1/9, 1/16, 1/10, 1/9, 1/16, 1/16, 1/9" Newton (1730). Then, he describes points p,q,r,s,t,u,x to be the "centers of gravity" of the arcs, with circles drawn around them proportional to the amount of color used. Then, finding the center of gravity (through a method he developed in *Principia*) between each of these seven circles yield a point z which gives the information of the final color. The wedge z is in dictates its color, while the distance from the center dictates its whiteness. If z is on the center of the circle, the light is white. The point Y is the point on the intersection of the circumference and the radius drawn from the center of the circle through z. The position of Y tells the information about color more precisely than z might. This method is merely an estimate, as Newton explains: "This Rule I conceive accurate enough for practice, though not mathematically accurate" Newton (1730). Despite this, it was still a useful tool to estimate the color of a product beam of light.

In one of the last propositions in Book I, Proposition IX Problem IV, Newton explores the optics behind rainbows. First, treating water droplets as perfect spheres, he shows the refractive and reflective processes occurring to bounce light toward an observer. He uses the diagram in Fig. 3 to explain the geometry behind the problem.

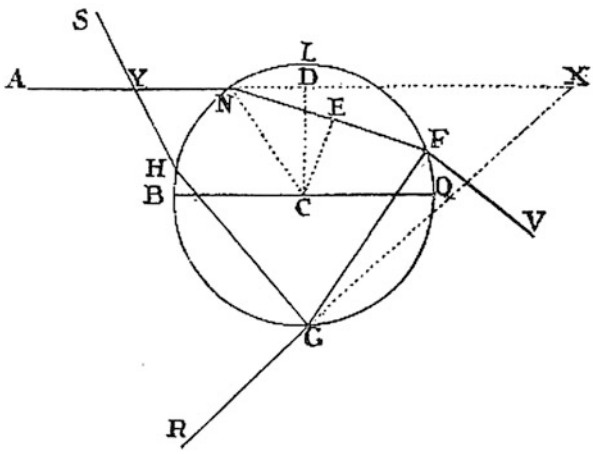


Fig. 3 Light bouncing in a water droplet as part of a rainbow Newton (1730)

First, he defines AN as a ray from the sun entering the raindrop. The ray refracts to F where it either refracts out of the drop to V or reflects to G where it either refracts out to R or reflects again to H. From H, the ray refracts out to S. AN and GR are produced to an intersection point, X. He examines the ratio of sine of incidence, I, and the sine of refraction, R. By moving the intersection point, B, toward L he notices that the arc QF, and thus the angle AXR will first decrease, then increase in size. He argues that the maximum of this angle and arc will be the biggest when $\frac{ND}{CN} = \frac{\sqrt{I^2-R^2}}{\sqrt{3R^2}}$. He also notes that the angle AYS will be smallest when $\frac{ND}{CN} = \frac{\sqrt{I^2-R^2}}{\sqrt{8R^2}}$. From there he is able to find a pattern about the angles of the emerging rays after n refractions inside the water droplet. The emergent angle will "continue to its limit" Newton (1730) when the ratio of ND to CN is $\frac{\sqrt{I^2-R^2}}{\sqrt{3R^2}}, \frac{\sqrt{I^2-R^2}}{\sqrt{8R^2}}, \frac{\sqrt{I^2-R^2}}{\sqrt{15R^2}}, \frac{\sqrt{I^2-R^2}}{\sqrt{24R^2}}$, etc. with the bottom number "being gather'd by continual Addition of the Terms of the arithmetical Progression 3, 5, 7, 9, & c" Newton (1730). He uses this series to predict where additional rainbows will appear in relation to the primary rainbow. He then goes out and measures the angle of the two rainbows he sees and shows why the colors appear in the order they do, one with an inverted order of colors compared to the other (Fig. 4). While Newton only sees two rainbows during his experiment, there are higher-order rainbows which get progressively fainter due to the high number of reflections needed, given that some light refracts out at each reflection point, so any rainbow above a secondary rainbow is hard to detect. These higher-order rainbows impressively follow the predictions Newton made with his infinite series.

In just Book I, Newton already heavily uses ratios and proportions to prove his theories on colors and reflections. Furthermore, he expands upon the ratios to construct infinite series to expand his theories into predictions that were later confirmed. Much of his work from Book I is still used today in geometrical optics Frank L. Pedrotti et al. (2007).

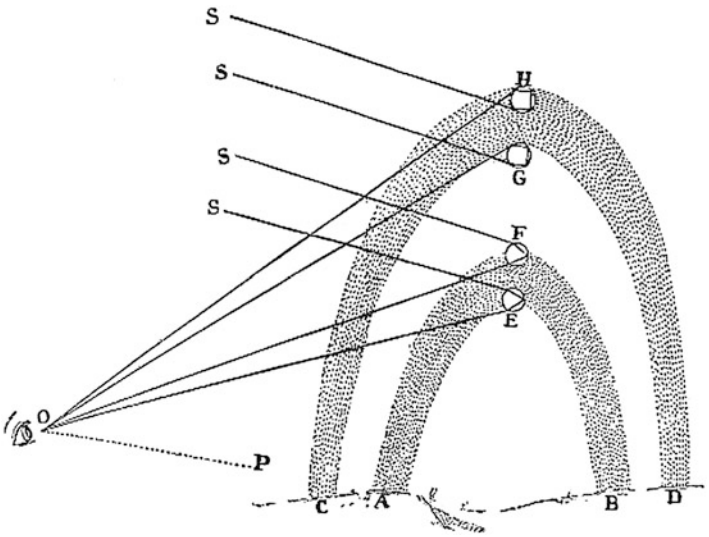
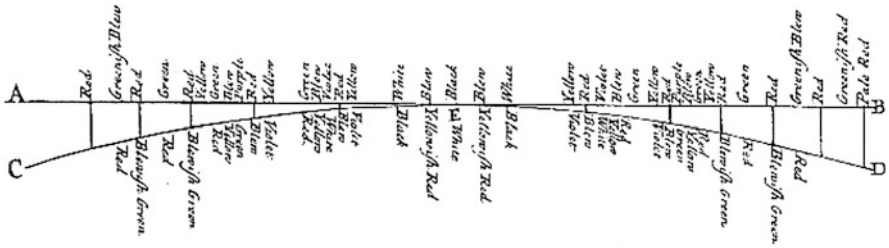


Fig. 4 A diagram of a rainbow Newton (1730)

3 *Opticks*: Books II and III

Newton's second book primarily addresses his work with thin films. Specifically, he seeks to find and quantify the color of thin films. First, however, Newton needed a way to measure the thickness of these thin films (so he could put a definition on "thin" to work with). Measuring thin films is difficult, and it was the key difference between his work and that of his contemporary Robert Hooke's Kumar (2006). To measure the thickness, Newton placed the films between a flat surface and a spherical surface dish resting on the flat surface. Then, knowing the diameter, the thickness of the film could be calculated as $d = \frac{D^2}{8R}$, where d is the thickness of the film, D is the diameter of the dish, and R is the radius of the sphere that the disk would be a part of Kumar (2006). In this setup, he shines light directly onto the center of the dish. The light refracts and reflects and forms circles of color around the center, which were periodic (Fig. 5).

Newton found that his corpuscular nature of light alone did not explain this phenomenon. So, he expanded his theory to include the ether (Kumar 2006). The ether was a hypothetical medium believed to exist that would have been the medium light propagated through (as waves needed a medium to travel through according to physics at the time). However, since Newton believed that light was not a wave, he did not use the ether as a medium through which light travels. Instead it served as a medium or field which only light interacts with and that interacts with the particles of light (Newton 1730; Kumar 2006). He explained that particles of light were different in size and strength and so interacted either weaker or stronger with the ether. This explained why different color bands seemed to be larger than other colors (Kumar


A
C
B
D
Red
Blue
Violet
Green
Yellow
Orange
Indigo
Black
White
Blue
Red

and which has many parts still used today, marking a turning point away from traditional natural philosophy and toward the more modern mathematical and scientific methods.

References

Albert DM (1997) Notes on Voltaire's 'the elements of Sir Isaac Newton's philosophy'. Doc Ophthalmol 94:59–81. https://link.springer.com/content/pdf/10.1007/BF02629681.pdf

Davidson MW (2015) Sir Isaac Newton. https://micro.magnet.fsu.edu/optics/timeline/people/newton.html

Evans MG (1955) Aristotle, Newton, and the theory of continuous magnitude. J Hist Ideas 16(4): 548–557. https://doi.org/10.2307/2707510

Joyce DE (1996) Euclid's elements. http://aleph0.clarku.edu/djoyce/elements/

Kumar A (2006) Newton's contributions to optics. Reson J Sci Educ 11(12):10–20. https://www.ias.ac.in/article/fulltext/reso/011/12/0010-0020

Kothapalli K, Simoson AJ, Bezuidenhout GA (2023) Light as a dipole: A tortoise and hare model. In: Sriraman, B. (eds) Handbook of Visual, Experimental and Computational Mathematics. Springer, Cham. https://doi.org/10.1007/978-3-030-93954-0_38-1

Marshall OS (1950) Alhazen and the telescope. Astronom Soc Pacific Leaflets 6:4

Newton I (1730) Opticks or, a treatise of the reflections, refractions, inflections, and colours of light, 4th edn. Project Gutenberg, Urbana, Illinois

Pedrotti FL, Pedrotti LM, Pedrotti LS (2007) Introduction to optics, 3rd edn. Pearson Prentice Hall, Upper Saddle River

Wawrzycki J (2021) Multiplication in Newtons principia. arXiv

Westfall RS (2023) Isaac Newton. https://www.britannica.com/biography/Isaac-Newton

Olry Terquem's Forgotten Problem and an Enumerative-Combinatorial Perspective on the Euclidean Algorithm

Robert G. Donnelly, Molly W. Dunkum, and Rachel McCoy

Contents

1 Introduction 1190
2 A Forgotten Problem 1191
3 The Networked-Numbers Game 1197
4 A Version of the Extended Euclidean Algorithm 1202
5 A Generic Game 1208
6 Is Bézout's Lemma a Constructive or Nonconstructive Result? 1212
7 Conclusion 1215
References 1216

Abstract

"Terquem's problem" is a name, given in the twentieth century, to the problem of enumerating certain integer sequences whose entries alternate in parity. In particular, this problem asks for the count of strictly increasing length m sequences of positive integers bounded above by some integer n whose odd-indexed entries are odd and whose even-indexed entries are even. This problem and its generalizations have been well-studied. However, the putative original source for this problem, an 1839 paper by Olry Terquem, is subtly different from the problem that is now attributed to Terquem. In this chapter, we highlight this distinction and also make connections between Terquem's "forgotten" problem, the Fibonacci sequence, a one-player game of numbers, continuant polynomials, (an extended version of) the Euclidean Algorithm, and Bézout's Lemma.

R. G. Donnelly
Department of Mathematics and Statistics, Murray State University, Murray, KY, USA
e-mail: rob.donnelly@murraystate.edu

M. W. Dunkum (✉) · R. McCoy
Department of Mathematics, Western Kentucky University, Bowling Green, KY, USA
e-mail: molly.dunkum@wku.edu; McCoy.800@osu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_146

Keywords

Terquem's problem · Fibonacci array · Networked-numbers Game · Continuant polynomials · Euclidean Algorithm · Bézout's Lemma

1 Introduction

The personages to whom this volume is dedicated – Chandler Davis, Reuben Hersh, and Yuri I. Manin – held many consonant views on the nature of mathematical philosophy and practice, but perhaps none more salient than this: Mathematics is, above all else, a *human* experience to be shared with one another across time and space. The mere accumulation of new results might seem, to those with casual or cynical interest in our subject, to be the sole driving purpose of mathematical practice, but such advances only represent real progress to the extent that they move within the currents of our ever-cycling global and intergenerational mathematical colloquy. Indeed, the discovery of new results is almost never an isolated event: concepts are innovated, patterns are discerned, conjectures are proffered, methodologies are developed, and (often multiple) proofs are obtained in a great collaborative movement.

In the ebb and flow of these currents over time, ideas might slip beneath the surface, or idly drift beyond our ken. Their rediscovery might be momentous (such as George E. Andrews' recovery of Ramanujan's "lost" notebook) or only modestly diverting. The study we offer here, turning on what we call the "forgotten" problem of Olry Terquem, is of the modest variety, but also hopefully illustrative of the potential edification of renavigating well-mapped passages, even when relying on pedestrian, but perhaps mildly unconventional, observational accoutrement.

In this chapter, we revisit a problem well known to the ages – understanding greatest common factors and the concomitant tools for their calculation – from our perspective as enumeratively minded mathematicians and as teachers (and former students) of discrete-themed transitions courses for undergraduate mathematics majors at regional public universities in the Mid-South of the United States. This experience has seasoned us with an especial appreciation for "explicit" solutions to existential problems, i.e., solutions which are characterized by their transparency, by their immediate computational availability, and by what some might call their constructive nature. These dispositions therefore inform the discourse undertaken here on certain classical elementary number theory problems (e.g., how so-called continuant polynomials can be expressed explicitly as sums over certain integer sequences considered by Terquem). We invite the reader to consider how their own mathematical priorities might differently influence their understanding and presentation of the rudimentary topics we take up here.

Our narrative begins in medias res with a counting problem whose statement appears to go back to a nineteenth-century paper by Olry Terquem. But, before we "push off and smite the sounding furrows," as it were, we offer a word on the esthetics of enumeration. Richard P. Stanley – a former student of Reuben Hersh's

kindred spirit Gian-Carlo Rota – offers several notions of what comprises a satisfactory answer to a counting problem (Stanley 1986). The first is an explicit formula, such as the formula $\frac{n(n-1)\cdots(n+1-k)}{k(k-1)\cdots(1)}$ for the number, typically denoted '$\binom{n}{k}$,' of k-element subsets of an n-element set when n and k are nonnegative integers. The second is a recurrence relation, such as the Pascal recurrence $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, where we take $\binom{0}{0}$ to be 1 and, for integers a and b, we take $\binom{a}{b}$ to be 0 when $b \notin [0, a]_{\mathbb{Z}}$. (We denote by $\mathcal{I}_{\mathbb{Z}}$ the set of integers within a given real number interval $\mathcal{I}$.). Another satisfactory answer is to present a one-to-one correspondence between the objects of the given counting problem and objects whose enumerative solution is known. For instance, by the multiplication rule of counting, one can place k students, chosen from amongst n students, in a line in $n(n-1)\cdots(n+1-k)$ ways. But, each such linear ordering of k students can be realized as the pairing of an unordered selection of k students, chosen in $\binom{n}{k}$ ways, together with one of the $k(k-1)\cdots(1)$ ways of placing k students in a line. A consequence of this bijective correspondence is the identity $n(n-1)\cdots(n+1-k) = \binom{n}{k} \cdot k(k-1)\cdots(1)$, from which we obtain the explicit formula $\binom{n}{k} = \frac{n(n-1)\cdots(n+1-k)}{k(k-1)\cdots(1)}$. Finally, one might offer what we call here a *(polynomial) refinement* of the problem. In this context, an enumerative solution might be presented as a uni- or multivariate polynomial (aka "generating function") whose monomial coefficients are the desired counts – such as the identity $(x+1)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$ from the Binomial Theorem. Or, the solution might be a polynomial that specializes to the desired count – such as the identity $\binom{n}{k} = \left.\frac{[n][n-1]\cdots[n+1-k]}{[k][k-1]\cdots[1]}\right|_{x=1}$, where "$[m]$" is shorthand for the polynomial $\sum_{j=0}^{m-1} x^j = \frac{x^m-1}{x-1}$ for any $m \in [0, \infty)_{\mathbb{Z}}$ and where the rational function $\frac{[n][n-1]\cdots[n+1-k]}{[k][k-1]\cdots[1]}$ is a polynomial usually called a Gaussian binomial coefficient. These enumerative esthetics will guide much of our investigation of Terquem's forgotten problem.

Two asides before we move on: During our research on this topic, the third-listed author helpfully translated Terquem's original paper (Terquem 1839, in French) as well as Eugen Netto's rendering of Terquem's problem (Netto 1901, in German). Also, we use the symbol "□" throughout to demarcate the end of proofs as well as the completion of other formally presented ideas (such as definitions, etc.).

2 A Forgotten Problem

Olry Terquem was a nineteenth century polymath – a mathematician, historian, journal founder and editor, religious reformer, and officer of the French Legion of Honor (Singer 1906). He is known for mathematical contributions in geometry (for example, as a discoverer of properties of the nine-point circle, cf. MacKay 1892), but in combinatorics his name is associated with an enumerative question called "Terquem's problem." Contemporary renderings of Terquem's problem generally follow the classical introductory combinatorics text by John Riordan, who seems to have borrowed it from an early twentieth century combinatorics text by the German mathematician

Eugen Netto (see Netto 1901, §49 Ch. 3). Here is Riordan's version (Riordan 1958, Ch. 1 problem 15), with notation lightly modified to better fit our context:

> Terquem's problem. For combinations of n numbered things in natural (rising) order, with t(n, m) the number of m-combinations with odd elements in odd position and even elements in even positions, or, what is the same thing, with t(n, m) the number of combinations with an equal number of odd and even elements for m even and with the number of odd elements one greater than the number of even for m odd, show that t(n, m) has the recurrence $\mathrm{t}(n, m) = \mathrm{t}(n-1, m-1) + \mathrm{t}(n-2, m)$ [with] $\mathrm{t}(n, 0) = 1$, [that] $\mathrm{t}(n, m) = \binom{\lfloor \frac{n+m}{2} \rfloor}{m}$, [and that] $f_{n+1} = \sum_{m=0}^{n} \mathrm{t}(n, m) = f_n + f_{n-1}$.

The objects to be enumerated in this version of Terquem's problem are called "alternating parity subsets" in the entry A046854 of the OEIS (2023). (For example, when $n = 4$, we get $\mathrm{t}(4, 0) = 1$, $\mathrm{t}(4, 1) = 2$, $\mathrm{t}(4, 2) = 3$, $\mathrm{t}(4, 3) = 1$, and $\mathrm{t}(4, 4) = 1$, affording the refinement $1 + 2 + 3 + 1 + 1$ of the Fibonacci number "8.") Such alternating parity subsets have been well-studied and generalized; for a somewhat recent example, see the paper by Toufik Mansour and Augustine O. Munagi (2012) and references therein.

However, the above version of Terquem's problem omits a requirement originally imposed by Terquem (1839). That requirement was a *parity condition on the lengths* of the alternating parity subsets he considered. We will see, in the remainder of this chapter, that this somewhat subtle but forgotten difference in Terquem's original problem actually places his problem in close proximity to some very classical enumerative and number theoretic phenomena, namely, a right-triangular Fibonacci-type array of numbers, continuant polynomials, (a version of) the extended Euclidean algorithm, and Bézout's Lemma. (Indeed, Terquem was generally aware of such connections.)

Now, to Terquem's forgotten problem. Assume n is a positive integer and that m is a nonnegative integer. An *oe-sequence* $T \coloneqq (T_1, T_2, \ldots, T_m)$ *bounded by* n is a strictly increasing sequence of m positive integers *wherein* m *and* n *have the same parity* ("On suppose de plus que $[m]$ et $[n]$ sont tous deux simultanément pairs ou impairs," cf. Terquem 1839, part 1), T_1 is odd, $T_m \leqslant n$, and T_i and T_{i+1} have opposite parity for any $i \in [1, m-1]_{\mathbb{Z}}$. In this definition, the "oe" prefix is meant to suggest the alternating odd-even pattern of an oe-sequence. We say that T has length m and write $\ell(T) = m$. When $m = 0$, then T is the empty sequence, is denoted by $()$, is considered to have length zero, and is considered to be an oe-sequence. Let $\mathcal{OE}(n)$ be the set of oe-sequences bounded by n. For reasons to be made apparent shortly, we set $\mathcal{OE}(-1) \coloneqq \emptyset$ and $\mathcal{OE}(0) \coloneqq \{()\}$. Here are some other small sets of oe-sequences:

$$\mathcal{OE}(1) = \{(1)\}$$
$$\mathcal{OE}(2) = \{(12), ()\}$$
$$\mathcal{OE}(3) = \{(123), (1), (3), \}$$
$$\mathcal{OE}(4) = \{(1234), (12), (14), (34), ()\}$$
$$\mathcal{OE}(5) = \{(12345), (123), (125), (145), (345), (1), (3), (5)\}$$
$$\mathcal{OE}(6) = \{(123456), (1234), (1236), (1256), (1456), (3456), (12), (14), (16), (34), (36), (56), ()\}$$

For $0 \leq k \leq \left\lfloor \frac{n}{2} \right\rfloor$, let $\mathcal{OE}(n,k)$ be the set of sequences T in $\mathcal{OE}(n)$ with $\ell(T) = n - 2k$.

Definition 2.1 *Terquem's forgotten problem* is the problem of enumerating $\mathcal{OE}(n,\ k)$. □

Likewise, an *eo-sequence* has all the properties of an oe-sequence except that it begins with an even integer instead of an odd and its length has parity opposite that of n; we then define $\mathcal{EO}(n)$ to be the set of eo-sequences bounded by n. This time we set $\mathcal{EO}(-1) := \{()\}$ and $\mathcal{EO}(0) := \emptyset$. For $1 \leqslant n \leqslant 6$, we have

$$\begin{aligned}
\mathcal{EO}(1) &= \{()\} \\
\mathcal{EO}(2) &= \{(2)\} \\
\mathcal{EO}(3) &= \{(23), ()\} \\
\mathcal{EO}(4) &= \{(234), (2), (4)\} \\
\mathcal{EO}(5) &= \{(2345), (23), (25), (45), ()\} \\
\mathcal{EO}(6) &= \{(23456), (234), (236), (256), (456), (2), (4), (6)\}
\end{aligned}$$

For $0 \leqslant k \leqslant \left\lfloor \frac{n-1}{2} \right\rfloor$, let $\mathcal{EO}(n,k)$ be the set of $T \in \mathcal{EO}(n)$ with $\ell(T) = (n-1) - 2k$. From here on, we refer to oe- and eo-sequences generically as *oereo sequences* (pronounced "OH-ree-oh").

If we extend the above data just a bit, we see that the arrays of numbers $(\mathcal{OE}(n,k))$ and $(\mathcal{EO}(n+1,k))$ match the rows of the right-triangular array depicted in Fig. 1, at least for $n \in [0,\ 11]_{\mathbb{Z}}$. To solve Terquem's forgotten problem, we will investigate this array.

Now, it is well-known that the sequence of sums of certain diagonals of Pascal's triangle is exactly the Fibonacci sequence $f_0 = 1, f_1 = 1, f_2 = 2, f_3 = 3, f_4 = 5$, etc.

1					
1					
1	1				
1	2				
1	3	1			
1	4	3			
1	5	6	1		
1	6	10	4		
1	7	15	10	1	
1	8	21	20	5	
1	9	28	35	15	1
1	10	36	56	35	6
			ETC.		

Fig. 1 Depiction of row zero through row 11 of the right-triangular Fibonacci array $\boldsymbol{F}$

This is credited to Edouard Lucas; for discussion, see Ch. 12 of the text by Thomas Koshy (2018). The rows of the right-triangular array of Fig. 1 coincide with those diagonals from Pascal's triangle. We call this the *right-triangular Fibonacci array* (or simply the *Fibonacci array*) and denote it by $\boldsymbol{F}$. We label the kth entry of the nth row as $f_{n,k}$, so that

$$\boldsymbol{F} = \left(f_{n,k}\right)_{n\in\{0,1,2,\ldots\},k\in\left\{0,1,\ldots,\left\lfloor\frac{n}{2}\right\rfloor\right\}}.$$

Formally, let us build $\boldsymbol{F}$ by declaring $f_{0,0} := f_{1,0} := 1$ and then applying the recurrence

$$f_{n,k} := f_{n-1,k} + f_{n-2,k-1} \tag{1}$$

for integers k and n with $n \geqslant 2$ and with the understanding that $f_{n,k} := 0$ when $k < 0$ or $k > \left\lfloor\frac{n}{2}\right\rfloor$. The following lemma is easily proved using this defining recurrence and is left as an exercise.

Lemma 2.2 *For any integers n and k with $0 \leqslant k \leqslant \left\lfloor\frac{n}{2}\right\rfloor$, we have*

$$f_{n,k} = \binom{n-k}{k} \quad \text{and} \quad \sum_{i=0}^{\left\lfloor\frac{n}{2}\right\rfloor} f_{n,i} = f_n. \tag{2}$$

□

Therefore, the sequence of entries on the *nth* row of $\boldsymbol{F}$ can be viewed as a refinement of the *nth* Fibonacci number. Indeed, some collections of polynomials, such as those studied by Lucas (1878, p. 186) and Ernst Jacobsthal (see Koshy 2018, Ch. 39), have coefficients from the rows of $\boldsymbol{F}$. OEIS entry A011973 (2023) offers many enumerative interpretations of the numbers comprising these refinements. Immediately below (Interpretation 2.3), we recall one such interpretation of the Fibonacci array numbers as counts of certain sequences of positive integers. We then solve Terquem's forgotten problem in Theorem 2.4 using a one-to-one correspondence between the sequences of Interpretation 2.3 and oereo sequences. (Recently, we added a comment to OEIS entry A011973 (2023) providing our interpretation of Fibonacci array entries as counts of oereo sequences.)

Interpretation 2.3: A Classical Enumerative Interpretation of the Fibonacci Array Fix a positive integer n and a nonnegative integer m. A *consecutive-free sequence* $S = (S_1, S_2, \ldots, S_m)$ *strictly bounded by* n is a strictly increasing sequence of m positive integers wherein $S_m < n$ and $S_{i+1} - S_i \geqslant 2$ for any $i \in [1, m-1]_{\mathbb{Z}}$. We say that S has length m and write $\ell(S) = m$. When $m = 0$, then S is the empty sequence, is denoted by (), is considered to have length zero, and is considered to be a consecutive-free sequence. Let $\mathcal{CF}(n)$ be the set of consecutive-free sequences strictly bounded by n. So, for example,

$$\mathcal{CF}(5) = \{(), (1), (2), (3), (4), (13), (14), (24)\}$$
$$\mathcal{CF}(6) = \{(), (1), (2), (3), (4), (5), (13), (14), (15), (24), (25), (35), (135)\}$$

Further, let $\mathcal{CF}(n,k)$ denotes the subset of sequences in $\mathcal{CF}(n)$ having length k, when k is a nonnegative integer. Now, if $\mathcal{S}(n-k,k)$ denotes the k-element subsets of $\{1, 2, \ldots, n-k\}$, we can think of each such subset as a sequence of length k by listing its elements in increasing order. With this convention, observe that the function $\mathcal{S}(n-k,k) \to \mathcal{CF}(n,k)$ given by $(S_1, S_2, \ldots, S_k) \mapsto (S_1+0, S_2+1, \ldots, S_k+k-1)$ is a bijection. So, we have $|\mathcal{CF}(n,k)| = |\mathcal{S}(n-k,k)| = \binom{n-k}{k} = f_{n,k}$ for all positive integers n and all integers k with $1 \leqslant k \leqslant \left\lfloor \frac{n}{2} \right\rfloor$. In particular, $|\mathcal{CF}(n)| = f_n$, the nth Fibonacci number. □

Our bijective solution to Terquem's forgotten problem (Theorem 2.4) will utilize the following notions. When entries of a given sequence are distinct, we may regard the sequence as a set and apply set operations. When n is a positive integer, we observe that a strictly increasing sequence $(T_1, \ldots, T_m)$ of integers with $1 \leqslant T_1$ and $T_m \leqslant n$ is an oe-sequence if and only if $\{1, 2, \ldots, n\} \setminus \{T_1, \ldots, T_m\}$ is the union of even-length subsequences of consecutive integers. For a nonnegative integer $k \leqslant \left\lfloor \frac{n}{2} \right\rfloor$, we define a function $\phi : \mathcal{CF}(n,k) \to \mathcal{OE}(n,k)$ in the following way. Given $S = (S_1, \ldots, S_k) \in \mathcal{CF}(n,k)$, let

$$T := \phi(S) := \{1, 2, \ldots, n-1, n\} \setminus \left(\bigcup_{i=1}^{k} \{S_i, S_i + 1\} \right).$$

Clearly $\ell(T) = n - 2k$, and therefore $\ell(T)$ has the same parity as n. Since T is formed from $\{1, 2, \ldots, n\}$ by removing consecutive pairs of integers, then the set of numbers between any two successive entries of T is an even-length (and possibly length zero) string of consecutive integers; so, T is an oe-sequence. These processes reverse, allowing us to begin with an oe-sequence bounded by n and with length $n - 2k$ and convert it to a length k consecutive-free sequence strictly bounded by n. Next, we define a function $\psi : \mathcal{EO}(n,k) \to \mathcal{OE}(n-1,k)$ as follows. Let $m := (n-1) - 2k$. Declare that $() \overset{\psi}{\mapsto} ()$ Otherwise, when $T = (T_1, \ldots, T_m)$, we set $\psi(T) := (T_1 - 1, T_2 - 1, \ldots, T_m - 1)$. It is evident that $\psi(T) \in \mathcal{OE}(n-1,k)$. We can reverse ψ by taking an oe-sequence bounded by $n-1$ and with length $(n-1) - 2k$ and adding "1" to each of its entries to obtain an eo-sequence bounded by n and with length $(n-1) - 2k$.

Theorem 2.4 (Solution to Terquem's Forgotten Problem) *Let n and k be integers with n positive and $0 \leqslant k \leqslant \left\lfloor \frac{n}{2} \right\rfloor$. Then the functions $\phi : \mathcal{CF}(n,k) \to \mathcal{OE}(n,k)$ and $\psi : \mathcal{EO}(n,k) \to \mathcal{OE}(n-1,k)$ defined above are bijections. Therefore $|\mathcal{OE}(n,k)| = f_{n,k}$ and $|\mathcal{EO}(n,k)| = f_{n-1,k}$, so the arrays of numbers $(\mathcal{OE}(n,k))$ and $(\mathcal{EO}(n+1,k))$ match the right-triangular Fibonacci array.*

Proof It is observed above that the procedures employed by each of ϕ and ψ are reversible, thereby producing inverse functions $\phi^{-1} : \mathcal{OE}(n,k) \to \mathcal{CF}(n,k)$ and

$\psi^{-1} : \mathcal{OE}(n-1,k) \to \mathcal{EO}(n,k)$. That is, ϕ and ψ are bijections. The remaining claims follow from the fact that $|\mathcal{CF}(n,k)| = f_{n,k}$, cf. Interpretation 2.3 above. □

The bijective proofs of Theorem 2.4 establish that Eqs. (1) and (2) above – which are a recurrence and an explicit formula – apply to counts of oereo sequences, thereby fulfilling several of the esthetic criteria for enumeration proposed in our Introduction. But, what of polynomial refinements? It will be helpful, for the sequel, to render some collections of oereo sequences as generating functions. The definitions to follow are effected by the following general notion: For any strictly increasing positive integer sequence $T = (T_1, \ldots, T_m)$, set $\mathrm{x}_T := x_{T_1} x_{T_2} \cdots x_{T_m}$ (where the x_i's are commuting indeterminates), with $\mathrm{x}_{()} := 1$ for the empty sequence.

Definition 2.5 For any integer $n \geqslant -1$, set

$$g_n := \sum_{T \in \mathcal{EO}(n)} \mathrm{x}_T \quad \text{and} \quad h_n := \sum_{T \in \mathcal{OE}(n)} \mathrm{x}_T,$$

where of course an empty sum is zero. Note that g_n and h_n are polynomials in $x_1, \ldots, x_n$ with unit coefficients. Call g_n an *eo-polynomial* and h_n an *oe-polynomial*; together, these are generically referred to as *oereo polynomials*.

For the record, here are the first several eo- and oe-polynomials:

$$g_{-1} = 1$$
$$g_0 = 0$$
$$g_1 = 1$$
$$g_2 = x_2$$
$$g_3 = x_2x_3 + 1$$
$$g_4 = x_2x_3x_4 + x_2 + x_4$$
$$g_5 = x_2x_3x_4x_5 + x_2x_3 + x_2x_5 + x_4x_5 + 1$$

$$h_{-1} = 0$$
$$h_0 = 1$$
$$h_1 = x_1$$
$$h_2 = x_1x_2 + 1$$
$$h_3 = x_1x_2x_3 + x_1 + x_3$$
$$h_4 = x_1x_2x_3x_4 + x_1x_2 + x_1x_4 + x_3x_4 + 1$$
$$h_5 = x_1x_2x_3x_4x_5 + x_1x_4x_5 + x_3x_4x_5 + x_1x_2x_3 + x_1x_2x_5 + x_1 + x_3 + x_5$$

A consequence of Theorem 2.4 is that when $n \geqslant 1$, the Fibonacci number f_{n-1} counts the number of terms in the polynomial g_n, and the Fibonacci number f_n counts the number of terms in h_n. Moreover, the number of terms of degree $(n-1)-2k$ (respectively $n-2k$) for $0 \leqslant k \leqslant \left\lfloor \frac{n-1}{2} \right\rfloor$ (resp. $0 \leqslant k \leqslant \left\lfloor \frac{n}{2} \right\rfloor$) within g_n (resp. h_n) is $f_{n-1,k}$ (resp. $f_{n,k}$). Given that these polynomials have an underlying Fibonacci-like structure, it is not surprising that they can be constructed using simple recurrence relations, as in our next result. We use the notation $g_n(c_1, \ldots, c_n)$ to indicate that each x_j has been specialized to some given (possibly complex) quantity c_j, and similarly for h_n.

Proposition 2.6 *The eo-polynomials* $\{g_n\}_{n \geqslant -1}$ *are uniquely determined by the recurrence relations* $g_{-1} = 1$, $g_0 = 0$, *and* $g_{n+1} = g_{n-1} + x_{n+1}g_n$ *for* $n \geqslant 0$. *Similarly,*

the oe-polynomials $\{h_n\}_{n \geqslant -1}$ *are uniquely determined by the recurrence relations* $h_{-1} = 0$, $h_0 = 1$, *and* $h_{n+1} = h_{n-1} + x_{n+1}h_n$ *for* $n \geqslant 0$. *For a positive integer n and complex numbers* $c_1, \ldots, c_n$, *we have* $h_{n-1}(c_2, \ldots, c_n) = g_n(c_1, c_2, \ldots, c_n)$ *and* $c_1 h_{n-1}(c_2, \ldots, c_n) + g_{n-1}(c_2, \ldots, c_n) = h_n(c_1, c_2, \ldots, c_n)$. *Moreover,* $h_n(c_1, c_2, \ldots, c_{n-1}, c_n) = h_n(c_n, c_{n-1}, \ldots, c_2, c_1)$.

Proof The claims in the first two sentences of the proposition follow from routine induction arguments, which we leave to the reader. The claims of the penultimate sentence follow from the definitions of g_n and h_n as sums over sets of eo- and oe-sequences. The last sentence follows from the fact that when k and n have the same parity, then the mapping $\sigma : \mathcal{OE}(n,k) \to \mathcal{OE}(n,k)$ given by $\sigma(S_1, S_2, \ldots, S_k) = (n+1-S_k, \ldots, n+1-S_2, n+1-S_1)$ is a bijection. □

The recurrence in Proposition 2.6 allows us to identify oe-polynomials as the so-called "continuant polynomials" of §4.5.3 of Donald E. Knuth's canonical work on *The Art of Computer Programming* (1997). There, Knuth exposits how such polynomials are related to continued fractions (hence their name) and also to the Euclidean Algorithm. (For a concrete illustration of this connection, see Example 4.4 below.) Knuth attributes to Leonhard Euler (presumably referring to Euler 1764) the following iterative derivation of the terms of the oe-polynomial h_n: They are obtained from $x_1x_2 \ldots x_n$ by "deleting zero or more nonoverlapping pairs of consecutive variables x_jx_{j+1}." As observed above, these correspond exactly to the oe-sequences bounded by n.

3 The Networked-Numbers Game

Before we connect oereo polynomials to the Euclidean Algorithm, we briefly digress to consider another natural occurrence of the Fibonacci array. In particular, we will see how the (univariate) *sign-alternating Fibonacci polynomials* defined by

$$\overline{F}_n(x) := \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k f_{n,k} x^{\lfloor n/2 \rfloor - k}$$

for each $n \in [0, \infty)_{\mathbb{Z}}$ occur within the context of a certain "node-firing" game and lead us to an explicit characterization of a desired property of game play (Theorem 3.3). It is easy to see that these sign-alternating Fibonacci polynomials are uniquely determined by the recurrence relations $\overline{F}_{-1}(x) = 0$, $\overline{F}_0(x) = 1$, and $\overline{F}_{n+1}(x) = -\overline{F}_{n-1}(x) + x^{n \bmod 2}\overline{F}_n(x)$ for $n \geqslant 0$.

This section can be skipped or lightly browsed on a first reading. However, it informs some of the considerations of the next several sections and, as a bonus, presents perhaps the simplest entry point – a node-firing game – to what we believe is the most sublime classification theorem in all of mathematics: *La Florado Klasado*. Node-ring games, including the "Chips Game" of Anders Björner, László

Lovász, and Peter Shor (1991) and the Networked-numbers Game of this section are considered in more generality by Kimmo Eriksson (1994). Our analysis here of a certain two-node game borrows ideas from a paper by Donnelly et al. (2021).

The *Networked-numbers Game (NG)* is a one-player game in which the player begins by choosing a finite simple graph (so, no loops and no multiple edges) and an assignment of real numbers, one to each node (i.e., vertex) of the graph. The only move afforded to the player is to *fire* at a chosen node by adding the number at the chosen node to the number at each adjacent node and then changing the sign of the number at the chosen node. A firing move is only *legal* in NG play if the number at the node-to-be-fired is positive. Game play continues until there are no legal moves remaining. See Fig. 2 for an illustration of NG play on a three-node graph. With just a bit more NG play, one can see that for the graph and initial numbers of Fig. 2, all games terminate with the same terminal numbers and in the same number of moves.

Let us add one further wrinkle to NG play. Say the underlying graph Γ has nodes $\{\gamma_i\}_{i\in I}$ indexed by a finite set I. We assign to each pair $(i, j) \in I \times I$ a real number $M_{i,j}$ satisfying

$$M_{i,j}\begin{cases} = 2 & \text{if } i = j \\ = 2 & \text{if } i \neq j \text{ and } \gamma_i \text{ is no adjacent to } \gamma_j \text{ in } \Gamma \\ < 0 & \text{if } \gamma_i \text{ and } \gamma_j \text{ are adjacent in } \Gamma \end{cases}$$

We view the $|M_{i,j}|$'s as multipliers for firing moves in the following way: If $\{\lambda_j\}_{j\in I}$ are the numbers at nodes $\{\gamma_j\}_{j\in I}$ of Γ and if node γ_i is fired, then the number at node γ_j is changed to $\lambda_j - M_{i,j}\,\lambda_i$. (If all negative $M_{i,j}$'s are "–1," then NG play proceeds as in the previous paragraph.) We call the pair $\Gamma = \{\gamma_i\}_{i\in I}$ and $(M_{i,j})_{i,j\in I}$ a *graph-with-multipliers*. To visually indicate the presence of such multipliers, we depict an edge between adjacent nodes γ_i and γ_j as $\gamma_i \overset{p}{\underset{q}{\bullet\!\!-\!\!\bullet}} \gamma_j$, where $p = |M_{i,j}|$ and $q = |M_{j,i}|$. We will also use the picture $\gamma_i \overset{\text{m}}{\bullet\!\!-\!\!\bullet} \gamma_j$ to represent such an edge when $pq = M_{i,j}M_{j,i} = 4\cos^2(\pi/m)$ and m is an integer exceeding two; for convenience we omit "③" and just represent our edge as $\gamma_i \bullet\!\!-\!\!\bullet \gamma_j$ when $m = 3$. Now let $\varrho := \frac{1+\sqrt{5}}{2}$, which is, of course, the golden ratio satisfying the relation $\varrho^2 = \varrho + 1$. The latter observation helps

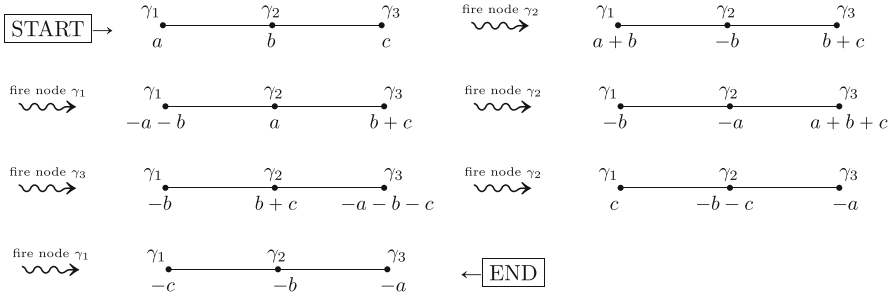


Fig. 2 Only nodes γ_1 and γ_3 are nonadjacent, and we assume a, b, and c are positive real numbers

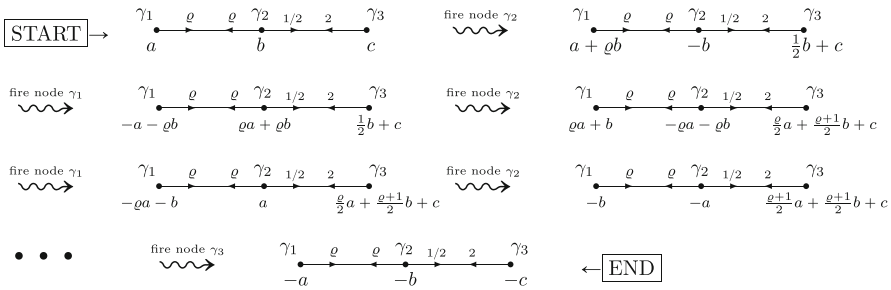


Fig. 3 Several of the node-firings of a networked-numbers game whose fifteen-node-firing sequence is (γ_2, γ_1, γ_2, γ_1, γ_2, γ_3, γ_2, γ_1, γ_2, γ_1, γ_3, γ_2, γ_1, γ_2, γ_3). (Here, a, b, and c are positive real numbers)

simplify certain quantities arising in NG play in Fig. 3 on a certain three-node graph-with-multipliers. As with the game illustrated in Fig. 2, one can (with patience!) see that all possible Networked-numbers Games played from the initial numbers in Fig. 3 terminate with the same terminal numbers and in the same number of moves.

Two natural questions arise concerning length of NG play: First, given a graph-with-multipliers, what conditions on the $M_{i,j}$'s will guarantee that any two games played from any given initial set of real numbers will require the same number of node-firings? Such a graph-with-multipliers is called *strongly convergent*. Second, for a given strongly convergent graph-with-multipliers, does there exist a game that terminates when all initial real numbers are nonnegative with at least one being positive? Call such a strongly convergent graph-with-multipliers *NG finitistic*.

The first question is answered by Eriksson (1996).

Eriksson's Strong Convergence Theorem *A connected graph-with-multipliers is strongly convergent if and only if on each edge of the graph, the product of its multipliers is in the set of real numbers* $\{4\cos^2(\pi/m) | m \in [3, \infty)_{\mathbb{Z}}\} \cup [4, \infty)$. □

In Theorem 3.3 below, we use sign-alternating Fibonacci polynomials to give a new proof of the "only if" part of the preceding equivalence.

The answer to the second of our questions is but one part of a broad and multifaceted classification result the authors (somewhat whimsically) call *La Florado Klasado*, which is Esperantese for "The Flowering Classification." We view *La Florado Klasado* as an assemblage of classification problems whose answers correspond to all (or some subset) of the graphs-with-multipliers in Fig. 4. Wilhelm Killing (1889) and Elie Cartan (1894) were the first to discern an aspect of this classification result (although not in the graphical form presented here) in their discovery of the finite-dimensional simple Lie algebras over the complex numbers; see Coleman (1989) for a lively recounting of this history. For exposition on and explicit statements of other parts of *La Florado Klasado* – including some very well-known and some less well-known results relating to Weyl groups, root systems, symmetric functions, Kac–Moody Lie algebras, posets, etc. – see §9 of the

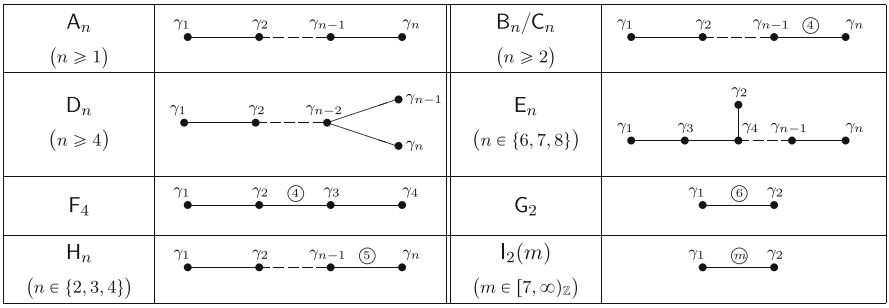


Fig. 4 Coxeter–Dynkin diagrams, rendered as graphs-with-multipliers, of types A through I. (Assume each edge has a "③" unless otherwise indicated)

manuscript by Donnelly (2022). Eriksson (1993, 1995) was the first to connect the NG and Coxeter groups. The "only if" part of the following equivalence is original to Donnelly (2008) and might be the most accessible of the many equivalences comprising this astoundingly rich classification result.

Theorem: An Equivalence from *La Florado Klasado* *A connected and strongly convergent graph-with-multipliers is NG finitistic if and only if it corresponds to one of the Coxeter–Dynkin diagrams associated with the finite irreducible Coxeter groups. That is, such a graph must be from one of the types of families of graphs-with-multipliers in* Fig. 4. □

For example, the graph-with-multipliers of Fig. 2 corresponds to the Coxeter–Dynkin diagram of type A_3 whose associated Coxeter group is the 24 element symmetric group on four letters, and the graph-with-multipliers of Fig. 3 corresponds to the Coxeter–Dynkin diagram of type H_3 whose associated Coxeter group is the group of symmetries of an icosahedron or dodecahedron.

And now to the main point of our Networked-numbers Game digression. We can demonstrate the necessity of Eriksson's strong convergence criteria by analyzing the Networked-numbers Game on two-node graphs-with-multipliers. The key to our analysis is to understand the role played by sign-alternating Fibonacci polynomials. To help make this connection, consider Fig. 5. Evidently, at each stage of game play, the numbers at the nodes are linear combinations of a and b with coefficients explicitly expressed in terms of sign-alternating Fibonacci polynomials evaluated at the product pq. We formalize this observation as the following fact, whose proof is deferred to §5 (see Corollary 5.2).

Fact 3.1 *Assume k is an odd positive integer. The formulas proposed in* Fig. 5 *are indeed valid expressions for the numbers at the nodes after the k^{th} and $(k+1)^{st}$ Firing moves of the two-node NG played in that figure.*

Explicit formulas for roots of certain univariate Fibonacci polynomials closely related to our sign-alternating Fibonacci polynomials were first obtained by Verner

Node-firing move	Progressive state of NG play on a two node graph-with-multipliers	
START	γ_1 p q γ_2 a b	
fire node γ_2	$a+qb$	$-b$
fire node γ_1	$-a-qb$	$pa+(pq-1)b$
fire node γ_2	$(pq-1)a+q(pq-2)b$	$-pa-(pq-1)b$
fire node γ_1	$-(pq-1)a-q(pq-2)b$	$p(pq-2)a+((pq)^2-3pq+1)b$
fire node γ_2	$((pq)^2-3pq+1)a+q((pq)^2-4pq+3)b$	$-p(pq-2)a-((pq)^2-3pq+1)b$
fire node γ_1	$-((pq)^2-3pq+1)a-q((pq)^2-4pq+3)b$	$p((pq)^2-4pq+3)a+((pq)^3-5(pq)^2+6pq-1)b$
$\vdots$	$\vdots$	$\vdots$
k^{th} node firing	$\overline{F}_{k-1}(pq)a+q\overline{F}_k(pq)b$ **?**	$-p\overline{F}_{k-2}(pq)a-\overline{F}_{k-1}(pq)b$ **?**
$(k+1)^{\text{st}}$ node firing	$-\overline{F}_{k-1}(pq)a-q\overline{F}_k(pq)b$ **?**	$p\overline{F}_k(pq)a+\overline{F}_{k+1}(pq)b$ **?**
$\vdots$	$\vdots$	$\vdots$

Fig. 5 Part of a networked-numbers game on a two-node graph-with-multipliers. (In this figure, the NG is played from positive starting numbers a and b on a generic two-node graph-with-multipliers whose nodes are $\{\gamma_1, \gamma_2\}$ and whose multipliers are the negative real numbers $M_{1,2} = -p$ and $M_{2,1} = -q$. Assume here that k is odd. Question marks "**?**" indicate that the formulas will be proved later)

E. Hoggatt and Marjorie Bicknell in 1973. From that work, one can obtain the following explicit description of the roots of sign-alternating Fibonacci polynomials (see Corollary 3.2 of Donnelly et al. (2021) for another approach).

Proposition 3.2 *The roots of the degree* $\lfloor n/2 \rfloor$ *sign-alternating polynomial* $\overline{F}_n(x)$ *are the* $\lfloor n/2 \rfloor$ *distinct real numbers comprising the set* $\left\{4\cos^2\left(\frac{j\pi}{n+1}\right) \middle| j \in [1, \lfloor n/2 \rfloor]_{\mathbb{Z}}\right\}$. □

Knowledge of these roots effects our new proof of the following result.

Theorem 3.3 (Necessity of Eriksson's Strong Convergence Criteria) *Suppose a connected graph-with-multipliers is strongly convergent. Then on any edge, the product of its multipliers is in the set of real numbers* $\{4\cos^2(\pi/m) | \ m \in [3, \infty)_{\mathbb{Z}}\} \cup [4, \infty)$.

Proof Without loss of generality, suppose there is an edge between nodes γ_1 and γ_2 with multipliers as in Fig. 5. Further, suppose $pq < 4$. Since $4\cos^2(\pi/2) = 0$ and the sequence $\left\{4\cos^2(\pi/m)\right\}_{m=2}^{\infty}$ strictly increases with limit 4, then there exists an integer $k \in [2, \infty)_{\mathbb{Z}}$ such that $4\cos^2(\pi/k) < pq \leqslant 4\cos^2(\pi/(k+1))$. If $pq = 4\cos^2(\pi/(k+1))$, we are done, so suppose pq is strictly between $4\cos^2(\pi/k)$ and $4\cos^2(\pi/(k+1))$. Now, for each $m \in [3, \infty)_{\mathbb{Z}}$, $4\cos^2(\pi/m)$ is the largest root of the monic polynomial $\overline{F}_{m-1}(x)$, and $4\cos^2(2\pi/m)$ is the second largest root, etc. (Proposition 3.2). Also, observe that since $2\pi/(k+1) > \pi/k$, then $4\cos^2(2\pi/(k+1))$

$< 4\cos^2(\pi/k)$. In particular, $\overline{F}_{m-1}(x) > 0$ when $x > 4\cos^2(\pi/m)$ and $\overline{F}_{m-1}(x) < 0$ when $4\cos^2(2\pi/m) < x < 4\cos^2(\pi/m)$. From these facts, we deduce that $\overline{F}_k(pq) < 0$ and that $\overline{F}_{m-1}(pq) > 0$ for each $m \in [3, k]_{\mathbb{Z}}$.

Suppose for the moment that k is odd. Let us begin a game by assigning to nodes γ_1 and γ_2 respectively an arbitrary positive number a and a positive number $b = 1 + \max\left\{\frac{-\overline{F}_{k-1}(pq)a}{q\overline{F}_k(pq)}, \frac{-p\overline{F}_k(pq)a}{\overline{F}_{k-1}(pq)}\right\}$, and by assigning to all other nodes negative numbers that are very large in magnitude (so as to be mostly unaffected by game play on γ_1 and γ_2). Since $b > \frac{-a\overline{F}_{k-1}(pq)}{q\overline{F}_k(pq)}$, then $\overline{F}_{k-1}(pq)a + q\overline{F}_k(pq)b < 0$. Then, by Fact 3.1 in concert with Fig. 5, we see that this NG terminates after the kth node-firing, if we begin by firing γ_2. To understand what happens when we fire γ_1 first, we can simply take the information from Fig. 5 and interchange the roles of a and b and of p and q. So, starting with γ_1, after the kth node-firing the numbers at nodes γ_1 and γ_2 are, respectively, $-q\overline{F}_{k-2}(qp)b - \overline{F}_{k-1}(qp)a$ and $\overline{F}_{k-1}(qp)b + p\overline{F}_k(qp)a$. Since $b > \frac{-p\overline{F}_k(pq)a}{\overline{F}_{k-1}(pq)}$, then $\overline{F}_{k-1}(qp)b + p\overline{F}_k(qp)a > 0$, in which case there is at least one more firing move available. Thus, when k is odd, the assumption that $pq \notin \{4\cos^2(\pi/m) \mid m \in [3, \infty)_{\mathbb{Z}}\}$ contradicts our strong convergence hypothesis. We can similarly obtain a contradiction in the case that k is even.

We conclude that if an edge of our connected and strongly convergent graph-with-multipliers has a multiplier product less than four, then the product must take the form $4\cos^2(\pi/m)$ for some $m \in [3, \infty)_{\mathbb{Z}}$. □

4 A Version of the Extended Euclidean Algorithm

The Euclidean Algorithm is an ancient procedure for computing the greatest common divisor of two integers without factoring them. This algorithm is Euclid's in that it famously appears in his great treatise, the *Elements*, as Proposition 2 of Book VII. However, most likely this result was known before Euclid, perhaps by the Pythagoreans, as contended by (among others) the mathematician and historian Bartel Leendert van der Waerden. From p. 115 of van der Waerden (1975):

> . . . Book VII is not a later reconstruction, but a piece of ancient mathematics. . . . Thus we have acquired an important insight: *Book VII was a textbook on the elements of the Theory of Numbers, in use in the Pythagorean school.*

Knuth writes that the Euclidean Algorithm is "the granddaddy of all algorithms, because it is the oldest nontrivial algorithm that has survived to the present day" (1997, §4.5.2). In this section, we explore commonalities between the Euclidean Algorithm and the Networked-numbers Game, a new perspective that hopefully adds something worthwhile to such a venerable piece of mathematics.

We begin by formulating a version of the Euclidean Algorithm as a pseudocode procedure that calculates the greatest common divisor and keeps track of some

important intermediate data. We will soon see how oereo polynomials allow us to recover from this data, any information that is yielded by what is often called the "extended Euclidean Algorithm." We take as our input a pair of positive integers a and b with $a \geqslant b$. (Positivity is assumed for the sake of simplicity.) The procedure applies the Division Algorithm in successive steps, producing at each step a quotient and a remainder, and returns the last nonzero remainder as the greatest common divisor. The motivating observation for the Euclidean Algorithm is the following simple lemma.

Lemma 4.1 *Let u and v be nonzero integers, and suppose $u = qv + r$ with $q \in \mathbb{Z}$ and $r \in [0, \mid v \mid)_{\mathbb{Z}}$. Then* $\gcd(u, v) = \gcd(v, r)$.

Proof From the relation $u = qv + r$, we see that any common divisor of v and r is a divisor of u, so $\gcd(u, v) \geqslant \gcd(v, r)$. From $r = u - qv$, we see that any common divisor of u and v is also a divisor of r, hence $\gcd(u, v) \leqslant \gcd(v, r)$. □

So, by applying the Division Algorithm to a pair of integers (u, v) as in the lemma statement, we obtain a "simpler" pair of numbers (v, r) with the same greatest common divisor. We codify this idea as follows.

Definition 4.2 The *Euclidean Algorithm (EA)* is the following pseudocode procedure:

Input: Positive integers a and b with $a \geqslant b$.
Initialization: Set $r_{-1} \coloneqq a$ and $r_0 \coloneqq b$, to be regarded as de facto remainders. Let i be our interation-counting index, initialized with $i \coloneqq 1$.
Step 1: Applying the Division Algorithm, we obtain unique integers q_i and r_i wherein $0 \leqslant r_i < r_{i-1}$ and

$$r_{i-2} = q_i r_{i-1} + r_i.$$

Step 2: If $r_i = 0$, then Return the following data: GCD $\coloneqq r_{i-1}$, NUM_STEPS $\coloneqq i$, REM_LIST $\coloneqq (r_{-1}, r_0, \ldots, r_i)$, and QUO_LIST $\coloneqq (q_1, \ldots, q_i)$.
Step 3: Otherwise, with $r_i > 0$, replace i with $i + 1$ and return to Step 1.
Output: A positive integer GCD (that will be shown to be the greatest common divisor of a and b), a positive integer NUM_STEPS that records the number of times we apply the Division Algorithm, the list REM_LIST of nonnegative integer remainders indexed from -1 up to NUM_STEPS, and the list QUO_LIST of positive integer quotients indexed from 1 through NUM_STEPS.

(We use the notation EA(a, b)[GCD] to denote the last nonzero remainder returned by the EA, EA(a, b)[NUM_STEPS] to denote the number of times EA applies the Division Algorithm, EA(a, b)[REM_LIST] to denote the list of remainders, and EA(a, b)[QUO_LIST] to denote the list of quotients.) □

As with any algorithm, we should certify that it actually terminates and prove that it returns the desired output. This is the content of the following theorem.

Theorem 4.3 (Boundedness and Correctness of the Euclidean Algorithm) *Let a and b be positive integers with $a \geqslant b$. With this input, the Euclidean Algorithm terminates, and the output value* EA(a, b)[GCD] *is the greatest common divisor of a and b.*

Proof Consider the sequence of remainders $r_{-1} \geqslant r_0 > r_1 > \ldots$. Now, $b = r_0 > 0$, $b - 1 \geqslant r_1 \geqslant 0$, $b - 2 \geqslant r_2 \geqslant 0$, etc. Since $b - b = 0$, the largest possible index of the remainder sequence is b, so the algorithm terminates in no more than b iterations. Let $n \coloneqq$ EA(a, b)[NUM_STEPS]. By Lemma 4.1, $\gcd(a, b) = \gcd(r_{-1}, r_0) = \gcd(r_0, r_1) = \cdots = \gcd(r_{n-2}, r_{n-1}) = \gcd(r_{n-1}, r_n)$. Since $r_{n-1} > 0$ and $r_n = 0$, then $\gcd(r_{n-1}, r_n) = r_{n-1} =$ EA (a, b)[GCD]. □

Example 4.4 For purposes of illustrating this version of the Euclidean Algorithm, we sought input numbers a and b such that the algorithm would terminate in exactly seven steps, i.e., seven applications of the Division Algorithm. How we found such a and b – and why we wanted this exact number of steps – will be revealed shortly. With this in mind, consider $a = 4449$ and $b = 935$. When the procedure call EA(a, b) is invoked, the following sequence of computations occurs:

$$\begin{aligned} 4449 &= 935 \cdot 4 + 709 \\ 935 &= 709 \cdot 1 + 226 \\ 709 &= 226 \cdot 3 + 31 \\ 226 &= 31 \cdot 7 + 9 \end{aligned} \qquad \begin{aligned} 31 &= 9 \cdot 3 + 4 \\ 9 &= 4 \cdot 2 + 1 \\ 4 &= 1 \cdot 4 + 0 \end{aligned}$$

EA output is as follows: EA(a, b)[GCD] = 1, EA(a, b)[NUM_STEPS] = 7, EA(a, b)[REM_LIST] = (4449, 935, 709, 226, 31, 9, 4, 1, 0), and EA(a, b)[QUO_LIST] = (4, 1, 3, 7, 3, 2, 4). Notice that from this data, we easily obtain an expression for $\frac{4449}{935}$ as a continued fraction:

$$\frac{4449}{935} = 4 + \frac{709}{935} = 4 + \frac{1}{935/709} = 4 + \frac{1}{1 + \frac{226}{709}} = 4 + \frac{1}{1 + \frac{1}{709/226}} = 4 + \frac{1}{1 + \frac{1}{3 + \frac{31}{226}}}$$

$$= \cdots = 4 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{7 + \cfrac{1}{3 + \cfrac{1}{2 + \frac{1}{4}}}}}}.$$

□

Let us now execute the Euclidean Algorithm from a generic starting point as if it were an arithmetic game akin to the Networked-numbers Game on two nodes. We

will call this the *Euclidean Remainders Game (ERG)*, although it is really a computational process rather than a game. We start with a pair of positive integers a and b such that $a \geqslant b$. The only move afforded to us in the Euclidean Remainders Game is to replace the larger of our two positive numbers, say u, by the remainder obtained when we divide the larger number by the smaller number, say v, using the Division Algorithm. That is, once we obtain $q \in [1, \infty)_{\mathbb{Z}}$ and $r \in [0, v)_{\mathbb{Z}}$ such that $u = qv + r$, then the pair (u, v) is replaced by the pair (r, v). This process is illustrated in Fig. 6 and carried out from a generic starting point.

With knowledge of the sequences of remainders and quotients, we can reverse the steps of the EA to recover the original input. Let us formulate this latter computational process as another diversion we call the *Dual Remainders Game (DRG)*. This game starts with a pair of nonnegative integers, exactly one of which is positive. At any step of the Dual Remainders Game, the only move allowed is to add to the smaller of the two numbers, the product of the larger number, and some chosen positive integer. We refer to the integers arising as pairs in DRG play as *dual remainders*. The DRG is illustrated in Fig. 7 and carried out from a generic starting point.

Quotients and remainders from Division Algorithm	Remainder pair	
START	a (also denoted r_{-1})	b (also denoted r_0)
write $a = q_1 b + r_1$ ⇝	$r_1 = a - q_1 b$	$r_0 = b$
$b = q_2 r_1 + r_2$ ⇝	$r_1 = a - q_1 b$	$r_2 = -q_2 a + (q_1 q_2 + 1) b$
$r_1 = q_3 r_2 + r_3$ ⇝	$r_3 = (q_2 q_3 + 1) a - (q_1 q_2 q_3 + q_1 + q_3) b$	$r_2 = -q_2 a + (q_1 q_2 + 1) b$
$r_2 = q_4 r_3 + r_4$ ⇝	$r_3 = (q_2 q_3 + 1) a - (q_1 q_2 q_3 + q_1 + q_3) b$	$r_4 = \begin{matrix} -(q_2 q_3 q_4 + q_2 + q_4) a \\ +(q_1 q_2 q_3 q_4 + q_1 q_2 + q_1 q_4 + q_3 q_4 + 1) b \end{matrix}$
$r_3 = q_5 r_4 + r_5$ ⇝	$r_5 = \begin{matrix} (q_2 q_3 q_4 q_5 + q_2 q_3 + q_2 q_5 + q_4 q_5 + 1) a \\ -\left(\begin{matrix} q_1 q_2 q_3 q_4 q_5 + q_1 q_2 q_3 + q_1 q_2 q_5 \\ + q_1 q_4 q_5 + q_3 q_4 q_5 + q_1 + q_3 + q_5 \end{matrix} \right) b \end{matrix}$	$r_4 = \begin{matrix} -(q_2 q_3 q_4 + q_2 + q_4) a \\ +(q_1 q_2 q_3 q_4 + q_1 q_2 + q_1 q_4 + q_3 q_4 + 1) b \end{matrix}$
$r_4 = q_6 r_5 + r_6$ ⇝	$r_5 = \begin{matrix} (q_2 q_3 q_4 q_5 + q_2 q_3 + q_2 q_5 + q_4 q_5 + 1) a \\ -\left(\begin{matrix} q_1 q_2 q_3 q_4 q_5 + q_1 q_2 q_3 + q_1 q_2 q_5 \\ + q_1 q_4 q_5 + q_3 q_4 q_5 + q_1 + q_3 + q_5 \end{matrix} \right) b \end{matrix}$	$r_6 = \begin{matrix} -\left(\begin{matrix} q_2 q_3 q_4 q_5 q_6 + q_2 q_3 q_4 + q_2 q_3 q_6 \\ + q_2 q_5 q_6 + q_4 q_5 q_6 + q_2 + q_4 + q_6 \end{matrix} \right) a \\ +\left(\begin{matrix} q_1 q_2 q_3 q_4 q_5 q_6 + q_1 q_2 q_3 q_4 + q_1 q_2 q_3 q_6 \\ + q_1 q_2 q_5 q_6 + q_1 q_4 q_5 q_6 + q_3 q_4 q_5 q_6 \\ + q_1 q_2 + q_1 q_4 + q_1 q_6 + q_3 q_4 + q_3 q_6 \\ + q_5 q_6 + 1 \end{matrix} \right) b \end{matrix}$
⋮	⋮	⋮
$r_{k-2} = q_k r_{k-1} + r_k$ ⇝	$r_k \stackrel{?}{=} g_k(q_1, \ldots, q_k) a - h_k(q_1, \ldots, q_k) b$	$r_{k-1} \stackrel{?}{=} -g_{k-1}(q_1, \ldots, q_{k-1}) a + h_{k-1}(q_1, \ldots, q_{k-1}) b$
$r_{k-1} = q_{k+1} r_k + r_{k+1}$ ⇝	$r_k \stackrel{?}{=} g_k(q_1, \ldots, q_k) a - h_k(q_1, \ldots, q_k) b$	$r_{k+1} \stackrel{?}{=} -g_{k+1}(q_1, \ldots, q_{k+1}) a + h_{k+1}(q_1, \ldots, q_{k+1}) b$
⋮	⋮	⋮

Fig. 6 The Euclidean remainders game: an interpretation of the EA as a "node-firing" process. (Start with the pair $r_{-1} = a$ and $r_0 = b$, where $a, b \in [1, \infty)_{\mathbb{Z}}$ and $a \geqslant b$. At the jth application of the Division Algorithm, replace remainder pair (r_{j-2}, r_{j-1}) with (r_j, r_{j-1}) when j is odd and replace (r_{j-1}, r_{j-2}) with (r_{j-1}, r_j) when j is even. Here, k is odd. Question marks "**?**" indicate that the formulas will be proved later)

DRG move	Dual remainder pair	
START	$r'_{-1} = 0$	r'_0
$r'_1 = q'_1 r'_0$	$r'_1 = q'_1 r'_0$	r'_0
$r'_2 = q'_2 r'_1 + r'_0$	$r'_1 = q'_1 r'_0$	$r'_2 = (q'_1 q'_2 + 1) r'_0$
$r'_3 = q'_3 r'_2 + r'_1$	$r'_3 = (q'_1 q'_2 q'_3 + q'_1 + q'_3) r'_0$	$r'_2 = (q'_1 q'_2 + 1) r'_0$
$r'_4 = q'_4 r'_3 + r'_2$	$r'_3 = (q'_1 q'_2 q'_3 + q'_1 + q'_3) r'_0$	$r'_4 = (q'_1 q'_2 q'_3 q'_4 + q'_1 q'_2 + q'_1 q'_4 + q'_3 q'_4 + 1) r'_0$
$r'_5 = q'_5 r'_4 + r'_3$	$r'_5 = \left(\begin{smallmatrix} q'_1 q'_2 q'_3 q'_4 q'_5 + q'_1 q'_2 q'_3 + q'_1 q'_2 q'_5 \\ + q'_1 q'_4 q'_5 + q'_3 q'_4 q'_5 + q'_1 + q'_3 + q'_5 \end{smallmatrix} \right) r'_0$	$r'_4 = (q'_1 q'_2 q'_3 q'_4 + q'_1 q'_2 + q'_1 q'_4 + q'_3 q'_4 + 1) r'_0$
$r'_6 = q'_6 r'_5 + r'_4$	$r'_5 = \left(\begin{smallmatrix} q'_1 q'_2 q'_3 q'_4 q'_5 + q'_1 q'_2 q'_3 + q'_1 q'_2 q'_5 \\ + q'_1 q'_4 q'_5 + q'_3 q'_4 q'_5 + q'_1 + q'_3 + q'_5 \end{smallmatrix} \right) r'_0$	$r'_6 = \left(\begin{smallmatrix} q'_1 q'_2 q'_3 q'_4 q'_5 q'_6 + q'_1 q'_2 q'_3 q'_4 + q'_1 q'_2 q'_3 q'_6 \\ + q'_1 q'_2 q'_5 q'_6 + q'_1 q'_4 q'_5 q'_6 + q'_3 q'_4 q'_5 q'_6 \\ + q'_1 q'_2 + q'_1 q'_4 + q'_1 q'_6 + q'_3 q'_4 + q'_3 q'_6 \\ + q'_5 q'_6 + 1 \end{smallmatrix} \right) r'_0$
$\vdots$	$\vdots$	$\vdots$
$r'_k = q'_k r'_{k-1} + r'_{k-2}$	$r'_k \overset{?}{=} h_k(q'_1, \ldots, q'_k) r'_0$	$r'_{k-1} \overset{?}{=} h_{k-1}(q'_1, \ldots, q'_{k-1}) r'_0$
$r'_{k+1} = q'_{k+1} r'_k + r'_{k-1}$	$r'_k \overset{?}{=} h_k(q'_1, \ldots, q'_k) r'_0$	$r'_{k+1} \overset{?}{=} h_{k+1}(q'_1, \ldots, q'_{k+1}) r'_0$
$\vdots$	$\vdots$	$\vdots$

Fig. 7 The dual remainders game: a "node-firing" process for reconstructing EA input from output. (Start with the pair $r'_{-1} = 0$ and $r'_0 \in [1, \infty)_{\mathbb{Z}}$. We take as given some positive integer sequence $q'_1, q'_2,$ etc. The jth move replaces the dual remainder pair $\left(r'_{j-2}, r'_{j-1}\right)$ with $\left(r'_j, r'_{j-1}\right)$ when j is odd and replaces $\left(r'_{j-1}, r'_{j-2}\right)$ with $\left(r'_{j-1}, r'_j\right)$ when j is even. Here, k is odd. Question marks "**?**" indicate that the formulas will be proved later)

Observe that at each stage of the ERG, each remainder seems naturally expressible as an integer linear combination of the initial integers a and b with each of the coefficients presented as the specialization of an oereo polynomial. Observe also that at each stage of the DRG, each dual remainder appears to be expressible as the product of the first nonzero dual remainder and some specialization of an oe-polynomial. Our proof demonstrating these two observations is enjoined with our proof of Fact 3.1 and is deferred to the next section (see Corollary 5.2). For now, we record these two verities as the following fact.

Fact 4.5 *Assume k is an odd positive integer. The proposed identities at steps k and $k + 1$ of the ERG played in* Fig. 6 *and the DRG played in* Fig. 7 *are indeed valid identities.*

Equipped with this knowledge, we can easily prove our next result, which has much overlap with what is known as Bézout's Lemma (more on this shortly). Moreover, our explicit understanding of outcomes in the Euclidean and Dual Remainders Games not only yields efficient proofs but also leads directly to the identities stated in Theorem 4.6.

Theorem 4.6 *Keep the hypotheses of Theorem 4.3, so* $\gcd(a, b) = \mathrm{EA}(a, b)$ [GCD]. *Now set* $n := \mathrm{EA}(a, b)$[NUM_STEPS], $(r_{-1}, r_0, \ldots r_n) := \mathrm{EA}(a, b)$[REM_LIST], *and* $(q_1, \ldots, q_n) := \mathrm{EA}(a, b)$[QUO_LIST].

1. *For any* $j \in [-1, n]_{\mathbb{Z}}$, *set* $\overrightarrow{g_j} := g_j(q_1, \ldots, q_j)$, $\overleftarrow{g_j} := g_j(q_{n+1-j}, \ldots, q_n)$, $\overrightarrow{h_j} := h_j(q_1, \ldots, q_j)$, *and* $\overrightarrow{h_j} := h_j(q_{n+1-j}, \ldots, q_n)$, *and momentarily allow* $r_{n+1} := \gcd(a, b)$. *For each such* j, $r_j = (-1)^{j+1}\overrightarrow{g_j}\, a + (-1)^j \overrightarrow{h_j}\, b$, $r_{n-j} = \overleftarrow{g_j} \cdot \gcd(a, b)$, *and* $r_{n-1-j} = \overleftarrow{h_j} \cdot \gcd(a, b)$.
2. *The set* $\{ax + by \mid x, y \in \mathbb{Z}\}$ *consists exactly of the integer multiples of* $\gcd(a, b)$. *In particular,* $\gcd(a, b)$ *is the smallest positive integer expressible as an integer linear combination of* a *and* b. *Moreover,* $\gcd(a, b) = (-1)^n g_{n-1}(q_1, \ldots, q_{n-1})a + (-1)^{n-1} h_{n-1}(q_1, \ldots, q_{n-1})b$, $b = g_n(q_1, \ldots, q_n) \cdot \gcd(a, b)$, *and* $a = h_n(q_1, \ldots, q_n) \cdot \gcd(a, b)$.

Proof In *(1)*, the assertion that $r_j = (-1)^{j+1}\overrightarrow{g_j}\, a + (-1)^j \overrightarrow{h_j}\, b$ follows from the ERG part of Fact 4.5. In Fig. 7, set $r'_j := r_{n-1-j}$ and $q'_j := q_{n+1-j}$. Then $r'_j = h_j\left(q'_1, \ldots, q'_j\right) \cdot r'_1 = h_j\left(q'_j, \ldots, q'_1\right) \cdot r'_1$, where the first equality follows from the DRG part of Fact 4.5 and the second equality follows from Proposition 2.6. So, $r_{n-1-j} = \overleftarrow{h_j} \cdot \gcd(a, b)$. By Proposition 2.6, $h_{j-1}(q_{n+1-(j-1)}, \ldots, q_n) = g_j(q_{n+1-j}, \ldots, q_n)$. So,

$$\begin{aligned} r_{n-j} = r_{n-1-(j-1)} &= h_{j-1}\left(q_{n+1-(j-1)}, \ldots, q_n\right) \cdot \gcd(a, b) \\ &= g_j(q_{n+1-j}, \ldots, q_n) \cdot \gcd(a, b), \end{aligned}$$

i.e., $r_{n-j} = \overleftarrow{g_j} \cdot \gcd(a, b)$.

For part *(2)*, note that every element of the set $\{ax + by \mid x, y \in \mathbb{Z}\}$ is a multiple of $\gcd(a, b)$; this set will consist of all the multiples of $\gcd(a, b)$ if we can show that $\gcd(a, b)$ is an integer linear combination of a and b. Of course, this latter fact is just the $j = n - 1$ case of *(1)*. □

To conclude this section, we recall a question from our earlier Example 4.4: How did we find integers a and b such that the Euclidean Algorithm would terminate in exactly $n = 7$ steps, and why $n = 7$? Well, in seeking examples to illustrate the algorithm to students in Discrete Mathematics courses, we thought that having students perform calculations that involve g_7 and h_7 – which have 13 and 21 terms respectively – would be sufficient to strongly suggest our proposed connection between the Euclidean Algorithm and our Fibonacci array. According to part 1 of Theorem 4.6, we can find such a and b simply by choosing a sequence $q_1, q_2, \ldots, q_7$ of quotients and then taking $b := g_7(q_1, q_2, \ldots, q_7)$ and $a := h_7(q_1, q_2, \ldots, q_7)$. Using the sequence 4, 1, 3, 7, 3, 2, 4 of quotients in Example 4.4, we got $b = 935 = g_7(4, 1, 3, 7, 3, 2, 4)$ and $a = 4449 = h_7(4, 1, 3, 7, 3, 2, 4)$. (Without executing the EA, how do we know a priori that these numbers a and b are relatively prime?)

Maarten Bullynck observes in §3.1 of Bullynck (2009) that certain number theory discourses written separately by Euler and Abraham Gotthelf Kästner in the later eighteenth century illustrated the Euclidean Algorithm with input integers requiring at most five or six applications of the Division Algorithm in order to return the gcd., although Kästner mentions a case "where one has to divide 54 times, before one finds the greatest common measure." Bullynck further notes that this latter question was a classic recreational problem originally due to Leonardo Pisano, whose alias we leave it to the reader to discern. (What is the simplest answer to this riddle?)

5 A Generic Game

The NG, as played in Fig. 5, and the ERG/DRG, as illustrated in Figs. 6 and 7, involve alternating moves of similar types. We capture this idea formally as follows. Fix complex numbers c_1, c_2, etc., and fix another number ε that will typically assume one of the values ± 1. Our *generic game* begins with a chosen pair (a, b) of real numbers. Let j be a positive integer. What we call the $(2j-1)$st *generic game move* changes an ordered pair of numbers (u, v) to $(u - \varepsilon c_{2j-1} v, \varepsilon v)$, and the $(2j)^{\text{th}}$ *generic game move* changes (u, v) to $(\varepsilon u, -\varepsilon c_{2j} u + v)$. That is,

$$\begin{pmatrix} u \\ v \end{pmatrix} \xrightarrow{(2j-1)^{\text{st}}\ \text{move}} \begin{pmatrix} 1 & -\varepsilon c_{2j-1} \\ 0 & \varepsilon \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u - \varepsilon c_{2j-1} v \\ \varepsilon v \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} \xrightarrow{(2j)^{\text{th}}\ \text{move}} \begin{pmatrix} \varepsilon & 0 \\ -\varepsilon c_{2j} & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \varepsilon u \\ -\varepsilon c_{2j} u + v \end{pmatrix}.$$

Notice that when $\varepsilon = -1$, $c_1 = c_3 = \cdots = q$, and $c_2 = c_4 = \cdots = p$, then the $(2j\text{–}1)^{\text{st}}$ (respectively, $(2j)$th) NG move played in Fig. 5 is exactly the $(2j\text{–}1)$st (respectively, $(2j)$th) generic game move. Also, when $\varepsilon = +1$ and $c_m = q_m$ for each index m, then the $(2j\text{–}1)$st (respectively, $(2j)$th) ERG move played in Fig. 6 is exactly the $(2j\text{–}1)$st (respectively, $(2j)$th) generic game move. And, when $\varepsilon = +1$ and $c_m = -q'_m$ for each index m, then the $(2j\text{–}1)$st (respectively, $(2j)$th) DRG move played in Fig. 7 is exactly the $(2j\text{–}1)$st (respectively, $(2j)$th) generic game move.

Our analyses of the NG, ERG, and DRG in the preceding sections led us to the following general result which, though somewhat tedious to formulate, is fairly easy to prove by induction. This result grants us some insight as to when and how certain computational processes involve oereo polynomials.

Theorem 5.1 *Let n be a positive integer, and let a and b be real numbers. For convenience, let $n_{\text{m}} := \lfloor (n-1)/2 \rfloor$ and $n_z := \lfloor n/2 \rfloor$. The result of applying the first n generic game moves to the column vector $\binom{a}{b}$ is*

$$\begin{pmatrix} \varepsilon^{n_z} \widehat{g}_{2n_{\text{m}}+1} & -\varepsilon^{1/2} \varepsilon^{n_z} \widehat{h}_{2n_{\text{m}}+1} \\ -\varepsilon^{1/2} \varepsilon^{n_{\text{m}}} \widehat{g}_{2n_z} & \varepsilon^{n_{\text{m}}+1} \widehat{h}_{2n_z} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \left(\varepsilon^{n_z} \widehat{g}_{2n_m+1}\right) a - \left(\varepsilon^{1/2} \varepsilon^{n_z} \widehat{h}_{2n_m+1}\right) b \\ -\left(\varepsilon^{1/2} \varepsilon^{n_{\text{m}}} \widehat{g}_{2n_z}\right) a + \left(\varepsilon^{n_{\text{m}}+1} \widehat{h}_{2n_z}\right) b \end{pmatrix},$$

where the quantities $\widehat{g}_k := g_k\left(\varepsilon^{1/2}c_1, \ldots, \varepsilon^{1/2}c_k\right)$ *and* $\widehat{h}_k := h_k\left(\varepsilon^{1/2}c_1, \ldots, \varepsilon^{1/2}c_k\right)$ *are specializations of oereo polynomials.*

Proof At $n = 1$, $\begin{pmatrix} \varepsilon^{n_z}\widehat{g}_{2n_m+1} & -\varepsilon^{1/2}\varepsilon^{n_z}\widehat{h}_{2n_m+1} \\ -\varepsilon^{1/2}\varepsilon^{n_m}\widehat{g}_{2n_z} & \varepsilon^{n_m+1}\widehat{h}_{2n_z} \end{pmatrix}$ becomes $\begin{pmatrix} \widehat{g}_1 & -\varepsilon^{1/2}\widehat{h}_1 \\ -\varepsilon^{1/2}\widehat{g}_0 & \varepsilon\widehat{h}_0 \end{pmatrix}$, which is exactly the matrix $\begin{pmatrix} 1 & -\varepsilon c_1 \\ 0 & \varepsilon \end{pmatrix}$. When $n = 2$, the matrix $\begin{pmatrix} \varepsilon^{n_z}\widehat{g}_{2n_m+1} & -\varepsilon^{1/2}\varepsilon^{n_z}\widehat{h}_{2n_m+1} \\ -\varepsilon^{1/2}\varepsilon^{n_m}\widehat{g}_{2n_z} & \varepsilon^{n_m+1}\widehat{h}_{2n_z} \end{pmatrix}$ becomes $\begin{pmatrix} \varepsilon\widehat{g}_1 & -\varepsilon^{3/2}\widehat{h}_1 \\ -\varepsilon^{1/2}\widehat{g}_2 & \varepsilon\widehat{h}_2 \end{pmatrix}$, i.e., $\begin{pmatrix} \varepsilon & -\varepsilon^{3/2}\left(\varepsilon^{1/2}c_1\right) \\ -\varepsilon^{1/2}\left(\varepsilon^{1/2}c_2\right) & \varepsilon\left(\varepsilon c_1 c_2 + 1\right) \end{pmatrix}$. The latter matrix also arises as the product $\begin{pmatrix} \varepsilon & 0 \\ -\varepsilon c_2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\varepsilon c_1 \\ 0 & \varepsilon \end{pmatrix}$. These computations confirm the theorem statement for $n = 1$ and $n = 2$.

Suppose that for some odd positive integer n, the matrix $\begin{pmatrix} \varepsilon^{n_z}\widehat{g}_{2n_m+1} & -\varepsilon^{1/2}\varepsilon^{n_z}\widehat{h}_{2n_m+1} \\ -\varepsilon^{1/2}\varepsilon^{n_m}\widehat{g}_{2n_z} & \varepsilon^{n_m+1}\widehat{h}_{2n_z} \end{pmatrix}$ represents the transformation corresponding to the first n generic game moves, with $n_m = \frac{n-1}{2} = n_z$. For the first $n+1$ moves, then, the representing matrix is

$$\begin{pmatrix} \varepsilon & 0 \\ -\varepsilon c_{n+1} & 1 \end{pmatrix}\begin{pmatrix} \varepsilon^{(n-1)/2}\widehat{g}_n & -\varepsilon^{1/2}\varepsilon^{(n-1)/2}\widehat{h}_n \\ -\varepsilon^{1/2}\varepsilon^{(n-1)/2}\widehat{g}_{n-1} & \varepsilon^{(n+1)/2}\widehat{h}_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} \varepsilon^{(n+1)/2}\widehat{g}_n & \varepsilon^{(n+2)/2}\widehat{h}_n \\ -\varepsilon^{(n+1)/2}c_{n+1}\widehat{g}_n - \varepsilon^{n/2}\widehat{g}_{n-1} & \varepsilon^{(n+2)/2}c_{n+1}\widehat{h}_n + \varepsilon^{(n+1)/2}\widehat{h}_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} \varepsilon^{(n+1)/2}\widehat{g}_n & \varepsilon^{(n+2)/2}\widehat{h}_n \\ -\varepsilon^{n/2}\left(\varepsilon^{1/2}c_{n+1}\widehat{g}_n + \widehat{g}_{n-1}\right) & \varepsilon^{(n+1)/2}\left(\varepsilon^{1/2}c_{n+1}\widehat{h}_n + \widehat{h}_{n-1}\right) \end{pmatrix}$$

$$= \begin{pmatrix} \varepsilon^{(n+1)/2}\widehat{g}_n & \varepsilon^{(n+2)/2}\widehat{h}_n \\ -\varepsilon^{n/2}\widehat{g}_{n+1} & \varepsilon^{(n+1)/2}\widehat{h}_{n+1} \end{pmatrix} = \begin{pmatrix} \varepsilon^{(n+1)_z}\widehat{g}_{2(n+1)_m+1} & -\varepsilon^{1/2}\varepsilon^{(n+1)_z}\widehat{h}_{2(n+1)_m+1} \\ -\varepsilon^{1/2}\varepsilon^{(n+1)_m}\widehat{g}_{2(n+1)_z} & \varepsilon^{(n+1)_m+1}\widehat{h}_{2(n+1)_z} \end{pmatrix},$$

as desired. Similar computations verify the result when n is even.

Here are specific consequences for the Networked-numbers and the Euclidean and Dual Remainders Games.

Corollary 5.2 *Let k be an odd positive integer. The formulas and identities in each statement below refer to* Figs. 5, 6, and 7.

(NG part) In the Networked-numbers Game, the numbers after k node-firings are:

$$\begin{array}{ll} k^{\text{th}} \text{ node firing} & \overline{F}_{k-1}(pq)a + q\overline{F}_k(pq)b \qquad -p\overline{F}_{k-2}(pq)a - \overline{F}_{k-1}(pq)b \\ (k+1)^{\text{st}} \text{ node firing} & -\overline{F}_{k-1}(pq)a - q\overline{F}_k(pq)b \qquad p\overline{F}_k(pq)a + \overline{F}_{k+1}(pq)b \end{array}$$

(ERG part) In the Euclidean Remainders Game, the remainder pairs after k applications of the Division Algorithm in the Euclidean Algorithm are:

$$\begin{array}{ll} r_{k-2} = q_k r_{k-1} + r_k & r_k = g_k(q_1, \ldots, q_k)a - h_k(q_1, \ldots, q_k)b \qquad r_{k-1} = -g_{k-1}(q_1, \ldots, q_k)a + h_{k-1}(q_1, \ldots, q_k)b \\ r_{k-1} = q_{k+1} r_k + r_{k+1} & r_k = g_k(q_1, \ldots, q_k)a - h_k(q_1, \ldots, q_k)b \quad r_{k+1} = -g_{k+1}(q_1, \ldots, q_{k+1})a + h_{k+1}(q_1, \ldots, q_{k+1})b \end{array}$$

(DRG part) In the Dual Remainders Game, the dual remainder pairs after k DRG moves are:

$$\begin{array}{ll} r'_k = q'_k r'_{k-1} + r'_{k-2} & r'_k = h_k(q'_1, \ldots, q'_k)\, r'_0 \qquad r'_{k-1} = h_{k-1}(q'_1, \ldots, q'_{k-1})\, r'_0 \\ r'_{k+1} = q'_{k+1} r'_k + r'_{k-1} & r'_k = h_k(q'_1, \ldots, q'_k)\, r'_0 \qquad r'_{k+1} = h_{k+1}(q'_1, \ldots, q'_{k+1})\, r'_0 \end{array}$$

Proof We observed, in the second paragraph of this section, that the NG, ERG, and DRG are special cases of our generic game. So, we only need to confirm that the generic outcome of Theorem 5.1 specializes to the outcomes described in the corollary statement for each of our three games.

For the NG, we have $\varepsilon = -1, p = c_2 = c_4 = \text{etc.}$, and $q = c_1 = c_3 = \text{etc.}$ There are two cases to consider: For some nonnegative integer ℓ, we have $k = 4\ell + 1$ or $k = 4\ell + 3$. To begin, suppose $k = 4\ell + 1$. Then $k_{\text{m}} = (k-1)/2 = k_{\text{z}}$ and $i^{k-1} = 1$ (for the imaginary number i). By Theorem 5.1, the numbers after k node-firings are $\left(\varepsilon^{k_{\text{z}}}\widehat{g}_{2k_{\text{m}}+1}\right)a - \left(\varepsilon^{1/2}\varepsilon^{k_{\text{z}}}\widehat{h}_{2k_{\text{m}}+1}\right)b$ and $-\left(\varepsilon^{1/2}\varepsilon^{k_{\text{m}}}\widehat{g}_{2k_{\text{z}}}\right)a + \left(\varepsilon^{k_{\text{m}}+1}\widehat{h}_{2k_{\text{z}}}\right)b$, which simplify to $\left(i^{k-1}\widehat{g}_k\right)a - \left(i^k\widehat{h}_k\right)b$ and $-\left(i^k\widehat{g}_{k-1}\right)a + \left(i^{k+1}\widehat{h}_{k-1}\right)b$. Now, $\left(i^{k-1}\widehat{g}_k\right)a - \left(i^k\widehat{h}_k\right)b = \widehat{g}_k a - i\widehat{h}_k b$. By definition $g_k = \sum_{j=0}^{(k-1)/2}\sum_{T\in\mathcal{EO}(k,j)} x_{T_1}\cdots x_{T_{(k-1)-2j}}$. When we set $x_{2m} := ip$ and $x_{2m+1} = iq$ for each m, then g_k specializes to $\widehat{g}_k = \sum_{j=0}^{(k-1)/2} f_{k-1,j}\, i^{(k-1)-2j}(pq)^{\frac{k-1}{2}-j} = \sum_{j=0}^{(k-1)/2}(-1)^j f_{k-1,j}\,(pq)^{\frac{k-1}{2}-j} = \overline{F}_{k-1}(pq)$. Similarly, $h_k = \sum_{j=0}^{(k-1)/2}\sum_{T\in\mathcal{O}\mathcal{E}(k,j)} x_{T_1}\cdots x_{T_{k-2j}}$, so that $\widehat{h}_k = \sum_{j=0}^{(k-1)/2} f_{k,j}\, iq\, i^{k-1-2j}(pq)^{\frac{k-1}{2}-j} = iq\sum_{j=0}^{(k-1)/2}(-1)^j \ f_{k,j}\,(pq)^{\frac{k-1}{2}-j} = iq\overline{F}_k(pq)$. Therefore, $\widehat{g}_k a - i\widehat{h}_k b = \overline{F}_{k-1}(pq)a + q\overline{F}_k(pq)b$, as desired. Also, $-\left(i^k\widehat{g}_{k-1}\right)a + \left(i^{k+1}\widehat{h}_{k-1}\right)b = -i\widehat{g}_{k-1}\,a - \widehat{h}_{k-1}\,b$. The eo-polynomial given by $g_{k-1} = \sum_{j=0}^{(k-3)/2}\sum_{T\in\mathcal{EO}(k-1,j)} x_{T_1}\cdots x_{T_{(k-2)-2j}}$ specializes to $\widehat{g}_{k-1} = \sum_{j=0}^{(k-3)/2} f_{k-2,j}\, pi^{(k-2)-2j}(pq)^{\frac{k-3}{2}-j} = -ip\overline{F}_{k-2}(pq)$. Similarly see that the

oe-polynomial $h_{k-1} = \sum_{j=0}^{(k-1)/2} \sum_{T \in O\mathcal{E}(k-1,j)} x_{T_1} \cdots x_{T_{(k-1)-2j}}$ specializes to $\widehat{h}_{k-1} = \overline{F}_{k-1}(pq)$. Then, $-i\widehat{g}_{k-1}a - \widehat{h}_{k-1}b = -p\overline{F}_{k-2}(pq)a - \overline{F}_{k-1}(pq)b$, as desired. To complete our analysis of the $k = 4\ell + 1$ case, we must also show that the numbers at nodes γ_1 and γ_2 after $k + 1$ node-firings are $-\overline{F}_{k-1}(pq)a - q\overline{F}_k(pq)b$ and $p\overline{F}_k(pq)a + \overline{F}_{k+1}(pq)b$. To confirm this, it suffices to verify the following matrix identity:

$$\begin{pmatrix} -1 & 0 \\ p & 1 \end{pmatrix} \quad \begin{pmatrix} \overline{F}_{k-1}(pq)a + q\overline{F}_k(pq)b \\ -p\overline{F}_{k-2}(pq)a - \overline{F}_{k-1}(pq)b \end{pmatrix} = \begin{pmatrix} -\overline{F}_{k-1}(pq)a - q\overline{F}_k(pq)b \\ p\overline{F}_k(pq)a + \overline{F}_{k+1}(pq)b \end{pmatrix}.$$

This follows easily from the facts that $\overline{F}_k(pq) = -\overline{F}_{k-2}(pq) + \overline{F}_{k-1}(pq)$ and that $\overline{F}_{k+1}(pq) = -\overline{F}_{k-1}(pq) + pq\overline{F}_k(pq)$. This completes our analysis of the case $k = 4\ell + 1$. Analysis of the $k = 4\ell + 3$ case is entirely similar.

For the ERG, we have $\varepsilon = +1$ and $c_m = q_m$ for each m. As in the preceding paragraph, $k_{\mathrm{m}} = (k - 1)/2 = k_{\mathrm{z}}$. Then $\left(\varepsilon^{k_{\mathrm{z}}}\widehat{g}_{2k_{\mathrm{m}}+1}\right)a - \left(\varepsilon^{1/2}\varepsilon^{k_{\mathrm{z}}}\widehat{h}_{2k_{\mathrm{m}}+1}\right)b$ and $-\left(\varepsilon^{1/2}\varepsilon^{k_{\mathrm{m}}}\widehat{g}_{2k_{\mathrm{z}}}\right)a + \left(\varepsilon^{k_{\mathrm{m}}+1}\widehat{h}_{2k_{\mathrm{z}}}\right)b$, which simplify to $\widehat{g}_k a - \widehat{h}_k b$ and $-\widehat{g}_{k-1}a + \widehat{h}_{k-1}b$ respectively, confirming the first two of the four ERG-related identities in the corollary statement. From the computation

$$\begin{pmatrix} 1 & 0 \\ -q_{k+1} & 1 \end{pmatrix} \quad \begin{pmatrix} \widehat{g}_k a - \widehat{h}_k b \\ -\widehat{g}_{k-1}a + \widehat{h}_{k-1}b \end{pmatrix} = \begin{pmatrix} \widehat{g}_k a - \widehat{h}_k b \\ -\widehat{g}_{k+1}a + \widehat{h}_{k+1}b \end{pmatrix}$$

we confirm that the numbers resulting from the $(k + 1)$st application of the Division Algorithm in the ERG are as claimed.

For the DRG, we have $\varepsilon = +1$ and $c_m = -q'_m$ for each m. We also regard $a =$ 'r'_{-1}' $= 0$ and $b = r'_0$. For our odd number k, $k_{\mathrm{m}} = (k - 1)/2 = k_{\mathrm{z}}$. Then $\left(\varepsilon^{k_{\mathrm{z}}}\widehat{g}_{2k_{\mathrm{m}}+1}\right)a - \left(\varepsilon^{1/2}\varepsilon^{k_{\mathrm{z}}}\widehat{h}_{2k_{\mathrm{m}}+1}\right)b$ immediately simplifies to $-\widehat{h}_k r'_0$. Specializing $h_k = \sum_{j=0}^{(k-1)/2} \sum_{T \in O\mathcal{E}(k,j)} x_{T_1} \cdots x_{T_{k-2j}}$ by exchanging each x_m with $-q'_m$ gives us $\widehat{h}_k = \sum_{j=0}^{(k-1)/2} \sum_{T \in O\mathcal{E}(k,j)} (-1)^{k-2j} q'_{T_1} \cdots q'_{T_{k-2j}} = -h_k(q'_1, \ldots, q'_k)$, so $-\widehat{h}_k r'_0$ is just $h_k(q'_1, \ldots, q'_k)r'_0$. Also, $-\left(\varepsilon^{1/2}\varepsilon^{k_{\mathrm{m}}}\widehat{g}_{2k_{\mathrm{z}}}\right)a + \left(\varepsilon^{k_{\mathrm{m}}+1}\widehat{h}_{2k_{\mathrm{m}}}\right)b$ immediately simplifies to $\widehat{h}_{k-1}r'_0$. Setting $x_m := -q'_m$ in $h_{k-1} = \sum_{j=0}^{(k-1)/2} \sum_{T \in O\mathcal{E}(k-1,j)} x_{T_1} \cdots x_{T_{(k-1)-2j}}$ yields $\widehat{h}_{k-1} = \sum_{j=0}^{(k-1)/2} \sum_{T \in O\mathcal{E}(k-1,j)} (-1)^{(k-1)-2j} q'_{T_1} \cdots q'_{T_{(k-1)-2j}} = h_{k-1}(q'_1, \ldots, q'_{k-1})$. This means that $\widehat{h}_{k-1}r'_0 = h_{k-1}(q'_1, \ldots, q'_{k-1})r'_0$. Our work so far in this paragraph confirms the first two of the four DRG-related identities in the corollary statement; the remaining two identities follow from entirely similar computations. □

6 Is Bézout's Lemma a Constructive or Nonconstructive Result?

The result traditionally known as Bézout's Lemma is the following assertion:

Lemma 6.1 (Bézout's Lemma) *Given integers a and b, not both of which are zero, there exist integers s and t such that* $\gcd(a, b) = as + bt$.

Integers s and t satisfying the existence claim of Lemma 6.1 are called *Bezout coefficients.* One of our motivations for presenting oereo polynomials is to showcase how they yield a straightforward and "constructive" proof of Bézout's Lemma by giving us immediate and transparent formulas for Bézout coefficients.

A Constructive Proof of Lemma 6.1 In the notation of part 2 of Theorem 4.6, we see that one pair of Bézout coefficients is $s = (-1)^n g_{n-1}(q_1, \ldots, q_{n-1})$ and $t = (-1)^{n-1} h_{n-1}(q_1, \ldots, q_{n-1})$. □

Although the result of Lemma 6.1 is attributed to Étienne Bézout (1766), a version of it was discovered more than a century prior by Claude Gaspard Bachet (1624). Working within the context of the Euclidean Algorithm from Euclid's Book VII, Bachet used symbolic algebra to describe a general solution to the Diophantine equation $ax - by = 1$, when a and b are co-prime (i.e., $\gcd(a, b) = 1$), by backtracking through the quotients and remainders produced by the steps of the algorithm. But it was with Bézout's work that the more general notion of solving $ax + by = \gcd(a, b)$ over the integers was recognized as a key result within elementary number theory (see, for example, the wonderfully readable account in Bullynck (2009) of the European history of such remainder problems in the centuries preceding Carl Friedrich Gauss).

But what might be practical about Bézout's Lemma? Here is a modern application. Cryptosystems often utilize modular arithmetic and, in particular, can require the computation of multiplicative inverses within this environment. Suppose, for example, we know that a and b are co-prime and wish to compute $\overline{b}^{-1}$ within the quotient ring $\mathbb{Z}_a := \mathbb{Z}/a\mathbb{Z}$, if such an inverse exists at all. (Here, $\overline{x}$ denotes the coset $x + a\mathbb{Z}$ in $\mathbb{Z}_a$.) Well, write $1 = as + bt$, where (s, t) is a pair of Bézout coefficients. Then $\overline{1} = \overline{as + bt} = \overline{a}\,\overline{s} + \overline{b}\,\overline{t} = \overline{b}\,\overline{t}$, so $\overline{b}\,\overline{t} = \overline{1}$. That is, $\overline{b}$ has a multiplicative inverse in $\mathbb{Z}_a$, and $\overline{b}^{-1} = \overline{t}$.

Given such computational contexts, it is curious that contemporary treatments of Bézout's Lemma are often nonconstructive. At least, this is our experience based on the five discrete-oriented textbooks we have used in recent years in undergraduate mathematics courses – by Burton (2010), Gallian (2015), Richmond and Richmond (2004), Rosen (2012), and Sundstrom (2022). Among these texts, only Gallian (2015) has no treatment of the Euclidean Algorithm. Only Burton (2010), Richmond and Richmond (2004), and Sundstrom (2022) consider linear Diophantine equations in two variables, which are worth noting since, of course, knowledge of a particular

pair of Bézout coefficients is needed to formulate a general solution to the Diophantine equation $ax + by = \gcd(a, b)$. Nonconstructive treatments of Bézout's Lemma can be found in Theorem 2.3 of Burton (2010), in Theorem 0.2 of Gallian (2015), and in Theorem 4.3.6 of Rosen (2012) (with a proof of the latter outlined in §5.2 Exercise #36). These nonconstructive approaches appeal to the Well-Ordering Principle of the positive integers, which asserts, without any computational specificity, that *any nonempty set of positive integers contains a smallest integer.* Here is a typical argument.

A Nonconstructive Proof of Lemma 6.1 Assuming a and b are not both zero, then $\{ax + by \mid x, y \in \mathbb{Z} \text{ and } ax + by > 0\}$ is nonempty and we may, by invoking the Well-Ordering Principle, declare d to be the smallest element in this set. Clearly $\gcd(a, b) \mid d$, so $\gcd(a, b) \leqslant d$. To see that d is a divisor of a, write $a = dq + r$ with $0 \leqslant r < d$, whence $r = dq - a = a(xq - 1) + b(yq)$, forcing $r = 0$. We similarly see that $d \mid b$, and we may conclude now that $d \leqslant \gcd(a, b)$. So, $d = \gcd(a, b)$. □

The "transitions" textbook by Sundstrom (2022) states a version of Bézout's Lemma in Theorem 8.8, but without proof (although an induction argument is indicated). Within our sample, only the text by Richmond and Richmond (2004) offers a constructive proof of Bézout's Lemma; this is done by providing an iterative procedure for computing Bézout coefficients as part of a correctness proof for the Euclidean Algorithm (see the proof of Theorem 3.2.1 in that text).

Curiously, none of our five sample texts gives a constructive proof of the Division Algorithm, and all of them follow the convention stated in Rosen (2012):

> We use the traditional name for this theorem even though the division algorithm is not actually an algorithm.

Theorem 6.2 below offers a typical statement of this theorem along with a typical nonconstructive proof of the theorem's existence claim (we take the uniqueness claim to be routine). While this theorem is conventionally known as the "Division Algorithm," we believe it ought to be called the "Quotient and Remainder Theorem."

Theorem 6.2 (Quotient and Remainder Theorem aka the Division Algorithm) *Given integers a and b with $b \neq 0$, there exist unique integers q and r such that $r \in [0, |b|)_{\mathbb{Z}}$ and $a = bq + r$.*

A Nonconstructive Proof of Existence in Theorem 6.2 Clearly the set $\{a - bk \mid k \in \mathbb{Z} \text{ and } a - bk \geqslant 0\}$ is nonempty. By the Well-Ordering Principle, it therefore contains a smallest element r of the form $a - bq$, from which we easily deduce that $0 \leqslant r < |b|$. □

It seems that a significant motivation for such nonconstructive approaches to the Division Algorithm is to present an early example of a use of the Well-Ordering Principle. By contrast, Sundstrom (2022) treats the Division Algorithm as an axiom,

in part for the reason that the natural numbers are not formally defined in his text and therefore the Well-Ordering Principle is regarded as unavailable (2022, §3.5, §4.1, and §8.1).

However, there seems to be merit in viewing the Division Algorithm as an actual algorithm. Such a process is intuitive and easy to articulate when a and b are positive integers with $a \geqslant b$: Repeat the process of subtracting b from a until the difference is smaller than b, and then the quotient q is the number of subtractions performed and the remainder r is the difference $a - qb$. Allowing the input integers a and b to be more generic makes the algorithm only slightly more involved, but this is relatively easy to handle. For any integer m, we let $\operatorname{sgn}(m) \coloneqq \begin{cases} |m|/m & \text{if } m \text{ is nonzero} \\ 0 & \text{if } m = 0 \end{cases}$. Note that $\operatorname{sgn}(|m| - m) = \mathbf{1}_{(-\infty,0)_{\mathbb{Z}}}(m)$, where $\mathbf{1}_{(-\infty,0)_{\mathbb{Z}}}$ denotes the so-called "indicator function" for the negative integers as a subset of $\mathbb{Z}$.

Definition 6.3 The *Division Algorithm (DA)* is the following pseudocode procedure:

Input: Integers a and b with $b \neq 0$.
Initialization: Let k be our iteration-counter, also to be viewed as a multiplier for $|b|$ and initialized with $k \coloneqq 0$. Let c be the difference between $|a|$ and the current multiple of $|b|$, initialized with $c \coloneqq |a|$.
Step 1: If $c < |b|$, let $q \coloneqq \operatorname{sgn}(ab) \cdot \left(k + \mathbf{1}_{(-\infty,0)_{\mathbb{Z}}}(a)\right)$ and $r \coloneqq a - bq$, and Return the following data: QUO $\coloneqq q$ and REM $\coloneqq r$.
Step 2: Otherwise, replace k with $k + 1$ and c with $c - |b|$, and return to Step 1.
Output: An integer QUO to be called the *quotient when a is divided b* and an integer REM to be called the *remainder*.

(We use the notation DA(a, b)[QUO] to denote the quotient returned by the Division Algorithm and DA(a, b)[REM] to denote the remainder. It must be shown that DA(a, b)[REM] is in the interval $[0, |b|_{\mathbb{Z}})$.) □

As we did with the Euclidean Algorithm, we certify in the following theorem that the DA terminates and prove that it returns the desired output. In particular, this theorem provides a constructive proof of the existence claim in Theorem 6.2.

Theorem 6.4 (Boundedness and Correctness of the Division Algorithm) *Let a and b be integers with $b \neq 0$. With this input, the Division Algorithm terminates, and for the output values $q \coloneqq$ DA(a, b)[QUO] and $r \coloneqq$ DA(a, b)[REM], we have $a = bq + r$ and $r \in [0, |b|)_{\mathbb{Z}}$.*

Proof The "current difference" measure c takes on, in succession, the values in the sequence $(|a|, |a| - |b|, |a| - 2|b|, \ldots)$. The algorithm terminates when $c < |b|$. Since the inequality $|a| < (|a| + 1)|b|$ is true for all integers a and b with $b \neq 0$, then $|a| - |a|$ $|b| < |b|$, confirming that there is some number in the sequence $(|a|, |a| - |b|, |a| - 2|b|,$

...) that is less than $|b|$. That is, the algorithm will terminate in no more than $|a|$ iterations.

Since $r \coloneqq \mathrm{DA}(a, b)[\textsc{rem}]$ is defined by the algorithm to be $a - b(\mathrm{DA}(a, b)[\textsc{quo}]) = a - bq$, then we get $a = bq + r$ automatically. So, we only need to check now that $r \in [0, |b|)_{\mathbb{Z}}$. It suffices to analyze the following four cases: $(++)$ $a \geqslant 0$ and $b < 0$, $(+-)$ $a \geqslant 0$ and $b < 0$,$(-+)$ $a < 0$ and $b > 0$, and$(--)$ $a < 0$ and $b < 0$. For case $(+-)$, let k be the value of our counter when $a - k|b| < |b|$ but $a - (k-1)|b| \geqslant |b|$. In this case, $q = \operatorname{sgn}(ab)\left(k + \mathbf{1}_{(-\infty,0)_{\mathbb{Z}}}(a)\right) = (-1)\ (k + 0) = -k$. Then $|b| > a - k\,|b| = a + q|b| = a + q(-b) = a - bq = r$. Also, $0 \leqslant a - (k-1)\,|b| - |b| = a - k\,|b| = a + q|b| = a + q(-b) = a - bq = r$. The $(++)$,$(-+)$, and$(--)$ cases can be analyzed similarly. □

7 Conclusion

There are many possible ways to extend the foregoing discussions. For example, we might play the Euclidean Remainders Game with three or more positive integers. What should be our requirement for termination of the game? Will every game terminate? What might "strong convergence" mean in this context? In another direction, we might consider oereo sequences and polynomials relating to the right-triangular *Lucas array* of Fig. 8, cf. OEIS entry A034807 (2023). To this end, we define *Lucasian oereo sequences* as the following sets, when $n \in [2, \infty)_{\mathbb{Z}}$: We let $\mathcal{OE}^{\mathrm{Lucas}}(n) \coloneqq \{T \in \mathcal{OE}(n+1) \mid T \text{ contains at least one of 1 or } n+1\}$, and we let $\mathcal{EO}^{\mathrm{Lucas}}(n) \coloneqq \{T \in \mathcal{EO}(n+1) \mid T \text{ contains at least one of 2 or } n+1\}$. Then define *Lucasian oereo polynomials* as

$$k_n \coloneqq \sum_{T \in \mathcal{EO}^{\mathrm{Lucas}}(n)} \mathrm{x}_T \quad \text{and} \quad l_n \coloneqq \sum_{T \in \mathcal{OE}^{\mathrm{Lucas}}(n)} \mathrm{x}_T.$$

What can be said about these Lucasian oereo polynomials?

$$\begin{array}{ccccc}
2 & & & & \\
1 & & & & \\
1 & 2 & & & \\
1 & 3 & & & \\
1 & 4 & 2 & & \\
1 & 5 & 5 & & \\
1 & 6 & 9 & 2 & \\
1 & 7 & 14 & 7 & \\
1 & 8 & 20 & 16 & 2 \\
1 & 9 & 27 & 30 & 9 \\
 & & \text{ETC.} & &
\end{array}$$

Fig. 8 Depiction of some rows of the right-triangular Lucas array

Our effort to better understand the historical record concerning Terquem's problem led us to connect his ideas with Fibonacci numbers, a one-player game of numbers, continuants, the Euclidean Algorithm, and Bézout's Lemma. In particular, we have a newfound, or at least now-unforgotten, interpretation of the numbers of the Fibonacci array and combinatorially explicit descriptions of continuants as oereo polynomials. All these ideas have antecedents in the literature, but, taken together as presented here, we believe they provide an explicit and combinatorially compelling account of these elementary number theoretic notions that might be advantageously accommodated in introductory discourses on these topics.

References

Bachet CG (1624) Problèmes plaisants et délectables, qui se font par les nombres. Partie recueillis de divers autheurs, partie inventez de nouveau avec leur demonstration, Second and enlarged edn, Rigaud, Lyon

Bézout E (1766) Cours de Mathématiques à l'usage des Gardes du Pavillon et de la Marine. Troisième partie: L'algèbre, Musier, Paris

Björner A, Lovász L, Shor P (1991) Chip-firing games on graphs. Eur J Comb 12:283–291

Bullynck M (2009) Modular Arithmetics before C.F. Gauss: systematizations and discussions on remainder problems in 18th-century Germany. Historia Mathematica (Elsevier) 36:48–72

Burton DM (2010) Elementary number theory, 7th edn. McGraw–Hill, New York

Cartan E (1894) Sur la structure des groupes de transformations finis et continus. Doctoral thesis,. Libraire Nony, Paris

Coleman AJ (1989) The greatest mathematical paper of all time. Math Intell 11:29–38

Donnelly RG (2008) Eriksson's numbers game and finite Coxeter groups. Eur J Comb 29: 1764–1781

Donnelly RG (2022) Finite diamond-colored modular and distributive lattices with applications to combinatorial lie representation theory. arXiv research monograph available at https://arxiv.org/abs/1812.04434v3. Accessed 27 June 2023

Donnelly RG, Dunkum MW, Huber M, Knupp L (2021) Sign-alternating Gibonacci polynomials. ECA 1(2):article # S2R15

Eriksson K (1993) Strongly convergent games and Coxeter groups. PhD thesis,. KTH, Stockholm

Eriksson K (1994) Node firing games on graphs. Contemp Math 178:117–127

Eriksson K (1995) The numbers game and Coxeter groups. Discret Math 139:155–166

Eriksson K (1996) Strong convergence and a game of numbers. Eur J Comb 17:379–390

Euler L (1764) Specimen algorithmi sigularis. Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae 9:53–69

Gallian JA (2015) Contemporary abstract algebra, 9th edn. Cengage, Boston

Hoggatt VE Jr, Bicknell M (1973) Roots of Fibonacci polynomials. Fibonacci Quart 11:271–274

Killing W (1889) Die Zusammensetzung der stetigen, endlichen Transformationsgruppen, zweiter Theil. Mathematische Ann 33:1–48

Knuth DE (1997) The art of computer programming, volume 2: Seminumerical algorithms, 3rd edn. Addison–Wesley professional (an imprint of Pearson Education), Carmel

Koshy T (2018) Fibonacci and Lucas numbers with applications, vol 1, 2nd edn. John Wiley & Sons, Hoboken

Lucas E (1878) Théorie des fonctions numériques simplement périodiques. Am J Math 1:184–196

MacKay JS (1892) History of the nine-point circle. Proc Edinb Math Soc 11:19–61

Mansour T, Munagi AO (2012) Alternating subsets modulo *m*. Rocky Mountain J Math 42: 1313–1325

Netto E (1901) Lehrbuch der Combinatorik. Verlag von B. G. Teubner, Leipzig

OEIS Foundation Inc. (2023) The on-line encyclopedia of integer sequences founded in 1964 by N. J. A. Sloane. Available at http://oeis.org. Accessed 27 June 2023

Richmond B, Richmond T (2004) A discrete transition to advanced mathematics. American Mathematics Society, Providence

Riordan J (1958) An introduction to combinatorial analysis. John Wiley & Sons, New York

Rosen KH (2012) Discrete mathematics and its applications, 7th edn. McGraw–Hill, New York

Singer I (1906) Terquem, Olry. In: Singer I et al (eds) The Jewish encyclopedia, vol 12. Funk and Wagnalls, New York, p 111

Stanley RP (1986) Enumerative combinatorics, vol 1. Wadsworth and Brooks/Cole, Monterey

Sundstrom T (2022) Mathematical reasoning: writing and proof version 2.1. Creative commons license BY-NC-SA. Available at the website below: https://www.tedsundstrom.com/mathematical-reasoning-writing-and-proof. Accessed 27 June 2023

Terquem O (1839) Sur un symbole combinatoire d'Euler et son utilité dans l'Analyse. J Math Pures Appl 4:177–184

van der Waerden BL (1975) Science awakening I, English translation by Arnold Dresden with additions of the author, 4th edn. Kluwer Academic Publishers, Dordrecht

The Ethics of Mathematical Practice

Paul Ernest

Contents

1 Introduction: A Vignette 1220
2 Philosophies of Mathematics 1221
3 Where Is Mathematics? (M)athematics Versus (m)athematics 1223
4 The Myth of Mathematics as a Unique and Unified Subject 1224
5 Epistemological Versus Social Models of Mathematics 1225
6 What Is Mathematical Practice? 1226
6.1 The Ethics of Mathematical Practices 1229
7 The Personal Ethics of Mathematicians 1230
8 Ethical Theory 1232
8.1 Medical Ethics 1232
9 Ethics in Science 1233
10 Social Responsibility of Science 1234
11 Social Responsibility of Mathematics 1235
12 Ethics in Mathematics 1236
12.1 Mathematics and Purity 1238
13 Virtue Theory and Mathematics 1242
14 Responsibility for Applications of Mathematics 1244
14.1 Mathematics in Public Communications 1244
14.2 Problematic Applications of Mathematics 1246
14.3 The Performativity of Mathematics 1247
15 Conclusion 1250
References 1251

Abstract

This chapter examines the role and need for ethics in mathematical practice. Mathematics is one of the few areas of study in which ethics is widely perceived as irrelevant. Many mathematicians and others resist the idea that we need to consider the ethics of both pure and applied mathematics. The foundations of this

P. Ernest (✉)
University of Exeter, Exeter, UK
e-mail: p.ernest@exeter.ac.uk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_9

resistance are analyzed and located in background philosophies and ideologies of purism and neutrality. The range of social practices is investigated, and different ethical problems and issues are brought to light. It is argued that virtuous mathematicians can legitimately pursue mathematics for its own sake, but as citizens they also have a responsibility to care about the social impacts of mathematics. A review of the literature on the social responsibility of science and mathematics reveals that, although long neglected, concerns about the ethics of mathematics are starting to emerge in publications and training practices. Some of the more ethically sensitive areas are explored and three problematic categories are distinguished and exemplified. These are (1) mathematics in public communications, (2) overt applications of mathematics with powerful social impacts, and (3) the hidden performativity of mathematics in restructuring society, institutions, and social practices.

Keywords

Ethics of mathematics · Absolutism · Social constructivism · Purism · Social responsibility · Mathematical practice · Social applications · Performativity

1 Introduction: A Vignette

An artist, a physicist, and a mathematician are interviewed for a late night TV show. The compere asks each of them two questions. What is the point of your practice? What are your social responsibilities?

The artist begins: "The point of art is to make artefacts or performances that explore the domain of beauty, and that push my craft to the limit in trying to achieve beauty. But it is also to explore the boundaries of what is art, and push them back without worrying about transgressing the limits of polite society approval. In short, to pursue art for art's sake. The social responsibility of art is merely to pursue these missions fearlessly without worrying about morality and ethics. Artistic expression is part of freedom, and this freedom should not be limited by bourgeois taste or the dead hand of the past."

The physicist answers: "The point of physics is to observe, describe and map the world in terms of its basic physical components and processes. We build theories that describe as accurately as possible the behaviour of matter, energy, space and time. Our overwhelming duty is to truth, and so we test all our theories with crucial experiments that try to falsify them, in order to approach truth more closely. Our social responsibilities arise from the power of our theories. Because we can predict and hence control the huge forces of nature and the world, including electromagnetic, nuclear, and gravitational energy, we have to be cautious and aware of the applications of physics and through it, technologies. We pursue basic theories to the limit, but it is right that scientists and society are aware of and control the impacts of applications on the social and material worlds. We have a duty to truth, but also to the good."

The mathematician answers: "The point of mathematics is to pursue mathematical truth. This is truth about the perfect world of mathematical objects, which are

timeless, perfect and exist in their own realm, untouched by material changes and decay. The truth we pursue can be exquisitely beautiful, but we are merciless in our rigour and only the most robust proofs can be accepted as demonstrating the truth of our results. Our responsibility is to pursue mathematical knowledge and beauty beyond all earthly bounds, knowing that we enrich human culture inestimably. Although our responsibilities end with the pursuit of mathematics for its own sake, it often turns out that the most abstract theories have surprising and very fruitful applications in science, technology and in modelling the physical world. But this is a happy accident, discovered only after we have revealed the truth. We pursue truth and beauty, but the only good we have to worry about is being good *mathematicians*."

Is this right? Does mathematics only have to concern itself with truth and beauty and not with the good of humanity in a broader sense? Has mathematics no social responsibilities, no debt to the society that nurtures and funds it? Is ethics irrelevant to pure mathematics, as many mathematicians think? In order to answer these questions I need to clarify the main philosophical answers to the question *what is mathematics*? There are also important subsidiary questions *where is mathematics*? And *what is mathematical practice?* which also have a profound bearing on ethical issues.

2 Philosophies of Mathematics

One of the major determinants of belief in whether ethics is relevant to mathematics is the philosophy of mathematics held by a mathematician. There is a range of opinions, beliefs, and schools of thought about what mathematics is that can be simplified into a dichotomy between absolutist and social constructivist philosophies. Here there is an enduring controversy over whether pure mathematics is objective and universal, or whether it is a humanly made construction (Hersh 1997; Ernest 1998). Both sides of this dispute have legitimacy, but they are incompatible philosophies of mathematics. The two sets of views are as follows.

First, there are what can be termed absolutists, who subscribe to the view that mathematics is an objective, universal, and absolute science.

Absolutists claim that the following are true about mathematics:

1. Mathematics is universal and unique, and all systems of mathematics can be translated into modern classical mathematics.
2. Mathematics is invariant and timeless and that mathematical concepts apply and mathematical results hold in all places and at all times. Even results that are several millennia old, such as Pythagoras' theorem and Euclid's proof of the infinity of prime numbers, are still valid.
3. Mathematics is objective, as there are fixed rules for determining its truths independent of subjective opinion or differing cultural outlooks.
4. Mathematics is discovered rather than invented. This is illustrated by simultaneous discoveries, and by the rich interconnectedness and overall unity of the disparate branches and topics of pure mathematics, no matter how far they are independently developed.

5. Above all else, Mathematics is based on an objective, universal, eternal, and absolute conception and definition of truth.
6. Values and ethics are irrelevant to pure mathematics, for they do not impinge on its zone of objectivity.

Absolutists and Platonists in the philosophy of mathematics, who adhere to some or all these views, are many and include the following notable mathematicians and philosophers: Paul Erdős and Kurt Gödel (Wikipedia 2020a), Hardy (1929), Penrose (1989), and Balaguer (1998).

In contrast, the opposing social constructivist philosophy, a minority view, takes the following position about mathematics:

1. Mathematics only exists as a complex product and element of social practices, and thus in definite material form comprises human activities and their productions: for example, inscriptions, texts, institutions, and agreements (Cole 2013).
2. Mathematics is not a unique and unified discipline, for mathematics is made up of many different and changing specialist social practices.
3. Classical mathematics as represented in standard texts is not the only possible science of mathematics, and there exist alternatives that are equally real-world applicable such as constructive mathematics.
4. Mathematics is not forced upon us through necessity, neither empirically nor logically. But once we accept the premises of a theory and its rules, then the results follow by necessity.
5. Mathematics rests upon definitions of truth, logic, reasoning, and proof that are not necessary, universal, or invariant across time and space, although they are regarded as so within mathematical practices.
6. Mathematics is a historical, social construction and is thus contingent on historical actions and human choices and interests, and is neither necessary nor arbitrary.
7. Mathematics is a human invention imbued with the values and interests of its makers and hence is not above or beyond ethical values (Ernest 1998; Hersh 1997).

As an example of the values mentioned in (7), pure mathematics may be claimed to be open and democratic because the calculations and proofs upon which its knowledge and applications depend must be open to scrutiny by others. Unlike authority-based areas such as religion, mathematical claims can only be founded on open, accessible, and persuasive demonstrations. Thus, mathematics is necessarily open and democratic, and these are values greatly prized in Western and other societies. (Harris (2015) argues that mathematics is caught on the horns of a dilemma. It is both democratic, as I argue here, and anti-democratic and elitist, because only a privileged few judge the correctness or value of new knowledge contributions.) However, this argument assumes that truth depends on its demonstrations. Absolutists argue that the truth of a theorem is absolute and precedes its human verification. Proofs are devices to demonstrate this truth to mathematicians. So this example does not represent a definitive way of choosing between the two competing philosophies of mathematics.

The contrast between the views of Absolutists and Platonists, on the one hand, and social constructivists, on the other hand, is brought into sharp focus in the controversy over whether mathematics is discovered or invented. For example, consider the field of Number Theory. This is a well-defined topic in mathematics that rests since Peano's formulation on a single element (0, in modern formulations), a single unary operation (S, the successor function), a single binary relation (=), and a handful of axioms (and a background logic and with a metatheory that allows the definition of further functions). On the basis of this rather limited foundation, the huge edifice of modern Number Theory has been erected, encompassing results stretching from Andrew Wiles' solution to the Fermat Conjecture to Gödel's Incompleteness Theorems. These results admittedly draw on concepts from beyond Number Theory, in the strict sense, in their proofs.

Results such as these are constructed with human ingenuity, including the invention of many subsidiary concepts, functions, and methods, within the constraints imposed by working within Number Theory with its axioms and rules of proof. Thus, they can be seen as marvelous products of spontaneous human invention. They can also be viewed as results discovered by an explorer mathematician within the pre-existing field of number once the gate to Number Theory has been opened for human access. They can be viewed as discoveries within the pre-existing domain of mathematical knowledge.

3 Where Is Mathematics? (M)athematics Versus (m)athematics

Some authors like to distinguish between what is termed *Mathematics* and *mathematics* (Bishop 1988). Mathematics with a capital *M* is the sum or body of formal mathematical knowledge and mathematical objects that exist independently of mathematicians. The lower case *m* in the term mathematics signifies that it refers to and comprises human mathematical practices. The distinction is an ontological one, about existence. Mathematics with a capital M is claimed to be real, made up of independently existing concepts, objects, proofs, truths, and theories. This is the traditional realist ontological position, as was traditionally opposed to Conceptualism (mathematics is made up of concepts that exist only in people's minds) and Nominalism (mathematics consists only of linguistic names, signs, and symbols).

Two forms of realism can be distinguished. First, there is Social Realism, which understands that Mathematics is something created by people but which has an existence beyond any individual or group of persons, but not beyond the sum of human culture and history. Thus, the answer to *where is Mathematics?* is that it is located in the cultural matrix of human activity embedded in the human society and history. Mathematics is culturally real, but does not exist independently of humans and human activity, and in particular, the social practices of mathematics (Ernest 1998; Hersh 1997).

Second, there is Platonism, which claims that Mathematics is real, but it exists independently of humanity and is to be found in some abstract, superhuman realm.

According to this view, mathematics is timeless, having always been true, and objective, not affected by human actions, let alone by human subjective beliefs. Thus, this abstract Platonic realm provides the answer to the question *where is Mathematics?* Platonists subscribe to the view that mathematical objects and truths are real and exist in an independent supraphysical realm to which mathematicians have access and from which they can gain knowledge.

It is of course possible to be a realist without subscribing to social constructivism or Platonism, by simply asserting that mathematics and its objects are real without specifying where mathematics is to be found.

It is clear that the question of where mathematics is has a direct bearing on the issue of ethics. For if mathematics is of another world, as in Platonism, then ethical and other human concerns are irrelevant. But if mathematics is socially located, then it is imbued with values and ethics like all other human activities and practices.

A further relevant question is whether mathematics is a single identifiable entity. Is it a unified collection of mathematical knowledge and mathematical objects (*Mathematics*), or is it a social practice or set of social practices (*mathematics*)?

4 The Myth of Mathematics as a Unique and Unified Subject

Ever since Plato, or before, humans have been inclined to view nouns as denoting some real or ideal object. London is a real place, albeit with vague boundaries. But to extend this view to all nouns – proper, general, or abstract – is a linguistic and philosophical fallacy. Apart from concrete particulars, far from only naming pre-existent objects, nouns create fictive objects that correspond to their names, such as Harry Potter, the Equator, or as I wish to argue here Mathematics. So what can the term "mathematics" denote? Epistemologically, it denotes a body of knowledge. But a body is a unified and connected material corpus, and there is no organized and unified system of mathematical knowledge. Within academic mathematics, just pure mathematics alone refers to hundreds of subfields and topics (Davis and Hersh 1980; Lerman 2010) that at best share a family resemblance (Wittgenstein 1953). There is no shared set of symbols, concepts, theories across most, let alone all, of mathematics. The study and theories of number constitute a central theme in the history of mathematics. But there are plenty of mathematical theories that do not use number in any essential way, such as Euclidean geometry.

A further question is whether all of mathematics uses the same methods of verification of its truth, the same system of proof? But not only have proof styles varied greatly over the past two millennia plus since Euclid, but many different proof styles and method coexist simultaneously in different branches of and specialisms within mathematics. Consider only the differences between Number Theory, Classical Analysis, Axiomatic Set Theory, Category Theory, Topology, and Non-standard Analysis. There is no central set of shared symbols, proof methods, or rhetorical styles (Knuth 1985). Thus, mathematics is made up of a family of overlapping but disparate practices.

This is bad enough. But the disparate collection of subfields, topics, and practices that make up pure mathematics far from exhausts mathematics. Beyond pure mathematics there are to be found practices of applied mathematics, theoretical physics, statistics, computer mathematics, school mathematics, accountancy mathematics, banking mathematics, the mathematics of trading and the markets, mathematics in economic theory, street mathematics, ethnomathematics, numerology, and so on. We cannot say that these all share the same basic ideas. Consider number again. There are many different systems and structures we label as number (this term itself is another fiction of unified referent). But even if we allow that number is a unified set of concepts for counting and ordering, there are areas of knowledge constituted by patterning, geometry, argumentation that are mathematics but do not involve number (Bishop 1988).

So there are no universally shared concepts of mathematics, and there is no central set of shared proof methods across all the specialisms within mathematics. Mathematics lacks epistemic and even epistemological unity. Looking at mathematics throughout history and across cultures, present and past, there are huge variations in what we term mathematics and its methods of justification.

Culturally, in terms of social practices, there is even less reason to claim that mathematics constitutes a unity. For in addition to all the above disparate knowledge fields, it cannot be said that mathematics originates in a single or unified set of mathematical practices. According to Bishop (1988), all of mathematics evolves from six basic areas of human activity which occur in all societies. These are counting, locating, measuring, designing, playing, and explaining. These practices do not even share a family resemblance, and since what we call mathematics originates in these activities, and no subsequent total unification takes place, mathematics remains a multiplicity, a plurality, and a motley of disparate knowledge fields, subspecialisms, and practices (Wittgenstein 1956; Lerman 2010).

But mathematics does exist. Instead of being a unified referent, a single body of knowledge, it is a vast and growing collection of different ideas and practices, some connected, some not, with a fuzzy boundary.

5 Epistemological Versus Social Models of Mathematics

Despite this critique, there is a traditional absolutist perspective of mathematics as a unity, a single entity, which is part of the absolutist ethos of mathematics. According to this perspective, the central paradigm for mathematics is represented by a skyscraper model. This has successive, rising stories made up of even more complex theories, based on and supported by those beneath, and built on the unshakeable foundations of mathematical axioms and logic at the bottom. This is a beautiful, powerful, and seductive image, subscribed to by Logicists, Formalists, and other Foundationalists in the philosophy of mathematics. But it fails epistemologically, as history has shown. For technical reasons, it fails to represent mathematics, because there is no unique logico-mathematical foundation for the whole of mathematics. Even if a candidate for one were to be found, it is impossible for this would-be foundation to be demonstrated to be safe and beyond all doubt, in other words, to

provide a proper foundation (Ernest 1991, 1998). Gödel's (1931) Incompleteness Theorems demonstrate that formal mathematics from arithmetic onward, in terms of increasing complexity, can never be complete nor shown to be demonstrably consistent.

In terms of social representation, the skyscraper metaphor also fails. In the world, mathematics has many locations and contexts. Some of the many social practices that make up mathematics were mentioned above. Elaborating further in this vein, an alternative to the skyscraper model, and offering a better representation of mathematics is the image of a dynamic, growing modern city.

Within such a city, we will find the following architecture, businesses, and districts:

1. Multiple skyscrapers, comprising the different formal theories in mathematics (some enduring and some demolished and replaced);
2. Universities and academic centers, representing different branches and practices of research mathematics; as well as other departments applying mathematics in physics, biology, psychological and social science research methods, computer science, and so on;
3. A business district, representing various applications of mathematics and accountancy, banking, stock market, cryptocurrency, and business mathematics;
4. Computers and the city-wide Internet, illustrating the increasingly important presence of computer-based mathematics, both in theory and in practice;
5. Shops, street markets, and *favelas*, representing informal mathematical practices on the streets, in the shops, and ethnomathematics (Probability coupled with finance even has its own street presence in betting shops (legal gambling premises on high streets in the UK), although probability, statistics, finance, and financial models are also central to banking and brokering businesses, including high-level actuarial models.) (D'Ambrosio 1985);
6. Schools, representing school mathematics and mathematics education.

The city model works epistemically, for it represents the different domains of knowledge, including both academic and informal, as well as the associated practices. It also works socially, indicating the conceptual geography, the varied societal locations, and human practices making up mathematics as it is performed in social practices. Thus, the city model captures the shape of mathematics far more accurately than the skyscraper model. Within it, I have used the idea of a mathematical practice. But what is a mathematical practice, one of the most central concepts that hold this book together?

6 What Is Mathematical Practice?

A social practice is an organized social activity that has an identity, members, and agreed forms of participation, both implicit and explicit, and has, above all, goals shared by all participants as the overarching focus. All aspects of a social practice can change over its lifetime, including membership, rules of active participation, the overall

structure, and even the goals. But if the goals change significantly, it becomes a different social practice, for they determine the whole point and direction of the practice.

In the context of ethics, Alasdair MacIntyre defines a social practice as follows:

> By a 'practice' I am going to mean any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realised in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended. (MacIntyre 2007: 111)

Social practices are always situated in space and time, in living, historical traditions. Mathematical practices are specialized social practices, and the city model described above illustrates the wide range of mathematical practices, which are so broad as to incorporate most, if not all citizens, of humanity. The practices can be classified into pure and applied, theoretical and practical, and with the mathematics utilized explicit or implicit. Categorizing them in this way helps to distinguish the main differences among the range of social practices of mathematics that exist.

1. *Pure theoretical practice with mathematics explicitly the main goal*
   A pure theoretical mathematical research practice with mathematics foregrounded as the explicit focus of research (e.g., research into group theory) is perhaps the first and paradigmatic mathematical practice that springs to mind. It is largely academic, comprising university-based researchers exploring the problems, theorems, methods, and theories of pure mathematics, individually or in groups and sharing their findings formally through journal publications and conference presentations.
2. *Applied theoretical practice with mathematics explicit*
   An applied theoretical mathematical research practice with mathematics foregrounded as the explicit focus of research (e.g., applied statistics research) is probably very similar to the pure mathematical research practices except for the types of problems addressed. One possible difference apart from the subject matter is that direct users of applied mathematics may consult these researchers over problems of application, such as using statistics in research methods. A further difference is the existence of commercial research tanks developing applied theory outside of the academy, as well as research conducted within it.
3. *Applied and theoretical practices with mathematics backgrounded*
   These are the academic practices of theory building and development in advanced physics such as quantum and relativity theory in which advanced mathematics may be backgrounded, but it remains present as an important language and tool. Such practices will often use mathematical concepts, models, and methods at extremely high levels and may even extend the mathematical theory employed.
4. *Pure practical practices with mathematics explicit*
   Pure practical mathematical practices in which mathematics is foregrounded can include teaching and writing school and college mathematical text-books,

together with the related activities of mathematics curriculum and examination development. In such practices, no existing mathematical theory is extended, but new ways of expressing, representing, and recontextualizing mathematical ideas and knowledge are developed (Dowling 1998).

5. *Pure and applied theoretical work, mathematics backgrounded*
   Such work includes the development of Apps and computer applications where mathematics is submerged and backgrounded simply as a set of tools, a necessary language in which the developers are fluent. This includes the development of the tools behind commercial apps, including the technical underpinnings of surveillance capitalism (Zuboff 2019), as well as those of social surveillance, regulation, and governance (O'Neil 2016). Increasingly, the public interfaces of governmental departments with the public are a series of Apps and web-based programs.
6. *Applied and practical practices, with mathematics explicit*
   This includes the practices of practical applied mathematics with mathematical knowledge explicitly foregrounded. It can include business mathematics, computer mathematics, actuarial statistics, and so on, where the novelty lies in the development of new practical applications.
7. *Pure and applied practical work, with mathematics backgrounded*
   This includes the practices of everyday or street mathematics conducted in businesses, shops, street trading, bank mortgage and loan calculations, architectural practices, and so on. Since virtually all businesses and commercial transactions are based on money, the mathematics of money, expressed in customer purchases and loans, as well as in financial spreadsheets, double-entry accounting are all part of the day-to-day running of businesses. The mathematics is submerged behind the purposes and goals of the social practice. On a larger scale, this also includes the practices of data-based businesses, surveillance capitalism, as well as data capture and mining in social surveillance, regulation, and governance.

I do not claim that this typology is definitive, for there may well be overlaps and indeed significant distinctions that are glossed over. For example, categories 3 and 5 represent applied theoretical practices with the mathematics backgrounded in pursuit of their different main goals, one within the academy and one without. Furthermore, there may be additional mathematical practices imperfectly covered by the typology including recreational mathematics and the production of mathematical or mathematically inspired art. Nevertheless, the listing serves to illustrate the existence of, and to distinguish between, the wide variety of different mathematical and mathematics-related practices both within the academy and outside of the academy across the breadth of social and cultural life. Each of these types of practice potentially raises different ethical problems and issues. However, for the purposes of this chapter, I shall focus on practices in which specialist mathematicians take the lead.

6.1 The Ethics of Mathematical Practices

In each specialist mathematical practice, I wish to distinguish four main domains of application of ethics, although more are possible.

(A) **Ethics of the institution**. First, there is the ethics of the institution that supports the mathematical research or development activities; how it treats the participants, and how it structures and regulates the social practice? There is also the issue of institutional ownership of intellectual property rights; who has ownership of any products of research? It may be left to their creators, or the institutions may take ownership and do marketing of or otherwise directly benefit in material terms from the products. In addition to the explicit ethical framework of the institution, there is the ethics of how the work is conducted in practice, for example, openness, participation in decision making, explicitness of rules, regulations, and contracts. Are mathematicians represented on the institutional bodies that make internal research grants, hire or promote mathematicians? (There is virtually no mathematics-specific literature on these first issues, presumably because the ethics of departments is understood to be independent of the discipline involved. Thus, codes of ethical conduct in mathematics such as AMS (2019) and EMS (2012) focus on individual ethics and responsibilities.)

(B) **Personal ethics of the participants**. Second, there is the personal ethics of the participants in the social practice the individuals involved. This concerns how they interact ethically with each other and with further persons outside of the practice but still within their professional orbit. For example, many if not most university mathematics departments involve teaching with both personal and institutional ethical obligations to students. Researchers and mathematical practitioners also have personal ethical obligations to both colleagues in the home institution and other international-based scholars in the specific field of mathematical practice with whom they collaborate or interact. Ethical rules of conduct include giving fair credit to colleagues, not plagiarizing others' work, and reviewing and refereeing fairly without fear, favor, or other inequities. Such ethical norms and how they are respected, monitored, and maintained are vital, and affect good practices within the institutions as well as affecting the individual mathematicians' work.

(C) **Explicit ethics of the content or subject matter**. Third, there is the ethics of the subject matter, the content and focus of the research and development. This is the ethics of the mathematical products, pure or applied, that emerge from any research and development conducted. To what extent are the ethical implications of these products considered explicitly and do ethical values and considerations play a part in shaping them? Do pure theoretical developments have easy to anticipate uses or possible applications outside of the practices? What are the ethical dimensions of this? What are the potential social and practical outcomes of applied projects and developments?

(D) **Hidden ethical outcomes**. Fourth, there are the unforeseen, implicit, or hidden ethical outcomes of the products of research and development. Pure developments in mathematical theory may have unanticipated ethical applications. Advances in pure Number Theory such as prime numbers have been seen to have implications for national and personal security, cryptography, cryptocurrencies, and the like. Applied developments such as the 2020 algorithm for adjusting student grades during the Covid lockdown in the UK (Ofqual 2020) had a significant negative impact on the grades of students from traditionally lower attaining examination centers, and were rejected after public outrage. Who is responsible for such unanticipated negative outcomes, causing disadvantage? Who takes responsibility for such outcomes, if anyone? Issues of social justice that may arise are of necessity also ethical issues.

This last is the most controversial area since it depends on philosophical views over whether mathematical knowledge products are value-laden including ethical values, as well as the ethics of applications of mathematics. If a mathematician is working on a faster algorithm or process for factorizing large numbers, can this be viewed as pure values- and ethics-free research? Given how codes based on large prime numbers and factorization contribute to military security, privacy, and financial security, can a procedure that enables faster identification of prime divisors still be regarded as ethically neutral?

7 The Personal Ethics of Mathematicians

Above, in subsection B, I raised the issue of the personal ethics of the participants in the social practices of mathematics. The issues mentioned go beyond the ethics of mathematics because they also have epistemic significance. Mathematics together with all the sciences depends on the trust and trustworthiness of mathematicians and scientists. The construction of mathematical and scientific knowledge goes through the phases or origination, testing and expression (write-up), submission to review and peer scrutiny, followed by publication in recognized and approved channels. At every phase, virtuous intention and honest procedures and actions are both expected and necessary for the good – that is the ethical as well as the effective – production and development of knowledge. Once knowledge productions are published, they gain the status of approved and warranted knowledge, at least provisionally, if not absolutely. Thus, trust and ethics are part of the knowledge warranting procedures.

Of course, there are epistemic and rhetorical safeguards built in. Scientific results must be accompanied by full accounts of the methodology and data gathering conducted so that the observations or experiments can be replicated. Mathematical claims must be substantiated by full and detailed proofs of the results so that they can be verified by expert mathematicians. Approved formats must be followed for the submissions, and these rhetorical requirements also help to maintain scientific standards and bring out the key steps in the warranting of knowledge.

Scientific write-ups, including mathematical papers, must be submitted to peer review and revised and ultimately evaluated so as to satisfy referees and editors, before their publication is sanctioned. The submitted account be as transparent and honest as possible, within the required rhetorical framework. In addition, the assessors and evaluators of any submission are expected to review and referee it fairly, and often blindly, to conceal the identity and status of authors. Normally, several referees judge submissions independently, and the editor weighs up the balance of their critical reports to deliver a verdict of rejection, required improvements, or acceptance. All such judgments are expected and required to be given dispassionately, without fear of the power or the desire to curry favor from established and senior mathematicians (and scientists). The judgments must be made without favoring family, friends, colleagues, or ideological compatriots, and conflicts of interest are widely required to be declared and reported (EMS 2012). Submitters are expected to give fair credit to colleagues, not to copy or plagiarize others' work, and to submit only what they are entitled to call their own (AMS 2019). These ethical norms are sometimes breached, and work that is not rightly that of the author may manage to get through the review process and be published. In such cases, other mathematicians are likely to recognize the misattribution and possible dishonesty and will contact the publishers and show disapprobation and censure toward the author, usually leading to a correction of the official record.

A related issue is what is termed the "theorem-credit economy in mathematics," whereby mathematicians obtain rewards and advancement from gaining credit for new results. Morris (2020) argues that in the competitive world of mathematics, some mathematicians try to rush their results into print in order to gain priority. She argues that some senior mathematicians are slow to publish exhibiting "intellectual generosity" in allowing others to gain the credit. From the virtue theoretic perspective of MacIntyre rushing for theorem-credit is pursuing external goods, not the internal goods of the mathematical practice. But as Morris (2020) indicates, withholding results is also ethically problematic as it may slow down advances in the mathematical practice.

The warranting of new mathematical results sometimes lets incomplete or even erroneous proofs, and thus, unproven results slip through into the realm of published knowledge. To avoid this, mathematicians prefer new results that have support from as many experts as possible (Andersen et al. 2020). However, the commitment of mathematicians and mathematical practices to the virtues of correct proofs and truth mean that such errors are almost inevitably uncovered and publicized. While it is a virtue to find and disseminate mathematical errors and infelicities, it is not a vice to publish genuine mistakes. The history of mathematics (and science) is strewn with mistakes, and indeed, they contribute positively to the growth and development of knowledge (Lakatos 1976).

In addition, most mathematicians employed by a college or research department have additional responsibilities because job also involves teaching mathematics or supervising students. These responsibilities include treating students with care and respect, teaching mathematics in an effective way that benefits the students, and engaging in self-evaluation and listening to feedback. It almost goes without saying that the responsibilities of mathematicians include being engaged with the profession

and keeping up to date with research and developments in their specialisms and maintaining their own interest and enthusiasm. Mathematicians have these ethical responsibilities because they are part of the contract accepted voluntarily in taking up employment in a mathematics department. That is, they should carry out their professional duties to the best of their abilities, including respecting clients, practicing their profession well, and enhancing their profession overall (Ernest 2019b). (Tanswell and Kidd (2020) suggest that possibly vicious teaching practices can prove to be epistemically productive, for example, the aggressive adversarialism of the Gelfand seminar in post-war Moscow. Without endorsing such practices, this illustrates that judgments about virtuous teaching need to take into account broader contextual factors including the goals of the social practices, rather than just the issues of ethics.)

Having introduced ethical aspects of mathematical practice and personal ethics of mathematicians, it is appropriate to look a bit more deeply into ethics and ethical theory itself.

8 Ethical Theory

Ethics is that branch of philosophy that treats the Good in a theoretical and systematic way. The most immediate relevant form is applied ethics, which treats the ethics of professional practices such as medicine. There are also further more theoretical schools of ethics, and I shall be considering one of them subsequently, namely virtue ethics.

8.1 Medical Ethics

Medical ethics is a good example of applied ethics since it has been developed over two millennia, from the time of Hippocrates (Wikipedia 2020b). It is based on a set of values that form the basis for ethical medical practice and which professionals can refer to in the case of any confusion or conflict. These values include the respect for autonomy, non-maleficence, beneficence, and justice (Beauchamp and Childress 2001). They can be expressed as the four following principles:

Beneficence. This is expressed in actions that promote the well-being of others, especially the others directly involved in any medical action or intervention.

Non-maleficence. This is the principle of not being the cause of harm. It can be interpreted in two ways. First, it is more important to do no harm than to do good, and it is also important to know how likely it is that your actions will be harmful even if intended to do good. Second, there is the utilitarian interpretation – to promote more good than harm.

Respect for autonomy. This is the principle that the rights of an individual person to self-determination should be maintained.

Justice. This concerns the distribution of outcomes and benefits of actions, and the decision of who gets what benefits. Ideally, a commitment to justice implies the unselfish wish to provide benefits equally for all, universally.

Now the question arises, do these apply to mathematical practices? If not, what principles would or should the mathematics community adopt as the basis for their ethics?

The principles of medical ethics all concern how to treat other humans and living creatures (such as laboratory animals) and how to safeguard them. Thus, medical ethics offers valuable examples for the ethics of the institutions and social practices of mathematics on the human side. It does the same for the personal ethics of the participants including mathematicians, students, and others in direct contact with mathematicians. But medical ethics in general does not offer direct guidance on the central activities of mathematical practices.

However, the principle of non-maleficence in its first sense does seem relevant. Mathematical developments and applications should not be the cause of harm. Judgments are involved as to first whether a mathematical application is the cause of an observed outcome, and second whether that outcome is in fact harmful. If applied mathematicians construct a range of models and one is used that turns out to be harmful, is the responsibility not that of the person or body choosing to deploy that particular model, rather than that of the mathematicians? Surely, the answer must involve examining the range of models provided and the mathematicians' narratives about their effects. In addition, determining what is harmful involves judgment related to a value-system. For example, has the development of the applied mathematics involved in nuclear physics proved harmful? This is open to debate because the huge power unleashed can be used for good or ill. Currently, different models of applying new Covid-19 vaccines in the population might result in varying probabilities in the balance of good and harm. There are judgments to be made about how soon such vaccines can be broadly used after their development and which sub-populations should be prioritized. Elsewhere I have argued that applications of mathematics both in education and in wider society cause harm to some, in addition to the undoubted good outcomes and benefits to many (Ernest 2018).

Thus, medical ethics, although well established, clear, and exemplary, offers limited suggestions for the main and central domains of ethics in mathematical practice beyond the considerations of beneficence and non-maleficence. If I was considering mathematics education here, where there are clients to whom one has responsibility, just as in medicine, then further values would apply. In addition to beneficence and non-maleficence, the principles of having respect for the dignity and autonomy of mathematics students, and justice in the fair distribution of teaching and benefits among students are relevant. Such issues can arise tangentially in the social practices of research mathematics where graduate students may be involved and working in collaboration with and under a mathematician's supervision.

9 Ethics in Science

Within science it is claimed that the principles of ethics are part of the bedrock of scientific activity. Society trusts that the results and the projected outcome of any scientific activity are based on an honest and conscientious attempt by the scientific

community to further knowledge disinterestedly. The moral principles of science require that it be conducted independently, without favoring or distorting any findings to satisfy outside, extrinsic or personal interests. In addition, ethics in science requires the fair treatment of all colleagues, especially junior colleagues and students (Sharma 2015). Such an account of ethics in science represents an internalist view of the ethics of science. That is, one that focuses on the ethics of the activities and processes within the scientific community, solely within scientific practices, without looking to broader ethical issues such as the social responsibility of science.

The concerns and principles of this ethics-in-science approach remain at the personal and institutional level, as does medical ethics, but bring in the additional issues of fairness and trust. It is widely assumed that ethical norms and principles will be followed during the conduct of scientific research so that it will be done in a replicable way with its reporting both honest and trustworthy. Fairness and trust are central to any scientific practice, and it is also assumed that the process of publication with its refereeing and editorial judgments will be done fairly. Scientific contributions including mathematical publications are critically scrutinized by gatekeepers, and the results of this process of warranting and validation are generally trusted. (However, the big financial stakes involved in medical and in particular pharmaceutics research and testing have resulted in unethical distortions of the objective scientific method to cherry-pick or exaggerate positive results and suppress negative results for financial gain (Goldacre 2012).) Because of the wide diversity of mathematical specialisms, most mathematicians have to take on trust the results and proofs published in areas outside of their own expertise. This requires trust in the institutions of mathematics including its conferences, journals, and book series. This trust covers not only the restriction of publication to those results and proofs that have been found sound and correct by expert mathematicians in the refereeing process. It also includes the publication of corrections and disclaimers when errors and falsehoods are found to have slipped through the net, for such is inescapable in all human processes and institutions.

Overall, the ethics-in-science approach I have described here offers principles for the ethical conduct within the institution and social practices of mathematics, as well as principles underpinning the personal ethics of the participants. But it does not treat the explicit ethics of mathematical knowledge and its applications, nor any hidden ethical consequences of mathematics itself or of the invisible or unacknowledged applications of mathematics outside of itself.

10 Social Responsibility of Science

A further area of applied ethics concerning science is the movement for the social responsibility of science. It is widely argued that the Promethean power of modern science and technology warrants an extended ethic of social responsibility. The Russell–Einstein Manifesto called for scientists to take responsibility for developing weapons of mass destruction and urged them to "[r]emember your humanity, and forget the rest" (Russell and Einstein 1955). A Hippocratic Oath for scientists has been proposed by Karl Popper (1969) and others (Wikipedia 2020c). Popper argues

that scientists should take professional responsibility to further the growth of knowledge; students of science should respect others engaged in science and their teachers, but the overriding loyalty of scientists is owed to humanity as a whole.

There are many publications and organizations stressing the social responsibility of science. For example, Mark S. Frankel (2013) from the American Association for the Advancement of Science argues that scientists need to think more deeply about their ethics beyond their "internal" professional responsibility of sticking to the research standards agreed upon by the scientific community. They must also address their "external" social responsibilities toward the larger community, because science affects policy decisions and offers risks and benefits to society. This should be done both through the ethical actions of scientists and through and being integrated into their graduate education. In particular, Frankel stresses that being a socially responsible scientist means looking beyond themselves to appreciate the global dimension of science. They should use their skills to help with global problems, not only applying their knowledge and skills, but also considering the value of what they do for others. This is an overall human responsibility and stems from a social debt and obligation because scientists' education and research are subsidized by society, he argues.

Mathematicians can learn a lot from such contributions on the social responsibility of science. Indeed, Bertrand Russell who co-authored Russell–Einstein Manifesto was a mathematician, logician, and philosopher, rather than a scientist. Although he knew mathematics played a big part in the development of nuclear weapons as a subscriber to a purist philosophy of mathematics, he was primarily calling for empirical scientists such as physicists to take responsibility for developing weapons of mass destruction (Russell and Einstein 1955).

A further development has been the proposal of the Archimedean Oath. This is an ethical code of practice for engineers and technicians, similar to the Hippocratic Oath. This includes a commitment to using one's knowledge for the benefit of humankind, respect for the environment, bearing the entire responsibility for all of one's actions and not putting one's skill to the service of people who do not respect human rights, not using one's knowledge for destructive purposes and not permitting considerations regarding religion, nationality, race, sex, wealth, and politics to harm people affected by one's actions (Wikipedia 2020d).

Overall, organizations and individuals supporting movements for the social responsibility of science and Hippocratic-type Oaths for science and technology add an external ethical concern with the impact on and responsibility to all of humankind to the internal concerns about maintaining ethics just within the social practice of science. But despite the widespread nature of such concerns, as yet, there is no widespread recognition of the social responsibility of mathematics.

11 Social Responsibility of Mathematics

Mathematicians have been slow to recognize the social responsibility of mathematics as a dimension of the applied ethics of mathematics. Although there is a British society for the social responsibility of science, and even a group called Radical Statistics (n.d.) concerned with social responsibility in statistics, there is no society

for the social responsibility of mathematics (pure or applied). Indeed, the very idea of the social responsibility of pure mathematics will seem to many a contradiction in terms, on account of mathematics being viewed as objective, neutral, and value free. There are notable exceptions. Hersh (1990) has long called for attention to ethics in mathematics, and Harris (2017) asks "Do Mathematicians Have Responsibilities?" raising many of the same issues as this chapter. Even in the present day, the American Mathematical Society (2019) interprets the social responsibility of mathematicians in an individualistic way that concerns only the professional behavior of mathematicians inside the mathematics community, with no ethical concerns about the impacts of mathematics on wider society.

It has been suggested that there should be a Hippocratic Oath for mathematicians (Davis 1988). Unsurprisingly, given the widespread views of the neutrality of mathematics, even of applied mathematics, this suggestion was ignored until very recently. However, Hannah Fry, a mathematician and popular media presenter on the subject, recently called for this very thing (Sample 2019). She goes on to assert that mathematicians should be taught ethics from day 1 (UCL 2019). Fry is an applied mathematician so she is not handicapped by the purist ideology widespread among pure mathematicians, discussed below.

A Hippocratic Oath on behalf of financial modelers, also applied mathematicians, was proposed over a decade ago (Derman and Wilmott 2009). This oath includes acknowledging that no equations can ever truly describe the social and physical worlds, explicitly committing the analysts to not giving the people who use the models any false ideas about their accuracy, and acknowledging that they may have enormous effects on society and the economy, many of them unanticipated. This position recognizes that the mathematical equations and models used by quants (quantitative researchers in finance) and financial modelers and analysts have a very great impact on financial markets. It is these mathematicians' responsibility to apprise decision makers in finance about the limited and provisional nature of all models and predictive tools, and their potential impacts.

12 Ethics in Mathematics

Courses on ethics and professional issues in undergraduate mathematics course remain uncommon, but one of the first was organized by James Franklin at the University of New South Wales, Sydney, around 2004 (Franklin 2005). After this early foray into the area, few further developments in this area can be found. But since 2016, a number of mathematicians under the leadership of Maurice Chiodo and Piers Bursill-Hall (2018) have instituted the Ethics in Mathematics Project at Cambridge University, UK. The project began with an undergraduate lecture course entitled "Ethics for the working mathematician," taught annually up to and including 2020. It includes case studies of activities that raise ethical questions including:

1. Financial mathematics and modeling
2. Cryptography, surveillance, and privacy
3. Fairness and impartiality in algorithms and artificial intelligence

Further lectures look beyond case studies and examine external forces and influences that shape how mathematicians behave including both ethical and professional issues.

The rationale for the project is as follows:

> [T]he majority of students who study mathematics as undergraduates do not continue in academia, but instead move into other industries. In the early 21st century it is surely a commonplace that mathematics and mathematicians play a fundamental role in the economy and in society, impacting sectors ranging from engineering and biotechnology to finance, information technology, data science, and public policy. Thus, it is important for them to have an understanding the ethical issues they may face as part of their work. (Chiodo and Vyas 2018, p. 1)

Overall, the course has all the ingredients of an excellent introduction to ethics for mathematical practitioners. In particular, it offers preparation for applied mathematical practices as well as for the practicalities of professional life for beginning mathematicians.

Beyond this course the project organizers held international Ethics in Mathematics conferences in 2018 and 2019 for mathematicians, historians, philosophers, and educationists, to explore and share conceptions of the ethics of mathematics (EiM 1 2018; EiM 2 2019). Many of the papers and talks presented are a valuable resource available online, together with an archive of publications supporting the project (CUEiM 2020).

The project is based on a common-sense model of increasing awareness of ethical issues and the mathematician's responses to them. The starting point is the proposition that generally mathematicians regard ethics as irrelevant to mathematics. The model proposes four stages of increasing involvement with the ethics of mathematics, comprising both awareness and action responses. A summary of the model is shown in Table 1.

The model serves both as a descriptive tool for classifying mathematicians' positions and as a prescriptive instrument specifying the desirable progress as a student moves up a hierarchy of increasing ethical awareness and action as a career mathematician.

Overall, this project can be seen as a valuable and exemplary first attempt to plug the hole left by the widespread neglect of ethics for mathematicians. It is primarily practice-driven and does not include any elements of theoretical ethics. This means that good and bad, vice and virtue are not analyzed. It is assumed that bad or harmful

Table 1 Levels of awareness and action concerning the ethics of mathematics (from Chiodo and Bursill-Hall 2018)

Level 0	Believing there is no ethics in mathematics
Level 1	Realizing there are ethical issues inherent in mathematics
Level 2	Doing something: speaking out to other mathematicians
Level 3	Taking a seat at the tables of power
Level 4	Calling out the bad mathematics of others (a) Calling out non-mathematicians using maths unethically (b) Calling out mathematicians using maths unethically

applications can be identified as such through everyday ethical common-sense. The organizers are probably right in making this judgment call with regard to developing an undergraduate course. That is, to focus on the ethics of problematic applications and the workaday practicalities of the beginning professional mathematician, as opposed to looking at ethics in a more philosophical manner. Students could choose to pursue additional studies in ethics if that interests them, but most will probably not take this step. It is very likely that most students will be interested in ethics only insofar as it impacts directly on the mathematics they are studying, including its applications and their own immediate professional practices.

This awareness raising course is relevant for mathematicians entering into both pure and applied practices in which the mathematics is explicit or high-level application practices where the necessary mathematics is subsumed behind the foregrounded goals of the application itself. It represents an important step forward.

12.1 Mathematics and Purity

It has been noted in the ethics in mathematics project and elsewhere including at the start of this chapter that many mathematicians believe there is no ethics in mathematics (Chiodo and Bursill-Hall 2018). This not simply due to a lack of awareness of ethical considerations, nor simply to an uneducated belief in the value-free nature of mathematics, but to adherence to a philosophy, the ideology of purity, often termed "purism" (Restivo 1985).

There are two meanings to the term "purity." First, there is its descriptive sense in which purity applies to something that is unmixed and undiluted or is in its basic or primal state, be it a substance or practice conducted solely for its own sake. Second, there is the normative, prescriptive, or evaluative sense of purity which applies to something that is unadulterated, unpolluted, or unbesmirched. In this sense to describe something as impure is to say that it is tainted or degraded either aesthetically (less beautiful, more ugly) or morally (less good, more bad). Even when purity is used in the descriptive sense, the ambiguity of the term risks smuggling in the connotation of impurity as a negative. Often the term "purity" is used by purists in its full, double-barreled sense, both descriptively and normatively.

Although the adjective "pure" in the term pure mathematics is descriptive, purism in mathematics is sometimes based on the normative or evaluative sense of purity and as such is a position concerned to keep out what are perceived as impurities and taint. Thus, from this position, any signs of human presence, interests, or values are excluded as detracting from the objectivism and absolutism of mathematics. Any attempt to attribute human interests or values to pure mathematics itself, not just to the practices and human foibles of mathematicians, is seen to be antithetical to the pristine purity of the subject. Values including those of ethics are repudiated by purists whose concern is to keep out the taint of ulterior motives such as utility and applicability and to defend pure mathematics as intrinsically valuable and self-driven; knowledge pursued and derived purely for its own sake.

Kline (1980) notes the tendency to purism in modern mathematics, and the mathematician Hardy (1940) celebrates the lack of utility of pure mathematics and

its ethical neutrality. Implicit in his praise is derogation of applied or impure mathematics. Part of the ideology of purism is that the pure is beautiful, the impure ugly: "Beauty is the first test: there is no permanent place in the world for ugly mathematics" (Hardy 1940, p. 14). My interpretation of Hardy is that he subscribes to both the positive and negative tenets of purism in mathematics.

1. *Positive tenet of purism*: Pure mathematics is the superior form of mathematics: it is virtuous, beautiful, and intrinsically valuable, standing in no need of justification. It is the product and expression of the virtuosity of the excellent mathematician.
2. *Negative tenet of purism*: Impure or applied mathematics may be a necessary evil for practical purposes, but it is an inferior form of mathematics driven by extrinsic and ulterior motivations.

Purity remains an important value within mathematics. Ancient Greek thinkers, including Plato, were the first to distinguish pure and applied mathematics and to exhibit purist values. Pure mathematics was said to concern "true being," the realm of pure thought concerning ideal objects and relationships, and thus was seen to be of supreme value (Plato 1941). Applications of mathematics were regarded as lowly activities performed by lesser beings for mundane practical purposes. In ancient Greek society, mathematics and philosophy were the preserve of the leisure class of free citizens, whereas applications of mathematics were the domain of tradesmen and slaves.

Purist values are reflected in the valuation of proof as a higher form of reasoning, and calculation as a lower form, from Plato onward (Ernest 2009). "Oriental mathematics may be an interesting curiosity, but Greek mathematics is the real thing" (Hardy 1940, p. 12). Purism values pure proof-based mathematics as being significant epistemologically, and pertaining to truth, wisdom, high-mindedness, and the transcendent dimensions of being. Equally, this ideology denigrates applied mathematics and calculation as technical and mechanical, pertaining to the utilitarian, practical, applied, and mundane – understood as the lowly dimensions of existence. Pure mathematics was the domain of philosophers and free thinkers. Applied mathematics (termed logistic) was the domain of slaves and tradespersons.

However, purity has only been a dominant value for a fraction of the whole history of mathematics. It was powerful in ancient Greek times and again in modern times, from the nineteenth century onward. Purism was not seen as significant for Galileo, Descartes, Newton, Leibniz, and others in the seventeenth century. At that time, according to the Neo-Platonist outlook, all of nature (the book of nature) was written in mathematics, the language of the divine, and pure and applied mathematics were not distinguished.

So why does purism come to dominate during these eras? In Ernest (2020a), I offer a number of reasons, including the following:

1. Hacking (2014) argues that the resurgence of purism in early nineteenth-century Germany can be traced directly back to the influence of Kant and his imposition of the philosophical categories of pure and applied reason. Following this fundamental distinction, the professionalization of university teaching and research

in the early to mid-nineteenth century, at first in Germany, led to boundary maintenance between professional mathematicians and outsiders alongside others holding an interest in mathematics. It furthered the demarcation of knowledge into disciplines reflecting university departments, helping to reinforce purism (Ferreirós 2016).

2. The dominant absolutist or Platonist philosophies of mathematics of modern times promote an idealized version of mathematical knowledge that is pure, superhuman, and values free. This philosophy supports purism (Erdős in Wikipedia 2020a; Hardy 1929; Penrose 1989; Balaguer 1998).
3. The standard presentation of pure mathematics in the modern era follows the Euclidean model and sequence, including, orders, definitions, axioms, lemmas, and theorems. Lakatos (1976) argues that this represents an ideological purification of proofs that serves to smooth out and eradicate marks of active human choices and labor in construction of mathematics, thus reinforcing purism.
4. The growth of interest in the foundations of mathematics with major work by Cauchy, Weierstrass, Cantor, Dedekind, Frege, Peano, Russell, Whitehead, and others turned the focus of many leading mathematicians inward, or reflected their already inward gaze, into the depths of pure mathematics and away from the world and applied mathematics.
5. During the years of the British Empire, Ancient Greece was "talked up" as the starting point of modern European thought, and the "Afroasiatic roots of Classical Civilisation" were denied and rejected (Bernal 1987). Such Eurocentrism promoted purism, but there is now a widespread literature repudiating the Eurocentric misrepresentation of mathematics, including Almeida and Joseph (2004), Joseph (2000), Pearce (n.d.), and Powell and Frankenstein (1997).
6. British and anglophone education until the nineteenth century was dominated by the classics. Many mathematicians would have been well versed in the Greek classics including Plato and Euclid and influenced by their idealist and purist tendencies. For example, graduated rules and measuring devices were disallowed in school geometry and were viewed as the inferior and imprecise tools of the engineer.
7. Mathematicians are taught and shaped as professionals by working with an older generation of established mathematicians, whose culture, including values and ideologies as well as knowledge and methods, is passed on from generation to generation (Kitcher 1984). Since the preparation of high school mathematics teachers is normally conducted at least in part by mathematicians, the purist values of the past are transmitted forward and sustained among many modern mathematicians and teachers.

For reasons such as these, purism remains a prominent value for many mathematicians up to and including the present day. Purism rejects any human features or influences on pure mathematics. Purism actively protects the disciplinary boundaries against any breaches from the taint of traces of human activity or values. Thus, from this perspective, ethics is perceived not only to have no bearing on pure mathematics, but any suggestion that it might have is repudiated as a threat to its purity.

Of course, purism also plays a valuable role because it maintains the purity of the disciplines and serves to defend the pursuit of knowledge for its own sake from outside interests and pressures (Restivo 1985). This can be useful in protecting the interests of a discipline such as mathematics from political interests that might seek to bend it toward more immediately useful applications, as was witnessed in Soviet Russia under Stalin (Grossman 1980).

Thus, a problem for relating ethics to mathematics is that from the perspective of purism mathematics is regarded as neutral and value free. Ethics is regarded as the set of values most irrelevant to mathematics. Any attempts to raise ethical issues with regard to pure mathematics are seen as possibly tainting or lowering the subject from its elevated state of purity. Purism strenuously rejects any suggestion of the relevance of ethics to pure mathematics and usually does so without serious consideration of ethics. On the face of it, to say that one mathematical concept is more vicious or virtuous than another seems absurd.

However, in Ernest (2016, 2018), I argue that even pure mathematics is value-laden and is imbued with ethical values. Similarly, Johnson (2012, 2017) argues that the formulation of probabilistic concepts and market equations has important and different ethical consequences for society, and MacKenzie (1981) shows how different statistical concepts by Pearson and Yule had profound effects on eugenics programs. So the ethical evaluation of mathematical concepts is not, in fact, absurd.

12.1.1 Values and Applied Mathematics

Like that of pure mathematics, the social practice of applied mathematics is widely regarded as neutral and free from social responsibility, although this is beginning to change. It is acknowledged to be value-laden, but the values identified are typically those of relevance, utility, efficiency, and so on. Applied mathematics is widely seen as neutral technical knowledge, brought in to solve scientific or real-world problems that originate outside of mathematics. Like a bag of tools, mathematics is used in the service of externally posed problems, and applied mathematicians are responsible for the means of solving the problems, not the ends, that is, the nature and significance of the problems themselves. Thus, in mathematics viewed as a tool "Control replaces explanation, and validation is accomplished by use" (Lenhard and Carrier 2015, p. 18). Within applied mathematics applications are generally evaluated in terms of their efficacy and utility in the narrow context of the problem posed, and any issues concerning the problem's wider social significance are not only ignored but viewed as beyond the business of mathematicians. Their role is simply to provide technical support. The research of Chiodo and Vyas (2018, p. 2) uncovered a widespread view that "while there are ethical issues in applied mathematics, these are imported from the disciplines that the mathematics in question is being applied to, and thus do not require a separate mention."

Although there are a growing number of exceptions, many applied mathematicians do not see themselves as having any social or ethical responsibility for their mathematical work. Chiodo and Bursill-Hall (2018, p. 5) summarize the beliefs of both pure and applied mathematicians as "[b]elieving there is no ethics in mathematics – This is where most mathematicians are today." If an applied mathematician

declines to work on developing a missile guidance system or an App that invades people's privacy on ethical grounds, it is seen as a personal choice. Likewise, a vegetarian might object to working in a slaughterhouse or meat packing plant on ethical grounds, but that is currently seen as a personal choice based on personal values rather than a society-wide ethical imperative.

Until recently, few voices have been raised calling for an ethic of social responsibility of mathematics, from within the field. In science, technology, and other areas of use across society, applied mathematics is seen as the servant, the under laborer whose job is to do the work, not to question the task. Thus, perceptions of applied mathematics do not challenge the denial of ethical responsibility associated with pure mathematics. For different reasons, applied mathematics is also regarded as neutral and free from social responsibility, merely a set of tools used by others. From this perspective, it is a technical activity, judged solely on its utility and effectiveness. Of course, as I have indicated above, this view is changing. Typically, it is applied mathematicians, and those involved with applied mathematics that are calling for ethics in mathematics.

13 Virtue Theory and Mathematics

Given the widespread views of the ethical neutrality of mathematics, the questions what is good mathematics and what is a good mathematician remain without satisfactory answers. There is a fundamental ambiguity at the heart of the question of what constitutes good mathematics. From an ethical perspective, good pure mathematics is that mathematics which benefits humankind and contributes to human flourishing. From an epistemic perspective, good mathematics is that pure mathematics which is clearly expressed, which is well justified, normally by means of mathematical proofs, and which generally conforms well to the epistemic standards of the social practice of mathematics and of the community of mathematicians. In the case of applied mathematics, ethically good applications are those that are beneficial to humankind and cause little or no harm. Applications that are good from an epistemic perspective are those that are correctly and rigorously formulated and that accurately predict outcomes or provide good explanatory models for their target domains.

On the face of it, ethical goodness and epistemic goodness are disjoint perspectives from which to consider mathematics. One draws on ethics and moral philosophy, while the other is epistemological. However, virtue theory, and in particular Alasdair MacIntyre's (2007) version of virtue theory, accommodates both of these perspectives. MacIntyre's theory rests on the concept of social practice and the first part of his theory concerns the goods of a social practice. The internal goods of a social practice are the things that are achieved in the course of pursuing the goals within the practice. They can only be achieved through the practice and relate to the particular intrinsic goals of that practice. In contrast, external goods are achieved or obtained in a variety of ways beyond the specific and particular social practice. External goods include fame, fortune, and other things that might extrinsically motivate us.

According to MacIntyre, mathematics is a social practice, or it might possibly be seen as a family of several overlapping practices. The virtues of a mathematician include those character traits that enable him or her to pursue the goods internal to the practice of mathematics, that is, to be a good mathematician. This is to seek to develop skill and ability to advance mathematical knowledge, in order to pursue the inner goods of mathematics, in other words, to develop the virtues of mathematical practice. Possessing such character traits is called virtuosity within that social practice. To be a good mathematician from MacIntyre's perspective is to do good work within the social practice of mathematics. This is to work well and effectively to achieve the goals of this social practice. These goals concern what is valued because it is useful, well formulated, original, powerful, beautiful, or otherwise desirable and acceptable within the social practice of mathematics.

Beyond this, MacIntyre's theory has two more layers or dimensions concerning the virtues that constitute the virtuous life. On the second level, MacIntyre argues that a virtuous person should also possess those traits of character that lead to and sustain a unified narrative of self and the overall good life. Third and lastly, MacIntyre claims that as social beings we humans all live within a cultural, moral, and historical tradition and that our virtues are also manifested in acknowledging, sustaining, and contributing to the development of our tradition in a way consistent with and conducive to our virtuous practices and unified virtuous lives (Ernest 2020a).

Applied to mathematics, MacIntyre's theory entails that every mathematician, every participant in the practice of mathematics, should not only cultivate their own individual virtues, be the best mathematician possible, but also take ethical responsibility for this practice and their overall lives. It is necessary but not sufficient to attain the goods internal to the practice of mathematics, that is, to develop worthwhile mathematical knowledge or applications. Our virtues must include taking responsibility for the practice as a whole and for its social and other outcomes. The humanistic moral tradition, of which I and most mathematicians and mathematics educators are a part, at least in the West, includes such virtues and values as rationality, truth, democracy, justice, fairness, autonomy, and freedom, as well as care for fellow humans, for living things, and for the environment. Thus, it is not enough to focus solely on the goods internal to the practices of mathematics research and teaching. We must also be concerned with the intended and unintended outcomes of these practices on other humans, on all living things, and on the planet as a whole.

According to MacIntyre's virtue theory, a good mathematician is one that addresses the goods internal to the practice of mathematics. That is, being conscientiously devoted to research, working on new proofs, problem solutions, and presentations of mathematics: being good as a mathematician. A good mathematics teacher is one devoted to the goods internal to the practice of teaching, namely working for student understanding manifested in their success at mathematical tasks and their ability to communicate mathematical ideas and reasoning (known within the practice as relational understanding, after Skemp 1976). In both types of mathematical practices, mathematics teachers (and most research mathematicians also teach) should also attend to the flourishing of students as developing human beings.

In any practice "human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended" (MacIntyre 2007, p. 279). The good mathematician will seek to improve his mathematical skills and creative products and contribute to the growth of the practice of research mathematics, just as a good teacher will seek to grow and learn as a teacher and contribute to the practice of mathematics teaching.

Overall, MacIntyre's virtue ethics provides a partial way of respecting the ideology of mathematical purism and even the neutrality of developing applications. According to his theory, the social practice of mathematics should focus exclusively on the internal goods of the practice. A virtuous pure mathematician will strive for these internal goods and eschew the external goods of pure mathematics, including uses, applications, personal rewards, and fame. Thus, MacIntyre's theory accommodates mathematical purism, as well as other ideologies of pure mathematics, by focusing on the internal goods of the social practice. A strength of this approach is that turning away from the purely internal goals within a mathematical practice could lead to a decline in the practice itself. What counts as good in mathematics is the mathematician's virtuosity, the virtues that give rise to their excellence as professional mathematicians working toward the inner goods of mathematical practices. Furthermore, I believe that the best judgments of what constitutes excellent advances in mathematics come from mathematicians themselves. Constraining them from pursuing the most intellectually exciting problems and most significant results in mathematics, as judged by themselves, seriously risks stultifying their activities and the growth of the discipline (Ernest 2020a). Harris (2015) offers a persuasive analysis and justification for this position.

14 Responsibility for Applications of Mathematics

The fact that a good mathematician focuses primarily on the internal goods of her social practice is not an excuse for ignoring the ethics of mathematical applications. As we have seen, MacIntyre's theory also requires mathematicians to take responsibility for each mathematical practice as a whole, for its social and ethical outcomes, as well as for our shared humanistic (or other) moral tradition. Above I distinguished four main areas of ethical concern for mathematicians: institutional and personal ethics for mathematicians, already discussed. I also distinguished the overt and visible ethics of mathematical products, pure or applied, and the implicit, hidden, and possibly unforeseen, ethical outcomes of applications of mathematics. Here I will consider these two areas. I divide them into three types of application: mathematics in public communications, ethically problematic applications of mathematics, and the performativity of mathematical applications. These are all part of the external ethical responsibilities of the mathematician.

14.1 Mathematics in Public Communications

It is widespread for governments and corporations to use mathematical signs or representations to lend support and an air of authority and certainty to

communications and pronouncements. Evans et al. (2014) conducted an investigation into the use of mathematical images in newspaper advertising in the UK. They found that this increased from under 2% in the period 1994–2003 to almost 5% in 2006–2008. They showed that advertisements use mathematics to lend the advertised products valued characteristics of mathematics including precision, rationality, and authority, rather than employing such characteristics to construct reasoned arguments. Mathematics is used in the public media as a rhetorical device to enhance persuasion, without necessarily using mathematically valid arguments.

Seife (2010) quotes the statistic published in a newspaper that only 0.027% of US felony convictions are wrongful. He points out that despite this assertion being false, for it was based on an incorrect back-of-the-envelope calculation, it was nevertheless used in a fallacious argument by the prosecution in a court case that ended with a prisoner being sent to his death.

A further example comes from the May 2020 broadcast by the UK Prime Minister on British national television about the Covid-19 virus threat and the UK Government response (Johnson 2020). In it the Prime Minister claimed that there is a central equation representing the danger from the Covid-19 virus and determining the UK government responses. The equation is: CAL = R + N, relating the *Covid Alert Level* (CAL), the *Rate of infection* spread (R), and the *Number* of infected cases (N). In Ernest (2020b), I analyze the equation and show that is it at best meaningless and at worst deliberately deceptive. The misleading use of mathematics to back up the presentation of the Government's Covid strategy was termed "number theatre" by David Spiegelhalter (Ahuja 2020). My conclusion is that this piece of number theatre is smokescreen designed to "blind the public with science," by giving a spurious sense of mathematical precision and scientific authority. The government chose to follow a politically expedient path in its policies. This led to the highest death toll in Europe at that time, in both absolute and comparative terms. Johnson wished to make it look as if the policy merely followed the route determined by science.

These examples illustrate the widespread use of mathematical signs or representations in communications in order to persuade and sometimes deceive (see also Hauge et al. 2019). Mathematics is used as a rhetorical tool intended to lend precision and force to presentations, that is to persuade without making any rational or mathematically buttressed arguments.

The term "proofiness" has been coined for the art of using bogus mathematical arguments to "prove" true something that is false. According to its coiner, "[o]ur society is now awash in proofiness. Using a few powerful techniques, thousands of people are crafting mathematical falsehoods to get you to swallow untruths" (Seife 2010: p. 8).

This is deeply problematic. In order to combat such deception and unethical manipulations of opinion and protect democracy, at a minimum two defenses are necessary. The first is a well-educated and numerate citizenry able to critically examine and evaluate published claims that rely on mathematical arguments. Secondly, we need mathematics-related professionals at all levels, from researchers to teachers that are willing and able to call out such misuses and who to see this as part of their public duty.

14.2 Problematic Applications of Mathematics

In the past half century, and particularly in the new Millennium, mathematics and mathematical applications have permeated Western society and probably most countries around the world. This is most notable in the universal and accelerating spread of information and communication technologies in governance, public services, and commerce. Such applications are often potentially benign, but there are an increasing number of problematic applications of mathematics, as are documented in O'Neil (2016), Skovsmose (2019), and elsewhere.

Many examples could be cited as ethically questionable. For example, early on Google found "detailed stories about each user – thoughts, feelings, interests – could be constructed from the wake of unstructured signals that trailed every online action" (Zuboff 2019, p. 53). Thus, covert profiles were constructed for millions of users and their friends without permission. This is a deliberate and unethical breach of data privacy principles. Furthermore, Google enables advertisers to target ads to people who search for racist terms, and Facebook enabled advertisers to direct their pitches to the news feeds of almost 2300 people who expressed interest in anti-Semitic topics (Zuboff 2019). Even search results were slanted by the Google search algorithm in ways that reinforced racist and other stereotypes. Noble (2018) found in searching online for resources for her daughter, the first hit for the keywords "black girls" was a pornography site, and that variant searches also turned up disturbing racial stereotypes and other distorted associations and links. Noble's first search took place in September 2011, and after publishing an article about it, the search algorithm was modified by Google. Nevertheless, many unwelcome and unethical misdirections remain in place based on commercial pressures. Thus, Google search results have been modified in "response to *pressure from businesses, outside interest groups and governments around the world*" (Grind et al. 2019, p. 1; original emphasis).

Another new and deeply problematic issue is the widespread circulation of "fake news" stories. Allcott and Gentzkow (2017) define "fake news" as distorted signals uncorrelated with the truth that impose private and social costs by making it more difficult to infer the true state of the world. They found that in the lead-up to the 2016 US election, there were 760 million instances of a user reading these intentionally orchestrated lies online or about three such stories for each adult American. This was part of huge personal data capture and exploitation affair surrounding Facebook and Cambridge Analytica (Cadwalladr 2017). Data of Facebook account holders' "friends" was captured and in combination with other data culled from Internet applications was used to build up a huge database characterizing millions of persons in terms of their social location, political leanings, commercial behavior, and preferences, by Google and others. This data was exploited in organized political disinformation campaigns and profit-driven fake news stories during the 2016 UK Brexit vote and US presidential election. It was partly done by means of individually shaped and targeted stories and posts designed to push a viewer's "hot button," an emotional and usually controversial issue or concern that triggers immediate intense reaction (Merriam-Webster 2020). Triggering a hot button causes the targeted person

to become emotional and irrational, blocking critical reflection and evaluation, and reinforcing strong attachments to particular social or political positions. This often strengthens commitments to radical social or political positions and elicits strong negative reactions to any contrary position. In consequence, through the use of false information and fabricated images, there is increased electoral support for polarized and populist leaders and causes. Woolley and Howard (2018) term this "computational propaganda" and argue that the large-scale automation of messaging and social media platform manipulation amounts to a new political communications mechanism of unprecedented power and threat. Not only is this deeply unethical, it also erodes the very fabric of a free and democratic society.

A further problematic application of mathematics is the building of AI software to help identify targets for lethal drone strikes. The Pentagon has been engaging in these developments with sub-contractors such as Google. However, many Google employees were unhappy with providing support for what they perceived to be an unethical application of mathematics and programming. A number of Google employees resigned and thousands of others signed a protest petition against taking part in the Pentagon project known as Maven. The outcome was that Google announced that it would not renew its contract to do this artificial intelligence work for the Pentagon (BBC 2018).

There are many other examples of applications of mathematics that have been attributed with negative, unethical, and sometimes even disastrous outcomes. For example, the Gaussian copula formula (a multivariate cumulative distribution function) was called "the formula that killed Wall Street" because its overextended use to price derivatives helped to trigger the Global Financial Crisis of 2008 (Salmon 2009).

So who is responsible for these ethically problematic applications? In my view, responsibility should be shared among all those involved, including applied mathematicians, information and communication technology professionals, computing specialists, and authors of algorithms and Apps, as well as those who fund, instigate, and control these applications across society at great social costs, for political and economic gain.

How can we avoid or reduce the harm or potential harm done by applications of mathematics in society like in these examples? Obviously, part of the answer must lie in requiring accountability from the tech and business corporations, as well as the political parties and pressure groups. We need national legislation and international concords to penalize cases of abuse involving from apps and applications. In addition, we need to better educate the persons involved, especially mathematicians, in the ethics of potentially harmful applications of mathematics.

14.3 The Performativity of Mathematics

In addition to the explicitly unethical outcomes that can come from certain mathematical applications and algorithms, there is the issue of performativity. Many mathematical models and algorithms go beyond describing and automating their target areas, but actively modify and transform their domains of application

(Skovsmose 2019, 2020; Skovsmose and Ravn 2019). Such transformation may be intentional and can even be perceived as benign, or they may be unanticipated outcomes.

An early example of this is the Victorian Payment by Results for teachers in British schools. An apparently sensible measure was to tie the annual salaries of teachers to the numbers of pupils in their classes and to their performance at particular exams. But the outcome was that once the inspector's visit was imminent, teachers drilled their pupils mercilessly on test items. Some falsified enrollment registers to keep numbers artificially high. Sick children were dragged along to school to satisfy attendance requirements (Hyndman 1978). Cramming, rather than teaching, became the means to ensure a teacher's livelihood – weakening pupils, teachers, and pedagogy (Welch 2009). This lesson of history has not yet been learned for English; schools are still graded and rewarded on the basis of student assessment scores. This results in many teachers "teaching to the test," distorting the curriculum and reducing the overall quality of students' learning experience.

In Ernest (2019a), I explore more widely the performativity of the mathematization of society as an ethical issue. For example, the widespread official and social identification of intelligence with IQ score not only limits the recognition of human capabilities but also demonstrably impacts on policies, lives, and human identities. Other measures such as evaluative scores in education, as remarked above, can change individuals, schools, and national and regional education systems (Gorur 2016). Warnings about this are not new; Campbell's law says, "The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor" (Goldstein 2014, p. 209).

There are many troubling but often hidden applications of mathematics needing ethical scrutiny and oversight right across the worlds of governance, banking, and financial markets. These are performative in shaping and driving policies and markets, and go well beyond being simply neutral tools (Johnson 2017). MacKenzie (2006) characterizes the mathematical models underpinning financial markets as engines, not cameras, because financial models shape and change markets performatively. Rather than just speeding up existing practices and making improvements in efficiency, such changes bring in whole new levels of risk and ethical jeopardy.

There is widespread use of programs and algorithms to ascertain if persons are worthy of loans or credit, or are likely to commit crimes. These can change those persons' lives without their knowledge or right of challenge. O'Neill (2016) documents case studies of how algorithms can distort the justice provided to minorities and provide artificial barriers to employment, credit worthiness, home ownership, and gaining insurance. They can even determine who are good employees through automated time on task efficiency measures. These and many more hidden control systems are currently at work, used by corporations and government bodies with virtually no oversight as to their efficiency and correctness let alone their ethicality. But "many of these models encode human prejudice, misunderstanding, and bias into the software systems that increasingly manage our lives" (O'Neil 2016, p. 3).

Marjanovic et al. (2018) term the many negative consequences of this "algorithmic pollution," but I think a more appropriate term is "algorithmic injustice."

As indicated, the mathematization of society is performative. Metrics and measures do far more than capture the concepts and qualities they are supposed to represent. They noticeably alter the social processes that they are implicated in, often with significant impacts on, peoples' overall lives. These changes can have major negative ethical effects, enacting discrimination and stereotyping and reducing opportunities and life chances. When measures are imposed by institutions and policies to manage, shape, and evaluate social practices, the measures become instruments of power to change the practice; they redirect it (Power 1999). Thus, for example, a hospital, prison, or school managed solely by the attainment of target measures of performance is at risk of compromising and dispensing with the care and professionalism traditionally provided by the key staff involved in the service.

In the UK, the reduction of professional competence and satisfaction to targets and measures of performance is often associated with "marketization" – the privatization of public services or the replacement of the professional ethos by a market orientation. The changed measures of satisfaction mean that a cheaper and less-skilled workforce can be brought in to deliver the service. The argument in favor is that this means that there are valuable cost savings to the public purse. However, in practice, the savings are often not realized, and furthermore, a reduced and degraded level of service is provided, directed solely at meeting the targets. Commissioning groups no longer monitor the quality of the social goods provided through professional expertise, but simply check whether the targets are met (Power 1999).

Where is the ethical scrutiny in this process? It is squeezed out, either by inattention or by design. The construction of a measure of some concept or of a model of an aspect of social practice is an epistemological or methodological project. (As an aside, I would like to conjecture a rule that "no domain of professional expertise and practice can be adequately represented by a finite set of measurable targets." Meeting such targets may be necessary but can never be sufficient for capturing the professional expertise, for two reasons. First, professional expertise and mastery can never be reduced to a set of rules. Expert practitioners and professionals always draw on tacit knowledge and intuition developed from their experiences during their path to expert status (Dreyfus and Dreyfus 1986). Second, any application of professional expertise involving other persons requires empathy and an ethic of care, and this cannot be legislated through targets to be fulfilled (Larson and Yao 2005; Slote 2007).) Technical skill, accuracy, and precision are the overarching values. At this point, in measure and model fabrication, there can be deliberate attempts to slant or even mislead, when measures are used for overtly political purposes (e.g., some government measures of unemployment or poverty, corporate designed Apps to conceal car pollution, such as the VW fraud). But even when done without intentional distortion, through their performativity, the forces of measurement and mathematics are radically reshaping and restructuring human practices and social reality, and even the futures and possibilities we can imagine (Skovsmose 2014). Beer (2016: p. 6, original italics) describes this process of performativity and its outcomes succinctly as the "relations that exist between

measurement, circulation, and *possibility*." Where the intensification of these changes will take us is hard to predict, but it is clearly toward a more totalitarian and less-free society. It is a society where there is less respect for the individual and their rights and where governments and corporations alike exercise increasing power and control over our lives. Without exaggeration, it is a future that looks like a blend of Orwell's *1984* and Kafka's *The Trial*.

Mathematics specialists are only one part of the surveillance state, surveillance capitalism, and the performatively reshaped social world, alongside actors from business, corporations, the state, and government. But specialist mathematicians play an essential role in these developments. Because of their privileged access, as well as their implication in the development of these measures, mathematicians have a special duty to acknowledge and identify these applications of their discipline, to scrutinize them ethically, and to call them out where ill-considered or harmful. More generally, as democratic citizens, we share with all other citizens the duty to take responsibility for the direction that social policy is taking and be cognizant of its social justice and ethical implications. There is a direct and inescapable line of responsibility for the mathematization of society that implicates mathematicians and leads to the necessity for political engagement and action. This might sound like an extreme demand of the normally quiet and neutral mathematical expert, but we are not only on the edge of a new world, we, above all others, are helping to construct it.

15 Conclusion

There is a long tradition of ethical neutrality and the purity of mathematics. This is deeply embedded in the culture of mathematics for both historical and philosophical reasons. Many mathematicians and others resist the idea that we need to consider the ethics of both pure and applied mathematics. However, the foundation of this resistance is analyzed above and located in background philosophies of mathematics and the ideologies of purism and neutrality. We are becoming increasingly aware of the powerful impact of mathematical applications on every aspect of modern life. Largely, this is due to the near universal take-up of information, communication, and control technologies across all aspects of modern society. These developments, including some well-publicized abuses, are causing a growing number of mathematical practitioners to wake up to the ethical responsibilities of the mathematician. Nevertheless, a belief in the ethical neutrality of mathematics is still widespread. It is argued above that virtuous mathematicians can legitimately pursue mathematics for its own sake, but as citizens, they also have a responsibility to care about the social impacts of mathematics.

A review of the literature on the social responsibility of science and mathematics reveals that concerns about the ethics of mathematics are starting to emerge in publications and training practices. Given the great power and concomitant risk of harm from the near universal applications of mathematics, a widespread teaching of the ethics of mathematics is called for among beginning mathematicians, just as is

done in science and medicine. This is needed both to question and call out questionable practices and applications, but also to raise mathematicians' awareness of their own responsibilities in developing these applications. Some of the more ethically sensitive areas are distinguished and examples from three problematic categories are explored. These include the use of mathematics in public communications, overt applications of mathematics that have powerful social impacts, and the hidden performativity of mathematics in restructuring society, institutions, and social practices. As I have shown, some applications threaten the very foundations of democracy as well as many of the rights and freedoms we enjoy. Because of the specialist expertise and their involvement in these developments, we mathematicians have a vital role to play in keeping governments and corporations ethical and honest.

To end on a more positive note, I should point out that mathematics is also being widely applied both to indicate the problems facing humankind and the Earth and to propose solutions. Without mathematical data and models, we would not understand the powerful impact we are having in degrading our environment, how pandemics spread and are checked, and the ways in which we can develop green technologies to counter global warming. Through mathematics we come understand, at least in part, for numerical knowledge can never equal the lived experience of discrimination and deprivation, such issues as racial injustice, gender imbalances in society, and the widening gap between the rich and poor. Once again, education in the ethics of mathematical applications should better sensitize mathematicians toward these problems and mobilize their efforts toward finding solutions.

References

Ahuja A (2020) Boris Johnson's Covid-19 threat alert system is a parody of mathematical precision. New Statesman, 11 May 2020. https://www.newstatesman.com/science-tech/coronavirus/2020/05/boris-johnson-s-covid-19-threat-alert-system-parody-mathematical. Accessed 16 May 2020

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31(2):211–236

Almeida DF, Joseph GG (2004) Eurocentrism in the history of mathematics: the case of the Kerala School. Race Class 45(4):5–59

AMS (2019) American Mathematical Society policy statement on ethical guidelines. http://www.ams.org/about-us/governance/policy-statements/sec-ethics. Accessed 1 Jan 2021

Andersen LE, Andersen H, Sørensen HK (2020) The role of testimony in mathematics. Synthese. https://doi.org/10.1007/s11229-020-02734-9

Balaguer M (1998) Platonism and anti-platonism in mathematics. Oxford University Press, Oxford, UK

BBC (2018) Google "to end" Pentagon Artificial Intelligence project. BBC online. https://www.bbc.co.uk/news/business-44341490. Accessed 22 Nov 2020. (This link does not work directly as of November 26, 2020. However, a search either with Google or BBC search via https://www.bbc.co.uk/search with the article title links to the required web page with this listed URL)

Beauchamp TL, Childress JF (2001) Principles of biomedical ethics. Oxford University Press, New York

Beer D (2016) Metric power. Palgrave Macmillan, London

Bernal M (1987) Black Athena: the Afroasiatic roots of classical civilisation, vol 1. Free Association Books, London

Bishop AJ (1988) Mathematical enculturation. Kluwer Academic, Dordrecht
Cadwalladr C (2017) The great British Brexit robbery: how our democracy was hijacked. The Guardian, 7 May 2017. https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy. Accessed 12 Oct 2019
Chiodo M, Bursill-Hall P (2018) Four levels of ethical engagement, ethics in mathematics, discussion paper 1/2018. https://www.ethics.maths.cam.ac.uk/assets/dp/18_1.pdf. Accessed 19 Nov 2020
Chiodo M, Vyas R (2018) The role of ethics in a mathematical education: a whitepaper. https://ethics.maths.cam.ac.uk/assets/dp/19_1.pdf. Accessed 20 Oct 2019
Cole J (2013) Towards an institutional account of the objectivity, necessity, and atemporality of mathematics. Philos Math 21(1):9–36
CUEiM (Cambridge University Ethics in Mathematics Project) (2020) The Cambridge University Ethics in Mathematics Project. https://www.ethics.maths.cam.ac.uk/. Accessed 19 Nov 2020
D'Ambrosio U (1985) Ethnomathematics and its Place in the History and Pedagogy of Mathematics. For The Learning of Mathematics 5(1):44–48
Davis C (1988) A Hippocratic oath for mathematicians? In: Keitel C (ed) Mathematics, education and society. UNESCO, Paris, pp 44–47
Davis PJ, Hersh R (1980) The mathematical experience. Birkhauser, Boston
Derman E, Wilmott P (2009) The financial modelers' manifesto. https://www.uio.no/studier/emner/sv/oekonomi/ECON4135/h09/undervisningsmateriale/FinancialModelersManifesto.pdf. Accessed 15 Nov 2020
Dowling P (1998) The sociology of mathematics education. Falmer (Routledge), London
Dreyfus SE, Dreyfus HL (1986) Mind over machine. Free Press, New York
EiM 1 (2018) The first meeting on ethics in mathematics. https://www.ethics.maths.cam.ac.uk/EiM1/. Accessed 19 Nov 2020
EiM 2 (2019) The second meeting on ethics in mathematics. https://www.ethics.maths.cam.ac.uk/EiM2/. Accessed 19 Nov 2020
EMS (2012) Code of practice of the European Mathematical Society. https://euro-math-soc.eu/system/files/uploads/COP-approved.pdf. Accessed 1 Jan 2021
Ernest P (1991) The philosophy of mathematics education. Routledge, London
Ernest P (1998) Social constructivism as a philosophy of mathematics. SUNY Press, Albany
Ernest P (2009) The philosophy of mathematics, values and Keralese mathematics. In: Ernest P, Greer B, Sriraman B (eds) Critical issues in mathematics education. Information Age Publishing, Charlotte, pp 189–204
Ernest P (2016) Mathematics and values. In: Larvor B (ed) Mathematical cultures. Springer, Cham, pp 189–214
Ernest P (2018) The ethics of mathematics: is mathematics harmful? In: Ernest P (ed) The philosophy of mathematics education today. Springer, Cham, pp 187–216
Ernest P (2019a) Privilege, power and performativity: the ethics of mathematics in society and education. Philos Math Educ J 35:1–19. http://socialsciences.exeter.ac.uk/education/research/centres/stem/publications/pmej/pome35/index.html. Accessed 13 Oct 2020
Ernest P (2019b) The ethical obligations of the mathematics teacher. J Pedagog Res 3(1):80–91. https://doi.org/10.33902/JPR.2019.6
Ernest P (2020a) Mathematics, ethics and purism: an application of MacIntyre's virtue theory. Synthese. https://doi.org/10.1007/s11229-020-02928-1
Ernest P (2020b) Unpicking the meaning of the deceptive mathematics behind the Covid alert levels. Philos Math Educ J 36:1–7
Evans J, Tsatsaroni A, Czarnecka B (2014) Mathematical images in advertising: constructing difference and shaping identity, in global consumer culture. Educ Stud Math 85:3–27
Ferreirós J (2016) Purity as a value in the German-speaking area. In: Larvor B (ed) Mathematical cultures: the London meetings 2012–2014. Birkhäuser, Basel, pp 215–234
Fineman S (2004) Getting the measure of emotion – and the cautionary tale of emotional intelligence. Hum Relat 57:719–740

Floridi L, Sanders JW (2004) On the morality of artificial agents. Mind Mach 14(3):349–379
Frankel MS (2013) Scientists' social responsibilities in an evolving global research environment. Address to American Association for the Advancement of Science, annual meeting, Boston, 16 February 2013. https://www.aaas.org/sites/default/files/Social-Responsibility_Frankel.pdf. Accessed 26 Nov 2020
Franklin J (2005) A "professional issues and ethics in mathematics" course. Aust Math Soc Gaz 32(2):98–100. http://www.austms.org.au/wp-content/uploads/Gazette/2005/May05/franklin.pdf. Accessed 26 Nov 2020
Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. English edition (1967): From Frege to Gödel: a source book in mathematical logic (trans and ed: van Heijenoort J). Harvard University Press, Cambridge, MA, pp 592–617
Goldacre B (2012) Bad pharma. Faber and Faber, London
Goldstein D (2014) The teacher wars: a history of America's most embattled profession, 1st edn. Doubleday, New York
Gorur R (2016) Seeing like PISA: a cautionary tale about the performativity of international assessments. Eur Educ Res J 15(5):598–616
Grind K, Schechner S McMillan R, West J (2019) How Google interferes with its search algorithms and changes your results. The Wall Street Journal, 15 November 2019. https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753. Accessed 1 Jan 2021
Grossman V (1980) Life and fate. New York Review Classics, New York
Hacking I (2014) Why is there philosophy of mathematics at all? Cambridge University Press, Cambridge, UK
Hardy GH (1929) Mathematical proof. Mind 38(149):1–25
Hardy GH (1940) A mathematician's apology. Cambridge University Press, Cambridge, UK
Harper D (n.d.) Online etymology dictionary. http://www.etymonline.com/index.php?term=justification. Accessed 14 Sept 2013
Harris M (2015) Mathematics without apologies: portrait of a problematic vocation. Princeton University Press, Princeton
Harris M (2017) Do mathematicians have responsibilities? In: Sriraman B (ed) Humanizing mathematics and its philosophy. Springer, Cham, pp 115–123
Hauge KH, Kacerja S, Lilland IE (2019) Xenophobia and numbers in the media – discussing mathematics education in the post-truth era. Philos Math Educ J 35:1–23. http://socialsciences.exeter.ac.uk/education/research/centres/stem/publications/pmej/pome35/index.html. Accessed 28 Nov 2020
Hersh R (1990) Mathematics and ethics. Humanist Math Netw J 1(5):20–23. http://scholarship.claremont.edu/hmnj/vol1/iss5/9. Accessed 20 Dec 2020
Hersh R (1997) What is mathematics, really? Jonathon Cape, London
Hyndman M (1978) Schools and schooling in England and Wales: a documentary history. Harper & Row, London
Johnson TC (2012) Ethics and finance: the role of mathematics. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2159196. Accessed 15 Sept 2013
Johnson TC (2017) Ethics in quantitative finance: a pragmatic financial market theory. Springer, Cham
Johnson B (2020) UK Prime Minister's broadcast on British national TV, 10 May 2020 about the Covid-19 virus threat and the UK Government response. https://www.youtube.com/watch?v=P8HC5sOHzus. Accessed 22 May 2020
Joseph GG (2000) The crest of the peacock, non-European roots of mathematics. Princeton University Press, Princeton/Oxford, UK
Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press, New York
Kline M (1980) Mathematics: the loss of certainty. Oxford University Press, Oxford, UK
Knuth DE (1985) Algorithmic thinking and mathematical thinking. Am Math Mon 92:170–181

Lakatos I (1976) In: Worrall J, Zahar E (eds) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge, UK
Larson EB, Yao X (2005) Clinical empathy as emotional labor in the patient-physician relationship. J Am Med Assoc 293(9):1100–1106
Lenhard J, Carrier M (2015) Mathematics as a Tool. ZiF-Mitteilungen, no. 2/2015, Center for Interdisciplinary Research, University of Bielefeld, pp 10–19. https://www.uni-bi.de/(en)/ZiF/Publikationen/Mitteilungen/Ausgaben/2015-2.pdf. Accessed 30 June 2020
Lerman S (2010) Theories of mathematics education: is plurality a problem? In: Sriraman B, English LD (eds) Theories of mathematics education: seeking new frontiers. Springer, Heidelberg/London/New York, pp 99–109
MacIntyre A (2007) After virtue: a study in moral theory, 3rd edn. Bloomsbury Academic, London
MacKenzie DA (1981) Statistics in Britain, 1865–1930: the social construction of scientific knowledge. Edinburgh University Press, Edinburgh
MacKenzie DA (2006) An engine, not a camera, how financial models shape markets. MIT Press, Cambridge, MA
Marjanovic O, Cecez-Kecmanovic D, Vidgen R (2018) Algorithmic pollution: understanding and responding to negative consequences of algorithmic decision-making. In: Schultze U, Aanestad M, Mähring M, Østerlund C, Riemer K (eds) Living with monsters? Social implications of algorithmic phenomena, hybrid agency, and the performativity of technology. Springer, Cham, pp 31–47
Merriam-Webster (2020) Hot button. Merriam-Webster.Com dictionary. Merriam-Webster, Springfield. https://www.merriam-webster.com/dictionary/hot%20button. Accessed 22 Nov 2020
Morris RL (2020) Intellectual generosity and the reward structure of mathematics. Synthese. https://doi.org/10.1007/s11229-020-02660-w
Noble SU (2018) Algorithms of oppression: how search engines reinforce racism. New York University Press, New York
O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy. Crown Books, New York
Ofqual (2020) Awarding GCSE, AS, A level, advanced extension awards and extended project qualifications in summer 2020: interim report. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909368/6656-1_Awarding_GCSE__AS__A_level__advanced_extension_awards_and_extended_project_qualifications_in_summer_2020_-_interim_report.pdf. Accessed 13 Nov 2020
Pearce IG (n.d.) Indian mathematics: redressing the balance. http://www-history.mcs.st-andrews.ac.uk/history/Projects/Pearce/index.html. Accessed 26 Nov 2020
Penrose R (1989) The emperor's new mind. Oxford University Press, Oxford, UK
Plato (1941) The republic of Plato (trans and annotated: Cornford FM). The Clarendon Press, Oxford, UK
Popper KR (1969) The moral responsibility of the scientist. Encounter (March 1969), pp 52–56. Republished as Popper KR (1971) The moral responsibility of the scientist. Bull Peace Propos 2(3):279–283. https://doi.org/10.1177/096701067100200311
Powell AB, Frankenstein M (eds) (1997) Ethnomathematics: challenging eurocentrism in mathematics. SUNY Press, Albany
Power M (1999) The audit society – rituals of verification. Oxford University Press, Oxford, UK
Radical Statistics Group (n.d.) About us. http://www.radstats.org.uk/about-radical-statistics/. Accessed 3 Aug 2017
Restivo SP (1985) The social relations of physics, mysticism and mathematics. Reidel Publishing Company, Dordrecht
Russell B, Einstein A (1955) The Russell-Einstein manifesto. http://www.pugwash.org/about/manifesto.htm. Accessed 23 Apr 2015
Salmon, F. (2009). Recipe for disaster: the formula that killed Wall Street. Wired, 23 February 2009. https://www.wired.com/2009/02/wp-quant/. Accessed 23 Nov 2020

Sample I (2019) Maths and tech specialists need Hippocratic oath, says academic. The Guardian, 16 August 2019. https://www.theguardian.com/science/2019/aug/16/mathematicians-need-doctor-style-hippocratic-oath-says-academic-hannah-fry. Accessed 25 Nov 2020
Seife C (2010) Proofiness: the dark arts of mathematical deception. Viking, New York
Sharma OP (2015) Ethics in science. Indian J Microbiol 55(3):341–344. https://doi.org/10.1007/s12088-015-0532-x.
Skemp RR (1976) Relational understanding and instrumental understanding. Math Teach 77:20–26
Skovsmose O (2014) Foregrounds, opaque stories about learning. Springer, Cham
Skovsmose O (2019) Crisis, critique and mathematics. Philos Math Educ J 35:1–26. http://socialsciences.exeter.ac.uk/education/research/centres/stem/publications/pmej/pome35/index.html. Accessed 28 Nov 2020
Skovsmose O (2020) Mathematics and ethics (Matemática e Ética). In: Bicudo M (ed) Revista Pesquisa Qualitativa. Edição Especial: Filosofia da Matemática 8(18). https://editora.sepq.org.br/rpq/issue/view/19. Accessed 28 Nov 2020
Skovsmose O, Ravn O (2019) Connecting humans to equations – a reinterpretation of the philosophy of mathematics. Springer, Cham
Slote M (2007) The ethics of care and empathy. Routledge, London
Tanswell FS, Kidd IJ (2020) Mathematical practice and epistemic virtue and vice. Synthese. https://doi.org/10.1007/s11229-020-02664-6
Tymoczko T (ed) (1986) New directions in the philosophy of mathematics. Birkhauser, Boston
University College London (2019) Mathematicians should be taught ethics "from day one". UCL News, 19 August 2019. https://www.ucl.ac.uk/news/headlines/2019/aug/mathematicians-should-be-taught-ethics-day-one. Accessed 25 Nov 2020
Welch A (2009) Mammon, markets, and managerialism. In: Cowen R, Kazamias AM (eds) International handbook of comparative education. Springer, New York, pp 597–600
Wikipedia (2020a) Philosophy of mathematics. Wikipedia. The free encyclopedia. https://en.wikipedia.org/wiki/Philosophy_of_mathematics. Accessed 4 July 2020
Wikipedia (2020b) Medical ethics. Wikipedia. The free encyclopedia. https://en.wikipedia.org/wiki/Medical_ethics. Accessed 26 Aug 2020
Wikipedia (2020c) Hippocratic oath for scientists. Wikipedia. The free encyclopedia. https://en.wikipedia.org/w/index.php?title=Hippocratic_Oath_for_scientists&oldid=951799973. Accessed 30 Aug 2020
Wikipedia (2020d) Archimedean oath. Wikipedia. The free encyclopedia. https://en.wikipedia.org/w/index.php?title=Archimedean_Oath&oldid=935595355. Accessed 31 Aug 2020
Wittgenstein L (1953) Philosophical investigations (trans: Anscombe GEM). Basil Blackwell, Oxford, UK
Wittgenstein L (1956) Remarks on the foundations of mathematics. Basil Blackwell, Oxford, UK
Woolley SC, Howard PN (eds) (2018) Computational propaganda: political parties, politicians, and political manipulation on social media. Oxford University Press, Oxford, UK
Zuboff S (2019) The age of surveillance capitalism: the fight for a human future at the new frontier of power. PublicAffairs, New York

Grothendieck: A Short Guide to His Mathematical and Philosophical Work (1949–1991)

Fernando Zalamea

Contents

1 Introduction 1258
2 First Period: 1949–1957 1258
2.1 Produits Tensoriels Topologiques et Espaces nucléaires (1949–1953) 1259
2.2 Résumé de la théorie métrique Des Produits Tensoriels Topologiques (1953) 1260
2.3 Sur Quelques Points d'algèbre Homologique (1955–1956) 1262
2.4 Classes de Faisceaux et théorème de Riemann Roch (1955–1957) 1264
3 Second Period: 1958–1970 1266
3.1 The Cohomology Theory of Abstract Algebraic Varieties (1958) 1266
3.2 Techniques de Construction en géométrie analytique (1961) 1268
3.3 Éléments de Géométrie Algébrique (EGA) (1959–1964) 1269
3.4 Séminaire de Géométrie Algébrique (SGA) (1960–1969) 1272
3.5 Standard Conjectures on Algebraic Cycles (1968) 1274
4 Third Period: 1981–1991 1276
4.1 La Longue Marche à Travers la théorie de Galois (1981) 1276
4.2 Pursuing Stacks (1983) 1278
4.3 Esquisse d'un Programme (1984) 1280
4.4 Les dérivateurs (1991) 1282
4.5 Récoltes et Semailles (1983–1986) 1284
4.6 La Clef Des Songes (1987–1988) 1286
5 A Synthetic Account of Grothendieck's Work 1288
5.1 Mathematical Heart and Categorical Reason 1288
5.2 Étalé and étale 1290
5.3 Relativity and Universality 1291
5.4 Types and Archetypes 1292
References 1294

F. Zalamea (✉)
Universidad Nacional de Colombia, Bogotá, Colombia
e-mail: fzalameat@unal.edu.co

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_27

Abstract

We present a short survey of Grothendieck's "known" work (1949–1991), before his seclusion in the Pyrenees (1991–2014). Three main periods are distinguishable: (1) 1949–1957, with his work in functional analysis, homological algebra, categories, sheaves, and *K*-theory; (2) 1958–1970, with the foundations of algebraic geometry, schemes, toposes, and motives; and (3) 1981–1991, with topological algebra, anabelian geometry, stacks, derivators, on the one hand, and long philosophical reflections, on the other hand.

Keywords

Grothendieck · Functional analysis · Category theory · Algebraic geometry · Topological algebra

1 Introduction

The immensity and depth of Grothendieck's work has prevented the community from obtaining a fair appraisal of his contributions. This is somewhat surprising, considering that Grothendieck himself often insisted in the "childish" flavor of his principal ideas. In this survey, organized in chronological order, we will emphasize those "simple" conceptual forces inscribed in his many mathematical creations, leaving some detailed technicalities for a wider monograph (see **Zal20**). Contrary to other approaches of Grothendieck's work, this survey will emphasize *(i)* the coherence and *continuity* of the entire *opus*; *(ii)* the *multilayered* levels of Grothendieck's thought, as it moves between abstraction and concretion, universality and particularity, archetypes and types, unity and multiplicity, structure and calculation, transits and obstructions; *(iii)* a permanent search for *smoothness* in Grothendieck's mathematical practice; *(iv)* an *equilibrium* of inventive power in the three main periods, 1949–1957, 1958–1970, and 1981–1991, against the usual reception, where the IHES period (1958–1970) is unduly overemphasized.

2 First Period: 1949–1957

Grothendieck's first period positions him as a mathematical *genius* (a term to be used sparingly in a century), with an incredible power of concentration, capable to navigate both the higher abstractions and the most concrete calculations, and to transit between widely different fields (functional analysis, homological algebra, categories, complex variables). After becoming the world's greatest specialist in topological vector spaces, Grothendieck turns away to algebraic geometry and introduces sheaves at the very heart of his mathematical practice. In the space of just a few years, Grothendieck opens up entire new perspectives of research and begins to deploy a *grand vision* and a *free inventiveness*, unheard of in his generation.

2.1 Produits Tensoriels Topologiques et Espaces nucléaires (1949–1953)

The *Ph.D. Thesis* [**g**1949–1953] situates Grothendieck among the most brilliant mathematicians of his generation. (The references of the form [**year(s)**, page] send to the bibliography of "Grothendieck Works," at the end of this article. These years mark the *periods of creation* of the works, sometimes before publication.) The story according to which the rebellious young man who had spent the year 1948–1949 in the *Séminaire Cartan* is sent by an annoyed Cartan in 1949 to Nancy, where his mentors Schwartz and Dieudonné assign him 14 open problems to calm his frenzy, is well known. (**Scharlau1** pp. 182–183; **Bringuier** pp. 38–39.) Some months later, he has solved them all, and possesses, according to Schwartz's testimony, many results which would suffice for multiple papers. The *Thesis*, a development of those researches, is defended in 1953 (February 28, doctoral committee consisting of Henri Cartan (*Président*), Laurent Schwartz (*Rapporteur*), and Gustave Choquet (*Examinateur*)), and published in 1955 by the *American Mathematical Society.*

Tensor products and nuclear spaces (introduced by Grothendieck) are studied at the core of the work. In some subsequent accounts of the *Thesis*, Grothendieck emphasizes a characterization of *nuclearity* through a *smooth* coincidence of projectivity $\hat{(\,)}$ and injectivity $\check{(\,)}$ (later denoted $(^{\vee})$), that is, E is nuclear if and only if $E\hat{\otimes}F \approx E\check{\otimes}F$ for all spaces F locally convex (or, alternatively, for all Banach [**g**1952a 100], [**g**1952b 196]). Nuclearity thus corresponds to a stability or *fixed-point* condition, but the *deviations* to the fixed-point situation provide, on the other hand, a very interesting expansion of intermediate variations (an illuminating case will be the 14 natural tensor norms in the case of Banach spaces, intermediate between the projective and the injective norms, see Sect. 5.4 below).

The *combination* of diverse mathematical techniques is exemplary. Topologization, functionalization, algebrization, linearity, ordering, are within the *Thesis*. The technical exuberance is remarkable, if one further takes into account that, a mere couple of years before, Grothendieck did not know anything about modern mathematical structures (in *Récoltes et semailles*, he reckons "ignoring what was a topological space" in 1948, and "not having yet heard uttered, at least in a mathematical context, such strange or barbaric words as group, field, ring, module, complex, homology" [**g1986** 1.18–19]). The entirely novel practice of topological tensor products, the structural behavior of Fredholm's theory, the use of topological projective and injective limits, the original connections with spaces of integrable functions, summable functions, or distributions, the emergence of incisive counter-examples, are all strong expressions of the fascinating *technical struggle* of the young *genius* with his more conventional mathematical surroundings.

The *Thesis* reveals a *multilayered strategy* which will become a trademark of Grothendieck's *style*: a *back-and-forth* between *(i)* definitions, *(ii)* mathematical characterizations, *(iii)* permanence properties, *(iv)* derived theorems, *(v)* examples and counterexamples, *(vi)* specific applications, *(vii)* returns to definitions in ascending or descending levels, etc. In particular, the *art of definition* (of which

Grothendieck will be very proud, having introduced more than a thousand definitions!) stands as one of the central features of a very careful Grothendieckian practice. In a wider context, Grothendieck's *art of language*, an extremely chiseled French sometimes close to Proust's elaborate sentences, is something that will be stressed by Serre and Cartier. The *clearness* of the architectonics depends in fact on the definitional adjustments in the edifice. In that sense, the *Thesis* is paradigmatic, with its long chains of equivalences or with its progressive subdeterminations between concepts, which may wait hundreds of pages before being resolved.

2.2 Résumé de la théorie métrique Des Produits Tensoriels Topologiques (1953)

The *Résumé de la théorie métrique des produits tensoriels topologiques* [**g**1953c] can be seen as a concretion of the *Ph.D. Thesis*, where the general ideas for tensor products of locally convex spaces are specified to the case of Banach spaces. The restriction to the class of Banach spaces produces some outstanding *structural* results on the intermediate norms between the projective and injective norms, described systematically through topological, algebraic, and lattice-theoretic properties. Grothendieck achieves a *characterization* of (the category of) Hilbert spaces inside Banach spaces, through injective and projective properties, solving the Banach-Mazur problem. On the other hand, he inverts correlations between the projective norm and the other intermediate norms (Grothendieck's inequality), thanks to the emergence of a deep constant, still not completely known today. (For a complete presentation of the work, see **DiestelFS**.) The article's inventiveness, mathematical and symbolical, constitutes one of the most accomplished and harmonious texts of Grothendieck's output, and may be recommended as a first entry to his world.

After finishing his *Thesis*, the young apatrid has a difficult time finding a job in France, and Schwartz and Dieudonné help him get a postdoc in São Paulo. (Between the available biographies (Scharlau, Schneps, Bringuier, Douroux, Pradeau), only Schneps investigates with care Grothendieck's stay in Brazil. See **Scharlau1** pp. 185–186, **Schneps2** Chap. 4, pp. 1–7, **Bringuier** p. 41, **Douroux** p. 126, and **Pradeau** p. 74.) In the first year, between March and October 1953, he produces his monograph on topological vector spaces [**g**1953d] and the *Résumé* [**g1953c**], yielding another period of splendor. (The text *Espaces vectoriels topologiques* appears as a "Curso de extensão universitária da Facultade de Filosofia, Ciências e Letras da Universidade de S. Paulo." It would be fully published only in an English translation in 1973: see [**g1953d**].) The *Résumé* is submitted a year later to the *Boletim da Sociedade de Matemática de São Paulo* and published in that journal in 1956. (The reception acknowledgement reads, "Trabalho recibido pela Sociedade de Matemática de São Paulo em Junho de 1954.") The second year in São Paulo will be, instead, a time of crisis (the only one reckoned by Grothendieck in his mathematical life), which will propel him to algebraic geometry (see Sect. 2.3 below). ("At the moment in which I was still doing Functional Analysis, that is around 1954,

it happened to me that I persisted without end on a question that I could not solve, I didn't have ideas and was content with turning around a circle of old ideas which, visibly, did not 'bite'. It was like this for a year, around the "approximation problem" for topological vector spaces, which was going to be solved only twenty years later by methods of a totally different order (...) It was a painful year - the only moment in my life that doing mathematics turned out painful for me! (...) From 1954 on, I took the habit in mathematics to have many irons on the fire at the same time" [**g1986** 1.163].)

A natural continuation of the *Thesis* [**g1949–1953**], the *Résumé* provides a new understanding of classes (not yet called "categories") of Banach and Hilbert spaces, from a triple perspective attentive to inequalities between tensor norms, analytic structures, and algebraic factorizations. Bourbaki's three central *structures-mère* enter simultaneously in play – order, topology, algebra – to solve difficult open problems (Banach-Mazur) and to open unexpected new roads (Grothendieck's inequality). (For an excellent introduction to Grothendieck's inequality and its unexpected uses in quantum mechanics, noncommutative geometry, graph theory, computation theory, etc., see **Pisier**.) The *back-and-forth* between the abstract – Banach and Hilbert spaces *at large* – and the concrete – spaces of type C (continuous functions, including c_0), L (integrable functions, including l^1), λ (L-subspaces), γ (C-quotients) – helps to anchor a general (archetypical) category around the particular spaces (types) which govern it (for precisions on "types" and "archetypes," see Sect. 5.4 below). In these processes, a *harmonious smoothness* prevails, yielding naturally a solution of Banach-Mazur: every Hilbert space can be characterized as being *simultaneously* a subspace of type L and a quotient of type C, that is, simultaneously a λ and γ space.

The methodological strategy of the *Résumé* [**g1953c**] is fascinating: *(i)* describe comparable canonical norms, *(ii)* transform those norms, and *(iii)* cover the full spectrum thanks to a structural hierarchy. Those three general levels possess various sublevels of understanding: (A) following Bourbaki, we have *(i)* order, *(ii)* topology, and *(iii)* algebra; (B) semiotically, we have *(i)* syntax, *(ii)* semantics, and *(iii)* pragmatics; (C) phenomenologically, we have *(i)* separation, *(ii)* continuity, and *(iii) entrelacs*; and (D) musically, we have *(i)* theme, *(ii)* variations, and *(iii)* fugue. (The "entrelacs" in Merleau-Ponty summons the interweaving of multiple perception levels, something also reflected in the "chiasme," interconnection of vision perspectives in the mind. See **Merleau1**; **Merleau2**.) In particular, at level *(i)*, we can recognize one of Grothendieck's major obsessions. In fact, from his *Thesis* until his generalized proof of Riemann-Roch, through the creation of abelian categories, Grothendieck detected ubiquitous appearances of *projectivity* and *injectivity* in many mathematical regions: projective and injective products (*Thesis* [**g1949–1953**]), projective and injective norms (*Résumé* [**g1953c**]), existence of enough injectives in abelian categories with a generator and infinitude axioms in order to reconstruct homology from derived functors (*Tôhoku* [**g**1955–1956]), inductive subspecializations (= "spatializations") of the generalized Riemann-Roch proof in the projective and injective cases (*Rapport Riemann-Roch* [**g1957**]).

2.3 Sur Quelques Points d'algèbre Homologique (1955–1956)

Sur quelques points d'algèbre homologique (nicknamed *Tôhoku* for its place of publication) [**g**1955–1956] insures the *takeoff* of category theory, with a new arsenal of ideas, techniques, examples, and theorems, which go well beyond the initial emergence of the theory in Mac Lane and Eilenberg (1945–1948). Grothendieck introduces central categorical concepts (subobjects, representable functors, equivalences, adjunctions), defines additive and abelian categories, uses many examples of categories of sheaves, axiomatizes infinitary processes, elucidates a common frame for cohomology and derived functors, provides representation and exactness theorems, proves (using Cantor's transfinite recursion!) the existence of enough injectives in abelian categories with a generator and suitable infinitary axioms. (A (topological) sheaf is the simplest and deepest mathematical object to study the *transits* and *obstructions* between the *local* and the *global*. A *sheaf* is formed by two topological spaces E and S, and a projection $p : E \to S$ which is a local homeomorphism (i.e., for each point $e \in E$, there exists a neighborhood $V \ni e$ such that p is a homeomorphism from V into its image). In essence, the space E is then deployed as a series of "layers" locally similar to fragments of S: S is a sort of "basis" which *unfolds* along E, and E is a "multiple" space which is *folded* over a "unitary" space S. The unfolding of a point $s \in S\ (= p^{-1}(s))$ is by definition the *fiber* over s. We have then an image of a space E *"étalé"* (in the terminology of Cartan's Seminary in the 1950s), that is, a deployed space over its basis S. One should not confuse here *étalé* with *étale*, basically opposed concepts, distinguished by only a final accent (see Sect. 5.2 below). The local inverses of p (over neighborhoods in S) are the *local sections*; the global inverses of p (over all S) are the *global sections*. The fundamental objective of *sheaf theory* consists in finding gluing conditions (or proving their impossibility) for local sections, in order to produce global sections (or not). Sheaves emerge from Riemann's ideas on *analytical continuation* in complex variables, and the paradigm of a sheaf is given by the germs of holomorphic functions (for other crucial examples, see notes 18 and 22 below).) An *imaginative richness* and an extremely *precise technique* go hand in hand along a never-ending "article" (102 pp.), where the *invention* of a language (categorization, abelianization) and the *discovery* of deep mathematical structures (sheaves, monoid/group actions, injective objects, derived functors) reveal *natural* connections between algebraic geometry, topology, (co)homology, and complex variables. Still today, the *Tôhoku* article may be recommended as an enlightening (and mostly readable) introduction to the richness of Category Theory. Only *patience* is needed, as always with Grothendieck, while he explains *everything*, in painstakingly humble and accurate detail.

A postdoc in Kansas, in 1955, will be the occasion for Grothendieck's immersion in algebraic geometry. ("The year 1955 marks a crucial turn in my mathematical work: the shift from 'analysis' to 'geometry'. I still recall that strong impression, as if I were leaving arid and aloof steppes, to find suddenly myself in some sort of 'promised Paradise' with luxurious richnesses, multiplied to the infinite (...) There is something in mathematics that (forever without doubt) fascinates me beyond any

other thing, not 'number', nor 'magnitude', but always form. And between the thousand and one faces that form chooses to be revealed to us, the one that has fascinated me most, is the hidden structure of mathematical things" [**g1986** P.26–27].) In that *crucial turn*, the role of Jean-Pierre Serre is fundamental. The correspondence with Serre [**g**1955–1987] presents the progressive emergence of the *Tôhoku* [**g1955–1956**]: (1) February 25, 1955, Grothendieck writes that he is exploring an analogy between the theory of derived functors and the cohomology with coefficients in a sheaf, without knowing yet "if it works well in the case of a non separated space (*Non-separation*, linked to general forms of *smoothness*, will become one of the pillars of Grothendieck's work. In particular, the need to consider non-separated sheaves will be one of the roads that will lead naturally to the appearance of the *topos étale of a scheme*, key in the resolution of Weil's conjectures. See Sects. 3.1, 3.4 below.)" [**g1955–1987** 13–14]; (2) July 13, 1955, Serre informs that "your article on homological algebra was read with care [in Bourbaki's congress] and converted everyone (even Dieudonné, who seems completely functorialized!)" [**g1955–1987** 17]; (3) a year later, September 191,956, Grothendieck indicates that he "has passed the best of last month redacting my *multiplodocus* of homological algebra (...) more than 100 pages (80 already drawn up) in big format" [**g1955–1987** 43]; (4) November 13, 1956, Grothendieck announces that he is going to propose the manuscript to Tannaka's *Tôhoku*, where "it seems that the river-flowing articles do not discourage them" [**g1955–1987** 49]. (The "multiplodocus" (*multiplodoque*), a Bourbaki joke, evokes a gigantic treaty, worthy of the size of a diplodocus, long held to be the longest of the dinosaurs.) *Sur quelques points d'algèbre homologique* is finally received March 1, 1957, in the *Tôhoku Mathematical Journal*, and published in Japan in that same year. We are here in presence of new *global and extended mathematical practices*, which overcome strong local obstructions and guarded traditions (in France and the USA). (For a precise description of the intermediate events related to the publication, in particular some difficulties with Eilenberg and the American journals (*American Journal of Mathematics, Transactions of the AMS*), see **Kromer2** pp. 120–128.)

Tôhoku's technical inventiveness pushes out into many levels of mathematical complexity: *(i)* the first construction, *ex nihilo*, of a general axiomatic categorical environment (abelian categories with additional axioms) where sophisticated mathematics could be achieved (See the following judgments (taken from Krömer, **Kromer2** p. 158): Mac Lane (1981): "The first papers on categories had no immediate sequels, because for this period they provided just a language. The notion of category theory as a subject of study in its own right appears only in the third phase of the abstract algebra movement (...) [*i.e.* the period 1957–1974 under Grothendieck's influence]"; Barr & Wells (1985): "[The *Tôhoku*] demonstrated that categories could be a tool for actually *doing* mathematics and from then on the development was rapid."); *(ii)* the elevation and axiomatic proof (via AB5 + generator) of the existence of enough injectives in a universal frame, a result which would allow a reconstruction of cohomology as a particular case of an even more *plastic, smooth, flexible* technique (derived functors); *(iii)* the use of a *multiplicity of functors* (representable, exact, sheaf-like, derived, resolvent) to establish a precise

architectonics of *transits and obstructions* between mathematical regions (topologies, algebras, differential varieties, arithmetical fragments); and *(iv)* the beginning of a *new vision* of algebraic geometry (connections with Serre's work [**g1955–1956** 120, 137, 171], intuition of new perspectives on Riemann and "arithmetical varieties" [**g1955–1956** 165–166, 199]).

The *definitions* are clean and simple, the *examples* are new and multifaceted, the *theorems and corollaries* are finely hierarchized and phased, and the *proofs* are economical and perfectly fitted in. The *genius is at work*, and a (non-exhaustive) list of the definitions introduced by Grothendieck is really bewildering: subobjects [123], (Numbers between brackets refer here to pages in [**g1955–1956**].) equivalence (and implicitly adjunction) [125], kernels and cokernels [126], additive and abelian categories [126–127], infinitary axioms (AB3-AB6) [128–129], categories of diagrams [130], categories of actions [132], (co)generators [134], injective and projective objects [135], quotient categories [137], (co)bording functors [139], derived functors [143], filtrations [145], presheaves and sheaves (functorial characterizations) [153], spectral sequences (functorially) [148, 171], representable functors [185], *G*-sheaves [195]... As mentioned by Barr and Wells, an *actual practice of mathematics inside Category Theory* became a reality with the *Tôhoku*.

2.4 Classes de Faisceaux et théorème de Riemann Roch (1955–1957)

Continuing with Grothendieck's amazing creativity in his first decade, his generalization of the Riemann-Roch theorem introduces another *deep projection from the archetypical to the spectra of types*. (Riemann-Roch (1857) unifies two very different approaches to obtain a natural invariant for a complex surface. On one hand, the genus g of the surface is obtained by counting the number of holes of the surface, or, equivalently, the number of cuts (minus one) with which the surface becomes disconnected (e.g., the genus of the sphere is 0, the genus of the torus is 1, etc.). On the other hand, one can think about the "good" functions (holomorphic) and the "bad" functions (meromorphic) definable on the surface. If we fix n points on the surface, consider the (vector) space *Hol* of holomorphic functions with zeroes on those points, and the (vector) space *Mer* of holomorphic functions with poles on those points. The Riemann-Roch theorem says that $n - g + 1 = dim\,(Mer) - dim\,(Hol)$. In this way, an intrinsic geometric invariant (the genus) is related to an extrinsic complex variable invariant (harmonic difference between *Mer* and *Hol*). At the very heart of profound connections between geometry, topology, complex variables, differential geometry, and algebra, Riemann-Roch marks the beginning of modern algebraic geometry.) In this case, the *K*-theory group invented by Grothendieck ($K(X)$ = free group of coherent sheaves over a variety X) [**g1957**] is projected onto the cohomology groups, providing a new, higher understanding of *forms*.

(Given a sheaf of rings over a topological space X (that is, where the fibers are rings, and where we denote the total space of those fibers by $\mathcal{O}_X$), one can then take $\mathcal{O}_X$ as a basis for a new sheaf, whose fibers are modules over each $R \in \mathcal{O}_X$. The resulting object (with due initial topologies) is a *sheaf of modules over a sheaf of rings*. If good properties of finite type moreover hold in the vertical transitions (modules of finite type) and in the horizontal transitions (kernels of finite type), we have the notion of *coherent sheaf* (Serre, FAC 1955). The article **Serre2** had a strong influence in Grothendieck. A fundamental example of coherent sheaf is the sheaf of germs of holomorphic functions (everything becomes thus connected around the complex variables, see Sect. 3.2 below).) Using previous tools by Serre (sheafification of Riemann-Roch's statement) and Hirzebruch (topologization of the proof), Grothendieck profits from a quantization (Todd) along the *functorial transit* (Chern) between K-theory and cohomology, from which the classical Riemann-Roch can be derived. The *geometrization* of holomorphic and meromorphic functions in complex variable, the *topologization* present in Riemann surfaces and sheaves, the *arithmetization* of the associated algebraic curves, the *linearization* obtained through group theory, the *categorization* of the structural objects and processes in play, are intertwined in a powerful *unitary vision*.

Even if Grothendieck's interest in Riemann-Roch goes back to Nancy (the "Themes for Bourbaki's Seminar 1951–1952" mention a "Kodaira proof of Riemann-Roch" that Grothendieck was to expose; see *Tribu* 25 (1951), p. 6), the deep new abstract conception is developed between 1955, in Kansas, and 1957, in Paris. The correspondence with Serre shows (1) the initial intuition, February 18, 1955 [**g1955–1987** 5], (2) the final proof, November 1, 1957: "Adjoined a very simple proof of Riemann-Roch, independent of the characteristic" – "I point out to you that 'morally' the new method rests on the determination of $K(X)$ and $A(X)$" [**g1955–1987** 57], underscoring thus the emergence of the group $K(X)$, foundation of K-theory, (3) the corrections and expansions, November 12, 1957: "Following your rectifications (...)" – "With the definition of Chern classes for sheaves (...) Riemann-Roch's theorem can also be stated in the case of non projective varieties (...)" [**g1955–1987** 59].

Grothendieck's main contributions in the generalization of Riemann-Roch occur in the elucidation and *natural understanding* of what seemed before a mysterious "*coup de génie*" by Riemann: *(i)* profound *linearization* of the environment (algebraic geometry, complex variable, arithmetic), thanks to the emergence of the fundamental structures of K-theory (group $K(X)$, ring $A(X)$); *(ii) relativization* of Serre's and Hirzebruch's results (sheaves, cohomology, topology, done for a fixed variety) to the case of a morphism $f : X \to Y$ (*variation of the basis*); *(iii)* ensuing *naturalization* of the Chern transformation between K (K-theory groups) and H^* (cohomologies), adjusted (*quantized* we would say today) through Todd's classes. If f_*^K denotes the image of $f : X \to Y$ under the functor K, and $f_*^H : H^*(X) \to H^*(Y)$ denotes the image of $f : X \to Y$ under the cohomological functors, (We follow here the excellent account **Karoubi2**.) we have that the diagram

$$\begin{array}{ccc} K(X) & \xrightarrow{f_*^K} & K(Y) \\ {\scriptstyle ch_X}\downarrow & & \downarrow{\scriptstyle ch_Y} \\ H^*(X) & \xrightarrow[f_*^H]{} & H^*(Y) \end{array}$$

is *not* commutative, and the Chern morphisms do not work *directly* as a natural transformation between K and H^*. Instead, *indirectly*, through Todd's classes in a generalized Riemann-Roch formula, we do obtain commutativity: $f_*(ch_X(x)T(X)) = ch_Y(f_!(x))T(Y)$ [**g1957** 54–68]. In the particular case in which *Y is reduced to a point*, Grothendieck's formula yields a previous formula by Hirzebruch, $f_*^H(ch_X(x)T(X)) = \sum(-1)^k dim\left(H^k(X,x)\right)$ (where x is fibered over X), which, in turn, yields Serre's formula, (The left side of equality provides the Euler characteristic of the fibration x, that is, the expression $m - gen(f) + 1$ if x is the fibration of f over its Riemann surface, while the right side, restricted to $k = 0, 1$, provides the harmonic cohomological difference between the meromorphic and holomorphic functions definable over the surface of f.) a direct expression of the classical Riemann-Roch in cohomological terms. In this way, the *elevation* of a concrete problematic (original Riemann-Roch), through new mediations (Serre's sheafification and cohomologization, Hirzebruch's topologization), leads to a superior abstract environment (sheaves, K-theory) where one can find the *structural archetypical keys* of the transformations which relate number (differential algebraic dimension) and space (topological genus). *Beyond the fundamental concrete types of mathematical knowledge (number, space), some abstract archetypes (sheaves) govern the situation*: a trademark of Grothendieck's *back-and-forth* understanding of mathematical practice, along the *penduli* universal-particular, abstract-concrete, and global-local.

3 Second Period: 1958–1970

Grothendieck's second period is the epoch of conformation of a strong school, which converts the IHES in the *mecca* of mathematical research in the 1960s. Dieudonné, with his proverbial generosity, already Bourbaki's official writer, disposes himself to the orders of the young *genius* (at only 30 years of age), who becomes surrounded with some of the best colleagues and students of his generation (Serre, M. Artin, Deligne, Gabriel, Verdier, Giraud, Illusie, etc.). The gigantic results (EGA) and (SGA), covering close to 7500 pages in a decade (!), set completely new foundations for algebraic geometry, introduce Grothendieck's main generalizations of number (schemes) and space (toposes), and open the way to solve the Weil conjectures.

3.1 The Cohomology Theory of Abstract Algebraic Varieties (1958)

The Edinburgh lecture, *The cohomology theory of abstract algebraic varieties* [**g**1958], at the International Congress of Mathematicians, constitutes another

major turn in Grothendieck's vision. (Grothendieck presents one of the plenary lectures at the Congress, already a sign of distinction and recognition (for the moment, mainly for the Riemann-Roch achievement). The Fields medals are delivered in Edinburgh to Roth and Thom; Grothendieck will receive it 8 years later, in Moscow, a meeting that he will not attend in objection to the Soviet violations of human rights (Dieudonné will obtain the medal for him): in fact, his anarchist and fearless mind will always be on the side of the oppressed, regardless of any political dogma.)It is the place where *schemes* emerge, allowing the extension and smoothening of many rigid techniques of algebraic varieties. The movements of schemes *over their basis* turn out to be fundamental, as well as the appearance of *nilpotent* elements, related to smooth differential calculations. In an extremely synthetic *pair of lines* (see below), already in the second page of the conference, Grothendieck offers an amazing view of the *étale* world, a deep weaving of Galois and Riemann, which will be extensively developed along schemes (EGA [**g1959–1964**]) and toposes (SGA [**g**1960–1969]). Under the construction of the *étale topos of a scheme*, it will offer the cohomological key to solve Weil's conjectures, 16 years later than the Edinburgh lecture and thousands of pages ahead!

The *heart* of the lecture is directed to one of the central kernels of mathematical thought: the dialectics between the discrete and the continuous. Weil's conjectures (1949) constituted a precise expression of that dialectics, using (after Artin, Hasse, Hecke) a generalized Riemann ζ function along curves over Galois finite fields, and proposing three main properties on the behavior of ζ: rationality, functional equation, and zeroes distribution. After the first emergence of schemes (Serre, Chevalley, Nagata, 1955–1956), where affine spaces were glued, the notion of abstract scheme proposed by Grothendieck glues local rings over the spectrum of *prime* ideals of an *arbitrary* commutative unitary ring (unifying thus the *Riemann trend*, on rings of holomorphic and meromorphic functions, and the *Galois-Dedekind trend*, on rings of algebraic integers). The generalization, a profound enlargement of the notions related to algebraic-topological spaces, turns out to be the needed instrument to study algebraic varieties over a field of characteristic p (discrete environment of the first schemes by Serre-Chevalley-Nagata) through adequate cohomology groups with coefficients in a field of characteristic 0 (continuous environment where one can use Lefschetz techniques to count points). Beyond this, Grothendieck's *genius* appears: the *deep intuition* that *étale* morphisms between schemes form the basis for a (Grothendieck) topology, (An *affine scheme* is the representation sheaf of a (commutative unitary) ring, that is, the sheaf $\pi : E \to X$ formed on the lower space X by the spectrum of prime ideals P of the ring A, with the Zariski topology (closed sets are algebraic sets), and on the upper space E by local ring fibers $\pi^{-1}(P) = A(A - P)^{-1}$. *Schemes* are sheaves which locally behave as affine Schemes. A morphism between schemes is *étale* if its fibers are of finite type (Galois flavor) and non-ramified (Riemann flavor). A formal definition of *étale* appears only much later, see [**g1960–1969** 1.4]) which will allow precisely the construction of the cohomological tools to solve Weil's conjectures. (The rationality will be proved by Dwork (1960, independently of Grothendieck), the functional equation by Grothendieck (1966), and the zeroes distribution by Deligne (1974). Deligne's very hard and ingenious

proof, worthy of a Fields medal (1978), even if using all of Grothendieck's machinery, will depart from the main road designed by the master (see Sect. 3.5 below), something that Grothendieck will never accept. It will be the main mathematical reason for the diatribes against his disciple in *Récoltes et semailles* [**g1986**].)

In the second page of the lecture, Grothendieck writes: "(...) it seems certain now that the Weil cohomology has to be defined by a completely different approach. Such an approach was recently suggested to me by the *connections between sheaf-theoretic cohomology and cohomology of Galois groups on the one hand, and the classification of unramified coverings of a variety on the other* (as explained quite unsystematically in Serre's tentative Mexico paper) (...)" [**g1958** 104, Grothendieck's emphasis]. If the *systematization* of that "revelation" would take many years and many many pages, in the *two underlined lines* one catches the essential kernel of the connections between Galois and Riemann, number and space, algebra and geometry, schemes and toposes. It is a *combined control*, both on the *internal* aspect of the structures given on the fibers of the scheme ("Galois groups": witnesses of finitude and separability) and on the *external* aspect of superposition ("unramified coverings") on the fibers. That *double process of smoothening* on sheaves, internal and external, allows to transfer to the cohomology of new abstract constructs, the counting Lefschetz formulas. If the statement of Weil's conjectures (1949) can be understood as a culmination of *modern* mathematics (tools from abstract algebra and set-theoretic topology to *approach* number theory), its progressive resolution under Grothendieck's school (1958–1974) can be seen as the beginning of *contemporary* mathematics (tools from homological algebra and categorical topology to *deepen* number theory). A complete methodological and mathematical revolution is thus under way in the higher realms of mathematical practice.

3.2 Techniques de Construction en géométrie analytique (1961)

Grothendieck's interest in complex variables was present all along his First Period: spaces of holomorphic and entire functions [**g1949–1953**] [g1953a] [g1953b], Riemann sphere [g1955a], and coherent complex sheaves [**g1955–1956**] [**g1957**]. His understanding of the holomorphic/meromorphic realm is taken further in his lectures for the *Séminaire Cartan* (December 19, 1960–May 8, 1961), *Techniques de construction en géométrie analytique* [**g1961**], where a thorough exploration of the Teichmüller tower is advanced. (Following Henri Cartan, *analytic geometry* referred now to the geometry of analytic (holomorphic) functions. This was very close to Serre's *GAGA*, **Serre1** which was a major influence on Grothendieck.)

The *container* of the lessons seems to reflect its mathematical *content*. Teichmüller's space T_g is first approached axiomatically, then it is described categorically through categories of analytic spaces, then differential local techniques are applied, and, finally, everything becomes integrated thanks to linearizations, coverings, and quotients. (A good introduction to Teichmüller spaces (1937–1939) can be found in **ImaTan**. The *Teichmüller space* over a given surface provides a

parameterization of *all* the complex structures over that surface (it is thus a *holistic, universal, natural* object, of immediate fascination to Grothendieck). If M_g is the *moduli* space of genus g (that is, the set of isomorphism classes of the closed Riemann surfaces of genus g), M_0 is reduced to a point (class of the Riemann sphere), M_1 is reduced to the complex plane (through elliptic curves) and M_g ($g \geq 2$) can be parameterized through $3g - 3$ complex parameters (Riemann 1857). The *Teichmüller space* T_g of genus g is constructed as an adequate covering of M_g, with ramifications associated to the fundamental group of some (= any) closed Riemann surface of genus g. Then T_g acquires globally a *canonical structure* of complex variety, produced from the local fragments of the action of the fundamental group over M_g. The *structure of all* T_g ($g \geq 0$) constitutes what Grothendieck will call later the *Teichmüller Tower* [**g1981**].) The very *text* of the lectures takes on a structural quality – an architectonic *maniera* – with many levels stratified in towers of definitions, lemmas, observations, examples, that reveal the hidden *contrapuntal harmony* of the complex variables environment.

Some of the strategies adopted at the *Séminaire Cartan* highlight the *smoothness* of complex variable spaces: *(i)* release ("*dégager*") functorial mechanisms for the use of coherent sheaves, applied in particular to complex variables, *(ii)* formulate *moduli* space problems in terms of representability of functors, *(iii)* weave *projectivity* properties with existence theorems, and *(iv)* bring closer schemes (algebraic geometry) and holomorphic varieties (analytic geometry), in an extended *GAGA* program which profits from the infinitesimal properties of *nilpotent elements in local rings*. On the other hand, the implicit *dialectics* taking place in complex variables act with all their strength: *(v) ascent* (functoriality, globality) and *descent* (projectivity, modularity, locality, nilpotency), *(vi) back-and-forth* between the One (categorical universality) and the Many (concrete categories of schemes and analytic spaces), *and (vii)* hierarchical stratification of mathematical knowledge along *circulating levels* (differential combinatorics). A highly interesting feature, both mathematical and methodological, of the *Techniques de construction en géométrie analytique* [**g1961**], turns out to be the revelation of *multiple concrete mathematical forms of smoothness ((i)–(iv))* thanks to a *unity of abstract methodological dialectics ((v)–(vii))*.

3.3 Éléments de Géométrie Algébrique (EGA) (1959–1964)

EGA [**g1959–1964**] can be regarded as one of the major treatises of twentieth century mathematics. Grothendieck's tireless inventivity and Dieudonné's not less inexhaustible writing produce in a few years a huge architectonics which offers, around schemes, new foundations for algebraic geometry. The manuscripts of EGA are written between 1959 and 1964, and published between 1960 and 1966. (EGA is divided into IV volumes (and 7 parts): *volume I* (228 pp., finished October 17, 1959, published 1960); *volume II* (222 pp., finished July 22, 1960, published 1961); *volume III.1* (167 pp., finished June 28, 1961, published 1961), *volume III.2* (91 pp., finished December 15, 1962, published 1963); and *volume IV.1* (259 pp.,

finished March 15, 1963, published 1964), *volume IV.2* (231 pp., finished December 15, 1963, published 1965), *volume IV.3* (255 pp., finished November 15, 1964, published 1966). It constitutes an amazing *tour de force* (deepness, discipline, effectiveness) which sums up around 1500 pages and can only be understood under the literal meaning of the *monumental*. We are here in presence of an *ideal of mathematical practice* which has seldom been carried out, in such immensity and concreteness, in the history of mathematics.) The environment of work is the *Institut des Hautes Études Scientifiques* (IHES), created specifically for Dieudonné and Grothendieck. (As Grothendieck recalls, "During the heroic years of the IHES, Dieudonné and me were its only members, and the only ones also to give it credibility and audience in the scientific world, Dieudonné editing the "Publications Mathématiques" (whose first volume appeared already in 1959, the following year of the foundation of the IHES by Léon Motchane), and me directing the "Seminars of Algebraic Geometry" [**g1986** 1.169].) Around them, a School gradually emerges, with brilliant young doctoral students under Grothendieck (some ten students, among which Deligne, Demazure, Giraud, Illusie, the Raynauds, and Verdier). A *community of mathematical practice* propels then the IHES to the forefront of research.

Some of the strategies present in EGA [**g1959–1964**] (and further extrapolated to SGA [**g1960–1969**], see Sect. 3.4 below) include: *(i)* an extension of the notion of *spectrum*, which allows the development of *smoothness and freeness* along structures and sheaves, and opens the way to a natural "algebraic differentiability" thanks to nilpotent elements; *(ii)* the construction of a *new path* for the understanding of number theory, which, in a parallel way to the classical reading (numbers → points → space → separation → discrete geometry), proposes a new reading (numbers → prime ideals → spectrum → non-separation → continuous geometry), and studies a *unification* of the discrete and the continuous; *and (iii)* the elaboration of a *conceptual filtering sieve*, in order to pass from algebraic varieties of characteristic p (environment of the Weil conjectures) to differential forms in a characteristic 0 environment, thanks to finer topologies on sheaves, schemes, and cohomologies. Then, many of the main *smoother* structures in EGA occur (regular rings, flat modules, smooth algebras, ringed spaces, (pre)schemes, coherent sheaves), and a *free* calculus on them is developed (functors, adjunctions, representation theorems, noetherianity, constructibility).

One of the central objectives of schemes consists in exploring worlds where the *indiscernibles* may become *discerned* (one of the many Galois "Resurrections" in Grothendieck's output). This happens, for example, when new infinitesimals are introduced to capture differences. Three main techniques enter into play (see **Dieud2** vol. 2, VIII, §§ 27–47): *(a) relativization*: given a scheme S, study of the *comma category* Sch/S of morphisms $X \to S$, *(b) basification*: study of $X \to S$ through $S' \to S$, which produces (via pullbacks) $X \times_S S' \to S'$, *and (c) representation*: study of $X \to S$ through the representable functor $h_X : Sch/S \to Set$. The techniques help to elucidate – beyond our "blind eyes," as Tarkovsky would later say – some of the

transits between the discrete and the continuous, between the finite and the infinite. Given a morphism between schemes $f: X \to S$, and given $s \in S$, if $\chi(s) = \mathcal{O}_s/\mathfrak{M}_s$ is the residual field over s (quotient of the local ring by its unique maximal ideal) and if one considers the basis change $Spec(\chi(s)) \to S$, the fiber $f^{-1}(s)$ can be seen as the underlying space of $X_s = X \times_S Spec(\chi(s))$ (see **Dieud2** pp. 198–200). Thus, X appears parameterized by the X_s, and a *smoothness* strategy consists in determining f from the properties of the X_s. How much does the punctual basis $Spec(\chi(s))$ determine in fact the general basis S? The *limits* in the category Sch/S can solve the question: the properties of schemes over arbitrary extensions may be reduced to properties over finite type extensions, and the properties of schemes over arbitrary rings can then be understood through properties of finite type algebras over $\mathbb{Z}$. (Dieudonné calls this the "realization of Kronecker's old dream," **Dieud2**, p. 200.)

Lefschetz's formula (valid in characteristic 0) is the fundamental (co)homological tool to count points in varieties (one of the objectives of the Weil conjectures in characteristic p), and thus the *controlled transit* between cohomologies in characteristic p and cohomologies in characteristic 0 becomes central in Grothendieck's program. (If $f : X \to X$ is a continuous function over a triangulable compact topological space X, Lefschetz's formula *counts* the number of fixed points of f thanks to some traces over the homology groups: $\sum_{x \in FIX(f)} i(f, x) = \sum_{k \geq 0} (-1)^k Tr(f_* \downarrow H_k(X, \mathbb{Q}))$ (where $i(f, x)$ is the index of f in x, intuitively, the number of f windings around x). Through Lefschetz's formula, one can reckon how the counting of points in a variety (left side of the equation) is related to the understanding of its homology (right side).) The construction of that *conceptual filtering sieve* (approaching characteristic 0 through "limits" of prime characteristic) is obtained thanks to the general *étale* environment, where many cohomologies blossom (*étale*, crystalline, l-adic, etc.), more akin to the finer topologies needed to approach the discrete obstructions of number theory. An ample *smoothness/softness theory to capture discrete structures*, or, which is the same, a far-reaching *topologization of algebra*, is at the very heart of the program. We can discover here many of the seeds of Grothendieck's works in his Third Period ("topological algebra," see Sect. 4 below).

It is interesting to compare the two extended decades 1958–1970 and 1981–1991: while the first *ascends* to general forms of smoothness (schemes, toposes) which then descend to the particular (Weil's conjectures), the second *grows* from particular expressions of smoothness (*moduli* spaces, Teichmüller tower, *dessins d'enfants*) which then produce general frames of unification (n-categories, stacks, derivators, looking for a deep entangling of homology and homotopy). With some perspective, one can thus observe that *all Grothendieck* is looking for a very deep *back-and-forth* between universal-general-archetypical forms and concrete-particular-typical forms, with many *zigzag* movements in the middle of its creative spectrum: vertical (EGA [**g1959–1964**], SGA [**g1960–1969**]), horizontal (*Longue marche* [**g1981**], *Esquisse* [**g**1984], see Sects. 4.1, 4.3 below), or diagonal (*Stacks* [**g**1983], *Dérivateurs* [**g**1991], see Sects. 4.2, 4.4 below).

3.4 Séminaire de Géométrie Algébrique (SGA) (1960–1969)

The *Séminaire de Géométrie Algébrique* (SGA) [**g1960–1969**] is one of the greatest *collaborative* efforts in the history of mathematical thought. (SGA is divided into 7 volumes (and 12 parts): *SGA1. Revêtements étales et groupe fondamental* – "Un séminaire dirigé par A. Grothendieck, augmenté de deux exposés de Mme. M. Raynaud" (447 pp., 1960–1961). *SGA2. Cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux* – "Alexander Grothendieck (rédigé par un groupe d'auditeurs), augmenté d'un exposé par Mme. Michèle Raynaud" (287 pp., 1962). *SGA3. Schémas en groupes* – "Un séminaire dirigé par M. Demazure et A. Grothendieck, avec la collaboration de M. Artin, J.E. Bertin, P. Gabriel, M. Raynaud, J.P. Serre." *Tome 1: Propriétés générales des schémas en groupes*; *Tome 2: Groupes de type multiplicatif et structure des schémas en groupes généraux*; *Tome 3: Structure des schémas en groupes réductifs* (564 + 654 + 529 pp., 1962–1964). *SGA4. Théorie des topos et cohomologie étale des schémas* – "Un séminaire dirigé par M. Artin, A. Grothendieck, J.L. Verdier, avec la collaboration de N. Bourbaki, P. Deligne, B. Saint-Donat." *Tome 1: Théorie des topos (Exposés I à IV)*. *Tome 2: (Exposés V à VIII)*. *Tome 3: (Exposés IX à XIX)* (525 + 418 + 640 pp., 1963–1964). *SGA5. Cohomologie l-adique et fonctions L* – "Dirigé par A. Grothendieck, avec la collaboration de I. Bucur, C. Houzel, L. Illusie, J.P. Jouanolou, J.P. Serre" (484 pp., 1965–1966). *SGA6. Théorie des intersections et théorème de Riemann-Roch* – "Un séminaire dirigé par P. Berthelot, A. Grothendieck, L. Illusie, avec la collaboration de D. Ferrand, J.P. Jouanolou, O. Jussila, S. Kleiman, M. Raynaud, J.P. Serre" (700 pp., 1966–1967). *SGA7. Groupes de monodromie en géométrie algébrique*. *Tome 1* "Dirigé par A. Grothendieck, avec la collaboration de M. Raynaud, D.S. Rim." *Tome 2* "Par P. Deligne et N. Katz" (523 + 438 pp., 1967–1969). From the point of view of *mathematical practice*, some remarks are in order: the continuous presence of Serre, the ever growing diversity of collaborators, and the funny appearance of Bourbaki (Borgesian autoreference, mention of a collective inside a collective). Just looking at the network of authors, one can sense the driving impulse of Grothendieck and the mystic of the Group. Compared with EGA [**g1959–1964**], SGA constitutes another *monumental* work which sums up nearly 6000 (!) pages.) In his *Esquisse* [**g**1972], Grothendieck underlines that, among his topological ideas, "The most fundamental seems to me the extension of general topology, in the spirit of sheaf theory (developed initially by Jean Leray), contained in the topos point of view. I introduced these toposes from 1958 on, for the necessity of defining an *l*-adic cohomology for algebraic varieties (more generally, for schemes), which would agree with a cohomological interpretation of the celebrated Weil conjectures. (For a corroboration of that early date, see [**g1986** P.24]. On the other hand, in the *Archives Grothendieck* at Montpellier, one doesn't find another temporal register associated with the beginnings of topos theory. For good histories of topos theory, see **Belanger** and **Kromer2**.) In fact, the traditional notion of topological space is not sufficient to treat the case of algebraic varieties over a field different of the complex numbers, since the Zariski topology does not provide reasonable 'discrete' cohomological invariants" [**g1972** 3–4]. (Again, and always in Grothendieck, the search to *understand the*

discrete through new continuous tools will propel many of his most revolutionary points of view. I thank here Alexander Cruz for many conversations around the dialectics (continuous ↔ discrete) in Grothendieck, and, particularly, around Grothendieck's very original inversion (continuous → discrete) expressed along the search for general forms of varieties, and universal topological algebra constructs.)

SGA [**g1960–1969**] provides a multilayered architectonics where many deep perspectives are brought together: a systematic work on foundations (schemes, sites, toposes), (Given a category C, a *Grothendieck topology* J on C is the synthetic analogue of an analytic (set-theoretic) topology: for each object X of the category, a collection $J(X)$ of morphisms over X is given ("coverings") such that the identity is a covering, coverings of coverings are coverings, and pullbacks of coverings are coverings [**g1960–1969** 4.1.219]. A *site* is a pair (C, J), with a category and a Grothendieck topology on it [**g1960–1969** 4.1.221]. A *sheaf* over a site is the synthetic analogue of an analytic sheaf over a topological space (through some appropriate categorical characterization [**g1960–1969** 4.1.223]). A *Grothendieck topos* (or just "topos" in this context) is a category equivalent to the category of *all sheaves over a site* [**g1960–1969** 4.1.302]. Grothendieck toposes were characterized by exactness properties (Giraud [**g1960–1969** 4.1.303]). A simplification of those exactness properties (limits, exponentials, subobject classifier) gives rise to Lawvere's *elementary topos* (1970), axiomatized over a first-order language. Any Grothendieck topos is an elementary topos, but not viceversa (e.g., topos of finite sets). Examples of toposes abound [**g1960–1969** 4.1.311–322]: topos of sets, topos of presheaves, topos of sheaves, topos associated to a topological space, topos associated to a space of operators, topos of actions, classifier topos of a group, initial and final topos, etc. The *étale topos of a scheme* (Grothendieck) is obtained by considering the *étale* morphisms over a given scheme X, which form a Grothendieck topology E (" *étale* topology"), and taking the topos of sheaves over the site $(Sch/X, E)$. The *étale topos* will turn out to be the central key in solving the Weil conjectures (Grothendieck-Deligne). On the other hand, the *arithmetic topos may* be a key to solve the Riemann hypothesis (**Connes**). In this way, toposes become deeply submerged into some of the most formidable mathematical conjectures of the last two centuries. For a good account of all the logical, *model-theoretic force* of Grothendieck toposes (beyond elementary toposes), see **Caramello**.) a new opening to abstract topological algebra (fundamental algebraic group), a construction of diverse cohomological theories (coherent, *étale*, l-adic, De Rham, crystalline... leading to the emergence of motives, see Sect. 3.5 below), a wide program of ramifications around the Weil conjectures (Lefschetz local and global theorems, Picard groups, study of homological and homotopical invariants of an algebraic variety [**g1972** 15]), and, more generally, an understanding of the archetypical connections between number and space. (Beyond Einstein's *space-time*, Grothendieck's *space-number* should offer new orientations for the mathematics of the twenty-first century. Time, just a *linear* dimension of number, should be enlarged to *ramified* magnitudes which may explain many of the intricacies of both arithmetical geometry and the cosmos.)

Grothendieck and the IHES school enter then in a new realm which serves as a "synthesis between two worlds, until then contiguous and tightly solidary, but nevertheless separated: the *'arithmetical' world*, in which reside the so called 'spaces' without principle of continuity, and the *world of continuous magnitudes*, in which reside the 'spaces' in the proper sense of the term (...) *In the new vision, those worlds separated before, form only one*" [**g1986** P.30]. Thus, *beyond separation*, thanks to the abstraction (*freeness, projectivity*) provided by sheaves, a new mathematical *smoothness* governs the interrelations between arithmetic and geometry. Grothendieck's revolutionary shift is as *simple* as deep: instead of *(i)* a topological space X, consider *(ii)* all sheaves over X, and study *(iii)* the topos T_X of all those sheaves; then *the understanding of the structure T_X turns out to be the true understanding of X*. A double differential and integral process (see Sect. 5.3 below) governs the situation: exploring *multiplicity* along all sheaves, and threading *unity* along the exactness properties of the topos.

3.5 Standard Conjectures on Algebraic Cycles (1968)

Around the mid 1960s, (From a semiotical and philosophical perspective, the general method is fascinating: to understand something, consider all possible points of view, and search for a *structural connection* between them. Only then you will be approaching *true knowledge*. A systematic use of the *liberty and multiplicity* of topos theory in our ordinary life would have huge consequences on tolerance, ethics, politics, and social action: an example of how *high mathematical practice* would transform our *everyday practice*.) Grothendieck imagines a "daydreaming" [**g1972** 17] where a *common form* hopes to unify the diverse cohomologies that have emerged until then. A *motif* (unity) is thought as a conceptual basis (or "heart of the heart" [**g1986** P.45]) over which many cohomologies can be represented (multiplicity). Contrary to his previous practices, Grothendieck publishes little on motives, but one of his best renderings is the article *Standard conjectures on algebraic cycles* [**g1968**]: "We state two conjectures on algebraic cycles, (An *algebraic cycle* over a variety X is a formal linear combination $\sum_i(n_i \cdot V_i)$ (coefficients n_i in $\mathbb{Z}$ or in $\mathbb{Q}$) of subvarieties V_i in X. The cycles, as *decomposition* processes of sophisticated mathematical objects, become central in the emergence of *motives*. In the Riemann-Roch strategy, the *equivalence classes* of cycles are obtained through *rational equivalence* (difference between cycles parameterized through rational functions) [**g1957** 38].) which arose from an attempt at understanding the conjectures of Weil on the ζ-functions of algebraic varieties. (...) The first is an existence assertion for algebraic cycles (considerably weaker that the Tate conjectures), and is inspired by and formally analogous to Lefschetz's structure theorem on the cohomology of a smooth projective variety over the complex field. The second is a statement of positivity, generalising Weil's well-known positivity theorem in the theory of abelian varieties. It is formally analogous to the famous Hodge inequalities, and is in fact a consequence of these in characteristic zero" [**g1968** 193].

Grothendieck delineates the consequences of the two conjectures (for their precise statements, see below), in order to obtain an eventual resolution of Weil's conjectures. The first step consists in recalling the rationality (obtained by Artin and Grothendieck in the SGA [**g1960–1969**]) that governs the Zeta function Z of a variety X with good *smoothness* properties (smooth, projective, irreducible) of dimension n over a finite field $\mathbb{F}_q$:

$$Z(t) = \frac{L'(t)}{L(t)},\ L(t) = \frac{\prod_{i=0}^{n} L_{2i}(t)}{\prod_{i=1}^{n} L_{2i-1}(t)},\ L_i(t) = \frac{1}{P_i(t)},\ P_i(t) = t^{dimH^i(\overline{X})} Q_i\left(t^{-1}\right)$$

where Q_i is the characteristic polynomial of the action of the Frobenius endomorphism of X over $H^i\left(\overline{X}\right)$, H^i denotes the i-th group of l-adic cohomology (l prime different of q), and $\overline{X}$ denotes the extension of X to the algebraic closure of $\mathbb{F}_q$ [**g1968** 193–194]. The second step consists in signaling that the first conjecture forces the P_i to have entire coefficients (independently of $l \neq q$), and in observing ("thanks to an idea essentially due to Serre" [**g1968** 194]) that the two conjectures *at the same time* force the eigenvalues of the Frobenius (that is, the reciprocals of the roots of the P_i) to have an absolute value $q^{\frac{i}{2}}$, which proves Weil's conjectures [**g1968** 194]. (Instead, the path followed by Deligne to prove Weil's conjectures is *completely different*. Grothendieck never really accepted that proof, in his view extremely ingenious but *artificial*, estranged from the *natural path* offered by the standard conjectures (and, in depth, by motives). That *mathematical discontent* is one of the many reasons, in a complex thread of disillusions, that alienated him from Deligne and the mathematical community in general.) Concisely and limpidly, in one page and a half, a vast panorama is drawn, still open today (2020) in many senses.

"The conjecture 1 (of Lefschetz type)" [**g1968** 195–197], states two equivalent forms ($A(X)$, $B(X)$) of the conjecture. If $C^i(X)$ denotes the set of i-th cohomology classes of algebraic cycles, Grothendieck indicates that one gets (through hyperplane projections: Lefschetz's operator) an homomorphism (*) $H^i(X) \to H^{2n-i}(X)$ ($i \leq n$) subject to the commutative diagram

$$\begin{array}{ccc} C^j(X) & \longrightarrow & C^{n-j}(X) \\ \downarrow & & \downarrow \\ H^{2j}(X) & \longrightarrow & H^{2n-2j}(X). \end{array}$$

Conjecture $A(X)$ postulates that (*) is always an isomorphism, and that, if $i = 2j$, (*) induces also an isomorphism (through the commutative square) at the level of algebraic cycles $C^j(X) \to C^{n-j}(X)$ [**g1968** 196].

"Conjecture 2 (of Hodge type)" [**g1968** 197–198], proposes to handle the good behavior of a bilinear form on "primitive cycles." In detail, if $\xi \in H^2(X)$ is the class of an hyperplane section, if $P^i(X)$ ($i \leq n$) is the "primitive part" of $H^i(X)$ (that is, the kernel of Lefschetz's morphism induced by ξ^{n-i+1}), and if one defines the set of (classes of) primitive cycles by $C^j_{Pr} = P^{2j} \cap C^j(X)$, one obtains then a bilinear symmetric form $C^j_{Pr} \to \mathbb{Q} : (x, y) \mapsto (-1)^j K\left(x \cdot y \cdot \xi^{n-2j}\right)$, where K is the

isomorphism $H^{2n}(X) \approx \mathbb{Q}_l$. The conjecture $Hdg(X)$ postulates that the form is positive-definite [**g1968** 197]. (If V is a real vector space, a bilinear symmetric form $B : V \times V \to \mathbb{R}$ is *positive-definite* if $B(x, x) \geq 0$ for all x. The positive-definite forms correspond to quadratic forms, basic instruments for arithmetical (Gauss) or metrical (Riemann) mediations, and are associated to central problems in representation and optimization. The emergence of a positive-definite form in algebraic geometry, close to Weil's conjectures, serves thus as a weaving between arithmetic, algebra, topology, and geometry: a *natural* mediating tool, very attractive to Grothendieck.)

Grothendieck underlines that, in characteristic zero, $Hdg(X)$ is obtained immediately from Hodge theory, and explains the great technical interest of having *simultaneously* the two conjectures in arbitrary characteristic: "$B(X)$ and $Hdg(X \times X)$ imply, by certain arguments of Weil and Serre, the following: if f is an endomorphism of X such that $f^*(\xi) = q \cdot \xi$ for some $q \in \mathbb{Q}$ (which is necessarily > 0), then the eigenvalues of $f_{H^i(X)}$ are algebraic integers of absolute value $q^{\frac{i}{2}}$. Thus, this implies all of Weil's conjectures" [**g1968** 197]. On the other hand, Grothendieck points out that the conjunction $Hdg(X)$ and $B(X)$ forces Hodge's Λ operator to arise from an algebraic cycle with rational coefficients, *independently of l* [**g1968** 198]. In this way, plumbing the depths, emerges the vision of some *global universal archetypes* (independent motives), which may then be projected over *local arithmetic strata* (l-cohomologies).

4 Third Period: 1981–1991

Grothendieck's third period escapes from the stringent exigences of architectural perfection (EGA, SGA), and allows the exploration of conjectural and error-prone wanderings around number and space. *The genius is set entirely free.* Inverting Poincaré's algebraic topology program, Grothendieck invents *topological algebra*, introducing new topological invariants for algebraic objects and new forms of *smooth* topology (tame topology). The emergence of *anabelian geometry* studies varieties through the action of their fundamental algebraic group. On the other hand, proceeding from low level geometrical objects, Grothendieck offers combinatorial descriptions for Riemann surfaces ("*dessins d'enfants*") and for their classification, level by level, along actions of the Grothendieck-Teichmüller group. Beyond this, in a powerful *back-and-forth* between concrete examples and universal theories, he introduces stacks and derivators, in order to study homotopy types and their homological counterparts from a categorical perspective. The 1980s become also the time of his very deep philosophical reflections on *mathematical creativity*, and, particularly, on his own work.

4.1 La Longue Marche à Travers la théorie de Galois (1981)

Grothendieck's retirement from the IHES in 1970 is due to a diversity of factors, not reducible to a single perspective: technical exhaustion (after 20 years of extremely

hard work, a point underlined by Cartier and Serre), individual liberation [**g1986** 1.65], social conscience [**g1986** 1.159], academic turmoil after May 1968 [**g1986** 1.156], disappointment with the mathematical community [**g1986** 1.65], rejection of the military support given to the IHES (the "standard" argument [**g1986** L.3, 1.169]). One often hears, with no reason, that Grothendieck abandons the world of mathematics. Nothing could be farther from the truth. He dedicates a great effort to the radical ecologist movement *Survivre et vivre* (1970–1975), but continues steadily to give lessons (*Collège de France* (1970–1972), *Université de Paris-Orsay* (1972–1973), long series of conferences in Romania, Canada, Germany, France, Switzerland, USA), and to foster the mathematical research of his students (*Université de Montpellier* (1973–1984), with even more Ph.D. students there than at the IHES). What seems to be true is that between 1970 and 1980 his own mathematical research is somewhat "muted." Its huge reemergence in the decade 1981–1991 appears then as a kind of "miracle," and one encounters again the inexhaustible Grothendieck at work.

The first expression of Grothendieck's "resurrection" is the unending (and unpublished) manuscript *La longue marche à travers la théorie de Galois* [**g1981**]. Written between January and June 1981, it covers close to 1600 pages in the principal text and close to 1000 pages in notes and addenda. Some sections of the manuscript have been transcribed (Malgoire, Schneps), but they are difficult to read. They will become integrated later in the synthetic *Esquisse d'un programme* [**g1984**] (see Sect. 4.3 below). Open to the careful observation of concrete examples, *La longue marche* [**g1981**] combines holistic perspectives (categorical treatments, towers) and case studies (*moduli* spaces, modular curves, *dessins d'enfants*). The study of *algebraic variation* through topological tools and the unveiling of *universal structures* (fundamental algebraic group, Grothendieck-Teichmüller group), behind Galois, Riemann, and Poincaré, become the foundations of Grothendieck's *topological algebra* in the 1980s.

The two central themes of *La longue marche* [**g1981**] are *(i) anabelian geometry* and *(ii) Grothendieck-Teichmüller theory.* Around *(i)*, the central point is to study a variety (algebra – number theory – Galois), or a scheme (Grothendieck), through its fundamental topological group (topology – geometry – Poincaré), or through its fundamental *algebraic* group (Grothendieck). Given a scheme X (automatically a topological space), one considers the fundamental homotopy group $\pi_1^{top}(X,x)$ (Poincaré) for some point $x \in X$ (if X is arcwise-connected, one can forget the base point x), and the *profinite completion* of $\pi_1^{top}(X)$ is by definition the *fundamental algebraic group* $\pi_1^{alg}(X)$ of X (Grothendieck). This group classifies the *étale*, since its finite continuous orbits are precisely the *étale* coverings of the Scheme. A fundamental example is the projective plane P except three points: in that case, the absolute Galois group $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$ acts over $\pi_1^{alg}(P)$ and hides deep arithmetical combinatorial structures (*dessins d'enfants*). The central problem of *topological algebra* becomes how to detect or describe those varieties X completely characterized by $\pi_1^{alg}(X)$. By definition, those are the *anabelian varieties*, whose geometry (space) and arithmetic (number) are codified in their (non-abelian) fundamental

algebraic groups. Grothendieck's main conjecture states that the *moduli* spaces M_g (see note 25 above) constitute the anabelian varieties over $\mathbb{Q}$, and, as a special case, that the hyperbolic curves over algebraic number fields are anabelian. Mochizuki solves in 1996 the particular conjecture (**Mochizuki1**) and launches then his ambitious program of new foundations for *arithmetic geometry* (**Mochizuki2**).

Around *(ii)*, the central object of study becomes the absolute Galois group $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$ (algebra – number theory – Galois) through its *actions over geometrical objects* (complex variable – geometry – Riemann). In that sense, as its very name shows, the "absolute" Galois group becomes an *archetype* which hopes to be understood through its actions over diverse *types*: curves, *moduli* spaces, fundamental groups (homotopic, algebraic), *dessins d'enfants*, and combinatorial towers. Given a variety $M \in M_{g,\,\nu}$ (g holes, ν selected points) and given its fundamental homotopy group $\pi_1^{top}(M)$, Grothendieck considers its profinite completion $\widehat{\pi_1^{top}(M)}$ and studies the profinite groups of loops, through automorphisms, interior automorphisms, quotients ($C_1(g, \nu)$), discretizations (basis of the group), and characters ($C_2(g, \nu)$). An strategy to understand the absolute Galois group $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$ consists then in presenting it as a quotient $C_1(g, \nu)/C_2(g, \nu)$ [**g1981** 26–28], *uniformly* over all (g, ν) with $(2g + \nu \geq 3)$. (Schneps indicates that this approximation fails at least when $(g, \nu) = (0, 3)$, due to a later result of Drinfeld. See **Schneps3** p. 62.) Nevertheless, Grothendieck himself observes that this seems to be a "very brutal" strategy [**g1981** 163], and that one should "proceed in a more inductive manner, from the presence of $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$ [in the Teichmüller tower] (for arithmetico-geometrical reasons), in order to clear up strong properties, and provide finally a purely algebraic characterization" [**g1981** 165]. (The program will be later developed by Drinfeld (1991), through what he will call the *Grothendieck-Teichmüller group* $\widehat{GT}$. A delicate description of $\widehat{GT}$ can be done through subgroups and products of $\widehat{F_2}$ (completion of the free group with two generators), $\widehat{K_5}$ (completion of Artin's braid group), and $\widehat{\mathbb{Z}}$ (completion of the integers). Drinfeld conjectures that $\widehat{GT}$ is also the *automorphism group of the Teichmüller tower*. An injective homomorphism $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right) \hookrightarrow \widehat{GT}$ holds (Drinfeld, Ihara, Belyi). A proof of the open conjecture $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right) \approx \widehat{GT}$ would provide the "purely algebraic characterization" dreamed by Grothendieck. See **Schneps3** pp. 64–66.) In all of this, we can see how Grothendieck *zigzags*, looking for a deep harmonization between vital polarities: visible/invisible, exterior/interior, discrete/continuous, local/global, layers/completion. A new form of *mathematical practice* emerges along a *new, open, free, tentative, plastic*, style. The conjectural, the inventive strata, the elastic heart, even if imprecise, *command* the theorematical, the solid architectonics, the strict reason. A myriad of questions are then opened for generations to come.

4.2 Pursuing Stacks (1983)

Between 1981 and 1991, Grothendieck will follow concentrated *impulses*, along *continuously flowing* manuscripts, either purely mathematical (*Longue marche*

[**g1981**], *Stacks* [**g1983**], *Esquisse* [**g1984**], *Dérivateurs* [**g1991**]), either methodological, philosophical, and mystical (*Récoltes* [**g1986**], *Songes* [**g**1987]). One can sense in the 1980s his strongly burgeoning spirit, never at rest. *Pursuing stacks* [**g1983**] is the longest Grothendieck manuscript written in English. Proceeding from an *epistolary mood*, that once initiated never stops, it includes three letters to Larry Breen (1975, 58 pp.), a letter to Daniel Quillen (1983, 23 pp.), a table of contents (1983, 10 pp.), and a deep *journal of discovery* (February 27, 1983 – November 4, 1983, 593 pp.) Around higher forms of sheaves and category theory, Grothendieck explores a "*vast synthesis* of homotopical algebra, (co)homological algebra, and topos theory," in a "*non precise and provocatively fragmentary*" manner (*Archives Grothendieck*, Montpellier, cote 109).

Stacks appear in different ways in [**g1983**] – as coefficients in noncommutative homological algebra, as strata in categorical towers, as unifier bridges between homology and homotopy – but Grothendieck never offers a precise definition of the concept. (A *stack* is considered today as a *generalization of a sheaf* in the context of 2-categories (or bicategories). In an iterated fashion, a n *– stack* is understood as a $(n+1)$ *– sheaf*. In some cases, *stacks* may be characterized through categorical *fibrations*. In a more precise approach, a *stack* over a site S is a pseudofunctor (preserves identity and composition with respect to isomorphisms) $S^{op} \rightarrow Cat$ (2-category) which satisfies an effective descent for coverings. A *stack of groupoids* takes values in the category of groupoids. A *stack of moduli* takes into account not only equivalence classes in *moduli* spaces but also the *automorphism groups* of those structures. In this way, the development of *stacks* is closely related to Galois' *ambiguity theory* applied to arithmetic. For a thorough guide of an ever growing panorama, see the many entries in the webpages *nLab* and *Stacks Project*.) This is due in part to the *work in progress maniera*, where the *vanishing lines* of the building become primordial, in detriment of its foundations. (The importance of the *style*, or *maniera*, in *Pursuing stacks* [**g1983**] is essential: "I sometimes find it difficult to find a proper balance in these notes between the need of working on reasonably firm ground, and working out suggestive terminology and notations for what is coming up, and on the other hand my resolve not to get caught again by the EGA style of work, when it was understood that everything had to be worked out in complete detail and in the greatest generality" [**g1983** 445].) Nevertheless, Grothendieck's *vision* may be broken down along some well defined steps: *(i)* acknowledgment of the existence of nontrivial automorphism groups (interest of Galois' ambiguity theory), in order to *structure certain structures* (first essential hierarchical iteration); *(ii)* classification of those structures (e.g., *moduli* spaces) thanks to continuous variations (interest of Riemann's representation theory), in order to *differentiate certain differences* (second essential hierarchical iteration); *(iii)* consciousness of the ubiquitous role of groupoids (algebraic or topologic) in the preceding themes (structuration, representation, differentiation); and *(iv)* integration, in higher sheaves, of all previous informations (variation, gluing, ascent/injection, descent/projection). The *coherence* of the program is solidified when one comes to see it as some sort of *methodological sheaf* applied to the very richness of mathematics.

Passages between diverse types of *stacks* become then entirely *natural*, in a *back-and-forth* between (1) *algebraic stacks* (algebraic varieties, schemes, toposes), (2) *differential and topological stacks* (differential varieties, topological spaces, toposes), and (3) $n-stacks$ ($0-stacks$ = sheaves, $1-stacks = stacks$, etc., leading to n-categories and n-groupoids). All these paths offer new horizons: (1) the fine theory of *moduli* spaces, oriented towards arithmetic (Langlands' program), (2) the classification theories in noncommutative algebra, oriented to anabelian geometry (Grothendieck-Teichmüller's program [**g1981**] [**g1984**]), (A central example of a *stack* is the (pseudo)functor $S^{op} \to M_{g,\nu}$ where S is the category of affine schemes over a fixed scheme with the *étale* topology and $M_{g,\nu}$ is the *moduli* space of Riemann surfaces of genus g with ν points. The particular case $(g, \nu) = (1, 1)$ (corresponding to elliptic curves) was first treated by Mumford (1965, introduction of the term "*stack*" in 1969), before Grothendieck considered the general case (g, ν).) and (3) the homotopical theories of types and categories, oriented towards an alternative foundation of mathematics in higher geometrical orders (Voevodsky's program, HoTT). (*Homotopy Type Theory* (HoTT) looks for a *fourth foundation* of mathematics (after set theory, intuitionistic species, and category theory). A strong geometric intuition (homotopy types) is postulated, *before* considering the combinatorial (sets), the constructive (species), or the relational (categories). An essential part of the program consists in *debilitating equality* (very restrictive for the natural structures of mathematics), and changing it with isomorphisms ("univalence axiom"). For an introduction to HoTT, see **VoevH**.) We observe here a *continuous thought* expressed through categorical unitary archetypes, *high* enough and *general* enough, to be able to cover the multiplicity of diverse particular types. (A comparison of Hilbert with Grothendieck, easily the two major mathematicians of the twentieth century, may be instructive: while, in Hilbert, each fragment of his production constitutes a well *separated and complete* conglomerate (algebra, number theory, geometry, functional analysis, logic, physics), in Grothendieck all fragments are always intertwined, from the implicit categorical *maniera* of his *Thesis* [**g1949–1953**] until the overarching "functional analysis" panorama of the *Dérivateurs* [**g1991**].)

4.3 Esquisse d'un Programme (1984)

Esquisse d'un programme [**g1984**] constitutes a research project for Grothendieck's admission to the CNRS, after his retirement from Montpellier University (1984). The *Esquisse* (dated January 1984 [**g1984** 51]) retrieves some of the central ideas of the *Longue marche* [**g1981**], and extends them to two new paths in *topological algebra*: on one side, *dessins d'enfants*, (A *dessin d'enfants* may be presented as a graph D finite, connected, bicolored, embedded in an oriented compact topological surface X, such that $X \backslash D$ is a finite union of disks (see **GirondoGonzalez**, p. 207). Some examples (posterior to Grothendieck) are striking: the (topological) sphere S^2, with an inscribed *dessin* formed by a black vertex, a white one, and another black, corresponds to the Riemann surface associated to z^2; the sphere with a *dessin* formed

by vertices black-white-black-white-black corresponds to the Riemann surface associated to $z^2(z-1)^2$, etc. On the other hand, proto-ideas around the *dessins* can be traced back to the nineteenth century (Hamilton, Klein, etc.)) through which Grothendieck combinatorially captures compact Riemann surfaces and studies their behavior under the actions of the absolute Galois group $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$, and, on the other side, "tame topology" (*topologie modérée*), which hopes to eliminate wild phenomena and artificial counterexamples in classical set-theoretic topology. An "ascent" to higher concepts is provided, in order to "descend" after to lower constructs. In this case, a systematic study of "high" algebraic actions (fundamental algebraic group, anabelian geometry, Grothendieck-Teichmüller group) is explored, in order to understand topological surfaces of "lower" dimension (compact Riemann surfaces, *moduli* spaces).

The *Esquisse d'un programme* [**g1984**] covers three main subjects: *(i)* Teichmüller's tower and the absolute Galois group (*Sect. 2* of the typescript), *(ii) dessins d'enfants* (*Sect. 3*), and *(iii)* tame topology (*Sects. 5, 6*). Two secondary topics are considered: regular polyhedra and two-dimensional geometry (*Sects. 4, 8*), *stacks* (*Sect. 7*). Finally, some considerations about the personal circumstances of the author are included: balance of research and teaching experience (*Sects. 1, 9, 10*). From *higher* perspectives, Grothendieck looks for a general theory of *dévissage* ("unscrewing"), in order to understand the architectonics of stratified mathematical architectures. This leads naturally to the iterative construction of *stacks* and to tame topology, hoping to capture *smooth combinatorial forms* along hierarchical transits. From *lower* perspectives, surfaces, complex algebraic curves over number fields, elliptic curves, modular subgroups, and regular polyhedra, offer many illuminating examples which help to adjust the architectonics. In the *middle ground*, modular multiplicities and their associated groupoids, general varieties over number fields and their associated fundamental groups, Riemann surfaces and their associated *dessins d'enfants*, stratified actions of the absolute Galois group and their invariants, *connect continuous and algebraic variations*, and become ciphered in the fundamental anabelian conjecture about the multiplicities $M_{g,\nu}$ (*stacks*, not varieties). The *Esquisse* [**g1984**] offers thus a typical Grothendieckian *vision*, conjugating abstraction and concretion, theory and examples, unity and multiplicity, universality and contingency, continuity and discreteness, globalization and localization. A *yin* drive (see Sects. 4.5, 4.6 below) towards particular forms of smoothness is complemented with a *yang* drive towards general forms of *dévissage*.

The *arithmetical heart* of the researches collected in the *Esquisse* [**g1984**] glues together (in a sort of methodological sheaf) the *continuous variations* present in the *moduli* spaces $M_{g,\nu}$ and the *algebraic variations* of $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$ acting on the Teichmüller tower. In particular, the appearance of the space of elliptic curves $M_{1,1}$ on the first floor of the tower, and its associated modular group $SL(2, \mathbb{Z})$, underlines the arithmetical richness of the environment. On the other side, the *natural* emergence of algebraic curves over number fields along compact Riemann surfaces (Belyi 1979), strikes Grothendieck very deeply. (Belyi's theorem characterizes the compact Riemann surfaces associated to complex algebraic curves

defined over number fields. More precisely, if S is a compact Riemann surface, the following are equivalent: *(i)* S is defined over $\overline{\mathbb{Q}}$, *(ii)* there exists a meromorphic function $f: S \to \mathbb{P}^1$ with at most three ramification points. Belyi's discovery, "that technically is reduced to such a small thing, exerted in me a very strong impression, and represented a decisive turn in my reflections, a shift of the center of my mathematical interests (...) I do not recall that another mathematical fact had striken me so strongly (...)" [**g1984** 12]. Thus, in Grothendieck's own words, the *decisive turn* towards algebraic geometry (Kansas 1955) and the *decisive turn* towards combinatorial topological algebra (Mormoiron 1984) seem to be two major illuminations. It is interesting to observe that the two circumstances occur far away from the official academic centers of mathematical thought, and after some important psychological crisis. *Mathematical practice*, through immense manuscripts (*Tôhoku* [**g1955–1956**], *Longue marche* [**g1981**]), well beyond just mathematical learning, will always be for Grothendieck a way to stabilize his soul.) Beyond this, thanks to Grothendieck's equivalence between compact Riemann surfaces and oriented topological surfaces with *dessins d'enfants* inscribed on them, the known action of $Gal\left(\overline{\mathbb{Q}} : \mathbb{Q}\right)$ over compact Riemann surfaces becomes transferred to a *new combinatorial action* over *dessins d'enfants*, and one can hope to obtain a *smooth combinatorial* description of the absolute Galois group. (This will lead to Drinfeld's *definition* of the Grothendieck-Teichmüller group $\widehat{GT}$ (see *Note 45* above).) Finally, the main goal of anabelian geometry, that is, to control curves over number fields through their fundamental algebraic groups (extending homotopy to the new configurations of topological algebra), achieves the architectonics of the *Esquisse* [**g1984**]. In this way, once more, Galois, Riemann, and Poincaré, become connected along the depths of Grothendieck's thought, but now under an arithmetic-algebraic-complex-geometric weaving to which are added an elementary combinatorics (*dessins d'enfants*) and a smooth universal understanding of topological forms (*topologie modérée*).

4.4 Les dérivateurs (1991)

Les dérivateurs [**g1991**] is Grothendieck's last known manuscript. (Between 1991 and 2014, the year of his death, Grothendieck wrote extensively on many subjects, embracing mathematics, physics, philosophy, mysticism, history, sociology, including a long history of evil in mankind which tracks in particular the monstruosities of concentration camps. Summing up to more than 50,000 pages (20 years is a very long time for a restless mind!), Grothendieck bequeathed his last writings (extremely well organized in careful arranged boxes, as reported by Georges Maltsiniotis) to the *Bibliothèque Nationale de France*. Nevertheless, his heirs have acted differently, and his heritage is being appraised for an eventual sale to a highest bidder (a hideous contradiction with Grothendieck's anarchist and free spirit). When those 50,000 pages will be made available to the public, our understanding of Grothendieck will acquire a new dimension, complementing his extremely complex intellectual portrait

through a *Fourth Period* (1991–2014), which should be full of surprises.) The gigantic text (around 2000 pages) was distributed to friends and colleagues, like the other texts of the decade, (In total, the output was immense, summing up circa 8000 pages (!): *mathematical*: *Longue marche* [**g1981**] ∼ 2500 pages; *Pursuing stacks* [**g1983**] ∼ 700 pages; *Dérivateurs* [**g1991**] ∼ 2000 pages; *philosophical*: *Récoltes et semailles* [**g1986**] ∼ 1400 pages; *La clef des songes* [**g1987**] ∼ 1000 pages. The *flow* of Grothendieck's writings, in solitary confinement, is almost equal in this *Third Period* to the production in the previous *First and Second Periods combined*, even if we count there the help of his multiple IHES collaborators!) and opens new roads for the eclosion of higher category theory. Grothendieck tries to axiomatize *at large*, at a universal global level in the category *Cat* of all categories, some canonical notions in topological algebra. Through a general axiomatics applied to adjunctions, localizations, and actions (oriented by the category *Hot*: *CW*-complexes with homotopy classes as morphisms [**g1983**]), Grothendieck seeks to recover *both homological algebra* (coefficients, abelian and derived categories) and *homotopical algebra* (limits, *n*-categories, topos constructs), as complementary *slopes* of a unitary whole ("prederivators," sort of localizations of Yoneda's Lemma, independently of basis changes). Beyond isolated efforts (Maltsiniotis, Cisinski, Groth), many of the ideas and techniques of *Les dérivateurs* remain yet to be assimilated and developed by the mathematical community.

Les dérivateurs [**g1991**] provides many topological techniques and perspectives to capture algebraic notions, through *smooth abstract coverings*: *(i) Hot* covered by *Cat*, *(ii)* types of homological algebra (coefficients) and types of homotopical algebra (limits) covered by general categorical archetypes, *(iii)* homotopical formalisms and (co)homological adjunctions covered by an axiomatics of (pre)derivators, (Fundamental examples of (pre)derivators are: *(i)* presheaves $A \to Hom(A^{op}, M)$ (A, M categories), *(ii)* localizations $A \to W_A^{-1}Hom(A^{op}, M)$ (W_A subset of arrows – weak equivalences" – in Yoneda's image of $A \subseteq M^{A^{op}}$), and *(iii)* a natural functor HOT: $Cat \to Cat$ associated to a localization of simplices in the category *Hot*. For precise definitions and developments, see **Cisinski**.) *(iv)* variations (functorialization of derivation processes) and invariants (intrinsic stability, independence of basis changes) covered by universal procedures in the category of all categories, and *(v)* resolutions to obstructions to non-canonicity in abelian and derived categories covered by *localizations*. (Grothendieck's problematic begins with observing some *canonicity failures* in the construction of the derived category of an abelian category [**g1955–1956**]. If A is an abelian category, $C(A)$ is the category of complexes over A, and W is the set of quasi-isomorphisms of $C(A)$ (morphisms $f : X \to Y$ such that $H_n(X) \approx H_n(Y)$ for all n), the derived category is defined by the (formal) localization $Der(A) = C(A)[W^{-1}]$. Then, many functoriality and adjunction problems emerge, in order to capture the original information of A. For an introduction, see **Maltsiniotis3**.)

In a letter to Thomason (Les Aumettes, 2.4.1991), (Available at www.grothendieckcircle.org, under a folder on derivators.) Grothendieck explains some of the essential points of *Les dérivateurs* [**g1991**]: universality of abstract axiomatizations in *Cat*, localizations *at large* in *Cat* and comparison with Quillen's model

categories, understanding of the additional structure of a derivator with respect to its underlying derived category, correlative study of adjunctions associated to a derivator. The *axiomatic* investigations (See [**g1991** I.23–25] (derivator axioms), *Archives Grothendieck*, cote 157–4, pp. 268–350 (localization axioms).) in the ever-evolving document (Again, as happens all along the 1980s, a *diary of invention* is proposed, and one can follow the many ideas, errata, changes, new elaborations, through the evolving manuscript. A detailed study of such *mathematical practice* in a great mathematician, with all its *residues of stratified invention*, is yet to be undertaken. It would provide a magnificent witness of *creativity at large*.) are witnesses of Grothendieck's practice to clear up ("*dégager*") the universal mechanisms that govern many regions of mathematics. As a consequence, Grothendieck obtains a sort of *general methodology of smoothening singularities* (strong conceptual connections with Hironaka's "resolution of singularities," 1964), where the *tameness* of the topology, the *n*-elevation of *stacks*, the unification at large of *derivators*, respond to the desired smoothness. The possibility of *transiting* in higher categories helps to overcome the many *obstructions* in lower regions. Grothendieck's work acquires thus a strong solidarity, going from his very *Thesis* [**g1949–1953**] until the *Dérivateurs* [**g1991**]. For developments on this *unity, continuity, and coherence*, see Sect. 5 below.

4.5 Récoltes et Semailles (1983–1986)

Récoltes et semailles [**g1986**] is perhaps the major analysis ever written on higher mathematical creativity. Beyond Poincaré's *L'invention mathématique* (1908) and Hadamard's *The Psychology of Invention in the Mathematical Field* (1945), (**Poincare Hadamard**.) Grothendieck's reflections surpass its predecessors both in extent (1500 pages), and in depth (detailed study of the emergence of Grothendieck's main mathematical ideas, a refined approach to one of the two greatest mathematicians of the twentieth century) (see *Note 50* above). The subtitle of the work, "Reflections and testimonies on a mathematical past," indicates its two main watersheds: on one side, reflections on *mathematical practice*, entering in detail inside Grothendieck's practices, and on another side, testimonies on his own life and his *mathematical school*, expressing a disenchantment with how the "seeds" (*semailles*) of his work had not been adequately "harvested" (*récoltes*) in the period 1970–1983. The *style* of the retrospective is typically grothendieckian: passionate, meticulous, exhaustive, irritating, astonishingly beautiful, both in his autoanalysis and in his accountability of the community. The harsh sincerity of the volume struck his mathematical colleagues, and ended up severing the few remaining ties that Grothendieck had with his previous intellectual environment.

Along this survey, we have had the opportunity to quote many times the manuscript. (The manuscript, not yet published in French, is accessible in many sites online. It has been translated in Japanese (Tsuji, also published), Russian (Fridman and Filkenberg), English (Lisker), and Spanish (Navarro). The best study of *Récoltes et semailles* [**g1986**] is **Herreman**.) The text is divided into seven main sections:

Walk Around a Work, or the Boy and the Mother (65 pp.): January–February 1986; *Letter and Introduction* (56 + xxii pp.): March 1984, May–June 1985, February 1986; *First part. Fatuity and Renovation* (171 pp.): June 1983, February–March 1984; *Second part. The Burial I, or the Clothes of China's Emperor* (249 pp.): April–June 1984; *Third Part. The Burial II, or the Key of Yin and Yang* (354 pp.): September 1984–January 1985; *Fourth Part. The Burial III, or the Four Operations* (478 pp.): February–June 1985; *The Doors over the Universe. Appendix to the Key of Yin and Yang* (127 pp.): March–April 1986. Acceleration and multiplication point to a *never-ending outpouring* of words, concepts, images, reminiscences. *Récoltes et semailles* [**g1986**] encompasses the inexhaustible richness of Grothendieck's persona, underappreciated and ill-understood by the mathematical community, in particular by his great peers, a Serre, a Dieudonné, who never accepted the hermit, the mystic, and the anarchist.

The *complexity* of *Récoltes et semailles* [**g1986**] captures the complexity of the Grothendieckian world: (1) disaggregation of mathematical thought, in its eternal fight between substance (e.g., spaces, numbers, structures) and form (e.g., cohomologies, motives, derivators); (2) reflections on the inroads to mathematics, and, in particular, on a naive entrance to creativity (freshness, smoothness, childish perception); (3) considerations on mathematical method, and, in particular, on the demands of constant and persevering work (tasks of the architect and the laborer); (4) systematic study of the yin-yang watersheds in creativity, in general, and in mathematical creativity, in particular; (5) journey around the intellectual biography of the author, his central passions, and his tasks in pursuit of alternative life movements (ecology, austerity, oriental philosophy); (6) detailed analysis of the historical and sociological circumstances where that biography is inserted; (7) strong criticism of the degenerescence processes occurring in the Western world in the 1970s and 1980s; (8) stylistic constructions, multi-spatial and multi-temporal, which mirror an inquisitive and open mind, attentive to contemporary trends.

The metaphor of the *rising sea*, "by immersion, absorption, dissolution" [**g1986** 3.553], acquires a fundamental technical value thanks to Category Theory, when an object ("nut") is submerged in an adequate category which absorbs it ("sea"). This is the case of the central theorem in the theory, Yoneda's Lemma, (In a parallel way to Yoneda (c. 1958), the Lemma also appears independently in EGA [**g1959–1964** III.1.7].) where a *double iteration* of the rising sea appears: *(i)* objects are first immersed in their "auras" (representable functors) and *(ii)* categories of objects are then immersed in presheaf categories. The result provides, for every small category, a *complete* extension through non-representable functors: these can be understood as *phantasmata* which come closer to the invisible. On the other hand, the rising sea occurs also naturally along toposes, when a topological space (or a site) is submerged in the category of *all* sheaves over it. In the topos emerges a powerful "superstructure" (completeness, exponentiation, subobject classifier), invisible from the initial space (or site). In these cases, the *rising sea* reveals a *general* mathematical richness, previously invisible, that can help to solve very difficult *particular* problems (e.g., Weil's conjectures).

The *Appendix* to *Récoltes et semailles* [**g1986**], *The Doors over the Universe* (abbreviated DU), 127 pages written in a month-long euphoria (March 17–April 14, 1986), adopts a sophisticated system of pairings yin/yang to understand human action, in general, and *mathematical practice*, in particular. On the yin side, one finds discovery processes, smoothness, a mother language of intuition. On the yang side, emerge inventive processes, architectonics, the superficial languages of description. Along yang resides reason ("razón" in Spanish), along yin lies the heart ("corazón" in Spanish, *duality* perfectly captured by the prefix "co"). Grothendieck proceeds to construct a *stratified combinatorics* of yin/yang pairs, having in a first level "polygamies" (a yang related to two yins, e.g., sun with moon and earth) and "polyanders" (a yin related to two yangs, e.g., earth with sun and sky) [**g1986** DU.7–8]. The first levels become then iterated along higher growing cycles. One of the major interests of (DU) is the use of Grothendieck's mathematical arsenal (sheaves, schemes, homologies, geometries, combinatorics) to provide an analysis of thought. In this way, *mathematical practice* orients our *common material practice*. In fact, many of the tasks assumed in (DU) refer to gluings, inversions, strata, breaks, thresholds: *(i)* variations, degrees, intensities, between warm (*yin*) and cold (*yang*) [**g1986** DU.16–17]; *(ii)* inversions, through subgroups, between continent and content, abstraction (*yang*) and concretion (*yin*) [**g1986** DU.19–20]; *(iii)* dialectics multiplicity/unity ("I feel myself as a multiple in search for unity") [**g1986** DU.23]; *(iv)* diagrams (hexagons, icosahedrons, trees) to capture tonalities yin/yang [**g1986** DU.28–32]; *(v)* dynamics between ideal (*yang*) and real (*yin*) [**g1986** DU.36–37]; *(vi)* zigzag iterations and homologies between unity/mystery (*yin*) and order/simplicity (*yang*) [**g1986** DU.37–40]; *(vii)* tension between discovery of structures (*yin*) and invention of languages (*yang*) [**g1986** DU.47–51]; and *(viii)* "accordion" between exterior (surface, light, *yang*) and interior (deepness, shadow, *yin*) [**g1986** DU.51–55], etc. In this way, there appear many implicit inversions and symmetries (Galois connections), and multiplications and unfoldings (Riemann ramifications). As an "Ahab after the White Whale" [**g1986** DU.66], Grothendieck immerses himself in the never-ending complexities of the world.

4.6 La Clef Des Songes (1987–1988)

La clef des songes, ou dialogue avec le Bon Dieu [**g1987**] studies the spiritual wanderings of the author, in an ever meticulous, extremely detailed manner, the Grothendieckian Way. After a precise register of his dreams in previous years (more than a "thousand" [**g1987** 6], Grothendieck's numbers always amaze), the author proceeds to a *dévissage* of the oniric realm, along extremely diverse strata and thresholds. The emergence of multiple types of spirituality (breath, faith, heart), intellectuality (eyes, language, reason), and sensibility (hands, passion, eros) offers a fascinating immersion in Grothendieck's mind. This leads him to the "revelation" of an archetype – a Dreamer, a Hand, an Eye, forms of the *Bon Dieu* – which is *projected* into his own dreams, writings, and visions. *Iterating* this self-reflexive mood, there emerges, naturally, a key (*clef*) to understand both Grothendieck's Self

and the World. (As in *The Doors over the Universe* (see above), it is enthralling to see how Grothendieck uses, beyond mathematics, the very central methodologies of his mathematical thought: *dévissage*, multiplication, projection, iteration, etc.) The central text of *La clef des songes* [**g1987**] (315 pages) is complemented with many notes (691 pages), which contain a long final development (circa 500 pages), *Les mutants*, where Grothendieck proposes a sort of genealogy of his spiritual Masters, and their influence on his own libertarian thought.

The manuscript is divided into six parts and many notes, summarized in what follows. *I. All the dreams are creation of the Dreamer* [**g1987** 1–46]: the dream and the knowledge of the self, discovery of the Dreamer, childhood, dreams proceed from the Dreamer, instants of truth, the key of the big dream, the voice of reason and the other voice, acts of knowledge and acts of faith, the spark and the flame, work and conception, the rhythm of creativity, cycles of Eros, emotion and thought. *II. God is the Dreamer* [**g1987** 47–79]: God is the Dreamer, lost knowledge, the incredible Good News, brothers in starvation, meeting the Dreamer, prohibited questions, reunion with God, only one Dreamer exists, the Creator and the Web, God is not defined nor proven, a new multiplication Table. *III. Voyage to Memphis (1): wandering* [**g1987** 80–136]: my parents, God's splendor, Rudi, a fall of wonders, God by sound reason, lost encounters, the calling, the turn, faith and mission, death, God speaks in low voice. *IV. Aspects of a mission (1): a freedom song* [**g1987** 137–182]: unthinkable convergence, testimony as a call for the discovery of one's self, Eros or power, Eye or sense, vision as testimony, soul and labor, man as a creator, creativity and repression, free creativity and interior work. *V. Aspects of a mission (2): spiritual knowledge* [**g1987** 183–268]: inclusive knowledge, beauty and contemplation, pain and shadows, the soul of things and the man without soul, the mentality of the flock, the use of Times, creativity and interior voice, creativity and listening, the tree of good and evil, archaic spirituality, law, ignorance, truth and creativity, resistance and suffering. *VI. Voyage to Memphis (2): seeds for a mission* [**g1987** 269–315]: acting, separation, beginning without end, hope, Tempest, the new man, surface and deepness, the call of silence, the messenger, crossing of the desert and revelation, seeds and harvest, tasks and gestation. *Notes* (notes numbered 1–57) [**g1987** N1-N176]: Eros, the One and the Infinite, wisdom of the body and action of God, the role of dreams, archetypes and God's manifestations, dream and free will, mystic experience and self-knowledge, miracles and reason, truth and knowledge, mathematics and imponderables, belief and faith, the boy and the mystic, my friend the "*Bon Dieu*," mission and creativity, Hell and Providence, eros and spirit, the great innovators, the creative boy, "unthinkable May 68," creativity and maturity, error and ignorance, doubts and security, work and desire. *The mutants* (notes numbered 58–145) [**g1987** N177-N691]: register and study of spiritual guides and freedom thinkers, Fujii Guruji, Gandhi, Walt Whitman, the mutants (1: spirit, justice, freedom), the place of Riemann, A. S. Neill, Edward Carpenter, Félix Carrasquer, the mutants (2: power, diversity, sex), Solvic (sic = Slovik), Steiner, Krishnamurti, the mutants (3: self-knowledge, reconciliation, crisis of civilization, hope), Darwin, Freud.

Through this list, one can see how Grothendieck is mainly concerned with *creativity*, and how the different layers of creativity *expand* from his own

experiences and dreams, towards a superior archetype (God), passing through many intermediate Masters (*mutants*: the visionary leaders of a "Great Mutation" needed for a redemption of mankind, a chronological list of Grothendieck's 18 *mutants* appears in [**g1987** N480-N482]). The *multiplicative richness* of Grothendieck's ideas covers history, phenomenology, metaphysics, psychoanalysis, sociology, religion, and mystic thought. As if it were a *universal sheaf* in action, Grothendieck looks for a common blowing, a unitary life, a universal harmonics ("global sections"), beyond the most particular concretions, blades of grass, daily dreams, sensible perceptions ("local sections"). *Content and form* are combined in a uniquely *smooth* state of grace in *La clef des songes* [**g1987**]: on one hand (content), the search for a common stem in the Tree of Life, found with passion in Whitman and Darwin, and reflected along Grothendieck's main mathematical unitary archetypes (nuclear spaces [**g1949–1953**], projective norms [**g1953c**], *K*-theory group [**g1957**], schemes [**g1959–1964**], toposes [**g1960–1969**], motives [**g1967**], anabelian varieties [**g1981**], derivators [**g1991**]); on the other hand (form), the development of a *free language*, disentangled from bonds and prejudices (types), and plastically open to "listen to things" (archetypes) which transcend language itself.

5 A Synthetic Account of Grothendieck's Work

Grothendieck's mathematical work shows a strong equilibrium between (1) high conceptualizations and (2) concrete examples. In the first trend, the use of Category Theory is ubiquitous; we will call *categorical reason* this tendency to the global and the general. In the second trend, the uses of complex variables, algebraic structures and topological forms pervade the work; we will call *mathematical heart* this support on the local and the particular. Grothendieck's work can then be understood as an *adjunction* between reason and heart, in Spanish, between *razón* and *corazón*. Since *corazón* = *co-razón*, we can sense a full *duality* ("co") between intelligibility and sensibility, something clearly stated in *Récoltes et semailles* and *La clef des songes*. The duality extends to a deep understanding of the polarities *étalé and étale* (all Grothendieck may be ciphered in an accent!), *relativity and universality*, *types and archetypes*.

5.1 Mathematical Heart and Categorical Reason

The *back-and-forth* between the particular and the universal, the concrete and the abstract, the typical and the archetypical, the Many and the One, the local and the global, the examples and the theories – paradigm of a *pendular mathematical practice* – pervades Grothendieck's work. We will call *mathematical heart* the environment of his local technical strengths (functional analysis, topology, complex variables, sheaves, co/homology, algebraic geometry, etc.), and we will call *categorical reason* the environment of his global conceptual visions (schemes, topos, motives, *stacks*, *n*-categories, derivators, etc.) Following a pendular movement

between "mathematical heart" and "categorical reason," a full dialectics ("co") is established between *co-razón* (= corazón = heart) and *razón* (= reason).

In Grothendieck's *First Period 1949–1957*, the mathematical heart appears in *topology* (A: topological vector spaces, B: sheaves, C: non-separation problems), *functional analysis* (D: Hilbert and Banach spaces, E: nuclear spaces, F: approximation problems), *complex variables* (G: Riemann-Roch, H: spaces of holomorphic functions, I: Riemann sphere), *abstract algebra* (J: tensor products, K: stability properties, L: co/homology). The twelve forces (A-L) support Grothendieck's originary intuitions and secure the *concrete richness* of his mathematics. A *calibration* of forces along the main productions of the decade ([**g1949–1953**], [**g1953c**], [**g1955–1956**], [**g1957**]) shows that themes (C) and (K) are ubiquitous, and that themes (B) and (J) reveal some of the favorite *universal* constructions of the young mathematician.

Along Grothendieck's *Second Period 1958–1970*, three new forces occur in the *mediation* between mathematical heart and categorical reason (X: schemes, Y: topos, Z: topos *étale* of a scheme). Emerging from the Edinburgh lecture [**g1958**], they propagate through the main IHES writings ([**g1961**], [**g1959–1964**], [**g1960–1969**]), converging finally in the standard conjectures [**g1968**]. A calibration of themes (A-L) and (X-Z) inside those texts shows how Grothendieck in this "communitary" period passes to work at high levels of abstraction ("categorical reason," X-Z), to the detriment of the initial concrete seeds ("mathematical heart," A-L) that fostered first his imagination.

Finally, Grothendieck's *Third Period 1981–1991* shows a strong return to classical themes at the heart of mathematics (α: *moduli* spaces of Riemann surfaces, β: absolute Galois group, δ: fundamental groups / groupoids, γ: homotopy types). A fascinating back-and-forth is constructed between the very concrete in arithmetic (anabelian varieties [**g1981**]) and the very abstract in homology and homotopy (derivators [**g1991**]), going through unexpected combinatorial mediations (*stacks* and *n*-categories [**g1983**], *dessins d'enfants* and tame topology [**g1984**]). In these years, a calibration of the themes (A-L), (X-Z), and (α-γ) shows how "corazón" and "razón" become tightly entangled, producing an extremely powerful integrative *co-reason* (= "co-razón").

The *continuity* of Grothendieck's thought becomes evident when one ponders his reflections on "gluing" *space-number*. In a *First Period*, *varieties of space* (topological vector spaces, spaces of holomorphic functions, Riemann surfaces) *resonate* with *varieties of number* (tensor products, rings, co/homologies), before becoming fully integrated at *two unification levels*, first as singular sheaves, then as plural sheaf categories. In this second level (*Second Period*), emerge (1) sheaf categories with *abelian fibers* (fibers are abelian groups, the case of *abelian categories*), (2) sheaf categories with *local fibers* (fibers are local rings, the case of *scheme categories*), (3) sheaf categories with *set-theoretic fibers* (fibers are sets, the case of *toposes*), and (4) sheaf categories with *arithmetic fibers* (fibers are finite and separable residuated fields, the case of *étale toposes*). In this last environment, the fine techniques of entangling space and number provide the adequate "rising sea" ("categorical reason") to solve Weil's conjectures ("mathematical heart"). On the other hand, those algebraic

topology or algebraic geometry forces become *inverted* in the *Third Period*, where the new constructs of topological algebra (fundamental algebraic group, anabelian varieties, Grothendieck-Teichmüller group, etc.) complement Grothendieck's deep explorations of the extremely complex and stratified *space-number* realm.

5.2 Étalé and étale

An *extreme synthesis* allows to condense a fair amount of Grothendieck's work in the *behavior of an accent*! In fact, the wide difference between the terms *étalé and étale* (very often confused, even between Grothendieck scholars) captures two of the main watersheds of Grothendieck's thought. On one side, (1) *étalé* summons the "*espace étalé*" of Cartan's school [**g1955–1956**], that is, a *sheaf* as an unfolded space. On the other side, (2) *étale* calls for the "*topologie étale*" of Grothendieck's school, that is, the smooth, folded category-theoretic topology used in the resolution of Weil's conjectures [**g1960–1969**]. The opposition is extreme: (1) the *étalé* alludes to multiplicity, ramification, unfolding of a given sheaf, while (2) *étale* alludes to unity, smoothness, folding in a category of schemes. The *pendular movement between the Many and the One* is double: in (1), over the particular (one sheaf), one studies the variation of topological sections, while in (2), over the general (multiple schemes in a category), one observes the stability of arithmetical fibers – (1) reveals the transit from the One to the Many (*étalé*) while (2) detects, inversely, the passage from the Many to the One (*étale*).

The *étalé* strategy (1) governs Grothendieck's world, traversed in all directions by sheaves. The main problem, understanding the *mixture space-number*, is explored through universal sheaf constructions (*K*-theory, topos), which explain the *types* of space and number as projections of sheaf-like *archetypes*. On the other hand, the *étale* vision (2) enters in harmonic conjugacy with the *flat or rising sea* (Victor Hugo's " *mer étale*"), (Grothendieck acknowledges that "in my work I am as yin, as sea and movement, as one can be" [**g1986** 3.595]. The " *étale* sea" appears in Victor Hugo's *Les travailleurs de la mer* (1865–1866): " *Le niveau uniforme du varech sur toutes les roches marquait la ligne de flottaison de la marée pleine et de la mer étale*," **Hugo1** p. 678.) where general categorical settings (abelian categories, topos, derivators) literally flood our understanding. The architectonic tensions between (1) *étalé* and (2) *étale* help to situate Grothendieck's major inventions. In the universal, abstract environment of the (1) *étalé*, live the main *strata and hierarchizations*, the staggered flats of the building: tensor norms in Banach spaces [**g1953c**], iterated sheaves [**g1959–1964**], Teichmüller's tower [**g1961**] [**g1981**], homotopy types and *n*-categories [**g1983**], etc. On the other side, in the universal, abstract environment of the (2) *étale*, live the main *coverings*, the high perspectives from the building's top: nuclear spaces [**g1949–1953**], abelian categories [**g1955–1956**], *K*-theory [**g1957**], toposes [**g1960–1969**], motives [**g1967**] [**g1968**], tame topology [**g1984**], derivators [**g1991**], etc.

Two strong varieties of *mathematical practice* are then combined in Grothendieck's work: (1) unfolding, unscrewing ("*dévisser*"), and towering (close

to a "mathematical heart"), complemented with (2) folding, screwing, and covering (closer to "categorical reason"). The *back-and-forth* methodology between (1) and (2) allows to explore a multitude of *apparently* wild and unlike situations, and to find *natural universal invariants* beyond the variable multitudes of mathematics and the World. A new *(1) differential* and *(2) integral calculus* is thus proposed at a very high conceptualization level.

5.3 Relativity and Universality

The detailed *practice* of "relative mathematics" is a permanent feature of Grothendieck's "doing." (We use the term "doing" ("hacer" in Spanish) in the sense of Javier de Lorenzo, who, since his *Introducción al estilo matemático* (1971), insisted in the importance of constructing a mathematical philosophy attentive to the advances of the discipline, and carefully oriented towards *mathematical practice*. See **Lorenzo1 Lorenzo2** De Lorenzo anticipated thus by 40 years (!) the present-day movement (fashion?) of the " *philosophy of mathematical practice.*" For a complete vision of De Lorenzo's work until the year 2000, see **Zal10**.) In fact, the *localization* and *variation* of the objects occur early in comma categories [**g1955–1956**], in local rings along coherent sheaves and in local projections of the *K*-theory group over (co)homology [**g1957**], or in local schemes over a given basis [**g1959–1964**]. These methods become then extended to any form of construction which may be deemed "Grothendieckian." *Relativity* constitutes one of the sine qua non forces of his thought. But on the other hand, its natural counterpart – to look for invariants beyond variation, universals beyond relativization, globals beyond localization – is what really defines his enterprise. His *systematic finding of archetypes beyond types* (see Sect. 5.4 below) distinguishes him from many previous great mathematicians.

The finding is not accidental and is due to the *systematic use of sheaves and categories* in Grothendieck's thought. Both an *epoch* (when sheaves and categories were emerging) and an exceptional *character* (vision, courage, solitude, discipline) were required to provide a deep new notion, a true revolution both in mathematics and philosophy: the ingenuous glimpse and the strenuous practice of the concept of a *relative universal*.

The notion immediately seems a contradiction in terms (how can something universal be relative and vice versa?), but it acquires a sense as simple as profound when one conjugates the dialectics *abstract-concrete* in Category Theory. The essential obstruction to overcome is to be able to imagine the idea of *the Universal without the Absolute*, something that can be obtained when universality is linked to a *practice of dynamical abstraction*. On one hand, *concrete categories* (classes of structured sets, e.g., groups, rings, ordered sets, topological spaces, differential varieties, etc.) provide very diverse distinctions, and capture the *relativity of mathematical practice*. But on the other hand, *abstract categories* (classes of morphisms with associative composition and identities) offer integral paths, and capture the *universality* of mathematical thought, thanks to simple and general definitions that can be provided through the extremely powerful quantifier "exists unique" (∃!). (For

example, an "initial object" 0 (defined by the property $\forall X \ \exists\ !\ 0 \to X$) is a relative universal in abstract categories, which incarnates in many ways in concrete categories: in the empty set 0/ in the category of sets, in the trivial group with one element $\{e\}$ in the category of groups, in the ring of integers $\mathbb{Z}$ in the category of (commutative unitary) rings, in a minimum *m* in the category associated to a partially ordered set, etc. On the other hand, 0 *does not* incarnate in the category of infinite sets (due to the infinitude, and thus non-unicity, of morphisms from ω, the only plausible candidate for an initial object). The *unifying structure* (A) of the abstract universal is thus *distributed* along the *variegated multiplicity* (B) of its very diverse concrete relativizations. We are here again in presence of a high conceptual *integral (A) and differential (B) calculus*. In a certain sense, our "blind eyes" (Tarkovsky) become then disposed to see: *beyond types emerge archetypes, beyond appearances we can grasp a "universal breath.").*

The *relative universals* are the *clef de voûte* of Grothendieck's cathedral. Thanks to them, a *cohesion* of multiple elements is obtained, and the cathedral's arches are well adjusted, without collapsing. The *topos* notion [**g1960–1969**] constitutes perhaps the richer relative universal of Grothendieck's opus. Its construction mixes many levels along the adjunctions between the pairs relativization/universalization and localization/globalization: *(i)* in an arbitrary category *C* with pullbacks, a Grothendieck topology *J* uses localization and variation techniques, giving way to a site (C, J); *(ii)* a sheaf on a site combines localizations over the base and a search for global sections in the unfolded (*étalé*) space; *(iii)* the collection of all sheaves on the site (Grothendieck topos) universalizes in a new unitary structure all possible types of relativizations. The result offers a generalization of space (topological space → site → topos), whose *relative universal* (topos) is extended well beyond the initial source (topological space). In the *ascending* injection to the universal, and, further, in the *descending* projection to the relative, many technical tools are invented (e.g., *étale* cohomology) which help to understand the very concrete (Weil's conjectures) thanks to the very abstract (topos of a scheme). In particular, in this example, the *conjunction of the étalé and étale visions* can be seen as a strong partial weaving between the universal (topological sheaf – *étalé* world) and the relative (arithmetical scheme – *étale* world).

5.4 Types and Archetypes

The *Tôhoku* [**g1955–1956**] introduced sheaves and categories into the landscape of "real" mathematics. (Along with **Corfield** we understand by "real" mathematics, the corpus of the hard-working mathematicians along their *mathematical practice*: combinatorics, number theory, abstract algebra, algebraic geometry, topology, complex variables, functional analysis, differential geometry, etc., well beyond just sets and logic (the reduced, impoverished, realm of analytical philosophy). Sheaves were invented in 1942 by Leray, and categories were invented in 1945 by Mac Lane and Eilenberg. Nevertheless, their actual use in "real" mathematics, as signaled by the very Mac Lane (see **Kromer2** p. 158), is due to Grothendieck.) Grothendieck's

contributions to Category Theory in this paper are impressive on their own (introduction of subobjects, adjoint functors, equivalences, representable functors, additive and abelian categories, generators, infinitary axioms), but the central points of the *Tôhoku* deal with two kinds of *archetypes related to mathematical practice*: (1) to prove that in an abelian category with a generator and a suitable infinitary axiom, any object ("type") can be embedded into an injective object ("archetype") and (2) to prove that the cohomology of a space, with coefficients in a sheaf ("types"), can be reconstructed as a projection of a formalism of derived functors ("archetypes"). In that way, "projective" (co)homology, one of the main mathematical investigations in the period 1895–1950, can be regarded as a seed of an even more fundamental theme: a theory of abstract derivators in suitable categories [**g1991**].

Grothendieck's *Tôhoku* strategy is very interesting in terms of the duality *abstraction-concretion*. First, categories abstract universal properties holding in the diversity of mathematical regions, through a series of *axioms*: associativity and identities (all categories), existence of an abelian group structure in the *Hom*-sets (additive categories, e.g., category of holomorphic fibered spaces over a Riemann surface), existence of kernels, cokernels, and adequate factorizations through them (abelian categories, e.g., category of sheaves with fibers abelian groups), etc. In this approach, the *axiomatic abstract definitions* have their expected *concrete models*. Second, in the abstract setting, that is in general categories which satisfy the axioms, not yet "reified" in concrete models, "archetypes" occur in a *natural way* (definitions and properties related to the quantifier "exists only" $\exists!$), and only afterwards they are projected into existential objects along concrete categories. Thus, abstract categories are "doubly archetyped" (*inside* categories, free objects appear; *outside* categories, free objects incarnate in apparently very different disguises), while concrete categories become just partial modelizations of general categorical transits. Third, in the abstract settings, many general theorems occur thanks to a *free machinery* which has escaped from the particularity of the concrete models.

Grothendieck's *La longue marche à travers la théorie de Galois* [**g1981**] sets a new research program, after his works of the First (1949–1957) and Second (1958–1970) Periods. His attention is now turned to "low-level" complexity objects, that is, surfaces in general and Riemann surfaces in particular, modular curves, complex towers, and combinatorial approximations to the absolute Galois group. As we have seen (Sect. 4.1 above), one of his main conjectures states that the anabelian varieties over $\mathbb{Q}$ are essentially the isomorphism classes of Riemann surfaces. In this way, an *archetypal* property (anabelianity) is expected to capture some of the main *types* of modern mathematical thought, the Riemann surfaces. On the other hand, *Pursuing stacks* [**g1983**] sets the stage for a profound connection between homotopy and homology, and a *hierarchy* of concepts, objects and techniques is explored through the many levels of homological and homotopical groups. In that sense, an n-stack can be understood as an n-truncated homotopy type, which in turn can be imagined as a $(n + 1)$-sheaf. The *iteration of types* gives rise to ∞-categories, where all information on n-morphisms (morphisms, functors, natural transformations, etc.) is piled-up. ∞-categories act then as *very large archetypes* which cover huge varieties

of models. Finally, *Les dérivateurs* [**g1991**] proposes several *axiomatic* treatments of the constructions first imagined in *Pursuing stacks*. Axioms are imposed on classes *W* of morphisms in *Cat*, to try to model some properties of weak equivalences in the category *Hot* of *CW*-complexes with homotopies between them. The program consists in trying to *elevate types in Hot to archetypes in Cat*. Axioms *free* the lower level constructions (say, in *Hot*), in order to get *smoothness* properties at higher levels (say, in *Cat*). In an *abstract, free, archetypal* environment, derivators capture some of the main *type constructions* of homotopy (coming from toposes and homotopy groups) and homology (coming from abelian categories and homology groups).

The *archetype/type dialectics* between abstract categorical constructions and structured, concrete models, considerably enriches the panorama of *mathematical practice*. The abstract and the concrete, the universal and the particular, categorical reason and mathematical heart, become all necessary to *nurture mathematical imagination*. Mathematics has to explore all possible worlds, beyond mere actuality, but, on the other hand, many of her greatest achievements occur when a general architecture illuminates an actual conjecture. The *back-and-forth between types and archetypes*, elevated to a method in Grothendieck's work, helps us to understand the irreducible complexity of mathematical thought.

References

Grothendieck, A. Produits tensoriels topologiques et espaces nucléaires. Providence: American Mathematical Society, 1955 [Doctoral Thesis, Nancy, 1953], 1949–1953

———"Résumé des résultats essentiels dans la théorie des produits tensoriels topologiques et des espaces nucléaires". In: Annales de l'Institut Fourier 4 (1952a), pp. 73–112

———"Produits tensoriels topologiques et espaces nucléaires". In: Séminaire Bourbaki 2 (1952b), pp. 193–200

———"Sur certains espaces de fonctions holomorphes I". In: Journal für die reine und angewandte Mathematik 192 (1953) (1953a), pp. 35–64

———"Sur certains espaces de fonctions holomorphes II". In: Journal für die reine und angewandte Mathematik 192 (1953) (1953b), pp. 77–95

———"Résumé de la théorie métrique des produits tensoriels topologiques". In: Bol Soc. Mat. São Paulo 8 (1956) (1953c), 1–79 [Manuscript, São Paulo, 1953]

———Topological Vector Spaces. New York: Gordon & Breach, 1973 [Seminar, São Paulo, 1953], 1953d

———"Sur quelques points d'algèbre homologique". In: Tôhoku Math. J. (1957) 9 (1955–1956), pp. 119–221

———"Classes de faisceaux et théorème de Riemann-Roch". In: SGA. Vol. 6. Paris: IHES, 1955–1957, pp. 20–77

———"Sur la classification des fibrés homolorphes sur la sphére de Riemann". In: American Journal of Mathematics 79 (1957) (1955a), pp. 121–138

———"The cohomology theory of abstract algebraic varieties". In: Proceedings International Congress of Mathematicians (Edinburgh). Cambridge: Cambridge University Press, 1958, pp. 103–118

——— "Techniques de construction en géométrie analytique". In: Séminaire Henri Cartan 13 (1961). Exposés 7-17, Paris: Secrétariat Mathématique ENS

——— "Motifs". 1967

——— "Standard conjectures on algebraic cycles". In: Algebraic Geometry, Oxford University Press (1969), 1968, pp. 193–199
——— Esquisse thématique des principaux travaux mathématiques de A. Grothendieck. Tech. rep. CNRS, 1972
——— "La longue marche à travers la théorie de Galois". 1981
——— "Récoltes et semailles". 1983–1986
——— "Pursuing Stacks". 1983
——— "Esquisse d'un programme". 1984
——— "La clef des songes ou dialogue avec le Bon Dieu". 1987
——— "Les dérivateurs". 1991
Grothendieck, A. and J. Dieudonné. Éléments de Géométrie Algébrique. IV vols., 7 parts. Paris: IHES, 1960–66 [manuscripts written between 1959 and 1964], 1959–1964
Grothendieck, A. and J.-P. Serre. Grothendieck-Serre Correspondence. Providence / Paris: American Mathematical Society, Société Mathématique de France, 2004 [letters between 1955 and 1987], 1955–1987.
Grothendieck, A. and students. Séminaire de Géométrie Algébrique. 7 vols., 12 parts. Paris: IHES, 1960–1969

Further Reading

Bélanger M (2010) Grothendieck et les topos: rupture et continuité dans les modes d'analyse du concept d'espace topologique. PhD thesis. Université de Montréal
Bringuier G (2015) Alexandre Grothendieck. Itinéraire d'un mathématicien hors normes. Toulouse: Privat
Caramello O (2017) Theories, Sites, Toposes. Oxford: Oxford University Press
Cisinski D-C (2003) Images directes cohomologiques dans les catégories de modéles. In: Annales mathématiques Blaise Pascal 10.2, pp 195–244
Connes A, Consani C (2014) The Arithmetic Site
Coréld D (2003) Towards a Philosophy of Real Mathematics. Cambridge: Cambridge University Press
De Lorenzo J (1971) Introducción al estilo matemático. Madrid: Tecnos
——— (1977) La matemática y el problema de su historia. Madrid: Tecnos
Diestel J, J. H. Fourie, and J. Swart. (2000) The Metric Theory of Tensor Products. Grothendieck's Résumé Revisited. Providence: American Mathematical Society, 2000
Dieudonné J (1974) Cours de géométrie algébrique. Paris: PUF
Douroux P (2016) Alexandre Grothendieck. Sur les traces du dernier génie des mathématiques. Mayenne: Allary Éditions
Girondo E, González-Diez G (2012) Introduction to Compact Riemann Surfaces and Dessins d'Enfants. Cambridge: Cambridge University Press
Hadamard J (1945) The Psychology of Invention in the Mathematical Field. Princeton: Princeton University Press
Herreman A (1999) Découvrir et transmettre. Une analyse de la dimension collective des mathématiques dans Récoltes et semailles. http://decouvrir-transmettre.alainherreman.fr, Pré-publications IHES
Hugo V (1969) Les travailleurs de la mer (1865-66). Vol. Oeuvres Complètes, XII. Paris: Club Français du Livre
Imagoshi Y Taniguchi M, An Introduction to Teichmüller Spaces. Tokyo: Springer-Verlag
Karoubi M (2014) L'inuence d'Alexandre Grothendieck en K-théorie. In: Alexandre Grothendieck. A Mathematical Portrait. Ed. by L. Schneps. Boston: International Press, 2014, pp. 13–23
Krömer R (2007) Tool and Object. A History and Philosophy of Category Theory. Basel: Birkhäuser

Maltsiniotis G Introduction à la théorie des dérivateurs, d'après Grothendieck. Preprint. url: https://webusers.imj-prg.fr/~georges.maltsiniotis/groth/Derivateurs.html
Merleau-Ponty M (1964) Le visible et l'invisible. Paris: Gallimard (reed. Folio 2004)
——— (1964) L'Oeil et l'Esprit. Paris: Gallimard (reed. Folio 2004)
Mochizuki S (1996) The Profinite Grothendieck Conjecture for Closed Hyperbolic Curves over Number Fields". In: J. Math. Sci. Univ. Tokyo 3:571–627
——— (2011–2012) Inter-universal Teichmüller Theory I-IV. Tech. rep. Kyoto University
Pisier G (2012) Grothendieck's theorem, past and present". In: Bull. Amer. Math Soc 49:237–323
Poincaré H (1908) L'invention mathématique. Paris: Institut Général Psychologique (cit. on p. 26)
Pradeau Y (2016) Algèbre. Éléments de la vie d'Alexandre Grothendieck. Paris: Éditions Allia, 2016
Scharlau W (2011) Who is Alexander Grothendieck? Anarchy, Mathematics, Spirituality, Solitude. A Biography. Part 1: Anarchy. Norderstedt: Herstellung und Verlag
Schneps L Grothendieck – Mathematics. Preprint. url: https://webusers.imj-prg.fr/~leila.schneps/grothendieckcircle/Mathematics/chap4.pdf
——— (1997) Grothendieck's "Long March through Galois Theory". In: Geometric Galois Actions. Around Grothendieck's Esquisse d'un Programme. London: Cambridge University Press, pp 59–66
Serre J-P (1955) Faisceaux algébriques cohérents". In: The Annals of Mathematics 61.2, pp 197–278
——— (1956) Géométrie algébrique et géométrie analytique. In: Ann. Inst. Fourier 6.1, pp 1–42
Voevodsky V et al. (2013) Homotopy Type Theory. Univalent Foundations of Mathematics. Princeton: Institute for Advanced Study
Zalamea F (2007) Javier de Lorenzo: por una filosofía dinámica de la praxis matemática". In: Mathesis 2.1, pp 1–35
——— (2019) Grothendieck. Una guía a la obra matemática y filosófica. Bogotá: Universidad Nacional de Colombia

An Uncertain Travel

Elijah Liflyand

Contents

1 Introduction 1298
2 What Is This About? 1299
2.1 Mathematics Versus Music 1300
2.2 Mathematics Versus Fiction 1301
3 Individual Work and Joint Work 1302
4 Applications: pro et contra 1304
5 Concluding Remarks 1305
References 1306

Abstract

In these notes, *my* personal reflections on practice and working philosophy of mathematics are discussed. Some of the reviewed issues definitely have a philosophical sense, or at least flavor. The generality of some of them is suited to similar observations of recognized specialists open-handedly quoted.

Keywords

Social side of mathematical activity · Music · Language · Fiction · Individual work · Joint work · Insignificant variables · Function of bounded variation · Hilbert transform · Stirling numbers · Applications · Seminar

E. Liflyand (✉)
Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_141

1 Introduction

The article by Etnyre (2019) is concluded with the statement

> Good introductions can elevate the community's appreciation to your work and quite possibly you, so make sure to write great ones.

Not throwing even a shadow on this instructive article, one may (half)jokingly ask whether that specific text needs a great introduction. Pronouncing this, we come to one of the borders between mathematical and philosophical paradoxes. In the sense that such a border may appear anywhere, making one hesitate what side of the border is more attractive that very moment.

I am not going to deeply discuss purely philosophical paradigms and their reflections in mathematical ways of thinking in full, so to say, philosophical depth. Instead, I shall try to recall the cases and aspects of my practice as an acting researcher in pure mathematics which, in my opinion, were of general nature common to all humanity, and which I dare, justified or not completely, call philosophical. Recall Littlewood's (1957)

> I asked Prof. Wittgenstein was this not a profound philosophical joke, and he said it was.

On the other hand, it is probably a moment for the first reference to Persson (2021), where not only the one who is interviewed sounds, but the interviewer as well; in particular, the latter said:

> Popper failed really to consider mathematics seriously, probably because like most modern philosophers, and here I very much include Wittgenstein, he did not know much about mathematics and had certainly done no work in mathematics, which is a prerequisite for understanding mathematics.

Arguing neither against nor for this statement itself, I cherish hopes for having the mentioned prerequisite, which may allow me to catch certain appropriateness in a less familiar area. However, I would like to begin with taking a roundabout approach to the subject. There was a famous Russian poet (of Georgian-Armenian origin) Bulat Okudzhava. Some of his rhymes converted to very nice and popular songs. However, he also was a successful and very interesting Prosaist. A part of his novels went back to certain events of Russian nineteenth-century history. Once being interviewed in a written form about how a writer like him works on historical issues, he preferred to present a metaphorical song instead of just writing down more or less standard words. Here is this song in my equi-rhythmical translation (possible to be sang – contrary to a non-rhythmical word for word translation as in Okudzhava [1982]), entitled

I am writing an historical novel
In a bottle, darkly brown,
once full of imported beer,
of a rose the scarlet crown

bloomed with soft and proud cheer.
I was writing of the past,
slowly letting it unravel;
from the first page to the last
one was my uncertain travel.
Refrain:
One's bequeathing as it's easing.
It is easing as one's breathing.
As one's breathing one's bequeathing,
never bothering to please.
This for nature is a withness.
What for — is not our business,
why — is not for us to quiz.
Distances were colored blue,
pure fiction was in plenty,
and from my own fate I drew,
any thread if it was handy.
Helped my heroes' ways to weave,
on the history inquired,
and an officer retired
was sometimes my make-believe.
Refrain:
Fiction is not a deceit;
nor a finish is the matter.
Let me keep on writing it
to the very final letter.
And before the rose-leaves die,
which inside the bottle grow,
let me cry the words that lie
stashed away since long ago.
Refrain:

I am aware of the translation by Evgenia Sarkisyants, some of her phrasing inevitably coincide with mine, but mainly I tried to have the refrain first of all and some other pieces as I see it, either for better (closer to the source) meaning, for better rhythm, or just to my taste. And, of course, keeping in mind the present topic and how this illustration fits it. Back to our subject, mathematical work can be described similarly, in a sense. In fact, once I tried to parody Okudzhava's rhymes as **I am writing a mathematical paper,** but let us skip too much poetry and return to mathematics and philosophy.

2 What Is This About?

However, it is not that easy to switch to purely mathematical impressions because of the social side of mathematical activity. The point is that it is impossible to conceal oneself in the ebony tower, and very often a mathematician faces the need to explain to plain people (of course, here the people who lack certain mathematical education are meant as plain) what such a person is doing (and even gets money for). Unfortunately, an immense majority identify mathematics with numbers and arithmetic calculations (something like accountant) or figures and measurements (say,

land surveyor). No wonder that many people identify mathematicians and computer programmers despite that this is by no means the case and interesting exceptions are just interesting exceptions. On the other hand, getting acquainted, many people immediately start telling how good or bad they are (were) in mathematics.

2.1 Mathematics Versus Music

Some time ago, my former tennis partner, a musician, firmly told me that music is something creative, contrary to mathematics, which is a dry routine work. Laying aside the fact that he had no idea what mathematics is (back to the previous paragraph), that meaningless discussion had brought my attention, and not for the first time, to possible relations between mathematics and music. I am sure that far more advanced minds than mine have been thought on these interrelations. However, my feeling is that only mathematicians have been troubled by such reflections. Well, many mathematicians are fond of music; some of them play music on a professional level. For example, Per Enflo and Louis Rowen used to give concerts, playing piano and violoncello, respectively. Closer to our topic, it seems that research in pure mathematics can justifiably be compared with composing music. Of course, the means are different as well as the ways of perception; however, some similarities are striking. Does a theorem (a theorem is an intentional simplification; the question may concern a new approach, a new notion, or a new theory) exist "somewhere" unless formulated, proved, and presented (published) by a mathematician? By the way, these steps may sometimes be fulfilled (or a more musical word – performed) by different people or groups. In fact, involved mathematicians evoke from "somewhere" a "melody," fixing it by means of "notes" – notations and words. The observation that the number of notes and other musical signs is essentially smaller than the number of notions and notations should not be a serious objection. After all, there is a diversity of musical instruments, performers, etc. Or, which is simpler and maybe more correct, music can be compared with a specific area of mathematics rather than with mathematics as a whole. Combinatorics, for instance. Sound extracted should not be considered as a determining factor. Indeed, there is a notion of inner hearing of music. Or, there are "professional" music lovers who come to a concert and listen to a music with the score in hands. Anticipating possible objections, let us mention that both manifestations can be emotional.

Looking at the matter the other way round, recall that mathematics is frequently identified with language. To my taste, this is almost a perfect definition (at least, to some extent). Maybe it is more precise to say that being a language for various phenomena, mathematics is imprescriptible from the language itself in which the results, proofs, and conclusions are given. But music is not less frequently called a language of soul, feelings, or the like. A different feature of music is a need of its performance, with masterly performance as a characteristic of this phenomenon. However, mathematicians are very often on the stage, even not mentioning their teaching activity. As V.K. Beloshapka wrote in his reminiscences about Vitushkin (posted on https://arzamas.academy/mag/1051-math?fbclid=IwAR04xUpED2FSKfnEOkCkyz7ZZPFV9PWhFgtpK3P8SkzwvvhR13PRuzo8rrQ):

> Life of a mathematician takes its shape as life of seminars. It is something like concerts, where mathematicians gather and tells something to each other.

And there are coryphaei, masters, and apprentices in this activity. We shall return to this issue in the sequel.

2.2 Mathematics Versus Fiction

One more activity which, at a glance, seems having nothing in common with mathematics is fiction. In fact, this is sometimes exactly the same process, or at least the published result of the corresponding process is, concerning maybe crime fiction or detective novels more precisely. Let us analyze this assertion. There is something intriguing and hidden, it is searched all the time, with certain clues found, real or false, and the proof of the "main result" as an apotheosis. One may object that the audience is very different, especially in amount. That's obviously correct, and the reason is similarly obvious: special terms and language in general as well as much or less symbolics. However, special language and symbolics, important and sometimes decisive, can be found in modern fiction. Here are a couple of examples immediately coming to my mind. In the famous "Jurassic Park" by Michael Crichton, the action could not move further without reading and understanding a computer print. In fact, not only the hero must understand the printing but the reader too. In the not less famous "The Flanders Panel" by Arturo Prez-Reverte, the reader must understand chess, including chess notation. Clearly, most of the readers do not understand these issues, just skip them and, to be honest, get the main pleasure in killing the time. The main objection might be that such readers just enjoy pleasant and talented reading. However, along with talented authors, there is an immense of a middle or low level. They more or less successfully publish their writings and find their gratifying readers. But the same is among mathematicians: there are very talented authors (in the sense of writing; not all the authors of great results are such) and many more not great in writing. Therefore, the main difference is the specially prepared audience, but the same is with the mentioned listeners of music with the score or spectators of the so-called "elite" movies.

To be precise, what I have compared with fiction concerned a plot, and research papers were mostly understood. But mathematicians do write books, and quite many authors write pretty many books. The process is discussed, and various recommendations on this subject are given in Silverman (2019). A wish to make one's book a delicious and popular reading, with many demanded copies, is apparently the same for any author, regardless of the topic. To mention a recent successful example, Iosevich (2007) is a very impressive one. However, a personal experience and conclusions should be mentioned in order to be complete and honest. Already publishing three books and having one more in process can be considered as a material for possible generalizations. First of all, a strong wish to present a readable text claims for certain demands, self-demands. There must be a plot, a scenario, and a thread going through the whole text. There must be something beyond a collection of facts and results. My

first book Iosevich and Liflyand (2014) was an attempt to overview the results where the Fourier transform can be characterized by its asymptotic behavior. The second one, Liflyand (2019), appeared when the separated facts on the behavior of the Fourier transform of a function of bounded variation seemed to be united in a comprehensive theory. In other words, when the puzzle was almost completed. The third book, Liflyand (2021), is more a textbook than a research monograph. The idea was to extend the first chapter of Liflyand (2019), given as a necessary brief toolkit to the further material, to a sort of a textbook on harmonic analysis on the real line in full value. In each of these cases, the goal was to present the author's view of a certain topic in a readable way and more or less full generality. In short, each time this is a comprehension of the topic as a whole, with slight abuse of terminology a philosophical comprehension. On the other hand, moving within the classification of mathematicians as theory builders and problem solvers, the one who writes a monograph (not a textbook) switches from the latter to the former, in a sense.

3 Individual Work and Joint Work

The problem of the general and the individual is considered as one of the fundamental problems of philosophy and psychology. I believe that in mathematics, it expresses itself in research. Except a very clear individual manner of thinking on a mathematical problem, there are different models of collaboration. Luckily, the publications authored by 20 or 40 names are really marginal in mathematics and mainly are not of research nature if appear. Two and three authors are the most typical collaborative groups in mathematics; by the way, the Research in Pairs program in Oberwolfach assume from two to four collaborators. I do have an experience with three other coauthors, but it is unique and probably not enough for conclusions. So, let us discuss the events where two or three people struggle with a problem or write a survey text.

The most natural and the most fruitful collaboration seems to be the one of two colleagues. Very often such collaborators are the friends possessing experience in discussing the problems of mutual interest or even in joint work on other problems. On the other hand, very often an accidental collaboration becomes a basis for a long-standing friendship. In each of the cases, this new or renewed collaboration starts with mutual interest in a certain problem and, what is maybe more beneficial, with understanding that they have an idea or ideas and tools for a successful attack on the problem. In fact, in most cases there is no much difference if three people participate. In many situation, two of them could be characterized as a pair described above, while the reasons appeared for the third one to join them are the same reasons as those above for two. To be honest, sometimes such an activity of two or three (or more) is accidental, valid for one occasion only, and does not result in further collaboration. It sometimes turns out that the participants or some of them are incompatible, sometimes they just fail to find more problems and tools of really mutual interest. In all of these situations, successful or not, the issues of psychology or social philosophy turn out to be more important than those professional.

Last but not least, what collaboration may protect better than individual work is the novelty. The possibility of working on something already discovered might be very bothering. Similarly, there exists an option on elaborating something formally new but trivial in essence. This is well illustrated in the underestimated book Lem (2013) (it was translated into English only in 2013, decades after its appearance):

> Not everything constitutes scientific truth: an ocean of insignificant variables is larger than an ocean of stupidity – and this is already saying something.

Let me describe a case in my practice, where some concrete mathematics is in order. There is a well-known theorem of Hardy and Littlewood. It asserts that if both a periodic function and its conjugate are of bounded variation, then the Fourier series of each converges absolutely. One day a desire appeared to extend this statement to the non-periodic case. However, the start of this project had been postponed for a distant period. This time was spent in asking far more educated colleagues and looking through the literature for a possible existence of such an extension. There was a feeling of how it can be that nobody had tried it. Finally, the work was done with one of my colleagues and published in Liflyand and Stadtmüller (2013). And only then, in a while I was terrified finding a non-periodic extension in a recognized paper, Hille and Tamarkin (1935), which I looked through many times. Luckily, that story had a happy end: in that paper a quite restrictive additional condition was assumed not needed in our work and in general. The point was that we benefited from using the modified Hilbert transform not known to Hille and Tamarkin (1935). Hence, the only negative feature of Liflyand and Stadtmüller (2013) turned out to be not referring to Hille and Tamarkin (1935); the result however remained safely new. In a different occasion, coming to a conference, I planned to devote free time to a particular problem in harmonic analysis. However, already during the flight, I started thinking on a problem of combinatorial nature and eventually spent all the week in cracking some relevant obstacles. Flying back, I already had a well-composed theory, with results and proof. Luckily, I believed in the novelty of my speculations for not a single moment. Indeed, in a while one of my more experienced and educated colleagues enlightened me in the regard that what I "discovered" were the Stirling numbers (of the second kind). One can read about them in every book on combinatorics. However, I was delighted in the coincidence of my calculations with the corresponding piece of the general theory. Unfortunately, similar troubles used to happen from time to time, and nobody knows when such an event may occur in the future.

The fears of false novelty or trivial novelty are already mentioned. There is one more fear, the one of possible mistakes. The whole third chapter of Littlewood (1957) is devoted to a collection of various misunderstandings and misprints, but mostly of logical or linguistic nature. However, in §13 of this chapter a misprint of wrong sign resulted in the use and application of a wrong statement. This is closer to what is meant as a fear: the use of a false argument, omission of an argument, and the like.

Back to the problem of the general and the individual, probably the most common form of constant generality is seminars. It has already been mentioned in

Beloshapka's saying above. In fact, the complete citation is continued with a funny statement.

> Mathematics is probably transmitted through airborne droplets. Some possess immunity from it since their school-days, while few others are ailing – a sort of a limited epidemics.

There are two roles on seminars that everybody plays, frequently or rarely: the speaker and the listener. I have also experienced a third role of the long-standing organizer of a seminar. It seems that having a continuously working and constantly attended seminar is completely beneficial. Sharing one's own results with colleagues and learning from them new facts and approaches – what can be more fruitful for a mathematician in addition to the main issue of searching new results and solving challenging problems. As of possible mistakes, a good and erudite audience sometimes saves the speaker from certain (not all, alas) mistakes in time, prior to the publication. Of course, there are enough people who avoid constantly playing the mentioned roles. Well, social philosophy will apparently cluster them as a group enjoying equal rights with those mentioned before.

4 Applications: pro et contra

Einstein's observation (Bloch 1995, p. 243)

> Inasmuch as the mathematical theorems are related to reality, they are not sure; inasmuch as they are sure, they are not related to reality.

adds fire to this everlasting contradiction. One moment in my personal experience seems to be very instructive. It concerns the work which finally resulted in Liflyand and Trebels (1998). Here more concrete mathematics comes into play. Considered are radial functions in the n-dimensional Euclidean spaces, that is, the functions depending on the Euclidean norm of the variables. As is well known, their Fourier transform is also radial and can be expressed in a more sophisticated form than the standard Fourier transform via the Bessel functions. The problem is under what conditions on functions (or for what classes of radial functions) the multidimensional Fourier transform can be expressed asymptotically in such a way that the leading term is a purely one-dimensional Fourier transform of a function related to the given one. (By the way, a priori such results have numerous obvious applications in the way that many important one-dimensional results can "for free" be extended to the radial case.) One day my collaborator claimed for examples concerning a certain assertion. I would be happy to construct some but failed. My statement instead was that the Fourier transform is a very important and recognized object that any new result on it sounds by itself and does not need immediate illustration. After a few days of this conflict, none of us was able to deliver an example, and my collaborator had to agree with me (well, the circumstances forced him to do so). This particular situation has, in my opinion, a philosophical side. Not only in the sense of

interrelations of theory and practice. There is one more feature of new results which may justify them regardless of their immediate applicability. What is meant is the aesthetical aspect, of course, a specific mathematical aesthetics is understood. It cannot substitute for the proof but can be a landmark on the road to a possible proof or just to an interest or lack of that in the issue in question. Along with the place of the considered problem, the aesthetical side may give an a priori belief that certain applications will appear in due time.

This is well expressed in (Stewart 1995):

> Mathematics is a harmonious whole: except that the harmony is incomplete, because there are always gaps in our knowledge and vaguely understood hints of new interrelations. In this sense, an application of any mathematics from this central body is an application of the whole. If you insist that mathematics justify its existence by providing applications, then an application of one part will justify the whole. We do not cut off a violinist's feet just because he doesn't use them in playing the violin: in the same way we ought not to dismiss group theory just because it won't pay the rent.

(A remark aside: Does pay as far as I know!)

There is an additional point in this controversial issue: whether a person has a right to do whatever one is interested in, not thinking much on possible applications, or in a wider sense, on possible relations with practice and other areas? Of course, I mean a professional side of this arbitrariness, not a discussion in Dostoyevsky's style. I believe that these problems are related to such social issues as the existence of middle class contrary to the elite group, its weight, and role.

5 Concluding Remarks

Of course, many of the above analogs and comparisons, like those with music and fiction, are relative and serve more for dialogs with strangers. All these similarities and differences are disputable, with numerous pro et contra in each of them. Last but not least, one may ask where philosophy is here. The answer is: everywhere. Even pondering over these problems makes a serious difference and becomes a decisive step toward philosophical reflection. This as well as possible getting down to philosophy in mathematics can be well illustrated by Hersh's explanation in Persson (2021):

> I had gotten hooked on philosophy of math when I volunteered to teach a course that was listed in my department's catalogue as Foundations of Mathematics. No one had ever offered it, before or since. I expected to just do my usual thing when teaching a subject I know nothing about – pick the best textbook I can find, and stay a chapter ahead of the students. Not this time! All the textbooks I found simply presented three viewpoints – logicist, intuitionist, and formalist, and left it plain that all three were inadequate, unsatisfactory, failures. End of course!! As a teacher I found that situation deeply unacceptable. After all, I ought to at least know what was my own personal philosophy of math. But I found that I simply didn't know. So I had to find out where I stood, what was my understanding of the nature of the subject to which I devoted my life. On my part, The Mathematical Experience

> was a stage in my struggle to figure out my own answers. Then also, my career as a mathematician had given me a special kind of experience, which had not been much exploited in a literary way. I was very lucky to find Phil Davis as a collaborator. We never dreamed that the book would make such a splash. It was far short of our original intentions, but we were desperate and submitted what we had. It seemed only a rag-bag at the time, but nevertheless, it worked, after all!

Hoping that the presented notes will also work to a certain extent and mentioning that the topic is inexhaustible, both theoretically and on the level of individual experience, what can be recommended for further reading and thinking (in addition to one's main research work) is to look through the publications that attracted my attention in general and while working on these notes, Bochner (1966), Gromov (2018), Manin (2007), Shapiro (2005), and Stewart (1995), to mention some. And, of course, numerous Hersh's books and articles. I would also like to mention Wigderson (2019): it is not less philosophical than "technological." For sure, these by no means will be boring.

References

Bloch A (1995) Murphy's law complete. Mandarin Paperbacks, Berkshire

Bochner S (1966) The role of mathematics in the rise of science. Princeton University Press, Princeton

Etnyre J (2019) The art of writing introductions. Notices AMS 66:361–362

Gromov M (2018) Great circle of mysteries. Mathematics, the world, the mind. Birkhäuser/Springer, Cham

Hille E, Tamarkin JD (1935) On the absolute integrability of Fourier transforms. Fundam Math 25:329–352

Iosevich A (2007) A view from the top. Analysis, combinatorics and number theory, Student mathematical library, 39. American Mathematical Society, Providence

Iosevich A, Liflyand E (2014) Decay of the Fourier transform: analytic and geometric aspects. Birkh

Lem S (2013) Summa technologiae. University of Minnesota Press, Minneapolis

Liflyand E (2019) Functions of bounded variation and their Fourier transforms. Birkhäuser

Liflyand E (2021) Harmonic analysis on the real line – a path in the theory. Birkhäuser/Springer, Cham

Liflyand E, Stadtmüller U (2013) On a Hardy-Littlewood theorem. Bull Inst Math Acad Sinica (New Series) 8:481–489

Liflyand E, Trebels W (1998) On asymptotics for a class of radial Fourier transforms. Z Anal Anwendungen (J Anal Appl) 17:103–114

Littlewood JE (1957) A mathematician's miscellany. Mandarin Paperbacks, London

Manin Y (2007) Mathematics as metaphor. Selected essays of Yuri I. Manin. With a foreword by Freeman J. Dyson. American Mathematical Society, Providence

Okudzhava B (1982) 65 songs, 2nd edn. Ardis, Ann Arbor

Persson U (2021) A conversation with Reuben Hersh. EMS Mag 121:20–35

Shapiro HT (2005) A larger sense of purpose. Princeton University Press, Princeton

Silverman JH (2019) To write or not to write ... a book, and when. Notices AMS 66:357–358

Stewart I (1995) Concepts of modern mathematics. Dover Publications Inc, New York

Wigderson A (2019) Mathematics and computation. A theory revolutionizing technology and science. Princeton University Press, Princeton

The Unreasonable Effectiveness of Mathematics to Generate New Spatial Imaginaries

Nathalie Sinclair

Contents

1 Introduction 1308
2 Mathematics in the History of Thought 1310
3 Mathematics as Charming and Mad 1312
3.1 Identify and Difference 1314
3.2 Negation and Indeterminacy 1317
3.3 Linear and Fractal 1318
3.4 Metric and Topological 1320
3.5 Parallel and Intersectional 1321
3.6 Devices for Producing New Relations 1322
4 Conclusion 1324
5 Cross-References 1326
References 1326

Abstract

Although not typically thought of as useful or relevant to social theory, this chapter proposes that mathematical ideas and tools can be used to generate new spatial imaginaries about the world, that is, new ways of imaging the relations that pertain to existence (who we are, as people, collectives, and societies) and to experience. I will describe some existing spatial imaginaries (for example: that the world is made of distinct parts) and draw on gender and race scholars to show how they function to maintain various forms of oppression. I also consider several examples of mathematical ideas and tools that trouble or transform normative relations in order to produce new ones – in particular, new spatial imaginaries for disrupting systems of oppression.

N. Sinclair (✉)
Simon Fraser University, Burnaby, BC, Canada
e-mail: nathalie_sinclair@sfu.ca

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_144

Keywords

Calculus · Feminism · Fiber bundles · Fractal · Intersectionality · Race · Spatial imaginary · Topology

1 Introduction

In his writing, Martinique poet and philosopher Édouard Glissant insists that in order to reckon with the consequences of colonialism, including the history of violence it has produced, we need new spatial imaginaries – indeed, he calls for an "insurrection of the imaginary faculties" (Glissant 2005, pp. 24–25). His own archipelagic philosophy, which draws inspiration from his Caribbean sensibility, strives to do just this. Whereas continental thought provides a kind of bird's eye view of the world "as a bloc, taken wholesale, all-at-once, as a sort of imposing synthesis" (2009, p. 45) that is fixed and static, archipelagic thought is multiple (one sees many islands) and relational (the islands are entangled in the spread of land and the movement of sea), as well as local "we know the rivers' rocks, without a doubt even the smallest ones" (ibid.). A spatial imaginary is not so much about imagination, and more a way of seeing/thinking/feeling life or, as Drabinski (2010) writes, "that precarious aesthetic sphere of knowing and being that structures a relation to the world" (p. 304). Spatial imaginaries that tend to the particular, the moving, the unpredictable, to the relational, the multiple, the compositional, are tending away from the abstract, fixed, and the defined and divided. Since many of the dynamics of oppression involve enclosing people into separate (abstract) categories and defining them in various modes of (fixed) "otherness," often in (neatly divided) hierarchies, it is evident that images of existence are significantly at play in systems of oppression.

I am interested in the possibility that mathematics can be a generative source for producing new spatial imaginaries, particularly ones that differ from more normative spatial imaginaries – those, like continent thinking, that guide everyday assumptions about who we are and how we relate to each other. This may seem counterintuitive, especially for those who see mathematics as being far removed from human concerns, and who further see it as a rigid, deductive, rule-based discipline that lacks the kind of poetic or speculative creativity that might disrupt existing spatial imaginaries (and the logics they entail). Indeed, few people would assume that the hermetic, symbolic, and predictive world of mathematics could have much to say about the messy, affective social world.[1] But those who do mathematics, I think, know that it is a playground for creating new objects and relations, even if they very rarely have anything to do with the physical or social world, and it is precisely these new

[1] Those who have studied the role of mathematics in the history of thought know that mathematical ideas can be both precursors to and accelerators of ways of worlding that reverberate in philosophy, architecture, art, literature, science, and so on.

objects and relations – which potentially exceed everyday experience – that I'm interested in.

In this chapter, I explore the different ways in which mathematics can participate in the creating of new spatial imaginaries. I won't argue that doing so absolves it from the many negative effects of mathematics in our society (as a gatekeeping discipline that serves to exclude many people from leading good lives, while causing extraordinary anxiety for a silent majority [see Ernest 2020]). Nor do I argue that it's the only way of producing new spatial imaginaries – as Glissant shows, so can poetry and philosophy. But there is something about mathematics – its freedom from the exigencies of the visible, physical world – that I believe makes it a particularly generative source for troubling normative relations and logics. We can use mathematics to imagine our worlds differently – not just as a model of how things *are*, but as a speculative apparatus for producing *new ways of being and becoming*.

In the sections that follow, I document several ways in which mathematics has already been used to produce new spatial imaginaries. I'm particularly interested in ones that disrupt some of the spatial imaginaries that have served to oppress, and which dominate capitalist, extractive, and colonial systems. A spatial imaginary structures ways of seeing/thinking/feeling relations that use some kind of spatial logic. When Deleuze and Guattari implore us to move away from tree-thinking to rhizomes, they are pointing to a spatial imaginary (the tree) that sets out relations as foundational (the trunk), teleological (toward a fixed goal), and hierarchical (always progressing). In contrast, rhizomes are decentralized, nonhierarchical, and non-teleological. Tree-thinking might lead us to assume that societies have evolved over time to be increasingly civilized, moving toward a goal of liberty and equality, based on a foundation of reason, laws, and sensibilities. But, as Graeber and Wengrow (2021) show quite convincingly, this narrative does not match current archeological evidence.

When Latour (2009) asks us to think about space as something that is made through relations rather than something that contains relations, he too is changing the spatial logic because space becomes inextricable from action, time, and objects, rather than being a container for them. Container thinking is a natural extension of set theory, where the only purpose of a set is to circumscribe its elements. What if the insertion of a new element somehow fundamentally changed the set so that it didn't just lasso around its members, but plugged back into a part of itself – so that the set was not just a whole, but a mechanism for making new parts? Space-as-container imagines a province or a state as being a collection of cities or regions and rivers, mountains, forests, people, animals, etc. There is a flatness to the collection, which merely encloses everything that can be named. But space-as-relation allows actual and imaginary space to be made – new spaces created by stories or events, spaces that cannot be captured by longitudinal and latitudinal coordinates. If my language is ripe with mathematics (sets, coordinates, topology, etc.), it is because these ideas – and the spatial relations they invoke – are not confined to the discipline of mathematics.

I am identifying as normative those ways of thinking that are really only named as such by those who want to imagine a different world, a world in which capitalist,

racial, and colonial forms of oppression are not as rampant. There are of course many other nonnormative spatial imaginaries that could be considered – and reading science fiction provides many examples, some of which might actually be more relevant or helpful than we now realize. But in this chapter, I draw primarily on research that offers critiques of persistent normative ways of thinking and, in particular, grapples with the oppressive structure of current ways of imagining the world as being made up of separable entities that operate in determinate and sequential ways.

Before turning to the examples, I take a short detour into the broader question of how mathematics figures into a history of thought in order to point to instances where the imagining in mathematics has already transformed the world.

2 Mathematics in the History of Thought

> I will not go so far as to say that the construct of history of thought without profound study of mathematical ideas of successive epochs is like omitting Hamlet from the play that is named after him. That would be claiming too much. But it is certainly analogous to cutting out the part of Ophelia. The simile is singularly exact. For Ophelia is quite essential to the play, she is charming—and a little mad. (Whitehead 1929, p. 31)

In his essay entitled *Mathematics and the Good*, the mathematician and philosopher Alfred North Whitehead (1941) undertakes the kind of study of the mathematical ideas of successive epochs mentioned in the quote above, ultimately arguing that mathematics is uniquely situated to attune us to the good because it forces us to reckon with the contingency of knowledge (and its relation to assumptions). While being a claim that one could agree with in principle, it is quite evident that current school mathematical practices likely do just the opposite, and that most people end up thinking that mathematics is about certainty more than contingency. But Whitehead's essay pays little attention to how mathematical ideas can provoke new ways of thinking. This is interesting, given that he himself used mathematics to work out a new set of relations that was better suited to this process ontology, which I will return to in the next section.

If we stick to the mathematics ideas that played a part in the history of thought, then Everdell's argument in his book *The First Moderns*, that the mathematical work around the time of Cantor, Dedekind, and Frege, particularly in relation to the question of "what is a number?," heralded – not on its own, but in a significant way – the late nineteenth-century cultural shift toward modernism.[2] Everdell locates the shift both in the turn toward rigour – exemplified in Frege's axiomatic approach

[2] Everdell is mainly interested in what it is that characterizes the twentieth century, to which he affixes the term "modernity." His list of characteristics of modernity though do not necessarily coincide with the way that modernism has been defined in relation to postmodernism. For my purposes, the actual tags and characteristics are less important than the fact that mathematics is part of the story.

to number – and in the turn to the digital, which can be seen in Dedekind's definition of the continuous through the discrete (the Dedekind cut). In fact, mathematics was just catching up with the positivist turn that had already taken hold in the sciences, one in which objectivity prevailed. In the words of Frege (1984):

> I understand objective to mean what is independent of our sensation, intuition and imagination, and of all construction of mental pictures out of memories of earlier sensations, but not what is independent of the reason, for what is independent of the reason? To answer that would be as much as [...] to wash the fur without wetting it. (p. 43)

Everdell (1998) writes: "Modernism began with mathematics, in a pattern closely resembling the way another revolution in thought had begun 250 years before with the profound mathematical meditations of Descartes" (pp. 351–352). What is interesting is the spatial imaginaries involved in the list of features of modernism that Everdell uses, spatial imaginaries that are still dominant today, including the idea of the discrete.

For Steingart (2023), if Frege's work marks a pivotal point in Modernism, it is the turn to axiomatics (or "high modernism") in the 1950s that entrains a whole new way of seeing the world, not only in sciences, but also in the social sciences. Following in the footsteps of mathematics, complex situations could be reimagined as being governed by a small number of axioms, which could be chosen so as to generate a complex phenomenon, not only in mathematics, but also in economics, psychology, and sociology. Even better was the promise that these mathematical systems would be fundamentally apolitical, working as sorts of creative abstractions that were blind to ethical considerations. This axiomatic use of mathematics, which is reductionist and abstract, does indeed entail certain assumptions about how the world is. It is the kind of thinking that is at play in Birkhoff's equation for measuring the aesthetic value of a certain object (the aesthetic measure, $M = O/C$, where O is order, and C is complexity). It reduces the aesthetic experience to a small number of variables that can then be used to produce a numerical quantity that appears to be "independent of reason," to use Frege's parlance.

Before leaving Everdell, it is worth noting that his history of thought takes the ideas and work of Cantor, Dedekind, and Frege to be of the same kind as that of artists, psychoanalysts, poets, architects, and dancers, in terms of contributing to epochal changes, which would have pleased Whitehead. But the ideas of Cantor, Dedekind, and Frege – thrilling as they are/were within mathematics – were more or less following a broader societal trend, though perhaps the perfect embodiment of those trends, thereby allowing mathematics to become a kind of model of (modernist) reason itself. This would not have pleased Whitehead.

Like most histories of thought, however, Everdell's is hopelessly linear. Discretization happened before and will happen again; other mathematical ideas were introduced in the twentieth century that were decidedly more about the continuous than the discrete. Indeed, there are also histories of "minor mathematics," which I will discuss in the next section, that interrupt the narrative of twentieth-century modernism. Nevertheless, more recently, it is physics rather than

mathematics that has provided the more potent new spatial imaginaries for the social sciences and humanities, through the work of feminist philosopher and physicist Karen Barad (2007). Their reading of the interpretations of Niels Bohr, along with the poststructuralist writings of Jacques Derrida, show how quantum concepts such as indeterminacy, entanglement, and nonlocality can give rise to fundamentally new ways of thinking about agency, causality, and time. The uses of quantum physics in the social sciences are many decades old, and have hitherto been mobilized as metaphors (perhaps sometimes fanciful), but Barad's work figures the quantum as an ontological reading of the world and, relevant to my interests, as an ethico-onto-epistemology. The new spatial imaginaries abound. Instead of positing two objects (people and things) as discrete entities and then investigating how they relate, Barad begins with the relation as the ontological primitive. Using graph theory imagery: Instead of two points, which can then be related (and studied for their individual agencies, for example), Barad posits a line, which is the relation, and which therefore is agential. A simple reorientation can have massive consequences for the social sciences: Instead of isolating the bully as the agential force that causes harm, for example, Zabrodska et al. (2011) locate the agency in the network of relations that produce bullying. Similarly, in mathematics education research, instead of positing the student and the tool as two distinct nodes in which the tool is a mediator to mathematical meaning, de Freitas and Sinclair (2014) study what mathematics the student-tool does (instead of what mathematics the students do with the tool or how do tool mediates certain mathematical concepts).

I bring up quantum physics for two reasons. First, it illustrates well what I mean by the production of new spatial imaginaries that are counterintuitive to normative ways of seeing/thinking/feeling the world. And second, Barad's use of quantum physics does not remain at the level of a metaphor, or even a model, and instead aims to rethink philosophical concepts such as ontology, epistemology, and ethics. A metaphor sets up a mapping between two things and it works because they are actually different. A spatial imaginary is more like a verb – it is the knowing and becoming of the world.

3 Mathematics as Charming and Mad

I pick up on Whitehead's description of Ophelia, as charming and mad, as it resonates with the mathematical ideas and practices that Deleuze and Guattari chose to mobilize in their own conceptualizing of political relationships. Indeed, they drew on the paradoxical, the temporal, and the problematic – sometimes unearthing and other times intensifying dangerous ideas or hidden practices. They contrasted the state-sanctioned mathematics, which silences and demotes unwieldy ideas and replaces them with rigor and axiomatics, with the nomadic or minor mathematics such as "becoming, heterogeneity, infinitesimal, passage to the limit, [and] continuous variation" (Deleuze and Guattari 1987, p. 363). Charming and mad, indeed.

Minor mathematics has obvious relevance in the production of new spatial imaginaries, if only because it often expresses the intuitive, eventful indeterminates that state mathematics attempts to control, and therefore forms of relating that might not be as rigorously defined or categorical or static. Of course, what is minor changes over time and, in fact, usually becomes major (or state) mathematics, as in the case of calculus, infinitesimals, and continuity. Indeed, while the infinitesimals were banned by Jesuits in the seventeenth century, though used surreptitiously by many, they became part of state mathematics through the work of Abraham Robinson. The radical and paradoxical nature of these new numbers, as Deleuze and Guattari worked with them, which had given rise to new, fertile ways of calculating, were eventually tamed (at the state level) – however, as they figure in mathematical practice, as they are sensed and transformed, their paradoxical nature never completely disappears.[3] But Deleuze and Guattari also turned to other mathematical concepts such as number and measurement and fractals. They drew particularly on topology (and associated concepts such as manifolds and singularities) as they reworked their political ontologies towards rhizomatic thinking.

In this section, I will discuss five different spatial imaginaries in which mathematics features significantly, and which generate relations that have sociopolitical import in that they deal with knowledge about existence and events. As Ferreira da Silva (2023) writes,[4] normative ontoepistemological (as well as judicial and economic) assumptions feature the following: *separateness*: that what can be known must be distinguishable from everything else (so, spatially separate, like points on a grid); *determinacy*: that knowing involves understanding the causes of observable effects (so, phenomena can be categorized into grids or tables); and *sequentiality*: that existents and events can be separated into discrete temporal units having positions in a sequence or plot points on a line, and often associated with images of progress (development of Homo sapiens from apes, of adults from children, and of civilized societies from primitive ones) (see also Mikulan and Sinclair 2023). Evidently there are spatial and temporal imaginaries at work here.[5]

The first example (*identity and*[6] *difference*), which draws on Deleuze and Guattari's work, concerns both separateness and sequentiality, where the new spatial

[3] We might even say that in all state mathematics, there is minor mathematics to be found, or rather excavated (see de Freitas and Sinclair [2022] for an example of this in relation to measurement).

[4] Ferreira da Silva is primarily concerned with Black racism and its ongoing relation to capitalism. However, as part of her project, she rereads post-enlightenment philosophy to uncover the "images of existence" that they instilled and promoted. Her own reliance on imagery (and diagrams) is part of what makes her analyses and philosophical insights so unique and persuasive. In addition, as the reader will see later in this section, she draws on mathematics in quite specific ways to produce new spatial imaginaries (while also naming the various ways in which mathematics is also co-involved in post-enlightenment thought).

[5] Importantly, for Ferreira da Silva, these three procedures of separateness, determinacy, and sequentiality (which are wrapped up in Blackness as cause and effect at the same time) operate as necessity.

[6] In order to avoid binaries, negations, and sequentiality, I am using "and" here in lieu of "or," or of the hierarchical connotation of "from. . . to. . .."

imaginary attempts to move from a philosophy of identity to a philosophy of difference – and which is interesting mathematically for the ways in which identity is taken as axiomatic, as in Euclid's axioms. The second example (*negation and indeterminacy*) principally concerns determinacy, reworking the classical binary categorization of A and not-A, which presumes to account for everything, in order to make space for opacity, for that which cannot be categorized and therefore treated causally. The third example (*linear and fractal*) chiefly concerns sequentiality, working in resonance with quantum time to rethink images of the mechanical, progressive passage of time, one that functions on a principle of replacement (the now replacing the past, which will then be replaced by the future). The fourth example (*metric and topological*) concerns all three, focused as it is on the move from spaces that can be measured from the outside, in which part-whole relations are stable and ordered, to spaces that produce their own measures, and do so in ways that respect spatiotemporal entanglements that are less prone to ascending lines and more to folds and inflections. The fifth example (*parallel and intersectional*) concerns both separateness and determinacy in that it shifts from multiple independent descriptors to entangled ones – entangled not only among themselves, but in broader sociomaterial processes. This in effect messes with determinacy because grids and tables are no longer adequate ways of accounting for effects. In the final section, I show how geometric inversion can function as an apparatus or device that makes things visible. To the extent that inversion works on relations of proximity and relatedness, inclusion, and exclusion, it becomes an interesting generator of new spatial imaginaries.

3.1 Identify and Difference

I have already mentioned Deleuze's interest in spatial imaginaries, which drew on biology, but also frequently on mathematics. One of his main concerns, as a philosopher, was to work out an alternative to the prevailing philosophy of identity, which he called a philosophy of difference, and which became crucial for his political ontology (de Freitas 2016). Following Leibniz, Deleuze found much inspiration for his philosophical work in calculus, particularly with concepts such as the differential relation, singularities, multiplicities, and the virtual. Of course, there is a long tradition of interaction between mathematics and philosophy in the West, going back to Plato – for whom Euclid's geometry provided a model for ideal forms. As Smith (2007) notes, though, philosophers of the nineteenth and early twentieth centuries, whose interests have primarily been foundations, tended to gravitate toward axiomatic set theory rather than to calculus. But Deleuze and Guattari were interested neither in foundations nor in ideals – they were interested in the conditions of real experience, and in formulating a philosophy of difference adequate to those conditions (see Smith [2007] for more details). Such a philosophy would break with the philosophy of identity, which defines a thing as essence or substance, which in turn has particular predicates or properties. In a philosophy of identity then, difference is secondary – I first define A and B and then compare them according to their

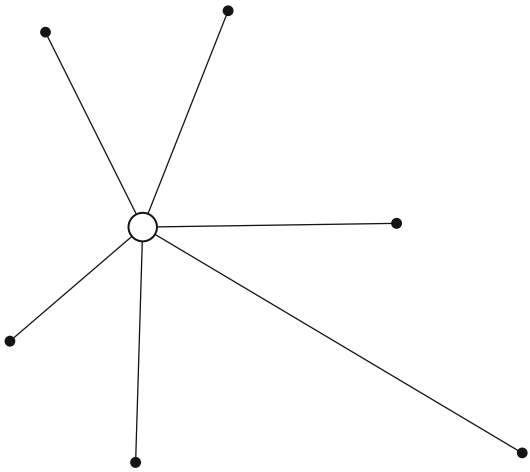


Fig. 1 Substantialist images: a thing and its properties

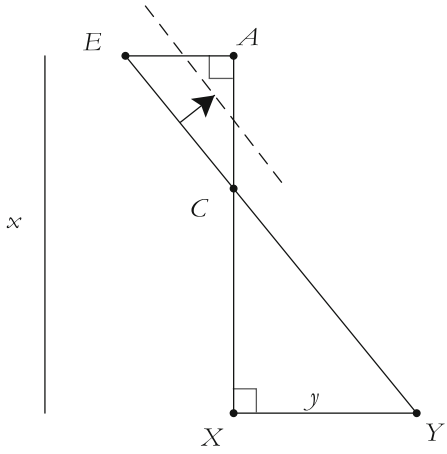
properties, which I have determined in advance.[7] Spatial imaginaries associated with the philosophy of identify include graph theoretical nodes and edges. A node is a thing (chair, human, triangle, and river) and it has edges that represent each of its properties (Fig. 1). With each node being a sovereign thing, equality can be achieved when the properties are the same, while difference arises when they are not. Deleuze's challenge was to think about difference as primary, so that A and B emerge out of pure differing (rather than out of preexisting differences).

Deleuze drew heavily on Leibniz for this work. I will run through the differential triangle example, also provided in Smith (2007), because it exemplifies so powerfully (and spatially) what difference as primary can look like.[8] Take two right-angled triangles CAE and CXY, as in Fig. 2. These two triangles are similar, which means that the ratio EA/AC in the top triangle is equal to XY/CX in the bottom triangle. Now imagine moving EY parallel to itself, toward A, as indicated by the arrow and dotted line. The lengths of EA and AC will decrease, eventually becoming zero, *yet*

[7] The idea of equality – which is so important in mathematics – is also based on identity. Mazur (2007) provides an excellent overview of the complexities of equality, its dependence on separateness (you can only say that two things are the same, if you separate out specific qualities, like white, or prime, or one-holed). Reading Mazur, I get the sense that even mathematicians want to move away from the overbearingness of equality, as evidenced by associated relations such as equivalent, congruent, and isomorphic!

[8] Indeed, it can be difficult to use language to give an adequate sense of Deleuze's philosophy of difference, particularly given the nounified tendencies of English, where things are already separate and distinct, rather than eventful. From a phenomenological point of view, imagine looking out at the ocean. As you look more closely, you begin to make distinctions *which arise from relations* – the water is higher here (on the wave) but not there, the color is a little darker over there, a bird swoops and catches a fish, etc. These are all distinctions that produce difference, where the differences are intrinsic to the eventful world, rather than external, as abstract categories. As Deleuze wrote, the problem with abstract categories is that they let too many fish pass through!

Fig. 2 Leibniz's triangle



their ratio will continue to exist and be a perfectly well-defined finite quantity *x*/*y*. This differential relation (*dx*/*dy*) is "a relation that persists even when its terms disappear" (p. 11). In contrast to differences that are between two things that have an a priori identity (and could therefore not exist without the things), Deleuze's concept of difference-in-itself does not require identity.

Taken as a spatial imaginary, both *motion* (the moving line EY) and *analogy* (that EA:AC as XY:CX) feature strongly. Motion helps move away from a substantialist view of things (static, determined through properties). Analogy suggests whatever EA and AC are at a given period is conditioned by the triangle XYC, so that the conditions of difference are determined with what they condition, and not more abstractly. This is an ontologically significant step in Deleuze's philosophy, but it is also one that offers new ways of thinking about identity and difference, which are crucial terms in gender, sex, class, ability, and race politics. A philosophy of identity takes categories such as Women and Man or White and Black or Rich and Poor, and substantializes them by loading them up with fixed (often stereotypical) properties that are usually determined in relation to a dominant identity (Man) which then subordinately determines the other (Woman). Within identity politics, it becomes difficult to attend to other differences (including ones not coded in the binary) or even to uniqueness, that is, the way in which we each exceed stereotypes. The differences upon which these categories arise are not seen as conditions, but as givens.

In a philosophy of difference, there are no natural categories that give rise to difference, but only differences of differences, which are always contingent. In other words, difference A with respect to condition *a* and difference B with respect to *b* are actualized situationally rather than universally. The conditions themselves must be plastic and mobile, "no less capable of dissolving and destroying individuals than of

constituting them temporarily" (Deleuze 1994, p. 38). Deleuze therefore replaces the traditional concept of substance with that of multiplicity.[9]

While Deleuze finds strong inspiration in calculus, we must recognize that calculus itself has also "[lent] credence to the classical view of determinism, that is, a clockwork universe without any novelty, in which the future was completely determined by the past" (Smith 2007, p. 15) through astronomical predicts and discoveries. Indeed, while the spatial imaginaries may be latent in the ideas of calculus, they also need to be recognized not just as techniques to solve problems (about astronomy), but as ways of understanding existence and events. The importance of the intentions associated with the wielding of new spatial imaginaries will become especially clear in the next example.

3.2 Negation and Indeterminacy

Brazilian and Black feminist scholar Denise Ferreira da Silva (2017) uses mathematics in a more radical way, using only the laws of arithmetic as abstract devices to "explore [the] potential of blackness to unsettle ethics" (p. 8). In her thinking about race, she works against the Hegelian dialectic logic that begins with the Aristotelian positing of A and not-A. The A, defined as what is taken to be the "natural" category, which in Ferreira da Silva writing is White, is the one against which other categories are measured, while not-A is Black, and therefore defined negatively, and only determined in relation to A.[10] She argues that continuing to fix a Black identity – by focusing on how Black lives matter, for example – maintains the category not-A, and will therefore never disrupt the logic in which Blackness is defined with respect to Whiteness (rather than on its own grounds).

A mathematical example might be helpful in exemplifying the unsettling that Ferreira da Silva is undertaking – and not that I am, with caution, following in Ferreira da Silva's footsteps by dangerously associating people with numbers.[11] Her project is like rethinking irrational numbers, which are usually introduced as the not-A of rational numbers, so that they are not fixed, passively, as a mere negation of the normal, natural category of rationals. Working towards making irrational numbers easier to understand or more represented in everyday life will not change the fact that they are ultimately subordinate to the category of rationals. How to think of

[9] Multiplicity here isn't about having multiple perspectives on the same thing, as an epistemological reading; it is about being multiple, ontologically – being different things at once. The image of the electron as *both* wave and particle is one way to think about this multiplicity.

[10] This dialectical habit of Hegel has been criticized by many, including Ferreira da Silva and Deleuze. Rendered visually, as the set A and then everything outside of A, but within the rectangle that is the "whole," there are three spatial assumptions being made about relations: that there is a whole that can be posited; that there is a strict distinction between A and not-A; and that the categories are static, unchanging, and fixed.

[11] And just as dangerously, bringing in language of rational and irrational to do so, much like she works with positive and negative, the racist connotations thus entailed.

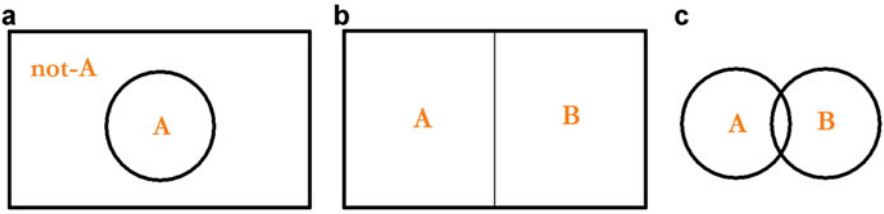


Fig. 3 (**a**) Opposing A and not-A; (**b**) separating A and B; and (**c**) distinguishing A and B without making them exclusive or exhaustive of other differences

irrationals in ways that somehow exceed the restricted, subservient space they have been given?

Ferreira da Silva mobilizes Blackness's capacity to unravel thought as dialectic (where this thought is always a racial dialectic) without reproducing the violence housed in Western concepts of knowledge, and she does this through arithmetic. Setting life (Whiteness) equal to 1 and not-life (Blackness) to 0 – which is, critically, not the Hegelian −1 of not-Whiteness – she proceeds to calculate. And so: 0 added to or subtracted from itself is still 0; 0 multiplied by anything is still 0. Therefore, addition, subtraction, and multiplication reproduce themselves, maintaining the categories. But division turns out to be a generative operation because dividing by 0 is undefined. This is how Ferreira da Silva achieves the indeterminacy that refuses to contain Blackness. She accords to Blackness the power to unsettle the ethical program governed by the determinacy of Aristotelian logic in which there is either A or not-A, a determinacy that captures all (the general, abstract) in one fell swoop.

The logic that Ferreira da Silva is working within and against, which she deploys in relation to Blackness (as a category of modern abstraction), is one that is relevant in many other settings. Feminists have also critiqued the way in which the category of Woman is defined in terms of that of Man; disability is similarly defined in relation to a normative category of ability. Glissant would argue that these habits of thought are not just discursive, but draw – perhaps subconsciously – on spatial imaginaries, by which he means the ways we assume things to be related. The dialectical (and often dualistic) imaginary is well expressed by the image in Fig. 3a. Its entailments may not be so obvious until we consider other ways of relating, as seen in Fig. 3b, c. Figure 3b operates under strict division, but does not function through negation, or through the abstract category that catches everything but A. Figure 3c does not posit the total space of possibility (there is no enclosing rectangle), and admits overlaps between A and B (and could perhaps allow the potential of more overlaps of C, D, etc.).

3.3 Linear and Fractal

Fractal images have been used widely in the popularization of mathematics, featuring on computer screens, in artwork, and in extracurricular or enrichment activities

with students. Few would object to the claim that fractals are more popular outside mathematics than within, where they tend to be seen as belonging to recreational mathematics. Ferreira da Silva (2016) mobilizes what she calls fractal thinking as an intentional way to disrupt the linearity of time. Clockwork time – which orders the past, present, and future, and which flattens temporality into extensive units of hours, days, and years – instills and maintains sequentiality. This sequentiality enables linear thinking in which the past is over, and doesn't intervene in the present or the future; it posits that the future is only yet to come, perhaps causally following from the past, but certainly not affecting it (see also Mikulan and Sinclair 2023). For Ferreira da Silva, fractal thinking is a way to account for how the past operates in the present, a way to reckon with the violence of history, to see similarities across space and time (see Fig. 4). She deploys these similarities in the specific context of violence against Black people, showing not only that these violences repeat, despite changes in law, in the skin colour of the president, in anti-racism education, but also that they repeat because they are the necessity by which capitalist extraction continues to accumulate the total value.

In contrast to linear thinking, which might posit a break – a "now we know, so we can be different" – fractal thinking does not erase the past. As Ferreira da Silva points out, and in resonance with Barad (2007), fractal thinking draws more on quantum rather than classical approaches to physics. In quantum physics, time does not operate linearly: the future can change the past; the past and future are symmetrical.

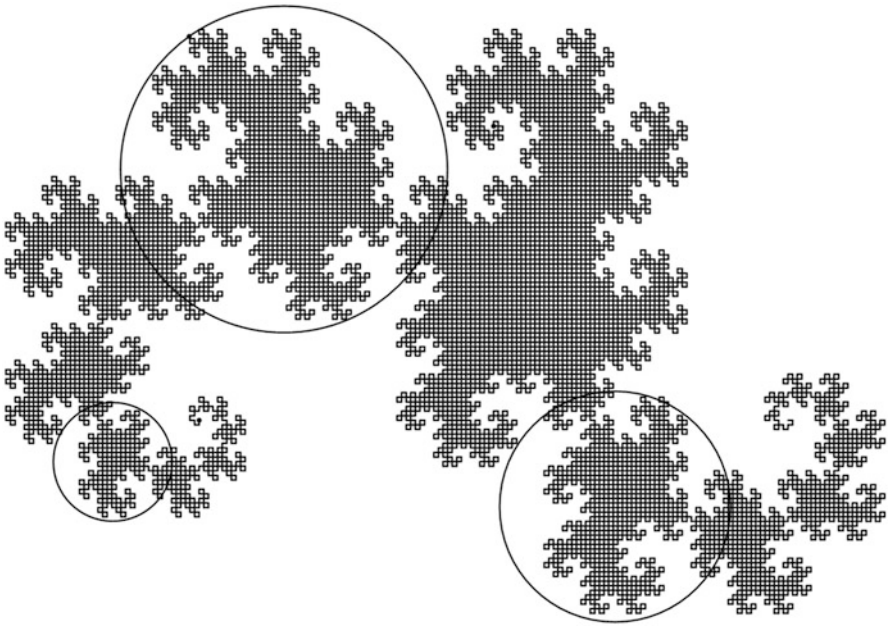


Fig. 4 The dragon curve with three circles to show similarity across different spatiotemporal moments

Without the sequential line, the whole assumption of progress dissipates, as we acknowledge the ways in which things repeat – similarities across scale. The world becomes not just the immediate present, and not even just the historical march toward the present, but also that which was never actualized – or, in the words of Ferreira da Silva, "a composite of so much of what never came to be and what never will be" (p. 265).

Although fractal images have significant spatial appeal, it is less the overall spatiality that interests Ferreira da Silva and more the spatiotemporality involved in the iterative similarity: that even though a particular event occurs in a particular time and place, it is literally present in future events (at other times and places) – that events are stacked in space-time rather than being lined up sequentially. Her injunction is for us to refuse to keep history in the past; to not be surprised when "the same thing" happens again "even though. . ."; and to cultivate different spatial imaginaries that can interrupt the neat, progressive flow of clockwork time. The repercussions are far-reaching. They change how we imagine our responsibility to the past and the future; they change assumptions about equality and liberty by acknowledging that humans cannot be – and never will be – equal across space-time (though certainly unique).

3.4 Metric and Topological

In some senses, Whitehead's interest in process philosophy arose from his own critique of the post-Enlightenment thinking of philosophers such as Kant and Hegel. Like Deleuze, he was interested in the conditions of real experience, and was therefore suspicious of fixed categories and static abstractions. While he did not share Glissant's or Ferreira da Silva's political agenda, there are aspects of his own spatial imaginaries that are consonant with theirs – particularly in relation to his mereotopology. Whitehead relied on a distinctively mathematical approach to generating the new relations that could better align with his process ontology. He found that set theory, the prevailing foundational theory, was inadequate to this philosophy, since it was based on the ideas of an abstract set (which we might say belongs to Glissant's continental thinking) and the relation of inclusion with its members. Whitehead abandoned the linear part – whole relation of set theory – where relations are either about inclusion or exclusion – in favour of an entanglement of superposed parts and whole, where each part *is* a whole *and* a part within that whole. While set theory invites thinking of the abstract collection, the "universe," or the container for the elements (the rectangle of Fig. 3a), Whitehead's mereotopology avoids this unified, abstract whole. It concerns the topological transformation of wholes and parts that do not bow to the fixed, spatial logic of set inclusion.

Mereotopology invites us to think of the complex network of open-ended activities, all in relations of resonance but also differentiated along multiple, spatiotemporal timelines. In other words, while one part–whole relation may hold at one particular time/space, the whole may become the part in another time/space. This is not simply a graph of points related by lines, since the points themselves are not

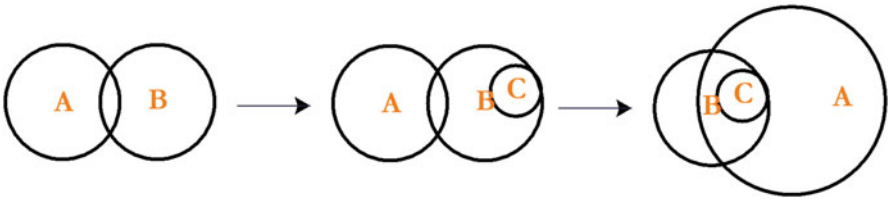


Fig. 5 Mereotopological relations of overlap, touching, partials inclusions, making space as relations change

fixed, and since the two- or even three-dimensional space in which the graph is embedded does not allow for simultaneous spatiotemporal timelines. Mereotopology, when operating as a spatial imaginary, conceives of relations as being overlapping (though still discrete), sometimes contradicting and always changing affiliations in the world – it is about events rather than discrete objects. Diagrammatically, this might involve multiplying the forms of relation (not just inclusion, but crossing, overlapping, tangent, straddling, etc., and iterating over time). These relations seem to invite sensings that are not just ocular – Kant's preferred sense of abstraction – but haptic (touching) and kinetic (moving) as well (Fig. 5).

3.5 Parallel and Intersectional

In his article entitled *Fiber Bundles and Intersectional Feminism*, Dagan Karp (2021) uses a mathematical concept (fiber bundles) as a means of explaining the sociopolitical concept of intersectionality. It is evident that the intended audience is mathematicians since most journal article–reading people would likely find it much easier to understand intersectional feminism than fiber bundles! Indeed, Karp provides a comprehensive, clear, and persuasive overview of gender inequity in mathematics, providing ample evidence of the underrepresentation of women in mathematics, and the even more significant underrepresentation of racialized women in mathematics – highlighting both institutional and personal biases, and therefore both psychological and sociological causes. He also provides an excellent overview of the different waves of feminism, including the more recent move toward intersectionality, coined by Kimberlee Crenshaw, which recognizes the multiple, intersecting identities that individuals have, including race, gender, ethnicity, sexuality, economic class, and ability, and how the dynamics of difference – and of inequity – affect certain people more than others when their multiple differences (female *and* poor, for example) are taken into account. An important insight of intersectional feminism is that "just as aspects of identity in individuals are intersecting and inseparable, so too are the systems of oppression, bias and privilege governing our lives" (p. 5). Given this multidimensional, entangled set of causes and consequences at play, Karp turns to fiber bundles to *model* this complex situation.

It can be instructive to consider other mathematical concepts that would not model intersectional feminism well. For example, while Venn diagrams certainly model the idea that individuals might be in different and sometimes intersecting categories, it fails to show how these multiple identities are shaped by and shape societal norms and histories – and, importantly, that systems of oppression affect some people more than others. On the other hand, the topological concept of fiber bundles enables Karp to make visible how the "fibers" of US society ("ableist cis-heteronormative white-supremacist capitalist patriarchy" (p. 8)) are maps between the base space (society, or the set of people) onto another topological space (certain aspects of identity, which are inseparable, or intersecting). Since these maps preserve the structure of the fiber, so too are systems of oppression preserved. In other words, the intersection of identities cannot be separated from these sociopolitical systems, and therefore both shape and are shaped by them.

By refusing the assumption of separability – both in terms of aspects of identity and in relation between individual and collective manifestations of them – Karp's use of fiber bundles breaks with discrete spatiality. Interestingly, the strength of the preservation of mapping, which highlights the inseparability of individual social forces, may also be its weakness if we are looking for spatial imaginaries that change systems of oppression. As such, this example shows that mathematics is quite good at modeling complex sociopolitical issues – and, as I've said before, it shapes and is shaped by existing spatial imaginaries – but the prior examples differed from this one in that they aimed to create new spatial imaginaries. For example, in the same way that Ferreira da Silva raises concerns about the reifying nature of identity politics, we can also ask how further narrowing down on intersectional identities operates both to highlight invisible forms of inequity *and* render static and abstract a category through a philosophy of identity. Do the structures of the fibers always have to be preserved? Can they be transformed through operations similar to Ferreira da Silva's division by 0?

3.6 Devices for Producing New Relations

Inversion is a geometric transformation that has the effect of turning the plane inside out about a given circle. Points that are infinitely far away curl into the finite middle, whereas points within the circle are mapped outside – with the center corresponding to the infinitely far away. The idea was invented by Jacob Steiner (1796–1863) and led to new branches of mathematics and to new ways of solving previously intractable problems. Inversion transforms any point (Q) inside a circle (center O, radius r) into one outside the circle (P), such that $OP/r = r/OQ$. The circle of inversion can be chosen arbitrarily, thus performing the crucial act of contingency. Figure 6a shows six points (labeled A, B, C, D, E, and F) and a circle of inversion. Once the inversion is performed, the points are mapped to A′, B′, C′, D′, E′, and F′ (Fig. 6b). Two points that were previously close together (A and B) are now extremely far apart, thereby upsetting proximity relations. Points that were previously peripheral (C and D) are not central. Points that were included (A, B, and E) are now excluded, and vice versa.

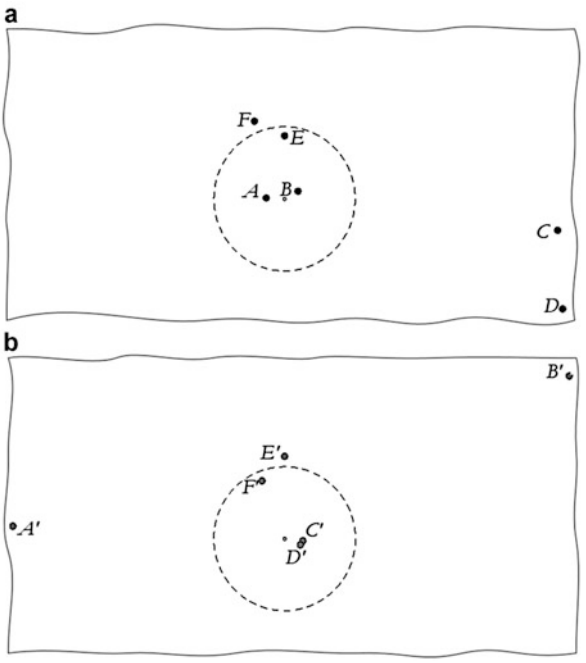


Fig. 6 (**a**) Points on the plane, along with a circle of inversion; (**b**) the inversion of the points

Points that are close to the circumference of the circle of inversion change little, which means that the transformation treats different points differently, depending on location.

de Freitas et al. (2022) use inversion to explore new geographic relations that can be created between cities (in their case, the locations of the different members of an international education project) that might disrupt normative associations. For example, cities that might have been previously related through associations of proximity (Calgary and Edmonton) can become wrenched apart if the circle of inversion is centered around them, scattering former alliances. Relatedly, previously far-flung locations, such as Vancouver and Manchester, can inherit new relations of proximity. With lines becoming circles, the usual bidirectional flows of goods (and historically, slaves) that comprise the trade triangle of the Atlantic Ocean now "devolve into vortexes, dethroned of their power to circulate in the same way" (p. 10). Inversion thus functions as a device for generating new ways of thinking spatial relations (less focused on "natural" categories of proximity and linearity), and speculating about new relations and new future movements. The new map they produce, through inversion, is seen as "folding mathematics into a critical geography," as "refus[ing] the typical tame and clean measures of Euclidean geometry" (ibid.).

Inversion thus functions as a generative device that is not intended to represent phenomena – as doing so would reify their existing relations – but to generate new ones for consideration. Much in the same way that one might perform a vertex

splitting on a graph or alter an axiom or transform a function, what is produced as a new object can then be explored for potential meanings. Inversion is a particularly powerful device because it works so closely with spatial imaginaries related to inclusion, boundaries, and proximity, which are central both in determinations of belonging (or not) to groups of people, tracts of land, or networks of trade.

Given that mathematics is rife with transformations – and since Klein's reworking of geometry, transformations have arguably become more important than objects – inversion is probably not the only device of potential interest in the generating of new spatial imaginaries. For example, sheafs are transformations used in algebraic topology that can be used to "enhance" a mathematical object such as graphs or topological spaces, functioning like a garden on a plot of land, to use Agrios's (2022) metaphor, growing stalks that can be used to produce new vegetables. Like fibers, sheafs remain faithful to the structure of the underlying object. However, what seems to me interesting about sheafs is the way they take up the move between local and global. In Agrios' example (of a sheaf of abelian groups of continuous functions on a topological space X), the sheaf section comes out as a single graph of a continuous function that is made up of the infinitely many pieces of graph, which are images of the group of continuous functions on specific domains of the x-axis (these domains are the open sets, or the chunks of dirt in the garden). This way of producing the global just by applying specific rules on the local – rather than having to use different rules for the global – seems like a compelling image to think with as we reckon with questions of how to address global problems such as climate and immigration at local scales (see Aït-Touati et al. 2019/2022 for a concerted discussion of the need for new ways of seeing the world). It also reified the quantum and fractal imaginaries from the earlier section.

4 Conclusion

The title of this chapter evokes the famed Eugene Wigner's paper *Unreasonable Effectiveness of Mathematics in the Natural Sciences*, written in 1960. According to both Platonic or sociocultural theories, which both adhere to a fundamental distinction between nature and culture, the effectiveness of mathematics in the natural world is indeed unreasonable. However, if we refuse to see mathematical concepts as belonging to some transcendental reality or as emerging uniquely from social foundations (including language, tools) and we acknowledge the material-discursive entanglement of world (Barad 2007) – including mathematics – then the effectiveness of mathematics in the natural world is not so unreasonable. Similarly then, it is not so unreasonable to think that mathematics shapes and is shaped by human and nonhuman experiences in the world. It is not unreasonable to imagine that our very ways of seeing and feeling the world and its relations are caught up in the various forms of logic that we then use in mathematics. Indeed, if one adopts a monist philosophical approach, then it's less clear that we should not be thinking about these examples I've offered as *mere* models or metaphors (relations that impose their own

forms of separateness). Mathematics, in its charming ways, *should* help us understand our (actual and potential) world.

As a mathematics educator, writing this chapter has reminded me of the very narrow scope of *possibilities for* new spatial imaginaries we offer in the K-12 mathematics curriculum. No topology, no fractals, no infinitesimals, no pathologies or paradoxes, and very little infinity and indeterminacy. I imagine that if we did include these topics, they would be handled as natural, continuous extensions of previous mathematical ideas (as we do when we introduce negative numbers or rational numbers), posing no threats to constant, inevitable progress (never addressing that which is lost along the way). Those are, after all, our dominant spatial imaginaries, inherited through modernist values, habits, and beliefs. Since generations of students, who grow up to be teachers, operate with them, it is not surprising that so little changes.[12] What if we moved back into an alliance between mathematics and philosophy, in education – one that was less focused on preparing students to join the workforce, and more intentional about helping them understand different ways of seeing/thinking/feeling?

Finally, readers might wonder why I focus on spatial imaginaries. In one sense, unlike language, these are often left unspoken and structure our assumptions, even if not at a conscious level. This is certainly cultural in nature, given the important role that eyes have been given in Western conceptions of reason – where, for Kant, for example, the eye is the noblest of the senses,[13] closest to providing abstractions. It is also technological in nature, surrounded as we are by images, which function so well on the flat surfaces that dominate our media (paper, screens, billboards, etc.). From a philosophical point of view, it would be interesting to consider what might happen were touch or sound to function in ways similar to sight. Might we consider tactile and aural imaginaries? If so, could mathematics still function as generatively in producing such imaginaries? Any study of the actual practice of mathematics reveals that, in fact, touch, sound, and image are central to mathematical thinking and communication (see, for example, Menz 2015). The current turn toward philosophy of mathematical practice might help expose the ways in which separateness, determinacy, and sequentiality may be good descriptors of published mathematics (including the mathematics of textbooks), but not of the actual doing of mathematics.

[12] Obviously, there are many other factors involved in the tenaciousness of modernist's spatial imaginaries, most important being the capitalist, colonial, individualistic society in which we live, which thrives on the discrete and the calculable, as well as in myths of progress. On one hand, it is quite mad to think that mathematics education could make a difference. On the other, as a discipline that straddles the sciences and the arts, that powerfully operates the gatekeeping functions in schools and universities, mathematics stands out as a compelling option.

[13] And for Plato, it was the voice. We know of the mathematical performances of Ancient Greek geometers, with their drawings in the sand, but what might we learn from speculating about the rhythms and tempos of their oral demonstrations? Could we find new audio-spatial imaginaries involving repetition, refrain, or harmony?

5 Cross-References

- ▶ Mathematical Practice as Philosophy, with Galois, Riemann, and Grothendieck
- ▶ Ontology in the History and Philosophy of Mathematical Practice: An Introduction
- ▶ The Ethics of Mathematical Practice

References

Agrios M (2022) A very elementary introduction to shears. https://arxiv.org/pdf/2202.01379.pdf. Downloaded 22 June 2023

Aït-Touati F, Arènes A, Grégoire A ([2019] 2022) Terra Forma (trans: Demarco A). The MIT Press, Cambridge

Barad K (2007) Meeting the universe halfway: quantum physics and the entanglement of matter and meaning. Duke University Press, Durham

de Freitas E (2016) Deleuze, ontology, and mathematics. In: Peters M (ed) Encyclopedia of educational philosophy and theory. Springer, Singapore. https://doi.org/10.1007/978-981-287-532-7_374-1

de Freitas E, Sinclair N (2014) Mathematics and the body. Material entanglements in the classroom. Cambridge University Press, Cambridge

de Freitas E, Sinclair N (2022) Proportion, analogy and mixture: unearthing mathematical measurement practices. In: Ingold T (ed) Knowing from the inside: cross-disciplinary experiments with matters of pedagogy. Bloomsbury, London, pp 63–80

de Freitas E, Sinclair N, le Roux K, Solares-Rojas A, Coles A, Ng O (2022) New spatial imaginaries for international curriculum projects: creative diagrams, mapping experiments, and critical cartography. Qual Inq 28(5):507–521

Deleuze G (1994) Difference and repetition (trans: Patton P). Columbia University Press, New York

Deleuze G, Guattari F (1987) Capitalism and schizophrenia: a thousand plateaus (trans: Massumi B). The University of Minnesota Press, Minneapolis

Drabinski JE (2010) What is trauma to the future? On Glissant's poetics. Crit Humanit Soc Sci 18(2):291–307

Ernest P (2020) The dark side of mathematics: damaging effects of the overvaluing of mathematics. In: Ineson G, Povey H (eds) Debates in mathematics education. Taylor and Francis, London, pp 29–42

Everdell W (1998) The first moderns: profiles in the origins of twentieth-century thought. The University of Chicago Press, Chicago

Ferreira da Silva D (2016) Fractal thinking. Accessions No 2. https://accessions.org/article2/fractal-thinking/

Ferreira da Silva D (2017) 1 (life) ÷ 0 (blackness) = $\infty - \infty$ or ∞ / ∞: on matter beyond the equation of value. e-flux J 79. http://www.e-flux.com/journal/79/94686/1-life-0-blacknessor-on-matter-beyond-the-equation-of-value/

Ferreira da Silva D (2023) Unpayable debt. Sternberg Press, London, UK

Frege G (1984) Collected papers on mathematics, logic, and philosophy (ed.: McGuinness B). B. Blackwell, Oxford

Glissant É (2005) La Cohée du Lamentin. Poétique V. Gallimard, Paris

Glissant É (2009) Poetics of relation (trans: Wing B). University of Michigan Press, Ann Arbor

Graeber D, Wengrow D (2021) The dawn of everything: a new history of humanity. Allen Lane, London

Karp D (2021) Fiber bundles and intersectional feminism. In Gerbelli-Gauthier M, Harris P, Hill M, Karp D, Riehl E (eds) A conversation on professional norms in mathematics. American Mathematical Society, pp 63–74

Latour B (2009) Spheres and networks: two ways to reinterpret globalization. Harv Des Mag 30(1). http://www.harvarddesignmagazine.org/issues/30/spheres-and-networkstwo-ways-to-reinterpret-globalization

Mazur B (2007) When is one thing equal to some other things? https://people.math.harvard.edu/~mazur/preprints/when_is_one.pdf

Menz P (2015) Unfolding of diagramming and gesturing between mathematics graduate student and supervisor during research meetings. PhD dissertation, Department of Mathematics, Simon Fraser University. https://theses.lib.sfu.ca/thesis/etd9092

Mikulan P, Sinclair N (2023) Time and education: time pedagogies against oppression. Bloomsbury, London

Smith DW (2007) The conditions of the new. Deleuze Guattari Stud 1(1):1–21

Steingart A (2023) Axiomatics: mathematical thought and high modernism. The University of Chicago Press, Chicago

Whitehead AN (1929/1978) Process and reality: an essay in cosmology. Macmillan, New York

Whitehead AN (1941) Mathematics and the good. In: Schilpp P (ed) The philosophy of Alfred North Whitehead. Northwestern University, Evanston, pp 666–681

Zabrodska K, Linnell S, Laws C, Davies B (2011) Bullying as intra-active process in neoliberal universities. Qual Inq 17(8):709–719

High Mathematics Meets High Finance

A Stylized History of Mathematics in Finance

Tom P. Davis

Contents

1 Introduction 1330
2 1202 Liber Abacci: Leonardo Pisano 1330
2.1 The Law of One Price 1331
2.2 The Value of a Firm 1331
2.3 The Time Value of Money 1332
3 1900, 1951 Brownian Motion and Stochastic Calculus: Louis Bachelier and Kiyosi Itō 1333
4 1952 The Efficient Frontier and Matrix Algebra: Harry Markowitz 1339
4.1 The Capital Asset Pricing Model 1341
5 1973 Option Pricing: Fischer Black, Myron Scholes, and Robert Merton 1344
6 October 19, 1993 1347
7 1947 Modern Techniques of Quantitative Finance: The Feynman-Kac Theorem 1348
8 1977 Monte Carlo Methods: Phelim Boyle 1350
9 1978 Partial Differential Equations: Eduardo Schwartz and Michael Brennan 1352
10 Conclusions 1358
References 1359

Abstract

Mathematics in finance has prehistoric origins, in fact it is argued that an accounting system of clay tokens used for prehistoric commerce were precursors to both letters and numerals. Furthermore, Hammurabi's code, the first known codified set of laws, has sections governing interest rates and banking operations. Throughout history, many mathematicians have turned their attention to finance. For instance, Fibonacci introduced the numerals to Western Europe in the thirteenth century in order to facilitate trade. Bernoulli first discovered Euler's

T. P. Davis (✉)
FactSet Research Systems Ltd., London, UK
e-mail: todavis@factset.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_135

constant e in the seventeenth century when investigating compound interest. However it was not until the turn of the twentieth century that quantitative finance began in earnest with the doctoral dissertation of Louis Bachelier where he used Brownian motion and stochastic calculus to determine the price of stock options. The late twentieth century saw the field burgeon into a widely studied discipline after two seminal events: the publication of the Black-Scholes-Merton paper in 1972 and the cancellation of the Texas Superconducting Supercollider by the US Congress in 1993.

Keywords

Quantitative finance · Mathematical finance · Stochastic calculus · History of mathematics

1 Introduction

This chapter will present the important ideas in mathematical finance and their historical origins, as well as how an influx of physicists and mathematicians to Wall Street accelerated the field. The chapter will focus on the ideas from mathematics that impacted finance, in so far as the ideas changed the way capital was allocated in Western society.

The first accounting system was discovered around 9000–8700 BCE, coinciding with the onset of agriculture (Schmandt-Besserat 1992). Clay tokens were used to facilitate widespread cereal production and even trade among different civilizations. For instance, the same clay tokens were found dating to the same times in ancient Syria as well as ancient Iran, in the Mesopotamian fertile crescent, near the Tigris and Euphrates river along the shores of the Mediterranean sea. These clay tokens predate both numerals and alphabets and were likely their precursors (Morley and Renfrew 2010). Furthermore, there is evidence that the onset of agriculture, trade, and commerce interest rates were levied, and hence ancient civilizations knew the time value of money. The amount of interest seems to be related to the ease of computations in the counting systems employed, and interestingly the interest rate became lower as the civilizations advanced (Hudson 2000).

2 1202 Liber Abacci: Leonardo Pisano

From the onset of written history until the thirteenth century, mathematical calculations for commerce in Western Europe were done at various times using pebbles as counters, Roman numerals, elaborate finger manipulations, and mechanical abaci – in fact the word “calculus” derives from the Latin word for pebble. This changed in the thirteenth-century CE when Leonardo Pisano, the son of a wealthy Pisan merchant of the Bonacci family, published the book *Liber Abbaci*. This book

introduced the Hindu-Arabic numerals 0–9, introduced arithmetic using these digits, and ushered in a fundamentally new way of conducting trade. Leonardo Pisano is today known as Fibonacci, a contraction of the latin *filius Bonacci* ("son of Bonacci") coined by the historian Guillaume Libri in 1838 (Devlin 2011).

In his early childhood, Fibonacci traveled to Bugia in North Africa, in modern day Algeria. There, he was introduced to the Hindu-Arabic numerals and realized their significance from a commercial point of view. On his travels, he studied mathematics and was influenced by Arabic scholars. Upon his return he wrote *Liber Abbaci*, which included many problems that introduced revolutionary business concepts, such as the law of one price, the value of a firm, and the time value of money.

2.1 The Law of One Price

There are three entire chapters devoted to the concept of the law of one price. Chapter 8, "On Finding the Value of Merchandise by the Principal Method," Fibonacci uses ratios to determine the price of a quantity of merchandise when the price of a different quantity is known. Clearly this type of calculation was done prior to *Liber Abaci* however not in the Hindu-Arabic numeral system where the calculation is straight-forward and can be done by anyone, not solely the dedicated computers who used abaci, Fibonacci had brought the power of computation out of secluded guilds and into the hands of the everyday businesspeople.

Chapter 9, "On the Barter of Merchandise and Similar Things," Fibonacci uses the principal method of valuing merchandise to teach merchants how to barter fairly. In the same chapter, he uses similar techniques to exchange different currencies, since in Italy the thirteenth century, there was no single currency in Italy, but rather dozens minted and in use. Therefore "foreign exchange" was necessary in order to trade with other city-states. Indeed, Chapter 10 is devoted to the alloying of monies in different coins in order to determine their equivalent values.

2.2 The Value of a Firm

The first sentence of the tenth chapter introduces a concept that seems obvious today; however, was an important leap forward for western business:

> It is proposed that certain men together are members sharing unequally in a company which will make some quantity of profit they will wish to divide moreover among themselves according to their shares.

Fibonacci thus ushered in the concept of owning a partial share of a company, in so far as a fair treatment of the profits was concerned, spending the entire chapter on how to use the Hindu-Arabic numbers to equally divide the company's profits.

2.3 The Time Value of Money

Although the time value of money was known since antiquity, Fibonacci understood its fundamental importance to finance, introducing the concept of present value analysis.

Fibonacci poses the following problem:

> A certain man proceeded to Lucca on business to make a profit doubled his money, and he spent there 12 denari. He then left and went through Florence; he there doubled his money, and spent 12 denari. Then he returned to Pisa, doubled his money, and it is proposed that he had nothing left. It is sought how much he had at the beginning.

Note that it is not clear in the problem formulation that the man also pays 12 denari in Pisa at the end of the trip. This is easily solvable by finding x in the equation

$$2(2(2x - 12) - 12) - 12 = 0 \tag{1}$$

However, Fibonacci has a more elegant solution. He notes that the capital doubles with each trip and therefore the initial capital will have to be sufficient such that it grows into the periodic payment in each city. The amount at the beginning – given the man is left with nothing at the end – is the periodic payment "discounted" by the rate at which it grows. There are three "parts" of the initial capital, the amount that grows into the payment in Lucca, Florence, and Pisa, each requiring a subsequent halving

$$x = \frac{12}{2} + \frac{12}{2^2} + \frac{12}{2^3} \tag{2}$$

giving the final answer 10.5 denari. The concept of time here is replaced by the physical movement of the man, the great insight that Fibonacci introduces in this problem is the concept of net present value or NPV. The traveling merchant problem describes an annuity which is a financial security that exchanges a fixed amount of money for a finite number of future fixed periodic payments. In modern parlance, the question would be posed as "what is the present value of a three period annuity that pays a 12 denari coupon at an interest rate of 100%?" The general formula the present value of all future cashflow payments, discounted at the quoted interest rate

$$A = \sum_{n=1}^{N} \frac{c}{(1+r)^n} \tag{3}$$

which can be simplified using the infinite $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ resulting in the closed form formula for an annuity.

$$A = c\left(1 - \frac{1}{(1+r)^N}\right). \tag{4}$$

Substituting $r = 1$, $c = 12$ denari and $N = 3$ leads to the same 10.5 denari as above.

The traveling merchant introduces the concept of NPV; however, Fibonacci goes further and applies it to the analysis of two streams of cashflows, ranking their value based on the higher NPV when he poses the following problem:

> A certain soldier because of his fief received from a certain king 300 bezants each year, and it is satisfied in four payments, and in each payment he takes 75 bezants; this is a payment for three months which by necessity is collected together; he asks for certain compensation in order to accommodate himself for interest because he accepts the 300 bezants instead of the 75 bezants of each payment, namely from payment to payment, of the capital and profit. Voluntarily acquiescing to this, he invests the bezants at a profit of two bezants per hundred in each month. It is sought how many bezants he makes in his investment.

Essentially, the king has decided to change the arrangement from 75 bezants every 3 months to 300 bezants annually, going from periodic payments to a lump sum. The soldier would like to be compensated for this change because he can earn interest quarterly, the question asks what is the difference in NPV of the two streams of cashflows. To solve this, note that the rate at which the soldier can earn is 6%, there are four periods and the payment is 75 bezants. Inserting into the annuity formula above reveals that this stream of cashflows is worth 259.88 bezants, whereas the NPV of a single lump sum payment is 237.62 bezants. Therefore the soldier should ask for an extra 22.25 bezants today, or that the lump sum be augmented by 28.09 bezants (since the NPV of 328.09 is 259.88).

3 1900, 1951 Brownian Motion and Stochastic Calculus: Louis Bachelier and Kiyosi Itō

Some say that March 29, 1900 should be recognized as the birthday of quantitative finance; it was the day that the French mathematician Louis Bachelier – a student of Henri Poincaré – defended his thesis at Sorbonne. The thesis, later published in one of the most influential French scientific journals *Annales scientifiques de l'École Normale Supérieure*, was called *Théorie de la Spéculation* (Bachelier 1900) and contained within it the first use of stochastic processes – although implicitly – to describe Brownian motions, which he applied to the pricing of stock options.

The calculus and analysis required to deal with stochastic processes were not fully developed into its modern form until Weiner and Itō in the 1950s. Stochastic calculus is the application of integral calculus to processes that are nowhere differentiable but everywhere continuous, a class of functions not thought to exist prior to 1872. That year, Karl Weierstrass introduced the pathological function

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x) \tag{5}$$

where $0 < a < 1$, b is a positive odd integer, and $ab > 1 + \frac{3\pi}{2}$ (Weierstrass 1872). Prior to the introduction of this function, it was believed, almost axiomatically, that if a function was continuous, it could only be nondifferentiable at a finite set of points, perhaps because there was a "proof" in one of André-Marie Ampère's papers – although this point is debated by some. These everywhere continuous but nowhere differentiable functions were met with disdain from the mathematical community, with Poincaré calling these functions "monsters" and Charles Hermite "a lamentable scourge." In fact, some mathematical proofs were upended as a result of the existence of such functions, and this leads to the careful treatment of all concepts and the birth of modern analysis (Gray 2015).

However, there are processes in both physics and finance that could be described by precisely such a function. In 1827, botanist Robert Brown observed both organic and inorganic particles under a microscope undergoing "jittery" motion, now named Brownian motion in his honor. The explanation for this behavior was not explained until 1905 when Albert Einstein wrote a paper attributing the phenomenon to the atomic theory of matter (Einstein et al. 1905) which was surprisingly still not generally accepted at the time. Bachelier realized that stock prices can be described by such processes.

A Brownian motion is a continuous function of time $W(t)$ (denoted W since it is known as a Weiner process (Wiener 1923)) starting at 0. The process up to time T is built up by realizations of normal random draws

$$W(T) = \sum_{i=0}^{N} [W(t_{i+1}) - W(t_i)] \tag{6}$$

where the increments are independent and identically distributed according to the normal distribution with mean 0 and variance Δt

$$W(t + \Delta t) - W(t) \sim \mathcal{N}(0, \Delta t). \tag{7}$$

These processes are martingales which means their expectation in the future is equal to the value today. This was embedded in Bachelier's thesis as the "Principle of Mathematical Expectation":

> *The mathematical expectation of a speculator is nil.* The generality of this principle needs to be appreciated: it signifies that the market, at a given instant, considers as having nil expectation not only the current trading operations, but also those that would be based on a subsequent movement of prices. The mathematical expectation of an operation that may be either positive or negative, if it produces a price movement, is *a priori* nil.

Mathematically this is shown by taking expectations

$$\mathbf{E}[W(t) \,|\, \mathcal{F}_s] = \mathbf{E}[W(t) - W(s) \,|\, \mathcal{F}_s] + \mathbf{E}[W(s) \,|\, \mathcal{F}_s] \tag{8}$$

where the $\mathcal{F}_s$ is technically a filtration, meaning "all of the information known up to the point s." Since this includes the value of $W(s)$, then this is a trivial expectation, and the first expectation is zero based on the property of Brownian motions. Therefore

$$\mathbf{E}[W(t)\,|\,\mathcal{F}_s] = W(s) \tag{9}$$

which is the definition of a martingale (and Bachelier's principle of mathematic expectation).

Stochastic processes differ from deterministic processes by virtue of the fact that they have nonzero quadratic variation, and responsible for the necessity of a new type of calculus in a stochastic setting. The quadratic variation of any process is defined as

$$\langle f,f\rangle[T] \equiv \lim_{\Delta t\to 0}\sum_{i=0}^{N-1}[f(t_{i+1}) - f(t_i)]^2 \tag{10}$$

where $t_i = \frac{i}{N}T$ (this can be relaxed and a nonuniform division of the time interval can be used, as long as the longest subinterval tends to zero). For differentiable functions, the mean value theorem can be used to simplify the summand, since there exists a point t_i^* such that $f'\left(t_i^*\right)(t_{i+1} - t_i) = f(t_{i+1}) - f(t_i)$

$$\langle f,f\rangle[T] = \lim_{\Delta t\to 0}\sum_{i=0}^{N-1}\left|f'\left(t_i^*\right)\right|^2[t_{i+1} - t_i]^2 \tag{11}$$

$$\langle f,f\rangle[T] \le \lim_{\Delta t\to 0}\Delta t\sum_{i=0}^{N-1}\left|f'\left(t_i^*\right)\right|^2[t_{i+1} - t_i].$$

In the limit of vanishing Δt, this expression becomes a Riemann sum

$$\langle f,f\rangle[T] \le \lim_{\Delta t\to 0}\Delta t\int_0^T |f'(t)|^2 dt = 0, \tag{12}$$

however still multiplied by Δt and hence vanishes. Therefore any function that is differentiable almost everywhere (it could be nondifferentiable at a countable number of points, which have mass zero and therefore do not alter the argument) has vanishing quadratic variation.

For Brownian motions, this argument does not hold as they are nowhere differentiable. The resulting quadratic variation can be calculated probabilistically by determining its mean and variance

$$\mu_{\mathrm{QV}} \equiv \mathbf{E}[\langle W,W\rangle[T]] = \lim_{\Delta t\to 0}\sum_{i=0}^{N-1}\mathbf{E}\left[(W(t_{i+1}) - W(t_i))^2\right], \tag{13}$$

the summand is the definition of variance and by 7 is equal to $t_{i+1} - t_i$. Therefore, the average quadratic variation of a Brownian motion is

$$\mu_{\mathrm{QV}} = \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} (t_{i+1} - t_i) = T. \tag{14}$$

The variance of the quadratic variation can be calculated by use of the fourth-order moment of the normal distribution

$$V_{\mathrm{QV}} = \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} 2(t_{i+1} - t_i)^2. \tag{15}$$

As above, this quantity is strictly less than the product of Δt times a Riemann integral giving the total time T, which vanishes as $\Delta t \to 0$. Since the variance is zero, the quadratic variation is not a random variable and is equal to the mean

$$\langle W, W \rangle [T] = T \tag{16}$$

"almost surely," which means that there are outcomes in the probability space where the quadratic variation is not equal to this mean; however, they comprise a set of measure zero, that is the probability of this happening is zero.

Tacitly, Bachelier identified the value of a stock with the Brownian motion $W(t)$, and therefore price movements with the increments of a Brownian motion $W(T) - W(t)$ (later Fisher Black and Myron Scholes will identify the returns of the stock with the increments of this process). Suppose an investor held $\Delta(t)$ shares of that stock, then the value is the product $\Delta(t)W(t)$, and has an amount of money $M(t)$ in an interest-free bank account. Their total portfolio is worth

$$V(t) = \Delta(t)W(t) + M(t). \tag{17}$$

At some future time $T > t$, they change the number of shares to $\Delta(T)$ by using the money available in the bank account, meaning $M(T) = M(t) - (\Delta(T) - \Delta(t))W(T)$ (this is known as a self-financing portfolio, where there is no cash injection or withdrawn, the future trades are funded solely by the movements of the securities in the portfolio). The change in the portfolio value can then be written to be

$$V(T) - V(t) = \Delta(t)\,(W(T) - W(t)). \tag{18}$$

Trading at infinitesimally spaced times t_i produces the result

$$V(T) - V(t) = \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)], \tag{19}$$

which may seem impossible in practice; however, now the Eq. 19 will yield to a similar analysis as the quadratic variation above. Equation 19 is suggestive of an integral, however not of the usual Riemann nor Lebesgue form, this is a new type called Itō integration for its discoverer Kiyoiso Itō between 1944 and 1951 (Itō 1944).

Symbolically an Itō integral is written

$$V(T) - V(t) = \int_t^T \Delta(s) dW(s) \tag{20}$$

even though there is no notion of a differential of a Brownian motion. Writing the integral is merely a short form, and the underlying limit in Eq. 19 is always implied (however the same can be said for Riemann integral). By letting T be infinitesimally greater than t, one can write

$$dV(t) = \Delta(t) dW(t) \tag{21}$$

which is known as a stochastic differential equation (SDE) (Itō 1951b; Oksendal 1992; Shreve 2004).

Itō integrals have zero mean, and quadratic variation given by

$$\langle V, V \rangle [T] = \int_0^T \Delta^2(s)\, ds \tag{22}$$

which can be shown with an argument closely following the quadratic variation proof for Brownian motions.

As previously mentioned, Black and Scholes took the return of the stock to be the normally distributed variable, so it is natural to ask how a function of a Brownian motion changes over time, such as $S(T) = S(t) \exp(W(T) - W(t))$. The first step is to break the time interval into infinitesimal intervals

$$S(W(T)) - S(W(t)) = \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} [S(W(t_{i+1}) - S(W(t_i))], \tag{23}$$

when Taylor expanding the term in bracket, the second-order term will not vanish due to the quadratic variation of the Brownian motion

$$\begin{aligned} S(W(T)) - S(W(t)) = {} & \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} S'(W(t_{i+1}))[W(t_{i+1}) - W(t_i)] \\ & + \frac{1}{2} \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} S''(W(t_{i+1}))[W(t_{i+1}) - W(t_i)]^2. \end{aligned} \tag{24}$$

The first term converges to an Itō integral, whereas care needs to be taken with the second. When the partition is fine enough that the function $S(W(t))$ does not vary significantly, then introducing another partition on each subinterval $[t_{i+1}, t_i]$ leads to

$$S(W(T)) - S(W(t)) = \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} S'(W(t_{i+1}))[W(t_{i+1}) - W(t_i)] + \frac{1}{2} \lim_{\Delta t \to 0} \sum_{i=0}^{N-1} S''(W(t_{i+1}))[t_{i+1} - t_i]. \tag{25}$$

where the second term now converges to a Riemann integral with ultimate result

$$S(W(T)) - S(W(t)) = \int_t^T S'(W(s))dW(s) + \frac{1}{2}\int_t^T S''(W(s))ds. \tag{26}$$

The differential form can be written with the understanding that it is symbolic only and always represents the integral form. The differential form

$$dS(W(t)) = S'(W(t))dW(t) + \frac{1}{2}S''(W(t))dt \tag{27}$$

is known as Itō's formula (sometimes referred to as Itō's Lemma) (Itō 1951a).

Itō's formula can be iterated, for instance a call option will depend on a stock price, which in turn will be a function a process defined by a stochastic differential equation (SDE). For example, the most general SDE for a stock is

$$dS(t) = \mu(t, S(t))dt + \sigma(t, S(t))dW(t) \tag{28}$$

and Itō's formula for a function of this stock $f(S(t))$ is

$$df(S(t)) = \frac{\partial f(S(t))}{\partial S} dS(t) + \frac{1}{2} \frac{\partial^2 f(S(t)}{\partial S^2} d\langle S, S\rangle(t). \tag{29}$$

Bachelier implicitly chose the simplest possible SDE to describe the stock

$$dS(t) = \sigma dW(t) \tag{30}$$

that its price changes are normally distributed with variance $\sigma^2(T - t)$.

In this simple case, the SDE (30) can be solved simply

$$S(T) = S(t) + \sigma[W(T) - W(t)], \tag{31}$$

and the price of a call option is simply the expectation of the hockey stick payoff of the stock price minus the strike

$$C(t, T, S(t), K) = \mathbf{E}[\max(S(T) - K, 0) \,|\, \mathcal{F}_t] \tag{32}$$

$$= \mathbf{E}[\max(S + \sigma[W(T) - W(t)] - K, 0) \,|\, \mathcal{F}_t]. \tag{33}$$

The Brownian increment is normally distributed with zero mean and variance $T - t$, and the expectation can be explicitly written as an integral

$$C(t,T,S(t),K) = \int_{-\infty}^{\infty} \max(S(t) + \sigma z - K, 0) \exp\left(\frac{-z^2}{2(T-t)}\right) \frac{dz}{\sqrt{2\pi(T-t)}} \quad (34)$$

leading to the final solution for the price of a stock option in the Bachelier model

$$C(t,T,S(t),K) = (S(t) - K)N(d_N) + \sigma\sqrt{T}n(d_N) \quad (35)$$

where $N(x)$, $n(x)$ are the cumulative normal density function and probability density function respectively and $d_N = \frac{S(t)-K}{\sigma\sqrt{T}}$.

4 1952 The Efficient Frontier and Matrix Algebra: Harry Markowitz

Prior to the 1950s, mathematics had little to do with asset selection, as investors tended to look at securities in isolation, as opposed to the portfolio as a whole. The fact that diversification was good had been known for a long time, evidence for diversification of business show up in Shakespeare's "The Merchant of Venice":

My ventures are not in one bottom trusted,
Nor to one place; nor is my whole estate
Upon the fortune of this present year:
Therefore my merchandise makes me not sad.

However, it was thought that diversification could reduce the risk of the portfolio completely, whereas Markowitz showed in the Nobel winning paper "Portfolio Selection" (Markowitz 1952) that this was not the case. An investor can reduce the nonsystematic risk inherent in individual stocks; however, they are still left with systematic risk present in the market as a whole. Markowitz showed that correlation between stocks can be used to create a portfolio that maximizes return for a given level of risk.

First, Markowitz defined "risk" as the standard deviation of the portfolio. It is instructive to look at a portfolio containing two securities to build up intuition before tackling the general case. Suppose there is a given amount of wealth that is to be distributed between these two securities. The amount in the first share is $w_1 = w$ and the amount in the second is $w_2 = 1 - w$. The expected return is additive

$$R_p = \mathbf{E}[w_1 R_1 + w_2 R_2] = w_1 \mathbf{E}[R_1] + w_2 \mathbf{E}[R_2] \quad (36)$$

however the risk is not additive

$$\sigma_p = \sqrt{\mathrm{Var}(w_1 R_1 + w_2 R_2)} \equiv \sqrt{\mathbf{E}\left[(w_1 R_1 + w_2 R_2)^2\right] - R_p^2} \quad (37)$$

$$= \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2(\mathbf{E}[R_1R_2] - \mathbf{E}[R_1]\mathbf{E}[R_2])}. \tag{38}$$

Assuming normally distributed returns, the cross term gives rise to a correlation term, and therefore the risk of the portfolio can be written

$$\sigma_p = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12}}. \tag{39}$$

Given the variances and correlations can be estimated from historical (observable) market prices, this equation relates the risk of the portfolio to w, the relative weight between the two securities. The risk-reward trade-off for this portfolio can be plotted with the risk along the abscissa and return along the ordinate. Allowable portfolios lie along the hyperbola connecting $(w_1, w_2) = (1, 0)$ and $(w_1, w_2) = (0, 1)$ corresponding to the entire portfolio in the first and second security respectively. This behavior is shown in Fig. 1. Note that this is a multivalued function, implying there can be two portfolios that have the same risk, but one has a higher return than the other. Markowitz's key insight was that no investor would (or should) purchase the portfolio with a lower return given the same amount of risk. He labeled these portfolios as *inadmissible*. Therefore in Fig. 1, the lower branch of the hyperbola are all inadmissible portfolios, and he defined the admissible portfolios as the efficient frontier. He reasoned that how an investor will choose their individual portfolio depends on where their individual utility function (a hypothetical function that, in this case, quantifies the investor's risk tolerance, i.e., are they risk adverse or risk seeking) touches the efficient frontier.

There are two special cases, which can be seen in the graph. If the two securities are perfectly correlated $\rho = 1$ (the lightest line in Fig. 1), then there are no diversification benefits, as can be seen by the straight line connecting the two extremes (only holding one of the securities), and all intermediate portfolios are admissible and the choice will be made solely by the amount of risk the investor is

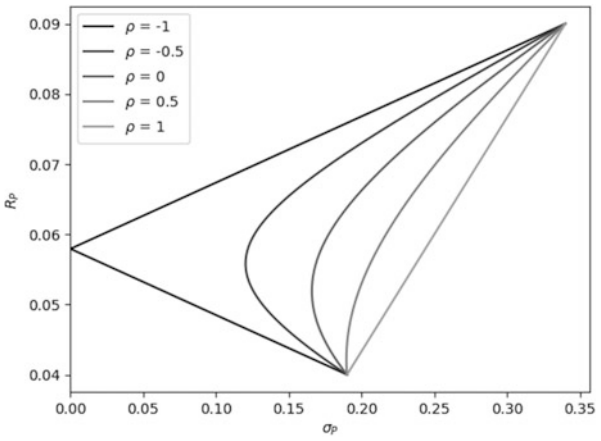


Fig. 1 Risk return profiles for a portfolio of two securities, with expected return [4%, 9%] and historical standard deviation [19%, 34%]. The correlation is varied to determine how it affects overall portfolio

willing to take. The second special case is when the securities are perfectly anti-correlated $\rho = -1$, which is the darkest line in Fig. 1. In this case, the total risk in the portfolio can be eliminated; however, this situation does not occur in reality.

Realistic portfolios contain more than two securities, sometimes upward of hundreds. The expected return of a portfolio with N securities is simply

$$R_P = \sum_{i=1}^{N} w_i \mathbf{E}[R_i] \equiv \mathbf{w} \cdot \mathbf{R} \tag{40}$$

and risk

$$\sigma_P = \sqrt{\mathbf{w}^T \Sigma \mathbf{w}} \tag{41}$$

where $\mathbf{w}$ is a vector of portfolio weights whose entries sum to one $\sum_{i=0}^{N} w_i = 1$, and critically Σ is the covariance matrix

$$\Sigma_{ij} = \mathbf{E}\left[\left(R_i - \mathbf{E}[R_i]\right)\left(R_j - \mathbf{E}\left[R_j\right]\right)\right] \tag{42}$$

whose diagonal is the vector of variances of the returns, and off diagonals are the covariance of the returns. The surface cannot be visualized for $N > 3$; however, the risk-return profile can still be plotted on the same axes as the $N = 2$ case above. In this case, there is not simple parametric form for the weights in the portfolio, the return and risk must be determined by sampling the hypersurface defined by the constraint

$$\sum_{i=1}^{N} w_i = 1 \tag{43}$$

Figure 2 demonstrates the efficient frontier for $N = 10$ stocks of varying risk and expected return.

The analysis of securities in this manner became known as modern portfolio theory (MPT).

4.1 The Capital Asset Pricing Model

Even though MPT was a great leap forward, it required the constant calculation of all variances and correlations in the market – a daunting task for the computation power available in the 1950s. In fact, the work of Markowitz did not permeate the investment community for about a decade, when William Sharpe and John Lintner furthered his ideas and developed the capital asset pricing model (CAPM) around 1964 (Sharpe 1964; Lintner 1965). The idea is based on the realization that not all risks in the market are equivalent. There is idiosyncratic risk, which can be hedged away from a portfolio as demonstrated by Markowitz, and systematic risk, which is

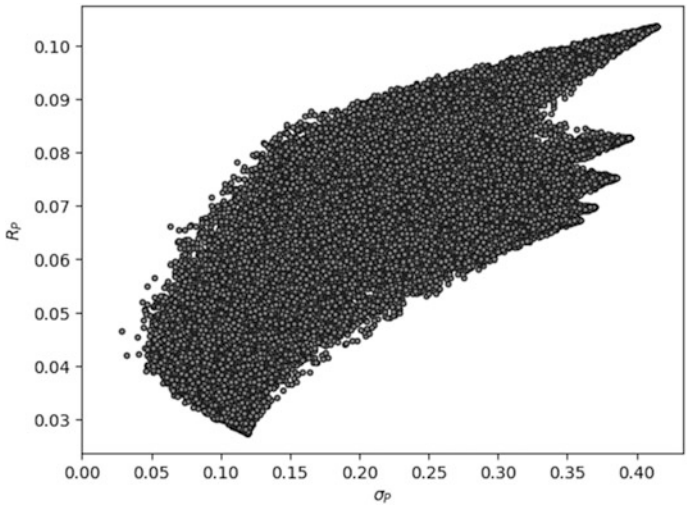


Fig. 2 The efficient frontier of a portfolio of $N = 10$ securities with expected returns [2.74%,4.15%,4.54%,6.57%,6.73%,6.74%,6.97%,7.53%,8.28%,10.38%] and standard deviations [11.95%,15.82%,16.07%,21.19%,35.30%,35.92%,36.98%,38.53%,39.54%,41.48%], using 50,000 sample portfolios, with random correlations between -1 and 1

the risk inherent in the market itself – a risk still present in a truly diversified portfolio. The first is associated with individual company idiosyncrasies, such as the CEO unexpectedly stepping down, or a drug not being approved for mass production. Taking on this type of risk should not be rewarded, since it can be eliminated from a portfolio.

Sharpe realized that there is another asset in the economy that has no risk whatsoever, the treasury bond. A combination of the risk-free asset and a specific portfolio on the efficient frontier will render all other portfolios inadmissible. This market portfolio is called the tangency portfolio, for reasons that will become clear. If the risk-free asset has expected return R_f and the tangency portfolio has return R_t then the expected return of the combined portfolio (investing the fraction ω of the total wealth in the risk-free bond and $(1 - \omega)$ in the tangency portfolio) is simply

$$R_P = R_f + \omega\big(R_t - R_f\big). \tag{44}$$

Sharpe called this the capital market line, the expected return of a portfolio containing both risk-free and risky assets.

> In effect, the market presents him with two prices: the price of time, or the pure interest rate . . . and the price of risk, the additional expected return per unit of risk borne (the. . .slope of the line).

The method of determining the tangency portfolio, as suggested by its name, is to determine the point at which a capital market line is tangent to the efficient frontier.

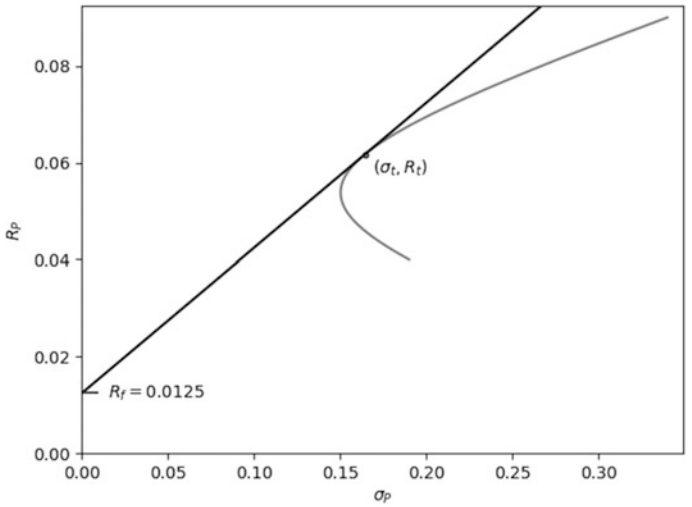


Fig. 3 The efficient frontier of a two security portfolio (expected returns [4%,9%] and standard deviations [19%,34%] with correlation −0.2) together with the tangency capital market line, where the risk-free asset has an expected return of 1.25%. Points past the tangency portfolio represent selling the risk-free asset short (in other words, borrowing money at an interest rate equal to the risk-free rate) and investing more than the investors total wealth in the market portfolio

This combination dominates all other portfolios, in that it has a higher expected return for the given amount of risk, and is shown in Fig. 3.

Sharpe defined the slope of this line as the "market price of risk," and today the ratio of a security's excess return (return above the risk-free rate) to its standard deviation $\frac{R_i - R_f}{\sigma_i}$ is called the Sharpe ratio. The Sharpe ratio can be used to determine the relative expected return per unit of risk for a given portfolio, the higher the Sharpe ratio, the more return an investor expects per unit of risk. The maximum of the Sharpe ratio for all portfolios on the efficient frontier is precisely the tangency portfolio.

In principle, every investor in the market would choose the same tangency portfolio, and for this reason it is also called the market portfolio, and can be thought of as being represented by a broad market index such as the Dow-Jones industrial average or the S&P 500.

So far, nothing has been said about how the individual security's return itself. The CAPM assumes that the Sharpe ratio of an individual security is equal to the market portfolio times the correlation of the security to that portfolio, which can be rearranged to give the security market line (SML)

$$R_i = R_f + \beta\left(R_M - R_f\right) \tag{45}$$

where $\beta = \frac{\rho\sigma_i}{\sigma_M}$. Therefore the excess return of an individual security is equal to the excess return of the market multiplied by the product of the ratio of standard

deviations and the correlation. This can be greater than 1.0 when the correlation is high and the standard deviation of the security is larger than the market, implying that this security's movements will be "amplified" as compared to the market. A $\beta < 1$ implies the opposite, that the security is more subdued than the market.

One could think of the β as a regression coefficient between the returns of the market portfolio and that of the security. In this sense, the CAPM is known as a "one-factor" model, as it ascribes all of the excess return of a security to the market. Fama and French (Fama and French 2004) extend this regression analysis to three factors

$$R_i = R_f + \beta_{iM}\left(R_M - R_f\right) + \beta_{iS}\mathbf{E}[\mathrm{SMB}] + \beta_{ih}\mathbf{E}[\mathrm{HML}] \tag{46}$$

where SMB (small minus big) is the difference between the returns on diversified portfolios of small and big securities (as defined by their market capitalization, or the total value of the shares outstanding) and HML (high minus low) is the difference between the returns on diversified portfolios with high and low book-to-market ratios – the ratio of the "book value" (accounting assets minus liabilities) and its market capitalization (number of shares outstanding multiplied by share price, which gives an estimation of how valuable the market perceives the company to be). This three-factor model performs quite well (Fama and French 2004), and this framework is the basis of modern models used in risk management.

5 1973 Option Pricing: Fischer Black, Myron Scholes, and Robert Merton

Prior to the seminal paper of Black and Scholes (Black and Scholes 1973; Merton 1973), the connection between option prices and differential equations was well known (Samuelson 1965; Sprenkle 1961; Thorp and Kassouf 1967). However, these predecessors typically have unknown parameters, such as the expected return of the underlying security, or a parameter quantifying the investor's risk tolerance, which the authors either attempt to estimate from market data or to derive from other models, such as the CAPM. Occasionally, the formulas were incredibly close to the Black-Scholes, as pointed out by Black and Scholes in their original paper.

Black and Scholes added a key insight based on the general equilibrium behavior of a continuously hedged portfolio with the following argument. Suppose an investor has a portfolio consisting of the option $C(t, S)$ (for ease of notation, the dependence on the maturity T and strike K have been suppressed) plus a short position in an uncertain amount of shares in the stock $\Delta(t)S(t)$

$$V(t) = C(t, S) - \Delta(t)S(t). \tag{47}$$

The investor will continuously change $\Delta(t)$ in order to keep this position instantaneously hedged. By Itō's formula, the change in the portfolio is given by the SDE

$$dV(t) = \frac{\partial C(t,S)}{\partial t}dt + \left(\frac{\partial C(t,S)}{\partial S} - \Delta(t)\right)dS(t) + \frac{1}{2}\frac{\partial^2 C(t,S)}{\partial S^2}d\langle S,S\rangle(t) \quad (48)$$

The distributional assumptions of $S(t)$ are encapsulated in the second condition, which states "the stock price follows a random walk in continuous time with a variance rate proportional to the square of the stock price. Thus the distribution of possible stock prices at the end of any finite interval is lognormal. The variance of the rate of the return on the stock is constant." Today this statement is encapsulated in the SDE

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \quad (49)$$

for some (unknown) drift μ of the underlying stock. The quadratic variation of this process is $S^2(t)\sigma^2 t$ and therefore can be used to simplify Eq. 48

$$dV(t) = \left(\frac{\partial C(t,S)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C(t,S)}{\partial S^2}\right)dt + \mu S\left(\frac{\partial C(t,S)}{\partial S} - \Delta(t)\right)dS(t) \quad (50)$$

the last term can be eliminated by taking $\Delta(t) = \frac{\partial C(t,S)}{\partial S}$, and therefore the stochastic part of the portfolio is gone, making this portfolio instantaneously hedged. This implies that the derivative of the option price is the correct number of shares to short in order to perfectly hedge the portfolio, which makes intuitive sense, and was actually posited in the Black-Scholes paper as opposed to derived.

The big insight that Black and Scholes had was that this portfolio is now economically equivalent to a money market account, and therefore by the law of one price, must accumulate interest at the instantaneous rate of interest $dV(t) = rV(t)dt = r(C(t, S) - \Delta(t))$, otherwise an investor could go long an option and short this hedged portfolio and make riskless profits at greater than the risk-free rate. This violates the law of no-arbitrage and therefore the option must have the same value as this hedged portfolio. By equating these and simplifying, the value of the hedged portfolio must be the solution of the celebrated Black-Scholes partial differential equation (PDE)

$$\frac{\partial C(t,S)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C(t,S)}{\partial S^2} + rS\frac{\partial C(t,S)}{\partial S} = rC(t,S). \quad (51)$$

Note that the above treatment is not precisely true, as Shreve notes in Exercise 4.10 in Shreve (2004) as $\Delta(t)$ is also a stochastic process and its quadratic variation and covariation with $S(t)$ need to be taken into account. Once this is done, the resulting equation is the same, and this simpler argument is presented here for its clarity of exposition, as well as being true to the original derivation. The argument is cleared up by Merton, which will be presented below.

Black and Scholes state that the solution of this PDE is the celebrated Black-Scholes formula

$$C(t,S) = SN(d_+) - Ke^{-r(T-t)}N(d_-), \tag{52}$$

where $N(x)$ is the standard normal cumulative density and

$$d\pm = \frac{\ln \frac{S}{K} \pm \left(r - \frac{1}{2}\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}. \tag{53}$$

The hedging argument in Black and Scholes (1973) is then followed by another derivation using CAPM described above. As explained by Merton (Merton 1973)

> . . .the portfolio weights are chosen to eliminate all "market risk." By the assumptions of the capital asset pricing model, any portfolio with a zero ("beta") market risk must have an expected return equal to the risk-free rate. Hence, an equilibrium condition is established between the expected return on the option, the expected return on the stock, and the risk-less rate.

Merton was initially skeptical, and therefore set out to relax some of the conditions set forth in the Black and Scholes paper. He added stochastic interest rates, dividends, and considered both puts and American options. However the most important insight in the paper was noting that the Black-Scholes formula implied that investors could disagree over the expected return of the stock but still agree on the price of an option, a characteristic not present in any previous option formula. For instance in Sprenkle (1961), the option price is tantalizingly close to the BS

$$C(t,S) = aSN(d_+) + (1 - P_e)KN(d_-) \tag{54}$$

$$d_\pm = \frac{\ln \frac{aS}{K} \pm \left(r - \frac{1}{2}\sigma^2(T-t)\right)}{\sigma\sqrt{T-t}}$$

where $a = \mathbf{E}\left[\frac{S(T)}{S(t)}\right]$, the investors expectation of the ratio of the stock price at expiry to the stock price today, and P_e quantified the risk-seeking behavior of the investor, related to how much the investor was willing to pay for leverage/financial risk.

Merton realized that despite the appeal of this characteristic, it must be understood how it comes about without appealing to homogeneous expectations explicit in the CAPM. Merton only assumed that the investors agreed on the variances and covariances of the stochastic stock and bond processes, and not their rates of return.

> This assumption is much more acceptable than the usual homogeneous expectations. It is quite reasonable to expect that investors may have quite different levels of information for current (and future) expected returns due to different levels of information, techniques of analysis, etc. However most analysts calculate estimates of variances and covariances in the same way: namely, by using previous price data. Since all have access to the same price history, it is also reasonable to assume that their variance-covariance estimates may be the same. (Merton 1973)

Merton then constructs a portfolio of the stock, the option, and a bond that matures at the expiry of the option. When this portfolio is made to be instantaneously riskless, a system of equations results that conspire to remove these expectations from the PDE. This concept would later be generalized into the fundamental theorem of asset pricing, formulated by Harrison and Pliska (1981).

This theorem connects arbitrage freedom with complete markets and martingale measures. A complete market is one in which there are enough securities to construct hedging portfolios for all possible contingent claims (otherwise known as derivatives) in the market. In the Black-Scholes discussion above, the option could be hedged exactly because it shares the same source of risk as the underlying stock. So a complete market has enough "sources of risk" to hedge the risks of all traded derivatives.

The fundamental theorem of asset pricing states that if a market is complete then there is no arbitrage and there exists a probability measure under which a discounted security price is a martingale, which is the probabilistic measure theory terminology for Bachelier's "principal of mathematical expectation," namely that the expectation of the future quantity is today's value. Therefore the fundamental theorem of asset pricing states that the discounted future value of the portfolio is a martingale

$$D(0,t)V(t) = \mathbf{E}[D(0,T)V(T) \mid \mathcal{F}_t]. \tag{55}$$

The probability measure under which this expectation is taken is known as the risk neutral measure.

While this is mathematically true, and extremely useful in finance, some disagree that this should be called a fundamental theorem.

> Finances objects of interest markets, money, assets, and securities are also abstractions. The aim of finance is finding true relationships between the abstractions they represent, and hence between the realities they should be pinned to. In that sense, like physics, finance cannot be a branch of mathematics. Finance could conceivably have a fundamental law *Money tends to flow into the hands of people who will let it reproduce most rapidly* might be one but it surely cannot have a fundamental theorem, for the world is not conditional. Derman (2013)

6 October 19, 1993

If there is 1 day that changed the future of quant finance, and therefore by consequence, the way in which capital gets allocated on Wall street, it is October 19, 1993. This is the day that the US Congress canceled funding for the Texas Superconducting Super Collider (SSC).

The SSC was a particle accelerator designed to probe energies higher than Large Hadron Collider at CERN and therefore explore physics at a smaller scale. The goal was noble, to determine whether the standard model of particle physics – the most successful physical theory ever conceived by humanity – was correct. It was

approved in May 1990 and construction began in early 1993 and thousands of academic positions were created and filled.

Once the project was canceled by Congress, these thousands of professional physicists were released into an academic marketplace that did not have the capacity to adsorb this magnitude of people. This army of people with advanced degrees started looking elsewhere to apply their skills. This was over 20 years after the Black-Scholes seminal paper and when options first started trading on exchanges at the Chicago Board of Options Exchange (CBOE) so there was definitely an understanding that mathematicians and physicists could add value to Wall Street firms and banks. And hence the first great influx of nonfinance PhDs migrated to Wall Street.

7 1947 Modern Techniques of Quantitative Finance: The Feynman-Kac Theorem

This chapter will now pivot from a historical perspective of the mathematics ideas that altered the way capital was allocated to the modern mathematical techniques used in quantitative finance, which will help explain why applied mathematicians and physicists find it a natural home. There is a deep link between stochastic calculus, Monte Carlo, and partial differential equations, all three being welded together in what is known as the Feynman-Kac Theorem.

At Cornell University in 1947, physicist Richard Feynman presented results from his thesis on the path integral approach to quantum mechanics, work for which he later shared the 1965 Nobel prize. Mark Kac, a Polish American mathematician was in the audience and approached Feynman afterwards remarking that the two of them were working on the same problem from a differing point of view. After a few days of calculation, Kac had connected his probabilistic approach to the Feynman path integral approach, thus discovering the Feynman-Kac formula. Kac was impressed by Feynman, stating "It is only fair to say that I had Wiener's shoulders to stand on. Feynman, as in everything else he has done, stood on his own, a trick of intellectual contortion that he alone is capable of."

The Feynman-Kac formula relates expectation values to the solutions of partial differential equations (PDEs) and thus gives the connection between random stock movements and the heat equation (Kac 1949). On one side is an expectation value – which can be calculated using Monte Carlo methods – whereas on the other is a PDE, whose solution is the expected value.

Although very difficult to prove, the formula can be motivated by topics already covered in this chapter. Suppose a random variable obeys the SDE

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dW(t), \tag{56}$$

which is Markov as can be seen by its integral form

$$X(T) = X(t) + \int_t^T \mu(X(s), s)ds + \int_t^T \sigma(X(s), s)dW(s) \tag{57}$$

since it only depends on the initial point $X(t)$ and not by any of the previous values.

Consider today's expectation of the time T value of a function of the stochastic variable

$$C(t,X(t)) = \mathbf{E}[\phi(X(T))\,|\,\mathcal{F}_t]. \tag{58}$$

This too will be Markov since it also has a SDE by Itō's formula, namely

$$dC(t,X(t)) = \left(\frac{\partial C(t,X(t))}{\partial t} + \mu(X(t),t)\frac{\partial C(t,X(t))}{\partial x} + \frac{1}{2}\frac{\partial^2 C(t,X(t))}{\partial x^2}\sigma^2(X(t),t)\right)dt + \frac{\partial C(t,X(t))}{\partial x}\mu(X(t),t)dW(t) \tag{59}$$

and therefore the expectation satisfies the Chapman-Kolmogorov equation. This equation states that, for Markov processes, the transition probability $\rho(xztT)$, which gives the probability that the random variable starting at $X(t) = x$ at time t transitions to the value $X(t) = z$ at time $T > t$ can be calculated by introducing an intermediate time s, where $t < s < T$, and summing over all possible intermediate states

$$\rho(xztT) = \int_{-\infty}^{\infty} \rho(xyts)\rho(yzsT)dz \tag{60}$$

By writing the expectation 58 as an integral, and introducing the Chapman-Kolmogorov equation (and letting $y = X(s)$)

$$C(t,x) = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\phi(z)\rho(yzsT)dz\right]\rho(xyts)dy. \tag{61}$$

The term inside the square brackets is precisely the expectation at the intermediary time $C(s,X(s)) = \mathbf{E}[\phi(X(T))\,|\,\mathcal{F}_s]$ and therefore the expectation itself is a martingale

$$C(t,X(t)) = \mathbf{E}[C(s,X(s))\,|\,\mathcal{F}_t]. \tag{62}$$

Since it is a Martingale, the drift term in Eq. 59 must vanish, therefore

$$\frac{\partial C(t,x)}{\partial t} + \mu(X(t),t)\frac{\partial C(t,X(t))}{\partial x} + \frac{1}{2}\frac{\partial^2 C(t,x)}{\partial x^2}\sigma^2(x,t) = 0 \tag{63}$$

with boundary condition $C(T, x) = \phi(x)$.

The original Feynman-Kac theorem was stated more generally than Eq. 58, specifically the expectation value

$$C(t,x) = \mathbf{E}\left[\int_t^T e^{-\int_t^s V(X(\tau),\tau)d\tau} f\big(X(s),s\big)ds + e^{-\int_t^T V(X(\tau),\tau)d\tau}\phi(X(T))\,|\,X(t)=x\right], \tag{64}$$

where the stochastic process obeys the same SDE 56, is the solution to the differential equation

$$\frac{\partial C(t,x)}{\partial t} + \mu(x,t)\frac{\partial C(t,x)}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 C(t,x)}{\partial x^2} - V(x,t)C(t,x) + f(x,t) = 0. \tag{65}$$

The function $f(x,\ t)$ can represent cashflows before maturity, if a bond pays $c \times P$ at times t_i then $f(x,\ t) = \sum_i cP\delta(t - t_i)$, and pays P at maturity $\phi(X(T)) = P$ then the expectation becomes

$$C(t,x) = \sum_i cP\mathbf{E}\left[e^{-\int_t^{t_i} V(X(\tau),\tau)d\tau}\right] + P\mathbf{E}\left[e^{-\int_t^T V(X(\tau),\tau)d\tau}\right] \tag{66}$$

from which $V(X(t),\ t)$ can be identified as the stochastic term structure of interest rates. Indeed the Black-Scholes Eq. 51 is identical to Eq. 65 with $V(x,\ t) = r$ and $f(x,\ t) = 0$.

The Feynman-Kac theorem connects expectation values to partial differential equations – simulating a process is equivalent to solving a PDE. As with any PDE, a closed form solution only exists in a small number of situations, which is fortunately the case in the Black-Scholes model. In other situations, the PDE must be solved numerically. However, when the dimensionality is high – which occurs for basket options where the derivative payoff depends on many securities and therefore many stochastic factors – solving the PDE numerically becomes too computationally expensive and therefore the left-hand side of the Feynman-Kac theorem can be used to simulate the stochastic processes using a numerical technique called Monte Carlo Simulation, a technique developed at Los Alamos during World War II (Metropolis et al. 1953).

8 1977 Monte Carlo Methods: Phelim Boyle

Phelim Boyle earned a PhD in physics from Queens University Belfast in 1960, before the watershed event in quant finance. Therefore there was no established path from physics to finance, and therefore Phelim became an actuary – a mathematical discipline that still attracts physics graduates to this day. Boyle published a seminal paper in quantitative finance (Boyle 1977), introducing Monte Carlo methods to the field.

The Black-Scholes model is one of a small number of models that permit a closed-form such a way that the ultimate payoff depends not only on the value of the stochastic factor at maturity, but also the entire path of the stochastic factor, then the price of this option is not Markov, and the mathematical machinery so far

described will not be applicable. The entire path of the stochastic process must be known, and this is precisely what Boyle described.

The Monte Carlo price of a derivative security is the simple average over all realizations of random variables

$$\mathbf{E}\left[f(S(T))\,|\,\mathcal{F}_t\right] \approx \frac{1}{N}\sum_{i=0}^{N} f(S_i) \tag{67}$$

as long as the random variables are sampled using the correct distribution, coming from the SDE of the underlying. For instance, the SDE for a lognormal distribution

$$dS(t) = S(t)rdt + \sigma S(t)dW(t) \tag{68}$$

by the transformation $X(t) = \ln S(t)$ and an application of Itō's formula

$$X(T) = X(t) + \left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W(T) - W(t)), \tag{69}$$

leading to

$$S(T) = S(t)\exp\left(\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W(T) - W(t))\right). \tag{70}$$

Recall that the increments of a Brownian process are normally distributed with zero mean and variance equal to the time increment. Therefore a simple Monte Carlo on a European call option can be computed

$$C(t,S) \approx \sum_{i=0}^{N} \max\left[S(t)\exp\left(\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}z_i\right) - K, 0\right] \tag{71}$$

where the $z_i \sim \mathcal{N}(0,1)$ are standard normal variates. Of course this integral can be performed exactly leading again to the Black-Scholes formula; however, if there are any "exotic" features of the option, the integral does not admit a closed form solution. These exotic features could be as simple as the ability to exercise the option at any time up until maturity – so-called American style options – or the payoff depending on an average of the stock price throughout the lifetime of the option – so-called Asian style options. In these cases, the Monte Carlo procedure can still be used, as it provides a method to sample not only the underlying at maturity, but also to create a path that the underlying takes. In this case the SDE is interpreted literally, for example the simplest approach to generating a path is the Euler method (a name borrowed from the simplest method to solve a differential equation)

$$S(t+\Delta t) \approx S(t)\left(r - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}z_0 \tag{72}$$

where $z_0 \sim \mathcal{N}(0,1)$ is a standard normal variate. This process can then be iterated to determine a possible path for the underlying $S(t)$ up until maturity.

In this way, the entire path of the security can be computed, and the price of path-dependent options can be calculated statistically by taking the average over a large number of simulations. The more simulations performed, the higher the accuracy as the standard deviation of the resulting mean is proportional to $N^{-1/2}$, where N is the number of simulations.

9 1978 Partial Differential Equations: Eduardo Schwartz and Michael Brennan

Pricing options with early termination features are not possible in closed form, as again the price of the option will not be a Markov process. In order to determine the arbitrage-free price of the option in this case, one could use the Monte Carlo technique pioneered by Philem Boyle; however, this method suffers from inaccuracies due to its random nature. Another method is commonly used when early exercise is possible, and the number of stochastic factors are low, namely solving the PDE numerically.

In order to discretize the Black-Sholes PDE 51, following Brennan and Schwartz (1978), a log transformation $y = \ln(S)$ is first applied to make the PDE have constant coefficients, and the time is transformed to be relative to maturity $\tau = T - t$. The resulting PDE

$$\frac{\partial C(\tau, y)}{\partial \tau} = \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial C(\tau, y)}{\partial y} + \frac{1}{2}\sigma^2\frac{\partial^2 C(\tau, y)}{\partial y^2} - rC(\tau, y) = 0 \tag{73}$$

is cast as an operator equation

$$\frac{\partial C(\tau, y)}{\partial \tau} = \mathcal{L}\, C(\tau, y) \tag{74}$$

where $\mathcal{L} = \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial}{\partial y} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial y^2} - r$. The "spatial" variable (the value of the underlying stochastic process) is discretized into an evenly spaced grid with spacing $\Delta y : y \to j\Delta y, j \in [0, N]$. The differential operator is replaced by a second-order discrete operator

$$\mathcal{L} \to \left(r - \frac{1}{2}\sigma^2\right)\left[\frac{\delta_{+1} - \delta_{-1}}{2\Delta y}\right] + \frac{1}{2}\sigma^2\left[\frac{1}{2}\,\frac{\delta_{-1} - 2\delta_0 + \delta_{+1}}{\Delta y^2}\right] + r, \tag{75}$$

where δ_n is an operator that discretely moves the value of y, $\delta_n C(t, y_j) = C(t, y_{j+n})$.

Time is also discretized with spacing $\Delta t\, \tau \to \tau_i = i\Delta t, i \in [-M, M]$, $\frac{\partial C(\tau, y_j)}{\partial \tau} \to \frac{C(\tau_{i-1}, y_j) - C(\tau_i, y_j)}{\Delta t}$. The index M is chosen by asymptotic considerations, and will depend on the boundary condition handling described later. There is a choice as to which time step appears in the function acted upon by the differential operator, which lead to three different methods of numerically solving the PDE. These γ "stencils," are (where $C_{ij} \equiv C(\tau_i, y_j)$).

1. Explicit

$$\frac{C_{i-1,j} - C_{ij}}{\Delta t} = \mathcal{L}C_{ij} \tag{76}$$

2. Implicit

$$\frac{C_{i-1,j} - C_{ij}}{\Delta t} = \mathcal{L}C_{i-1,j} \tag{77}$$

3. Crank-Nicolson

$$\frac{C_{i-1,j} - C_{ij}}{\Delta t} = \frac{1}{2}\left(\mathcal{L}C_{i-1,j} + \mathcal{L}C_{ij}\right) \tag{78}$$

The three stencil choices are depicted graphically in Fig. 4. Essentially, the time at which the τ derivative is taken is different in the three methods. The explicit uses the backwards finite difference approximation based at τ_{i+1}, the implicit the forward approximation based at τ_i, and the Crank-Nicolson uses the centered approximation based at the mid-point $\frac{1}{2}(\tau_{i+1} + \tau_i)$. The centered approach is best from the accuracy (in time) point of view, because centered finite differences are second-order approximations and have error $O(\Delta\tau^2)$, whereas the forward and backward have error $O(\Delta\tau)$, see Andersen and Piterbarg (2010) for details.

In the fully explicit case, the discretized equations become

$$\frac{C_{i+1,j} - C_{i,j}}{\Delta t} = \left(r - \frac{1}{2}\sigma^2\right)\left(\frac{C_{i,j+1} - C_{i,j-1}}{2\Delta y}\right) + \frac{1}{2}\sigma^2\left(\frac{C_{i,j-1} - 2C_{i,j} + C_{i,j}}{\Delta y^2}\right) - rC_{i,j}. \tag{79}$$

Since $C_{0,j}$ is the payoff, this equation can be cast entirely in terms of known quantities – hence the "explicit" title – and solved forwards in time (recall that $\tau = T - t$, so the option is being priced backwards in time).

Of course these techniques were well known prior to the Black and Scholes formula; however in 1978, Micheal Brennan and Eduardo Schwartz found an interesting connection between the numerical procedures and jump diffusion

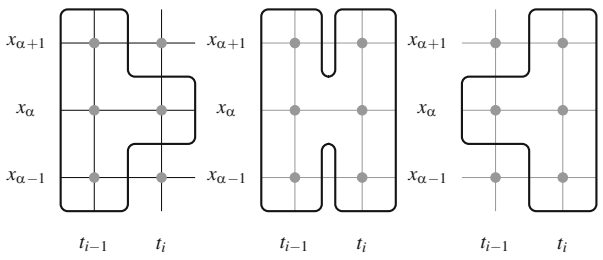


Fig. 4 Options for discretization for numerically solving a partial differential equation. From right to left, the implicit, Crank-Nicolson, and explicit methods

processes (Brennan and Schwartz 1978). They noted in this explicit case, the solutions to the equations are of the form

$$C_{i-1,j} = \frac{1}{1+r\Delta t}\left[p_+ C_{i,j+1} + pC_{ij} + p_- C_{i,j-1}\right] \tag{80}$$

where

$$p_\pm = \left[\frac{1}{2}\left(\frac{\sigma}{\Delta x}\right)^2 \pm \frac{1}{2\Delta x}\left(r - \frac{1}{2}\sigma^2\right)\right]\Delta t \tag{81}$$

$$p = \left[1 - \left(\frac{\sigma}{\Delta x}\right)^2\right]\Delta t. \tag{82}$$

Since $\sum_i p_i = 1$, Brennan and Schwartz point out that these equations are identical to a stochastic jump model that had been written down by Cox and Ross in 1976 (Cox and Ross 1976), and therefore the implicit PDE approach can be thought of as taking discounted estimations in a jump diffusion model, where the price can jump to one of three states.

Brennan and Schwartz also show that the fully implicit case, which is more complicated, also describes a stochastic jump diffusion process, however with infinite states. The ultimate price of a derivative security can be written in a form of a discounted expectation in a jump diffusion model

$$C_{i-1,j} \approx \frac{1}{1+rk}\sum_{q=-\infty}^{\infty} p_q^* C_{i,j+q} \tag{83}$$

where $\sum_{q=-\infty}^{\infty} p_q^* = 1$ and can be interpreted as the probability that the logarithm of the underlying security price jumps to y_q.

Brennan and Schwartz omit the Crank-Nicolson case, where the equations have the highest accuracy, without sacrificing any computational speed – as the matrices are still tridiagonal. To see this, the discretized equations

$$\begin{aligned}\frac{C_{i+1,j} - C_{i,j}}{\Delta t} = &\frac{1}{2}\left[\left(r - \frac{1}{2}\sigma^2\right)\left(\frac{C_{i,j+1} - C_{i,j-1}}{2\Delta y}\right) + \frac{1}{2}\sigma^2\left(\frac{C_{i,j-1} - 2C_{i,j} + C_{i,j}}{\Delta y^2}\right) - rC_{i,j}\right] \\ &+\frac{1}{2}\left[\left(r - \frac{1}{2}\sigma^2\right)\left(\frac{C_{i+1,j+1} - C_{i,j-1}}{2\Delta y}\right) + \frac{1}{2}\sigma^2\left(\frac{C_{i+1,j-1} - 2C_{i+1,j} + C_{i+1,j}}{\Delta y^2}\right) - rC_{i+1,j}\right],\end{aligned} \tag{84}$$

can be simplified to

$$\begin{aligned}&C_{i+1,j-1}\left(-\frac{a}{4} + \frac{b}{2}\right) + C_{i+1,j}\left(1 - \frac{r\Delta t}{2} - b\right) + C_{i+1,j+1}\left(\frac{a}{4} - \frac{b}{2}\right) = \\ &C_{i,j-1}\left(\frac{a}{4} - \frac{b}{2}\right) + C_{i,j}\left(1 + \frac{r\Delta t}{2} + b\right) + C_{i,j+1}\left(-\frac{a}{4} - \frac{b}{2}\right),\end{aligned} \tag{85}$$

where

$$\begin{aligned} a &= \frac{\left(r - \frac{1}{2}\sigma^2\right)\Delta t}{\Delta x} \\ b &= \frac{\sigma^2 \Delta t}{2\Delta x^2}. \end{aligned} \tag{86}$$

These equations can be cast in matrix form in the interior of the system

$$A\mathbf{C_{i+1}} = B\mathbf{C_i}. \tag{87}$$

However to implement this on a computer, the spatial boundary conditions $j \in \{-M, M\}$ must be handled, and there are many ways of approaching this. For very simple options, the value at the edge of the matrix can be replaced by the value of the payoff. However this method has two drawbacks, first it only applies to situations where there is no early exercise so cannot be applied generally, and second care must be taken that the equation is truly in the asymptotic region, and could therefore could be computationally expensive. Second, one could impose some smoothness of the solution, such as imposing that the second derivative vanishes

$$\frac{\partial^2 C_{i+1,\pm M}}{\partial y^2} = 0, \tag{88}$$

which is more general and can be applied in more situations, such as early exercise, however still imposes a careful asymptotic analysis is done and therefore could require a larger spatial grid. Finally, a method that overcomes both can be used where the equations themselves determine the boundary condition, with the drawback that the matrix is no longer tridiagonal and therefore less efficient computational methods are required to be used.

Since there are no nodes past the boundaries, the centered finite difference approximations 75 can no longer be used. In order to continue to have $O(\Delta y^2)$ accuracy at the boundaries, a higher order purely forward and backward finite difference approximation must be used. At the lower boundary, all translation operators δ_n must move back into the grid and therefore be positive

$$\begin{aligned} \left.\frac{\partial}{\partial y}\right|_{y=-(M+1)\Delta y} &\to \frac{1}{2\Delta y}\left[-3\delta_0 + 4\delta_1 - \delta_2\right] + O\left(\Delta y^2\right) \\ \left.\frac{\partial^2}{\partial y^2}\right|_{y=-(M+1)\Delta y} &\to \frac{1}{\Delta y^2}\left[2\delta_0 - 5\delta_1 + 4\delta_2 - \delta_3\right] + O\left(\Delta y^2\right). \end{aligned} \tag{89}$$

On the other hand, the maximum grid nodes must be translated downwards in order to remain in the grid

$$\begin{aligned}\left.\frac{\partial}{\partial y}\right|_{y=(M+1)\Delta y} &\to \frac{1}{2\Delta y}[3\delta_0 - 4\delta_{-1} + \delta_{-2}] + O(\Delta y^2)\\ \left.\frac{\partial^2}{\partial y^2}\right|_{y=M+1\Delta y} &\to \frac{1}{\Delta y^2}[2\delta_0 - 5\delta_{-1} + 4\delta_{-2} - \delta_{-3}] + O(\Delta y^2).\end{aligned} \tag{90}$$

The trick is to add one more point just outside the grid using this technique and keep track of this "ghost" node ($j = M + 1$) at each iteration. The boundary matrix elements in 87 are augmented to be

$$\begin{aligned}A_{i,-M} &= \left(1 - \frac{r\Delta t}{2} - b + \left(-a + \frac{5}{2}b\right)c_1\right)\\ A_{i,-M+1} &= \left(\frac{1}{4}a + \frac{1}{2}b + \left(\frac{1}{4}a - 2b\right)c_1\right)\\ A_{i,-M+2} &= \frac{1}{2}bc_1\\ B_{i,-M} &= \left(1 + \frac{r\Delta t}{2} + b - \left(-a + \frac{5}{2}b\right)c_1\right)\\ B_{i,-M+1} &= \left(-\frac{1}{4}a - \frac{1}{2}b - \left(-\frac{1}{4}a - 2b\right)c_1\right)\\ B_{i,-M+2} &= -\frac{1}{2}bc_1\end{aligned} \tag{91}$$

where

$$c_1 = \frac{\left(-\frac{a}{4} + \frac{b}{2}\right)}{1 - \frac{r\Delta t}{2} - \frac{3}{4}a + b}. \tag{92}$$

The matrix at the upper bound must be similarly altered

$$\begin{aligned}A_{i,M} &= \left(1 - \frac{r\Delta t}{2} - b + \left(a + \frac{5}{2}b\right)c_2\right)\\ A_{i,M-1} &= \left(-\frac{1}{4}a + \frac{1}{2}b + \left(-\frac{1}{4}a - 2b\right)c_2\right)\\ A_{i,M-2} &= \frac{1}{2}bc_2\\ B_{i,M} &= \left(1 + \frac{r\Delta t}{2} + b - \left(a + \frac{5}{2}b\right)c_2\right)\\ B_{i,M_1} &= \left(\frac{a}{4} - \frac{b}{2} - \left(-\frac{1}{4}a - 2b\right)c_2\right)\\ B_{i,M-2} &= -\frac{1}{2}bc_2\end{aligned} \tag{93}$$

where

$$c_2 = \frac{\left(\frac{a}{4} + \frac{b}{2}\right)}{1 - \frac{r\Delta t}{2} + \frac{3}{4}a + b}. \tag{94}$$

The ghost node at the lowest edge of the grid evolves according to

$$\begin{aligned} C_{i+1,-(M+1)}\left(1 - \frac{r\Delta t}{2} - \frac{3}{4}a + b\right) &= C_{i,-(M+1)}\left(1 - \frac{r\Delta t}{2} - \frac{3}{4}a + b\right) \\ &+ C_{i+1,-M}\left(-a + \frac{5}{2}b\right) + C_{i+1,-M+1}\left(\frac{1}{4}a - 2b\right) \\ &+ C_{i+1,-M+2}\frac{1}{2}b + C_{i,-M}\left(-a + \frac{5}{2}b\right) \\ &+ C_{i,-M+1}\left(\frac{1}{4}a - 2b\right) + C_{i,-M+2}\frac{1}{2}b \end{aligned} \tag{95}$$

and at the maximum

$$\begin{aligned} C_{i+1,M+1}\left(1 - \frac{r\Delta t}{2} + \frac{3}{4}a + b\right) &= C_{i,M+1}\left(1 + \frac{r\Delta t}{2} - \frac{3}{4}a - b\right) + C_{i+1,M}\left(a + \frac{5}{2}b\right) \\ &+ C_{i+1,M-1}\left(-\frac{1}{4}a - 2b\right) + C_{i+1,M-2}\frac{1}{2}b \\ &+ C_{i,M}\left(a + \frac{5}{2}b\right) + C_{i,M-1}\left(-\frac{1}{4} - 2b\right) + C_{i,M-2}\frac{1}{2}b. \end{aligned} \tag{96}$$

Finally, the evolution equation to solve including boundary conditions is

$$A\mathbf{C_{i+1}} = B\mathbf{C_i} + \mathbf{g} \tag{97}$$

where the vector $\mathbf{g}$ is zeros except for the maximum and minimum entries

$$\begin{aligned} g(M) &= C_{i,M+1}\left(-\frac{a}{4} - \frac{b}{2} - \left(1 + \frac{r\Delta t}{2} - \frac{3}{4}a - b\right)c_2\right) \\ g(-M) &= C_{i,-(M+1)}\left(\frac{a}{4} - \frac{b}{2} - \left(1 + \frac{r\Delta t}{2} + \frac{3}{4}a - b\right)c_1\right). \end{aligned} \tag{98}$$

Thus, the PDE can be solved easily by numerical means, as A is the only matrix that needs to be inverted, and this maintains its tridiagonal nature for all the boundary conditions described above.

10 Conclusions

This chapter has explored the mathematical ideas that changed how capital was allocated – although the history of financial risk was deliberately omitted despite it being central to modern portfolio construction, since it is a topic deserving of its own treatment and chapter.

Interest rates, and thus the time value of money, are a concept that predates writing – likely dating to the onset of the agricultural transition over 12,000 years ago. It is fascinating to note that the first discovered set of codified laws include a cap on interest rates and rules regulating borrowing and lending – the first banking regulations were included in the first historical set of laws.

The full power of the time value of money was not realized until 1202 when Fibonacci demonstrated that capital allocation decisions could be made using net present value analysis. Together with the law of one price and the value of the firm, Fibonacci revolutionized the way trade was done in the thirteenth century. It has been argued that a well-developed financial and commercial system was a key prerequisite for the Italian renaissance and enlightenment.

Between the thirteenth and twentieth centuries, mathematics evolved immensely, but curiously finance did not. It was not until Louis Bachelier's thesis, defended on March 29, 1900, that the machinery of advanced mathematics was applied to the financial markets. Even then it took another 50 years for the mathematics in Bachelier's thesis to be put on a solid analytical footing. Incredibly, Bachelier must have intuited many of the mathematical advancements in the twentieth century, including stochastic calculus and the analytical foundations of probability theory. Although many of the great mathematicians had worked on probability theory prior to the twentieth century, it was only put on a solid analytical foundation by Kolmogorov, much like Cauchy, Reimann, and Weierstrauss only put calculus on a solid mathematical foundation after hundreds (if not thousands) of years of use.

In 1953, Harry Markowitz came out with "Portfolio Selection" and a whole new quantitative method was introduced for deciding which investments enter into portfolios. From there, the capital asset pricing model was born and mathematical methods became a necessary tool for investing. Prior to Markowitz, correlations were understood as important; however, it was believed that all risk could be removed from a portfolio with enough diversity. Therefore, mathematical tools, such as covariance analysis, were not evident in the asset selection process. Although not used by asset managers in their original form, the modern portfolio theory and CAPM are still subjects that are taught, and many concepts deriving from these ideas, such as the Sharpe ratio, are still commonly used as portfolio metrics today.

In 1973, the mathematical toolkit used by asset managers to grow to include advanced mathematics, as the price of an option was shown to be the solution of a partial differential equation (although almost certainly this was known to Thorpe prior to the Black-Scholes paper), a mathematical object not regularly taught in finance nor economic curricula. The amazing thing about the Black-Scholes option price was not that it can be derived using the same mathematics that governs how

heat diffuses, but that it did not contain any unobservable quantities, such as the investors expectation, or their risk aversion. Incredulous, Robert Merton set to rederive the results and found that this was the case, introducing the ideas that would later lead to the fundamental theorem of asset pricing written down by Harrison and Pliska.

Tools from mathematical physics and applied mathematics are being applied constantly to the complex world of financial derivatives, and some early methods have been presented in this chapter. New mathematical ideas are continuously being applied to quantitative finance. As of the writing of this chapter, machine learning and artificial intelligence are just being tested as a way to allocate capital and hedge derivatives (Hutchinson et al. 1994; Buehler et al. 2019). Thus, the future of quantitative finance may look very different from its past.

References

Andersen L, Piterbarg V (2010) Interest rate modeling. No. v. 1. In: Interest rate modeling. Atlantic Financial Press. URL: http://books.google.com/books?id=oI2fcQAACAAJ

Bachelier L (1900) Théorie de la spéculation. Annales Scientifiques de L'Ecole Normale Supérieure 17:21–88. Reprinted in P. H. Cootner (ed), 1964, The random character of stock market prices, Cambridge, MA. MIT Press

Black F, Scholes M (1973) The pricing of options and corporate liabilities. J Polit Econ 81(3):637–654

Boyle PP (1977) Options: a Monte Carlo approach. J Financ Econ 4(3):323–338. URL: https://ideas.repec.org/a/eee/jfinec/v4y1977i3p323-338.html

Brennan MJ, Schwartz ES (1978) Finite difference methods and jump processes arising in the pricing of contingent claims: a synthesis. J Financial Quant Anal 13(3):461–474. URL: http://www.jstor.org/stable/2330152

Buehler H, Gonon L, Teichmann J, Wood B (2019) Deep hedging. Quant Finance 19(8):1271–1291. https://doi.org/10.1080/14697688.2019.1571683

Cox JC, Ross SA (1976) The valuation of options for alternative stochastic processes. J Financ Econ 3:145–166

Derman E (2013) *A young person's guide to pricing and hedging*. https://emanuelderman.com/wp-content/uploads/2013/06/The-Young-Persons-Guide.pdf. Accessed 24/10/2023

Devlin K (2011) The man of numbers: Fibonacci's arithmetic revolution. Walker Publishing Company, New York, NY

Einstein A et al (1905) On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat. Ann Phys 17(549–560):208

Fama EF, French KR (2004) The capital asset pricing model: Theory and evidence. J Econ Perspect 18(3):25–46

Gray J (2015) The real and the complex: a history of analysis in the 19th century, Springer Undergraduate Mathematics Series. Springer International Publishing. URL: https://books.google.co.uk/books?id=z-wmswEACAAJ

Harrison J, Pliska SR (1981) Martingales and stochastic integrals in the theory of continuous trading. Stoch Process Appl 11(3):215–260. https://doi.org/10.1016/0304-4149(81)90026-0. URL: http://www.sciencedirect.com/science/article/pii/0304414981900260

Hudson M (2000) How interest rates were set, 2500 BC-1000 AD: Má, tokos and fnus as metaphors for interest accruals. J Econ Soc Hist Orient 43(2):132–161. URL: http://www.jstor.org/stable/3632489

Hutchinson JM, Lo AW, Poggio T (1994) A nonparametric approach to pricing and hedging derivative securities via learning networks. J Financ 49(3):851–889
Itō K (1944) 109. stochastic integral. Proc Imp Acad 20(8):519–524
Itō K (1951a) On a formula concerning stochastic differentials. Nagoya Math J 3:55–65
Itō K (1951b) On stochastic differential equations. Am Math Soc 4:1–51
Kac M (1949) On Distributions of Certain Wiener Functionals. Trans Am Math Soc 65(1):1–13
Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Rev Econ Stat 47(1):13–37
Markowitz H (1952) Portfolio selection. J Financ 7(1):77–91. URL: http://www.jstor.org/stable/2975974
Merton RC (1973) Theory of rational option pricing. Bell J Econ Manage Sci:141–183
Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21(6):1087–1092
Morley I, Renfrew C (2010) The archaeology of measurement: comprehending heaven, earth and time in ancient societies. Cambridge University Press. URL: https://books.google.co.uk/books?id=TzrNgAsJY1MC
Oksendal B (1992) Stochastic differential equations. In: An introduction with applications, 3rd edn. Springer-Verlag, Berlin, Heidelberg
Samuelson PA (1965) Rational theory of warrant pricing IMR. Ind Manag Rev (pre-1986) 6(2):13
Schmandt-Besserat D (1992) Before writing, vol. I: from counting to cuneiform, vol 1. University of Texas press, Austin, TX
Sharpe WF (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. J Financ 19(3):425–442
Shreve SE (2004) Stochastic calculus for finance II: continuous-time models, vol II. Springer, New York, NY
Sigler L (2012) Fibonacci's liber abaci: A translation into modern english of Leonardo Pisano's book of calculation. Springer, New York
Sprenkle CM (1961) Warrant prices as indicators of expectations and preferences. Yale Econ Essays 1(2):179–232
Thorp E, Kassouf S (1967) Beat the market: a scientific stock market system. Random House, New York, NY
Weierstrass K (1872) Uber continuirliche functionen eines reellen arguments, die fur keinen worth des letzteren einen bestimmten differentailqutienten besitzen, akademievortrag. Math Werke 2:71–74
Wiener N (1923) Differential-space. J Math Phys 2(1–4):131–174

Section IV

Practical, Historical, and Philosophical Instances of Probability

Egan J. Chernoff

Introduction to Practical, Historical, and Philosophical Instances of Probability

Egan J. Chernoff

Abstract

The confluence of practical, historical, and philosophical instances of probability is represented, here, as a problem. A simple problem, really, but the investigation of said problem is purposely presented in a manner that raises practical, historical, and philosophical problems. In this instance, while solving the problem, different individuals differently become privy to different information. At first glance, no problem. Upon further inspection, probabilistic implications are revealed. After-all, while, yes, the Four of Hearts is the Four of Hearts, it is also a four of hearts. In other words, it is, concurrently, many things (not a queen, a four, a red card, not a spade, etc.). This information, while appearing mundane, becomes important to consider when examining different answers to the problem. As one might expect, having two people, standing right next to or across from each other, both staring at the same deck of cards, but with different probabilities that the next card drawn is a queen, can be a recipe for disaster. It can also be a window into how information about events is interpreted and shared. A probabilistically rooted answer to an enduring question: Do you see what I see? Well, let's do the math.

Keywords

History · Intuition · Neuralyzer · Philosophy · Playing cards · Probability · Problems

E. J. Chernoff (✉)
College of Education/Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK, Canada
e-mail: egan.chernoff@usask.ca

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_136

At the confluence of practical, historical, and philosophical instances of probability, for me, is a problem, technically, two problems. These two problems beget numerous other problems.

Here is the first problem:

> A card is drawn from the deck, the card is replaced, and a second card is drawn from the deck. What is the probability that the second card drawn is a queen?

The answer to the first problem is 4/52 or 1/13.

Here is the second problem:

> A card is drawn from the deck, the card is not replaced, rather, it is placed face down beside the deck, and a second card is drawn from the deck. What is the probability that the second card drawn is a queen?

The answer to the second problem is, of course, 4/52 or 1/13.

Beginning begotten problems, these two problems, taken in concert, present answers that are counterintuitive in nature, albeit not for everyone. Given the long history of counterintuitive probability problems (e.g., Problem of Points, Monty Hall Problem, The Two Child Problem, Birthday Problem, Bertrand's Box Paradox, and many, many others) running into a counterintuitive probability problem should come as no surprise. Given the prominent plethora of research questioning whether or not humans are intuitive statisticians (e.g., Kahneman and Tversky 1972 and many others) also adds to a lack of surprise. As Sir David Spiegelhalter succinctly noted in his commendation for *Probabilistic Thinking: Presenting Plural Perspectives* (Chernoff and Sriraman 2014), "I often get asked why people find probability so unintuitive and difficult. After years of research, I have concluded it's because probability really is unintuitive and difficult." Therefore, getting back to the counterintuitive situation at hand, and a little more specific, that the answer for the first problem, a probability question *with* replacement, is the same as the answer for the second problem, a probability question *without* replacement, rattles what Fischbein (1975) denotes as a *secondary intuition*, that is, an intuition built up through the process of schooling.

Typically, as documented in research regarding the teaching and learning of mathematics, *replacement* is a trigger warning for independent events, whereas *no replacement* triggers dependent events. Quasi-conversely, independent events get associated with replacement and dependent events with no replacement. Those grappling with their schooled intuition for 4/52 being the answer for both a problem with replacement, and also a similar problem without replacement, can defer to a calculation to help alleviate concerns that "obviously" the answer to the second problem is 4/52.

$$\begin{aligned}
P(Q2) &= P(Q1 \text{ and } Q2 \text{ or } \neg Q1 \text{ and } Q2) \\
&= P(Q1 \times Q2) + P(\neg Q1 \times Q2) \\
&= (4/52 \bullet 3/51) + (48/52 \bullet 4/51) \\
&= (4/52 \bullet 3/51) + (4/52 \bullet 48/51) \\
&= 4/52(3/51 + 48/51) \\
&= 4/52(51/51) \\
&= 4/52(1) \\
&= 4/52
\end{aligned}$$

Comforted with a cold calculation, further problems emerge from an extension of the second problem. Consider, for example, a card is drawn from the deck, the card is not replaced, rather, it is placed face down beside the deck; a second card is drawn from the deck, which is also not replaced, rather, it is placed face down beside the first card that was not replaced; then, a third card is drawn from the deck. What is the probability that the third card drawn is a queen? The answer is, which can also be confirmed by those who take a moment to attend to the more laborious calculation for this case, 4/52 or 1/13. Extending the problem along the same line of reasoning, that is, three cards drawn and placed face down beside each other, beside the deck, the probability that the fourth card drawn is a queen is 4/52 or 1/13. Not looking to continue the extension solely for the purpose of completing further and further crass calculations, we stop this particular extension path at the fourth version of the problem, as witnessed in Fig. 1, which provides the backdrop for the emergence of a new set of problems.

Up until this point, the largest hurdle to establishing that the fourth card drawn from the deck being a queen is 4/52 is a shaking off of a secondary intuition, one rooted in a school-based understanding of dependent and independent events, and keeping track of a bunch of fractions, whether with pencil and paper or some form of technology aiding the calculation. However, more problems emerge – practical, historical, and philosophical problems – when conditioning events get considered.

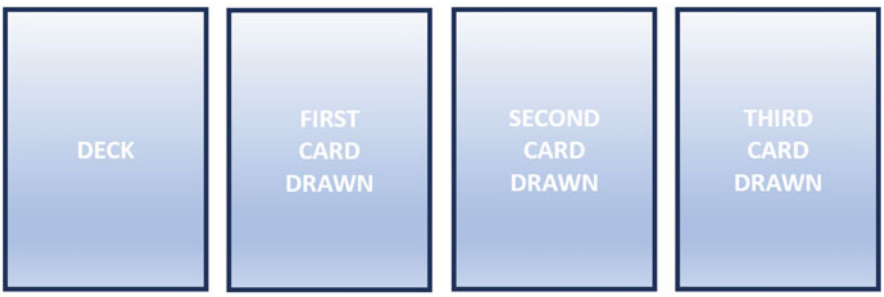


Fig. 1 Deck of cards (left). The three drawn cards and not replaced, rather, placed next to the deck and each other (right)

Imagine, as a new problem, two individuals taking part in the above scenario, as detailed in Fig. 1. Let us call them Thomas and Blaise. At the moment, for both individuals, the probability that the fourth card drawn is a queen is 4/52. However, let us now imagine a scenario, a practical scenario, were one of the two individuals, Thomas, curious about the cards placed beside the deck, casually picks up, say, the first card drawn and placed next to the deck, takes a look at the card, and then places it back where they got it from. At this point, the probability that the fourth card drawn is a queen remains 4/52 for Blaise; however, the probability that the fourth card drawn is a queen may now be different for Thomas, the one who picked up and took a look at the face of that first card drawn from the deck. Now there are two different individuals with two different probabilities for the fourth card drawn being a queen. Looking to extend this new thread of problems even further, similar scenarios, with more and different people peeking at cards, can be put into play.

Continuing from the scenario presented in Fig. 1, more different probabilities are possible, especially when Thomas and Blaise are joined in the room by Pierre-Simon, Bruno, Andrey, Abraham, Richard, and Jacob. Joining Thomas, who took a look at the first card drawn, imagine a scenario where Pierre-Simon snuck a peek at the second card drawn from and placed next to the deck, and Bruno became privy to the face of the third card drawn from and placed next to the deck. At that point, there are four different individuals each with a different probability that the fourth card drawn from the deck is a queen. Well, "potentially" four different probabilities would be safer to say at this point, but I do not wish to lose the forest for the trees at this moment.

To recap, the probability that the fourth card drawn is a queen is still 4/52, albeit for Blaise. For Thomas, Pierre-Simon, and Bruno, the probability that the fourth card drawn is a queen needs to take into consideration the information each of them obtained when looking at the face of their respective cards. Similarly, but slightly differently, Andrey, Abraham, and Richard in this scenario become privy to information regarding two cards. In particular, Andrey accustoms himself with the faces of the first and second cards drawn, Abraham acquaints himself with the faces of the first and third cards drawn, and Richard recognizes information associated with the faces of the second and third cards drawn. While we are it, Jacob gains access to the faces of the first, second, and third card that is drawn. Let us recap.

In the scenario being presented we have Blaise who is still not privy to any of the information contained on the faces of the three cards placed next to the deck. As for Thomas, Pierre-Simon, and Bruno, they have each gained information from seeing one card, albeit different cards. Whereas Andrey, Abraham, and Richard garnered information from two cards, albeit two different cards, for each of them. Lastly, Jacob is aware of the faces of all the cards that were drawn. In sum, now there are eight different individuals, each with knowledge of different, relevant information.

Like students in a math class, I imagine tasking each of the individuals with calculating the probability that the fourth card drawn is a queen, and to share their results with the rest of the group. Perhaps I am projecting a bit, but what I imagine happening next is all hell breaking loose.

One conversation theme emerging from the group discussion would be temporally based. I could see Thomas and Bruno having a discussion about "when" they learned about the face of their card, and if, and how, that impacts the probability the fourth card drawn is a queen. Differences between sneaking a peek at the face of the card after it has been laid down and seeing the face of the card, in real time, as the card is placed next to the deck would get hashed out, but temporal issues would remain. Worthy of note, should Thomas, in discussion with Bruno, share details associated with the face of his card, then, actually, Bruno would have a new probability where his calculation is no longer solely based on the information from the face of the card he saw, but also on now knowing what card Thomas saw, as well. To remedy this situation, we enter into the mix, of course, a neuralyzer.

For those of you well versed in the *Men in Black* film franchise, you know what a neuralyzer is and what it does. On the off chance you are not familiar with a *neuralyzer*, it is a cigar-sized metal tube that, when a particular button is pressed, emits a bright flash, like that of a camera. Unlike a camera, however, the flash erases memories for any person or persons who see the flash. The period to be erased, that is, how many seconds, minutes, hours, days, weeks, months, and years, is entered into the neuralyzer, via buttons and dials, before the flash takes place. Presently, the neuralyzer is not real. It is from a movie. A science fiction, action, comedy film. However, in what follows, as one would when watching a movie, I ask you to suspend reality (even if just for a few moments) and embrace the notion that seeing a flash results in a memory wipe.

Back to the discussion between Thomas and Bruno, where Thomas let slip information about the face of his card, to get the boys back on track, that is, working out the probability that the fourth card drawn is a queen, albeit with each of them having seen the face of a different card, we flash the neuralyzer for Bruno. (Flash!) Just like that, Bruno is no longer saddled with the information about Thomas's card. In other words, Bruno is back to his original calculation with one card.

A similar theme, that is, a temporally based one, could also be taking place between Andrey, Abraham, and Richard, the individuals who were able to take a look at two, albeit different, cards in their calculation of the probability of the fourth card being a queen. A discussion focused on the different probabilities that could emerge based on the cards they were assigned would take place. I also envision, in that discussion, a sidebar conversation about the differences between seeing one card then another versus seeing both cards at once. Andrey, for example, who sees the first and second card, could have garnered the information by seeing (1) the first card next to the deck and then the second card next to the deck; (2) the second card and then the first card; or (3) both cards at the same time. All of which causes Jacob to chime in, then, that he might have to rethink his calculation because of the different ways in which the faces for one, two, and three cards could be digested. It would probably be best, at this point, to give Andrey, Abraham, and Richard a quick flash, in case Jacob let something slip in their conversation before going back to work on his own in the corner. (Flash!) Before or after, left versus right, all at once or one at a time, one card, a couple of cards and on and on; and, the conversation, thus far, has not even addressed the faces of the cards.

Getting back to Thomas, let us say, for the sake of argument, the face of his card, that first card that was drawn and placed beside the deck, is the Four of Hearts. His reporting back to the group of the probability that the fourth card drawn was a queen would be a window into whether Thomas saw the Four or Hearts as the Four of Hearts.

The Four of Hearts, yes, is the Four of Hearts. However, the Four of Hearts is also a four of hearts. The Four of Hearts is multivalent, that is, concurrently, it is many things. The Four of Hearts is not a queen, for example. It is also not a seven. The Four of Hearts is a four. It is a red card. It is a four and a heart. This information, while mundane, is important when assessing Thomas's probability that the fourth card drawn is a queen. After all, to say that the four of hearts is not a queen is correct. Should Thomas interpret the four of hearts as not a queen, this would be sussed out from his probability that the fourth card is a queen. The probability of the fourth card being a queen, then, is a window into one's interpretation of an event, which, in this instance, is seeing the face of the first card placed next to the deck.

It could be argued that the face of the card, when revealed to Thomas, exists in a state of superposition. In other words, the card concurrently exists in all the different ways that it could be interpreted: not a queen, not a spade, four, heart, red card, four or hearts, rank, suit, rank without suit, suit without rank, and on and on. It is Thomas's answer to the task, then, the probability, that sheds light on Thomas's interpretation of the face of the card he saw. Thomas's answer, arguably, reveals his particular collapse of the superposition and, thus, how the card was interpreted.

In line with this reasoning, an exposed card, yes, does provide new information; however, how said information is interpreted and shared, how the card is actually seen, is garnered from the answer stemming from the calculation of the probability that the fourth card drawn is a queen. Attending to an old quandary, that is, do you see what I see, for example, when a card is flipped over for "you" and "I" to see, cue the calculations. Also, cue a discussion on whether or not the information is *in* the cards.

Taking temporal troubles and multivalent musings together, more problems may emerge. Having Thomas (who saw the face of the first card placed next to the deck) talking with Andrey (who saw the faces of the first and second) and Abraham (who saw the faces of the first and third) could make him privy to information concerning all three faces of the cards placed next to the deck. Before resorting to a quick neuralyzer flash, lest we forget, too, that all these people under discussion are from the human race.

Members of the human race suffer from many difficulties. Our memories are not like video camera recordings. Some of us have particular predilections for telling the truth or saving face. Not everyone is of keen eyesight. Some of us, myself included, live with color blindness. Others may not be intimately familiar with the particulars, such as rank and suit, of a deck of cards. Beyond just the issues with ourselves, communicating well with others is no easy task. As Red Auerbach (a famous basketball coach) once said, it does not matter what is said, what matters is what is heard. On and on I could go about members of the human race. Instead, the larger point attempting to be made here is that humans are fallible when interpreting and

sharing information garnered from events, like what one sees when seeing the face of a playing card or many playing cards, which could have probabilistic implications.

These implications hold true for other probabilistic situations and scenarios. Different probabilities may emerge from the timing of verbal information received from another. Similarly, different observers, those given access to the same circumstances, but with different underlying knowledge of the scenario, might emerge with different probabilities. Stated more generally, "probabilities quantify not events but our *information* about events" (Devlin 2014, p. xii). As demonstrated throughout the above problems, our information about events, how said information is interpreted, how it is shared, can be established probabilistically.

I began this introduction by declaring that I have a problem. Problems, really. My biggest problem, I see now, would be to end an introduction to a section on practical, historical, and philosophical instances of probability by aligning myself with a particular philosophical interpretation of probability. Fortunately, for me, I think I have a way out of this mess that I've gotten myself into. . .

Flash!

As noted by Gillies (2000), "The theory of probability has a mathematical aspect and a foundational or philosophical aspect. There is a remarkable contrast between the two. While an almost complete consensus and agreement exists about the mathematics, there is a wide divergence of opinions about the philosophy" (p. 1). Within the wide divergence of opinions about the philosophy of probability (e.g., logical, subjective, frequency, propensity, and others), there is one significant bifurcation that has been recurrently acknowledged since the emergence of probability around 1600 (Hacking 1975). Hacking describes this duality of probability as the "Janus-faced nature" (p. 12) of probability, explaining "on the one side it is statistical, concerning itself with stochastic laws of chance processes" (ibid.); and "on the other side it is epistemological, dedicated to assessing reasonable degrees of belief in propositions quite devoid of statistical background" (ibid.). Throughout history, when asking do you see what I see, with respect to philosophical interpretations of probability, individuals have not seen eye to eye on the matter, to put it mildly.

Given the tumultuous history over differing philosophical interpretations of probability (Hacking 2001), treading into said arena is not for the faint of heart. I thank, then, Professors Jeffrey Rosenthal, Jason Douglas Rosenhouse, and Christian Martin Hennig for helping me bring this section on practical, historical, and philosophical instances of probability to fruition.

References

Chernoff EJ, Sriraman B (eds) (2014) Probabilistic thinking: presenting plural perspectives. Springer Science, Berlin/Heidelberg

Devlin K (2014) The most common misconception about probability? In: Chernoff EJ, Sriraman B (eds) Probabilistic thinking: presenting plural perspectives. Springer, New York, pp ix–xiii

Fischbein E (1975) The intuitive sources of probabilistic thinking in children. Reidel, Dordrecht
Gillies D (2000) Philosophical theories of probability. Routledge, New York
Hacking I (1975) The emergence of probability: a philosophical study of early ideas about probability induction and statistical inference. Cambridge University Press, Cambridge, MA
Hacking I (2001) An introduction to probability and inductive logic. Cambridge University Press, Cambridge
Kahneman D, Tversky A (1972) Subjective probability: a judgment of representativeness. Cogn Psychol 3:430–454

How Markov's Little Idea Transformed Statistics

Jeffrey S. Rosenthal

Contents

1 Introduction 1372
2 Random Variables 1373
3 Markov Chains 1374
4 Markov Chain Monte Carlo (MCMC) 1376
5 The Metropolis Algorithm 1377
6 MCMC in Practice 1378
7 Conclusion 1380
References 1380

Abstract

We discuss Andrey Andreyevich Markov's early twentieth century idea of a Markov chain, which generalized the traditional notion of independent random variables to a model that was more general but still mathematical tractable. We then describe how that led to the hugely popular modern statistical approach of using Markov chain Monte Carlo algorithms to estimate complicated quantities such as Bayesian posterior distributions.

Keywords

Markov chain · Monte Carlo algorithms · Metropolis algorithm · MCMC · Andrey Andreyevich Markov

J. S. Rosenthal (✉)
University of Toronto, Toronto, ON, Canada
e-mail: jeff@math.toronto.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_104

1 Introduction

It was the beginning of the twentieth century, and Andrey Andreyevich Markov was not happy. The field of mathematical probability was flourishing, with seminal contributions from de Moivre and Fermat through to Gauss, Laplace, Cauchy, Poisson, Chebyshev, and many others. And yet, most of the key results – laws of large numbers, the central limit theorem, and so on – were always about independent random variables, i.e., random quantities which have no effect on each other. This assumption seemed very limiting and unrealistic. However, without this assumption, the random variables could behave however they wanted, and it seemed impossible to make mathematical progress. How could the independence assumption be generalized, while still allowing for valid theorems to be proven?

Markov eventually settled on a model in which each random variable depended on the others, but only through the adjacent ones. That is, a sequence of jointly defined random variables X_0, X_1, X_2,... form a Markov chain if the conditional distribution of X_n, conditional on all the other variables, is completely determined by X_{n-1} and X_{n+1}. This setup was more general than independence, but still tractable enough to analyze mathematically.

Markov published his results in 1906 (Markov 1906), adding a new model to the world of probability theory. He proved a version of the law of large numbers, and later of the central limit theorem, for his new concept of Markov chains. In the decades following, many mathematical papers about Markov chains were published. Moreover, Markov chains had some success as models of card shuffling, diffusions, branching processes, and later to stock prices. Further yet, genuine applications remained somewhat limited. In fact, Markov's original example had been the pattern of consonants and vowels in written language, but they surely do *not* really follow a Markov chain. For example, a vowel is much less likely to occur of the last three letters were all vowels, than if just the last one letter were a vowel. It was challenging to find real-world examples of random sequences in which X_n was indeed influenced by X_{n-1}, but completely conditionally independent of X_{n-2}.

Then, in the latter half of the twentieth century, a whole new application of Markov chains emerged, in the burgeoning field of computer simulation. In 1953, Metropolis et al. (1953) used a simple iterative computer simulation to study the behavior of interacting molecular systems, for up to 224 particles. This was an early application of a Monte Carlo algorithm which uses random simulation to estimate complicated quantities. In doing so, they created the Metropolis algorithm, the world's first Markov chain Monte Carlo (MCMC) algorithm. Their algorithm was later generalized to the Metropolis-Hastings algorithm (Hastings 1970), the Gibbs sampler (Gelfand and Smith 1990; Geman and Geman 1984; Tanner and Wong 1987), and many other variations; see e.g (Tierney 1994). The term "Markov chain Monte Carlo (MCMC)," emphasizing that these new algorithms applied Markov's model to the Monte Carlo paradigm, originated in (Geyer 1992). MCMC algorithms are now hugely popular in many branches of statistics, especially in Bayesian inference, as well as many other fields from molecular dynamics to engineering to machine learning to financial modeling. The volume (Brooks et al. 2011) contains

thousands of references, and a casual Google search for "Markov chain Monte Carlo" (in quotation marks) returns over two million web pages, indicating just how far Markov's little idea has come.

In this chapter, we review Markov chains and MCMC algorithms and explain how they have been so useful in modern statistical analysis.

2 Random Variables

A random variable is a mathematical model of an uncertain quantity. It is defined on some underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where Ω is an underlying sample space, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and $\mathbf{P}$ is a probability measure on $(\Omega, \mathcal{F})$. A random variable is then a measurable function from $(\Omega, \mathcal{F})$ to a measurable range space $(\mathcal{X}, \mathcal{G})$, most commonly to $(\mathbf{R}, \mathcal{B})$ (the real numbers with the Borel subsets). The probability that a random variable X takes a value within a subset $A \in \mathcal{G}$ is then given by

$$\mathbf{P}[X \in A] := \mathbf{P}\{\omega \in \Omega : X(\omega) \in A\}. \tag{1}$$

(If Ω is finite or countable, then we can let $\mathcal{F}$ be the collection of all subsets of Ω, and simply take $\mathbf{P}[X \in A] = \sum \{\mathbf{P}(\omega) : X(\omega) \in A\}$, i.e., define probabilities by a simple sum over individual points.)

Individual random variables can be studied in terms of their probability distributions $\{\mathbf{P}[X \in A] : A \in \mathcal{G}\}$, including such familiar notions as binomial, normal, exponential, Poisson, and other univariate distributions, perhaps extended to distributions of functions h of random variables as $\{\mathbf{P}[h(X) \in A] : A \in \mathcal{G}\}$.

The situation gets more intricate when there are multiple random variables defined jointly on the same underlying space $(\Omega, \mathcal{F}, \mathbf{P})$. If X_1 and X_2 are two such random variables, then we can define their joint probability distribution by

$$\mathbf{P}[X_1 \in A, X_2 \in B] = \sum \{\mathbf{P}(\omega) : X_1(\omega) \in A \text{ and } X_2(\omega) \in B\}, \tag{2}$$

for any $A, B \in \mathcal{G}$. This leads to questions about the relationship between the random variables.

The simplest case is independent random variables, meaning random quantities which do not affect each other, or formally for which

$$\mathbf{P}[X_1 \in A, X_2 \in B] = \mathbf{P}[X_1 \in A]\ \mathbf{P}[X_2 \in B] \tag{3}$$

for all $A, B \in \mathcal{G}$. More generally, a collection $\{X_i\}_{i \in I}$ of random variables are all independent if for all subsets $I_0 \subseteq I$, and all choices $\{A_i\}_{i \in I_0}$ of elements of $\mathcal{G}$, we have

$$\mathbf{P}[X_i \in A_i \text{ for all } i \in I_0] = \prod_{i \in I_0} \mathbf{P}[X_i \in A_i]. \tag{4}$$

Related, a collection $\{X_i\}_{i \in I}$ is identically distributed if for each fixed $A \in \mathcal{F}$, the value $\mathbf{P}[X_i \in A]$ is the same for all $i \in I$.

If a sequence $X_1, X_2, \ldots$ of random variables is both independent and identically distributed (i.i.d.), then many results are known. For example, the law of large numbers says that the sample averages $\frac{1}{n}\sum_{i=1}^{n} h(X_i)$ converge to their common mean $\mu := \mathbf{E}[h(X_i)]$, and the central limit theorem says $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[h(X_i) - \mu]$ converges in distribution to a Normal $(0, \sigma^2)$ distribution if $\sigma^2 := \mathbf{E}[(h(X_i) - \mu)^2] < \infty$. These two fundamental theorems, and many others, give a fairly rich and detailed picture of the behavior of independent random variables.

3 Markov Chains

As mentioned, Markov provided a generalization of the independence assumption. In particular, a sequence $X_0, X_1, X_2, \ldots$ forms a time-homogeneous Markov chain if there is a fixed transition kernel $\{P(x, A)\}_{x \in \mathcal{X}, A \in \mathcal{G}}$ with the property that

$$\mathbf{P}[X_n \in A | X_0, X_1, \ldots, X_{n-1}] = P(X_{n-1}, A), \quad A \in \mathcal{G}. \tag{5}$$

That is, given the history up to time $n - 1$, the probabilities for the value at time n depend only on the most recent value X_{n-1}, in an explicit way given by the transition kernel P.

For a concrete example, suppose $\mathcal{X} = \{1, 2, 3, \ldots\}$, with $\mathcal{G}$ the collection of all subsets of $\mathcal{X}$, and with transition kernel P defined by $P(x, \{x + 1\}) = 1/4$ for all $x \in \mathcal{X}$, $P(x, \{x - 1\}) = 1/2)$ for all $x \geq 2$, $P(x, \{x\}) = 1/4$ for all $x \geq 2$, and $P(1, \{1\}) = 3/4$, together with additivity. This Markov chain has, at each time, a probability 1/4 of increasing by 1, a probability 1/4 of staying the same, and a 1/2 probability of decreasing by 1 (unless it is already at the state 1, in which case the probability 1/2 is added to staying at 1); see Fig. 1.

Over time, a Markov chain might settle into a "stable pattern," or more precisely a stationary distribution. A probability distribution π on $(\mathcal{X}, \mathcal{G})$ is called *stationary* for the Markov chain $\{X_n\}$ if it has the property that whenever X_n has the distribution π, then X_{n+1} will also have the distribution π. That is,

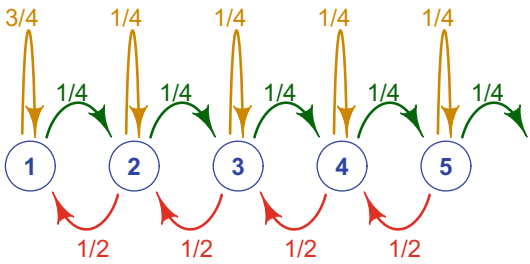


Fig. 1 The transition probabilities of the Running Example Markov chain

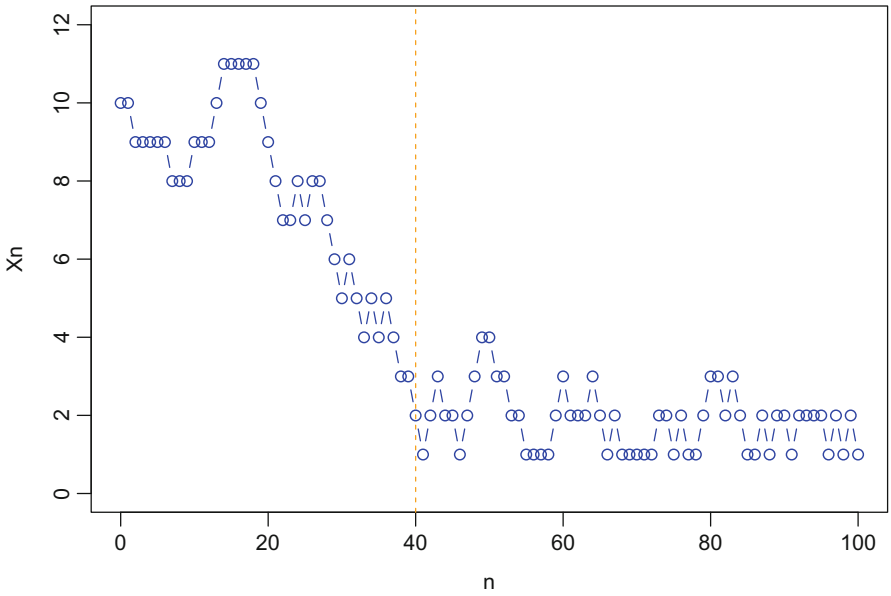


Fig. 2 A typical realization of the Running Example Markov chain

$$\int_{x\in\mathcal{X}} \pi(dx)P(x,A) = \pi(A). \tag{6}$$

On a discrete space, this can be written as

$$\sum_{x\in\mathcal{X}} \pi\{x\}P(x,A) = \pi(A). \tag{7}$$

Intuitive, the probability distribution π is invariant for the chain, i.e., its probabilities do not change while the chain runs.

The importance of stationary distributions is the Markov Chain Convergence Theorem which says that, under the mild assumptions of irreducibility and aperiodicity (nearly always satisfied and not considered further here), a chain will eventually *converge* to its stationary distribution. That is, if a Markov chain has a stationary distribution π, and is irreducible and aperiodic, then for all $A\in\mathcal{G}$,

$$\lim_{n\to\infty} \mathbf{P}[X_n\in A] = \pi(A). \tag{8}$$

Informally, the chain will settle down into its stationary distribution in the long run.

To see this more concretely, consider the above Running Example. Figure 2 shows a typical realization of that Markov chain. Although it starts with larger

values, from approximately time $n = 40$ onwards, it appears to be in a stable pattern, visiting approximately the same states with approximately the same probabilities.

This "settling down" Markov chain convergence property is enough to prove a law of large numbers and central limit theorem for Markov chains, similar to the i.i.d. case but under stronger assumptions; for a modern treatment, see, e.g., the monograph (Meyn and Tweedie 1993). Thus, the time-homogeneous Markov assumption helps to re-capture some of the power and utility of i.i.d. sequences, but in a more general context.

4 Markov Chain Monte Carlo (MCMC)

When they were first introduced, Markov chains served as models of certain natural scientific (and even social) phenomenon. However, in the latter half of the twentieth century, it was slowly realized that Markov chains could be used in an entirely different way, as computational algorithms.

The basic idea of a Monte Carlo algorithm is simple. Suppose π is a "target" probability distribution on $(\mathcal{X}, \mathcal{G})$, and $h : \mathcal{X} \to \mathbf{R}$ is some measurable function, and you want to estimate the expected value

$$\mu := \mathbf{E}_\pi(h) := \int_{x \in \mathcal{X}} h(x)\pi(dx) \tag{9}$$

(so, on a discrete space, $\mu = \sum_{x \in \mathcal{X}} h(x)\pi(x)$). In principle, μ could be calculated directly by an integral or sum, but in a complicated application on a high-dimensional state space that might be quite infeasible. However, if you could simulate random variables $X_1, X_2, \ldots, X_M$ which are i.i.d. with distribution π, then you could estimate μ by

$$E := \frac{1}{M} \sum_{i=1}^{M} h(X_i). \tag{10}$$

Indeed, the law of large numbers says that as $M \to \infty$, the estimator E will converge to μ.

The problem is that, if $\mathcal{X}$ is high-dimensional and π is complicated, there is no feasible computer program to simulate independent random variables $X_1, X_2, \ldots$ which each have the distribution π. However, it turns out to be surprisingly easy to simulate a *Markov chain* $X_0, X_1, X_2, \ldots$ which has π as a stationary distribution and will thus converge to π in distribution. This gives rise to Markov chain Monte Carlo (MCMC) algorithms which simulate a Markov chain $X_0, X_1, X_2, \ldots$ which converges to π and then estimate μ by

$$E := \frac{1}{M - B} \sum_{i=B+1}^{M} h(X_i). \tag{11}$$

If the burn-in period B is large enough that the distribution of X_B is approximately π, and $M - B$ is large enough that the Markov chain law of large numbers approximately holds, then this estimator E will again provide a good approximation to μ.

5 The Metropolis Algorithm

While there are many ways to construct Markov chains which leave π stationary, the oldest and simplest one is the Metropolis algorithm (Metropolis et al. 1953). Assume the state space $(\mathcal{X}, \mathcal{G})$ has a reference measure ρ (e.g., counting measure on a discrete state space, or Lebesgue measure on $\mathbf{R}$), with respect to which π has a density function f so that $\pi(dx) = f(x)\rho(dx)$. The algorithm uses proposal distributions $Q(x, dy) = q(x, y)\rho(dy)$ for $x, y \in \mathcal{X}$, which are symmetric (i.e., $q(x, y) = q(y, x)$). (For example, on a discrete state space, $Q(x, \cdot)$ might give probability 1/2 to each of $x + 1$ and $x - 1$. Or, on the real line, $Q(x, \cdot)$ might correspond to a Normal(x, v) distribution for some fixed variance parameter $v > 0$. Importantly, $Q(x, \cdot)$ should be easy to sample from on a computer.)

The algorithm then proceeds as follows. First choose some initial state $X_0 \in \mathcal{X}$. Then, iteratively for $n = 1, 2, \ldots$, given $X_{n-1} \in \mathcal{X}$, generate X_n as follows. First, generate a *proposal* state Y_n sampled from the probability distribution $Q(X_{n-1}, \cdot)$ and independently generate a random variable U_n sampled from the uniform (Lebesgue) measure on the interval [0,1]. Then define X_n by

$$X_n = \begin{cases} Y_n & U_n < f(Y_n)/f(X_{n-1}) \\ X_{n-1} & \text{otherwise} \end{cases} \tag{12}$$

That is, if $U_n < f(Y_n)/f(X_n)$, then the proposed new state Y_n is "accepted" so we set $X_n = Y_n$; otherwise Y_n is "rejected" so we keep $X_n = X_{n-1}$. Incredibly, this simple algorithm, which only barely mentions π via its density f in the accept/reject rule, actually makes π a stationary distribution to which it will converge:

Theorem 1 *The above Metropolis algorithm defines a Markov chain $\{X_n\}$ which has π as a stationary distribution. Furthermore, if the chain is irreducible and aperiodic, then it will converge in distribution to π.*

Proof. For $y \neq x$, the only way for the chain to move from x to y (or to a neighborhood dy of y) is to first propose such a move and then accept it. Thus, for $y \neq x$,

$$\begin{aligned} P(x,dy) &:= \mathbf{P}[X_n \in dy \,|\, X_{n-1} = x] \\ &= \mathbf{P}[Y_n \in dy \,|\, X_{n-1} = x]\ \mathbf{P}[U_n < f(Y_n)/f(x)] \\ &= Q(x,dy)\ \min[1, f(Y_n)/f(x)] \\ &= q(x,y)\rho(dy)\ \min[1, f(y)/f(x)]. \end{aligned} \tag{13}$$

It follows that the bivariate measure

$$\begin{aligned}\pi(dx)P(x,dy) &= \pi(dx)\ \mathbf{P}[X_n \in dy \,|\, X_{n-1} = x] \\ &= f(x)\rho(dx)q(x,y)\rho(dy)\ \min[1, f(y)/f(x)] \\ &= q(x,y)\ \min[f(x), f(y)]\rho(dx)\rho(dy).\end{aligned} \tag{14}$$

Since $q(x,y) = q(y,x)$, this measure is symmetric upon swapping x and y, i.e., $\pi(dx)\ P(x,dy) = \pi(dy)\ P(y,dx)$, which says that $\{X_n\}$ is a reversible Markov chain. (The above argument only shows this for $y \neq x$, but it is immediately true for $y = x$ too.) It follows that if X_{n-1} has the distribution π, then

$$\begin{aligned}\mathbf{P}[X_n \in A] &= \int_{x \in \mathcal{X}} \pi(dx)\ \mathbf{P}[X_n \in A \,|\, X_{n-1} = x] \\ &= \int_{x \in \mathcal{X}} \int_{y \in A} \pi(dx)\ P(x,dy) \\ &= \int_{x \in \mathcal{X}} \int_{y \in A} \pi(dy)\ P(y,dx) \\ &= \int_{y \in A} \pi(dy)\ P(y,\mathcal{X}) \\ &= \int_{y \in A} \pi(dy)\ \ (1) = \pi(A),\end{aligned} \tag{15}$$

so X_n also has the distribution π. Therefore, π is a stationary distribution.

The second statement follows immediately from the first one and the Markov Chain Convergence Theorem. □.

In summary, the Metropolis algorithm provides a simple method – propose a new state symmetrically, then accept or reject it according to a simple rule – which guarantees that π is a stationary distribution, to which (under very mild conditions) the chain will converge. This is what makes MCMC possible. For a graphical simulation of the Metropolis algorithm in action, see the web page at (Rosenthal 2020).

6 MCMC in Practice

To see how the Metropolis algorithm actually works, suppose the state space is $\mathcal{X} = \{1, 2, 3, \ldots\}$, so ρ is counting measure, and suppose the target probability distribution given by $\pi(\{x\}) = 2^{-x}$ for $x \in \mathcal{X}$. Suppose we choose the simple proposal distributions $Q(x,\ \{x - 1\}) = Q(x,\ \{x + 1\}) = 1/2$. Then, at each iteration, our algorithm proposes a new state $Y = X_{n-1} \pm 1$ which increases or decreases the current state by one (with probability 1/2 each), and then with probability $\min[1,\ \pi(Y)/\pi(X_{n-1})] = \min\left[1,\ 2^{-Y}/2^{-X_{n-1}}\right]$, it accepts the proposal and sets $X_n = Y$; otherwise it rejects the proposal and sets $X_n = X_{n-1}$.

So, what are the resulting transition probabilities? Well, since $\pi(\{0\}) = 0$, proposed moves from state 1 to state 0 are always rejected. However, for $x \geq 2$, since $\pi(\{x-1\}) > \pi(\{x\})$, proposed moves from $x-1$ to x are always accepted. Meanwhile, for any $x \in \mathcal{X}$, proposed moves from x to $x+1$ are accepted with probability $\min[1, 2^{-(x+1)}/2^{-x}] = 1/2$. Further, whenever a move is rejected, then the chain stays at x. A careful inspection shows that these transition rules are exactly the same as those in the diagram of Fig. 1. That is, this Metropolis algorithm exactly coincides with our Running Example, which therefore by Theorem 1 has stationary distribution π, given by $\pi(\{x\}) = 2^{-x}$ for $x \in \mathcal{X}$ as above.

Hence, a typical realization of this Metropolis algorithm will be something like that in Fig. 2. In particular, it might start at some larger state (e.g., $X_0 = 10$), but after some burn-in period B, it will send to settle into a pattern of sampling the different states with probabilities roughly equal to the target stationary distribution π. So, as above, if the burn-in period B and total run length M are both large enough, then $\frac{1}{M-B}\sum_{i=B+1}^{M} h(X_i)$ will be a good estimate of the expected value $\mu := \mathbf{E}_\pi(h)$.

For MCMC algorithms to be successful, we need to choose the burn-in value B large enough that the distribution of X_B is approximately equal to π. While there have been some attempts to provide precise mathematical bounds on this burn-in time (see, e.g., (Rosenthal 2002) and the references therein), such bounds cannot usually be obtained. Instead, convergence is usually assessed empirically, by monitoring the output from the Markov chain and using statistical or time-series approaches to decide if the chain seems to have reached a steady state. For example, for chain output as in Fig. 2, it might be determined empirically that roughly $B = 40$ iterations (the orange dotted line) are sufficient to approximately reach stationarity.

Of course, the distribution π in our Running Example is so simple that it can easily be sampled without using MCMC. However, MCMC is indeed very commonly required in more complicated situations. In particular, it is very often needed for Bayesian inference. The basic Bayesian paradigm is as follows. We wish to estimate some unknown parameters θ, given some observed data $D = d_0$, a prior distribution $\{\mathbf{P}[\theta \in A]\}_A$, and a statistical likelihood model $\{\mathbf{P}[D \in A|\theta]\}_{\theta,A}$. The all-important posterior distribution, which tells us the final probability distribution of the unknown parameters θ given the observed data and the model, is then given (using Bayes' Rule) by

$$\pi(d\theta) := \mathbf{P}[\theta \in d\theta | D = d_0] = \frac{\mathbf{P}[\theta \in d\theta]\ \mathbf{P}[D = d_0|\theta]}{\mathbf{P}[D = d_0]}, \tag{16}$$

which is proportional to the prior $\mathbf{P}[\theta \in d\theta]$ times the likelihood $\mathbf{P}[D = d_0|\theta]$.

In principle, the formula (16) gives precise information about the nature of the posterior distribution, usually expressible as a posterior density function. That should be sufficient to compute posterior probabilities $\pi(A)$, or more generally posterior expectations $\mathbf{E}_\pi(h)$, by computing an integral (or sum) with respect to π as in (9). However, in practical statistical applications, the posterior distribution π is often very complicated and high-dimensional, so that these integrals cannot be computed directly, nor even well approximated by numerical integration, so their

computation was thought to be infeasible. This challenge presented a major obstacle to the use of Bayesian inference in practical statistical problems and limited it for most of the twentieth century primarily to theoretical and philosophical and foundational interest rather than to genuine application. However, it was realized around 1990 (Gelfand and Smith 1990; Tanner and Wong 1987; Tierney 1994) that MCMC algorithms such as the Metropolis algorithm can provide good estimates of $\mathbf{E}_\pi(h)$ as in (11), often quite easily with good accuracy and limited computational effort.

These MCMC algorithms are now routinely used in Bayesian statistics, which accounts for their extreme popularity. They have transformed the entire field of Bayesian statistics, from a theoretical curiosity to one of the most influential areas of statistical analysis, with significant applications to physics and finance and machine learning and so much more.

7 Conclusion

This chapter has explained how Markov chains, introduced in 1906 as a tractable generalization of independent random variables, found tremendous application many years later as a tool for statistical computation, which completely transformed statistics by making the application of Bayesian inference computationally feasible, thus allowing its application to so many modern fields of study.

Andrey Andreyevich Markov died one hundred years ago, in 1922. If he were alive today, he would surely be astonished and thrilled by the incredible amount of progress that has been achieved in statistical computation and so many other areas, as a direct result of that little idea that he proposed so long ago.

References

Brooks S, Gelman A, Jones G, Meng X-L (eds) (2011) Handbook of Markov chain Monte Carlo. Chapman & Hall, New York

Gelfand AE, Smith AFM (1990) Sampling based approaches to calculating marginal densities. J Am Stat Assoc 85:398–409

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741

Geyer C (1992) Practical Markov chain Monte Carlo. Stat Sci 7(4):473–483

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109

Markov AA (1906) Extension of the law of large numbers to dependent quantities (in Russian). Izv Fiz-Matem Obsch Kazan Univ (2nd Ser) 15:135–156

Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E (1953) Equations of state calculations by fast computing machines. J Chem Phys 21:1087–1091

Meyn SP, Tweedie RL (1993) Markov chains and stochastic stability. Springer-Verlag, London. Available at: http://probability.ca/MT/

Rosenthal JS (2002) Quantitative convergence rates of Markov chains: a simple account. Electron Commun Probab 7(13):123–128

Rosenthal JS (2020) Metropolis algorithm javascript program. Available at: http://probability.ca/jeff/js/metropolis.html

Tanner MA, Wong WH (1987) The calculation of posterior distributions by data augmentation (with discussion). J Am Stat Assoc 82:528–550

Tierney L (1994) Markov chains for exploring posterior distributions (with discussion). Ann Stat 22:1701–1762

The Use and Abuse of Probability in Evolutionary Biology

Jason Rosenhouse

Contents

1 Introduction 1384
2 Biological Terminology 1386
3 The Basics of Darwinian Evolution 1386
4 Evolution as Part of the "Probabilistic Revolution" 1388
5 Biometry Versus Mendelism 1391
6 Some Equations of Population Genetics 1395
7 Coalescent Trees and Phylogenetic Reconstruction 1398
8 The Abuse of Probability by Anti-Evolutionists 1401
9 Conclusion 1407
References 1407

Abstract

Darwin's publication of *The Origin of Species* in 1859 situated biology firmly within the "probabilistic revolution" in the history of science. One consequence of this was an increased emphasis on probability and statistics in mathematical practice. The earliest attempts to test Darwin's theory relied on statistical analyses of phenotypic traits in various species. With the rediscovery of Mendel's work in the late nineteenth and early twentieth centuries, emphasis shifted to the development of genetical models based on combinatrics and discrete probability. The interplay between evolution and probability continues in the present, for example, in the development of coalescent theory and the study of Markov models on trees. We will survey this history, and we will discuss a few of the ideas underlying some of these mathematical models. We will also consider, and reject, certain probabilistic arguments presented by critics of evolutionary theory.

J. Rosenhouse (✉)
James Madison University, Harrisonburg, VA, USA
e-mail: rosenhjd@jmu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_107

Keywords

Evolution · Evolutionary biology · Probability · Darwin · Population genetics · Coalescent theory · Intelligent design

1 Introduction

Charles Darwin presented his theory of evolution in *The Origin of Species*,[1] published in 1859. In so doing, he fundamentally transformed the practice of biology. Prior to Darwin, biology was mostly a descriptive discipline concerned with dissecting organisms and with arranging species into coherent systems of classification. After his work, biology was a proper theoretical science, one worthy of a place right alongside physics and chemistry. In the light of evolution, the patterns of anatomical similarities and differences among species were no longer brute facts, but were instead the logical outcome of a comprehensible historical process. Classification now had an objective basis in the one true tree of life and no longer had to be based on arbitrary human preferences.[2] The empirical and philosophical questions raised by evolution gradually initiated a panoply of research programs that continue to bear fruit in the present.

There has been extensive scholarship exploring the historical effect of evolution on biological practice.[3] Less well-explored is the effect of Darwin's work on mathematical practice. I do not mean by this that evolution altered the way mathematicians viewed the philosophical underpinnings of their discipline. Rather, I mean that evolution offered surprising new opportunities for mathematical modeling, especially with regard to probability theory and statistics.

Darwin himself was famously ambivalent about mathematics, and the *Origin* contains no equations or mathematical models. That notwithstanding, the eventual ascent of his ideas to the centerpiece of biological thought was in large measure the result of abstract, probabilistic modeling undertaken first by Ronald Fisher and later by Sewall Wright and J. B. S. Haldane, in the early twentieth century. Probability is a major part of modern evolutionary theory, playing a central role in models of short-term allelic change in gene pools, as well as in methods of phylogenetic reconstruction.

[1] The full title of Darwin's book is *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. Since this title is rather long, it is customary to refer to it simply as *The Origin of Species*. We should note that in this context, "races" does not refer to human races, but instead refers to local populations of organisms.

[2] The only diagram in *The Origin of Species* illustrates Darwin's idea of a tree of life. The tree metaphor is based on the idea that genes are only transferred "vertically" from ancestors to descendants. Recent research has suggested that "horizontal" gene transfer, meaning transfers of genetic material among contemporary populations, was prevalent in the earliest stages of evolution. This suggests that the tree metaphor is inapt for early evolution (Martin 2011). However, most researchers continue to regard the tree metaphor as appropriate for most of natural history.

[3] Among many excellent references on the scientific, philosophical, and theological legacy of Darwin's ideas, I recommend the books by Ruse (1979), Bowler (2003), and Larson (2004).

Darwin's theory implied that chance was an ineliminable part of natural history and, therefore, that probability and statistics were essential to understanding it. From the mathematical side, many innovations in probability and statistics arose from a desire to model evolutionary processes.

However, for as long as there have been evolutionists, there have also been anti-evolutionists, meaning those who believe that no fully naturalistic process can explain the broad sweep of natural history and that it is necessary to appeal in some way to the actions of an intelligent designer. They make a variety of arguments in defense of their views, and some of these arguments rely explicitly on probability. Scientists are mostly dismissive of these arguments, finding them to be both biologically and mathematically fallacious.[4] I agree with this view. However, their superficial sophistication and their cachet among lay audiences in cultural disputes about the proper relations between church and state make them worth taking seriously. Science and mathematics should not be seen merely as ivory tower disciplines, and when their methods are abused to serve political agendas, it behooves scholars to take notice.[5]

This chapter is organized as follows: Section 2 provides definitions for certain key terms that may be unfamiliar to some readers. Section 3 describes those elements of Darwin's theory of evolution relevant to our later discussions. Section 4 situates Darwin's work in the broader historical context of the so-called probabilistic revolution and notes some interesting parallels between Darwinian natural selection and the introduction of statistical techniques in physics.

The decades following the publication of the *Origin* saw the emergence of two schools of thought on how to test Darwinian hypotheses. The so-called "biometricians" applied statistical methods to the study of naturally occurring variations, while the "Mendelians" preferred methods drawn from combinatorics and discrete probability. This dichotomy is explored in Sect. 5. While both approaches had merit, it was the methods of the Mendelians that had more relevance specifically to evolution. Their work led to the modern science of population genetics, and we present some of the basic models and equations of this discipline in Sect. 6. Section 7 presents a brief discussion of more modern uses of probability in evolutionary studies. Finally, we discuss the probabilistic arguments of evolution's critics in Sect. 8.

[4] In 2015, the Pew Research Center released polling data on the views of members of the American Association for the Advancement of Science (AAAS), one of the main professional societies for scientists in America. They found that 90% of all AAAS members agreed that human evolution was due to natural processes. This number ticked up to 92% for working Ph.D. biomedical scientists. A majority of the rest believe that evolution was "guided by a supreme being," but this should not be taken to imply that they agree with the specific anti-evolutionist arguments I will discuss. Moreover, we could note that thousands of research papers in evolutionary biology appear every year, compared to almost none that are supportive of creationism, intelligent design, or any other non-naturalistic approach to natural history.

[5] The cultural, sociological, and religious aspects of evolution, especially in the context of disputes over public school science curricula in the United States, are beyond the scope of this article. However, some useful references in this regard are the books by Nelkin (1982), Forrest and Gross (2004), Scott and Branch (2006), Miller (2008), Scott (2009), Berkman and Plutzer (2010), Ravitch (2011) and Rosenhouse (2011).

A short survey article of this sort must inevitably begin with an apology for everything that has been left out. Whole books get written on the subject matter of any of the major sections of this paper. It is hoped, however, that these brief remarks will serve as an introduction to a fascinating chapter in the history of both mathematics and biology.

2 Biological Terminology

As we go along, it will be necessary to employ some of the standard terminology used by geneticists and evolutionary biologists. It will be convenient to have all the needed definitions in one place, to avoid frequent disruptions to our story later on. Readers familiar with basic biology can skip this section.

It is customary to make a distinction between *genotype* and *phenotype*. The genotype of an organism is its collection of genes. Its phenotype is the sum total of its observable characteristics. When we are discussing a population of interbreeding organisms, the collection of all genes found in any of the individuals is called the *gene pool* for that population. A particular version of a gene that appears in a specific organism is called an *allele*. The location on a chromosome where an allele resides is called the *locus* of that allele.

The two main mechanisms of evolutionary change are *natural selection* and *genetic drift*. In both cases, we imagine that an allele initially appears with some frequency in a population and that, at some later time, we find that frequency has changed. If the change in frequency occurred because the presence or absence of the allele affects the probability that an organism will survive long enough to reproduce, then we say the change resulted from natural selection. If instead the change occurred because of random sampling bias or some other stochastic factor, then the change is attributed to genetic drift.

If we sequence a particular genetic locus in a large number of interbreeding organisms, we invariably find many variations of the same gene. This variation is referred to as *genetic polymorphism*. A common scenario among sexually reproducing organisms is that an individual inherits one copy of a gene from its mother and one from its father. If the two copies are the same, then the individual is said to be *homozygous* for this gene. Otherwise, it is said to be *heterozygous*.

Finally, *phylogenetics* refers to the branch of evolutionary biology devoted to working out the evolutionary relationships among modern species, or possibly extinct species in the context of paleontology. Such relationships are typically expressed in the form of trees of descent.

3 The Basics of Darwinian Evolution

Darwin's theory begins with two empirical observations:

1. Species are comprised of individuals that vary from one another, and at least some of these variations are heritable.

2. Species tend to reproduce at a geometric rate, implying that their numbers would increase quickly if this tendency is not checked by environmental constraints.

We then draw two conclusions from these observations:

3. There is a struggle for existence among the individuals of a species. Some individuals will fortuitously possess variations that give them an advantage in this struggle, and they will tend to leave more offspring than those not possessing those variations.
4. These variations will accumulate through the generations, to the point where they alter the average physical characteristics of the species.

Darwin referred to the process through which favorable variations accumulate over time as "natural selection."[6]

None of these points is controversial, and they have been readily agreed to by everyone from Darwin onward. Items (1) and (2) are simple empirical facts, and items (3) and (4) seem like undeniable conclusions from them. The controversy did not arise until Darwin extended these ideas beyond what many of his contemporaries felt were their appropriate scope.

Darwin believed that natural selection was so powerful that over time the accumulation of small variations could lead to new species. He argued that all modern species arose through a process of descent with modification from more ancient species, which themselves arose from more ancient species still. Eventually, this process traces back to the origin of life, or perhaps to a small number of very ancient original life forms. Darwin made no attempt to explain the origin of life, but instead took as given some relatively simple form of ancient life. He argued that through all the eons in which that ancient life was evolving, natural selection was an especially important mechanism of organic change.

To persuade others of this conclusion, Darwin began *The Origin of Species* with a chapter about pigeons. He noted the success of human breeders in producing pigeon varieties by selectively breeding from among those possessing characteristics the breeder found desirable. This was meant as proof of concept for the notion that small, naturally-occurring variations could add up to significant change in the average traits of a species. This chapter was followed by several others of a theoretical nature, in which Darwin laid out the central concepts underlying the process of natural selection and responded to possible objections to his ideas.

This led to the heart of the book, in which Darwin laid out the evidence for evolution. His argument took the form of a consilience of inductions, meaning that he showed how evolution neatly accounted for the facts in several seemingly distinct

[6]For a quick overview of modern evolutionary biology, try the book by Charlesworth and Charlesworth (2003). The textbook by Kardong (2007) strikes a good balance between rigor and readability. The more advanced textbooks by Herron and Freeman (2014) and Futuyma and Kirkpatrick (2017) are also very helpful. It is a testament to Darwin's genius and clarity of thought that *The Origin of Species* (Darwin 1859) is still worth reading today. He anticipated questions that continue to inspire modern research, even as he lacked the tools in his own time to answer them.

branches of the life sciences. In particular, he pointed to facts drawn from the fields of biogeography, classification, morphology, and paleontology. In each case, he argued that the facts were best explained by a process of descent with modification, rather than through a process of separate or special creations.

These ideas raise a host of scientific and philosophical questions, and Darwin's critics were quick to point them out. Was natural selection really so powerful that it could transform one species into another? Could complex structures like wings or eyes actually arise through a fully natural process? Can a scientific theory legitimately be defended by a consilience of inductions in the absence of direct experimental confirmation? Does evolution entail nominalism over essentialism with regard to species? If evolution plays out through natural selection, then what room is there for teleology, final causes, or divine action?

These are all good questions, and all have been the subject of intense discussion from Darwin's time right through to the present. For our purposes, however, the most important part of his theory, and arguably the most radical, is that he placed chance right at the heart of natural history.

4 Evolution as Part of the "Probabilistic Revolution"

A modern mathematical reader will quickly note the probabilistic machinery appropriate to Darwin's theory. Evolution has the structure of a Markov chain. Roughly, the states of the chain are the gene pools of organisms and their environments. The transition probabilities are such that states genetically close to what already exist are more likely to arise in the next generation than those that are far away. The process is "memoryless" in the sense that the future evolution of a species depends only on its current state and not on the historical vagaries leading to that state. More specifically, we see aspects of a random walk, as when genetic drift and neutral mutations lead to evolutionary change in the absence of natural selection, as well as branching processes, as when a local population becomes isolated from its ancestral stock and eventually becomes a new species.

However, this sort of mathematical machinery did not exist in 1859. Probability itself was in a somewhat rudimentary state at that time and was still mostly concerned with combinatorial problems related to games of chance.[7] Philosophically, probability was dominated by the work of Laplace, who took an epistemic view of the subject. Having accepted the law-bound picture of the universe inherent in Newtonian physics, he believed that statements of probability could only be expressions of human ignorance. This view was considered so fundamental that he opened his *Philosophical Essay on Probabilities* by writing:

[7] Todhunter (1865) remains the definitive reference for the state of probability at this time.

> All events, even those which on account of their insignificance do not seem to follow the great laws of nature, are a result of it just as necessarily as the revolutions of the sun. In ignorance of the ties which unite such events to the entire system of the universe, they have been made to depend upon final causes or upon hazard, according as they occur and are repeated with regularity, or appear without regard to order; but these imaginary causes have gradually receded with the widening bounds of knowledge and disappear before sound philosophy, which sees in them only the expression of our ignorance of the true causes.[8]

Laplace did not present this as a viewpoint in need of defense, but simply as an obvious conclusion of then-modern science.

Given the prevalence of this view, you can imagine the response of many of Darwin's contemporaries to his theory, which featured two distinct roles for chance: favorable variations were said to arise fortuitously, and selection was understood as a statistical tendency of certain members of a population to leave more offspring than their competitors. This ran into philosophical opposition from those who deplored the idea of chance playing an explanatory role in natural history, as well as theological opposition from those who feared that Darwin was ruling out a role for divine providence in the origin of species.

"Chance" can be a challenging notion to pin down since it means different things in different contexts.[9] To say that some event occurred by chance might mean that it occurred with no cause at all, as is sometimes suggested in modern discussions of quantum mechanics. It might mean that the event did not result from the plan of an intelligent agent. Sometimes the intent is that the event was inherently unpredictable beforehand on account of the limitations on human knowledge. Still another meaning is that the event was the result of a large number of independent causal chains, as when we consider all of the things that had to happen just so for a traffic accident to occur.

Darwin's notion of chance was not really any of these.[10] He certainly did not mean the variations were uncaused. Rather, his intention was that the variations among organisms are random with respect to the needs of the organisms. He argued that we should not imagine that an organism senses that a particular variation would be useful and then through some inner drive manifests just that variation. In a passage from his book *The Variation of Plants and Animals Under Domestication*, he offered a useful analogy:

> [I] have spoken of selection as the paramount power, yet its action absolutely depends on what we in our ignorance call spontaneous or accidental variability. Let an architect be

[8] Laplace (1951, 3). For recent discussions of the philosophical interpretations of probability most appropriate to evolutionary biology, see the two papers by Millstein (2003, 2016).

[9] Eble (1999) provides a lucid discussion of different uses of chance specifically in the context of evolution and paleobiology.

[10] For a discussion of Darwin's understanding of chance in biology, see the book by Johnson (2015) and the paper by Beatty (2006). For discussions of chance in evolutionary biology more generally, see the anthology edited by Ramsey and Pence (2016), and the papers by Beatty (1984), Shanahan (1991), and Millstein (2000).

> compelled to build an edifice with uncut stones, fallen from a precipice. The shape of each fragment may be called accidental; yet the shape of each has been determined by the force of gravity, the nature of the rock, and the slope of the precipice – events and circumstances, all of which depend on natural laws; but there is no relation between these laws and the purpose for which each fragment is used by the builder. In the same manner the variations of each creature are determined by fixed and immutable laws; but these bear no relation to the living structure which is slowly built up through the power of selection.[11]

However, with regard to scientists' openness to chance as an explanatory factor in nature, there was a fortuitous aspect to the timing of *The Origin of Species*. The mid-nineteenth century is a period in scientific history referred to by some scholars as the "probabilistic revolution."[12] Researchers in a variety of disciplines were coming to realize that statistical and probabilistic techniques were not just useful, but essential to solving practical problems.

In an interesting historical coincidence, Darwin published his book just a year before James Clerk Maxwell published his seminal work on the probability distribution appropriate to the velocities of individual particles in large ensembles of gas molecules, which later became foundational in the development of statistical mechanics.[13] This is especially noteworthy since Darwinian evolution has some interesting parallels with statistical mechanics.

At the heart of statistical mechanics is the idea that order can arise from randomness. The predictable, law-like behavior of matter arises from the random motions and collisions among large ensembles of microscopic particles. This approach led to tremendous advances in our understanding of thermodynamics. For example, the second law holds that entropy will never spontaneously decrease in an isolated system. In practical terms, this means that the energy of an isolated system tends to assume more dissipated forms and becomes less available for doing work. In classical thermodynamics, this was simply an empirical observation that was believed in general because it was seen to hold in any specific system in which the principle could be tested. Statistical mechanics, with its distinction between macrostates and microstates, provides insight into why energy dissipation in isolated systems is to be expected.

Darwin was essentially doing for natural history what statistical mechanics had done for thermodynamics. He showed there was no need to reference the master plan of an omnipotent engineer to explain the existence of modern species. Instead, they could arise from a chaotic jumble of random variations passed through the sieve of natural selection. Moreover, taking this view allowed tremendous insight into other aspects of biology.

Philosophically, the scientists of Darwin's time might have been more willing to countenance chance explanations than were their counterparts in previous eras.

[11] Darwin (1868, 236).
[12] Krüger et al. (1987).
[13] Maxwell (1860a, b).

However, that still left open a central question: How can we put Darwinian natural selection to empirical tests?

5 Biometry Versus Mendelism

Darwin was largely successful in convincing the scientists of his time that evolution had occurred. The numerous lines of evidence he adduced from various branches of anatomy, coupled with the evidence from paleontology and biogeography, were sufficient to convince scientists that modern species shared a far higher degree of relatedness with one another than had previously been appreciated.[14]

However, Darwin was far less successful in convincing people that natural selection was the primary mechanism of evolution. Skeptics wondered if the small, random variations central to Darwin's theory could really accumulate to the point of creating new species. Moreover, the favored theory of heredity at the time held that the characteristics of the offspring were a literal blend of the characteristics of the parents. If this were correct, then small variations would quickly vanish when the organism possessing the variation mated with those who did not. Regression to the mean would quickly occur, making significant directional evolution impossible.

A rival explanation of evolutionary change relied on so-called sports, meaning offspring that exhibited large variations relative to their parents. According to this school of thought, it was the variations themselves that led to significant evolutionary change. Natural selection, to the extent it was relevant at all, served only to affect minor changes in the characteristics of species. Advocates of this view came to be known as "mutationists."[15]

Plainly, the impasse could only be broken through a careful, quantitative study of naturally occurring variation among individuals. However, this presented a problem since the physical basis of heredity was entirely unknown at that time. The very existence of genes, much less their chemical structure, would not be confirmed until the early twentieth century. Consequently, scientists could only study observable, phenotypic variations in organisms. This led to an area of research known as "biometry."

The general philosophy underlying this approach was clearly expressed by Walter Weldon, an especially prominent biometrician, in a commentary on one of his papers:

> The questions raised by the Darwinian hypothesis are purely statistical, and this statistical method is the only one at present obvious by which that hypothesis can be experimentally checked.[16]

[14] This history is recounted in Ruse (1979), Mayr (1991), and Morange (2021, 157–222), among many other references.

[15] Provine (1971) and Gigerenzer et al. (1989, Ch. 4) are especially useful references here.

[16] Weldon (1895, 381).

Elsewhere he elaborated on what this meant:

> It cannot be too strongly urged that the problem of animal evolution is essentially a statistical problem: that before we can properly estimate the changes at present going on in a race or species we must know accurately (a) the percentage of animals which exhibit a given amount of abnormality with regard to a particular character; (b) the degree of abnormality of other organs which accompanies a given abnormality of one; (c) the difference between the death rate per cent. in animals of different degree of abnormality of parents, with respect to any organ; (d) the abnormality of offspring in terms of the abnormality of parents and *vice versa*. These are all questions of arithmetic; and when we know the numerical answers to these questions for a number of species we shall know the direction and the rate of change in these species at the present day – a knowledge which is the only legitimate basis for speculations as to their past history and future fate.[17]

In Weldon's view, the main problem of evolutionary research was to determine, for some measurable trait in an organism, the extent of its variation around the population's mean and its degree of correlation with the same trait in the previous generation. In other words, how much variation was there and how heritable was this variation?

Reading Weldon's papers today, it is impossible not to be impressed by the sheer patience and attention to detail this work required. In an example we can take to be representative of biometrical research generally, Weldon examined variations in the carapaces of the abundant crab species *Carcinus moenas*. By taking measurements on thousands of samples and then comparing those measurements with those taken on ancestral samples by prior researchers, Weldon was able to measure the degree of selective advantage of certain carapace designs over others. Here is a typical passage:

> The belief in which the work was undertaken was, that the law of variation would be found throughout to be that of the ordinary probability equation and this belief was tested in the following way: – In each of the thirty-five groups, the arithmetic mean of the frontal breadths, and the mean of all the deviations from it, were determined; and from the "mean error" found in this way the modulus of the probability function was calculated. Then, by calling the mean of each group zero, and expressing the deviations from the mean in terms of the modulus, a number of curves were obtained, in each of which the modulus was unity and mean zero; a similar curve of adults was constructed, and the corresponding ordinate of all the thirty-six curves so obtained were added together.[18]

To clarify, "the ordinary probability equation" is what we today call the normal or Gaussian distribution. However, in Weldon's time this distribution was called the "law of errors," from its historical roots in studying measurement errors in experimentation. This explains Weldon's references to "mean errors." We should note, however, that this term has some biological significance as well. Prior to Darwin, the general belief was that species had immutable essences and that individual variation

[17] Weldon (1893, 329).

[18] Weldon (1894, 363).

represented a distraction from understanding that essence.[19] Darwin reversed that thinking since it was a consequence of his theory that species were ephemeral and that it was the variations themselves that mattered. That Weldon was still thinking of variations as "errors" is suggestive that he had not fully absorbed Darwin's way of thinking, even while trying to collect data in support of evolution by natural selection.

The theoretical development of statistical methods in the late nineteenth century owes a lot to biometry and especially to the work of its most distinguished practitioner, Karl Pearson. Among other contributions, Pearson developed several techniques for determining curves of best fit, as well as the "chi squared" test for statistical significance. Pearson published a series of 18 papers under the title "Mathematical Contributions to the Theory of Evolution," and believe me when I tell you they do not make for light reading.[20]

Biometricians focused on traits that varied continuously, such as height, weight, and shell length. They sided with Darwin against the mutationists by emphasizing the importance of selection acting on small variations. They produced examples, such as Weldon's crab studies, showing that this process could lead to directional change.

Starting around 1900, however, they faced a strong challenge from researchers focusing on discrete character traits. The challengers took the pioneering work of Gregor Mendel as their inspiration. Mendel was a Moravian monk who undertook some of the first systematic investigations of inheritance. He focused on varieties of peas, distinguishing, in particular, those that were wrinkled from those that were smooth. He cross-bred various combinations of pea plants, keeping careful track of the frequency of each variety in subsequent generations. He discovered that a trait might be absent from one generation, only to reappear later on. Data of this sort eventually led him to the now standard notions that there were particles of heredity (today called "genes," of course), that offspring received one copy of each gene from its parents, and that some genes were dominant and others recessive.

Mendel's work was largely ignored in his own time, but it was reintroduced to the scientific community in the early twentieth century by researchers such as Hugo De Vries. The sort of mathematics appropriate to understanding particulate inheritance of discrete traits differs from the sort appropriate to studying continuous variation. Contrary to Weldon's assertions about statistics, the new mathematics of inheritance was combinatorial and probabilistic.[21]

An early success in this vein is the Hardy-Weinberg law, named for the mathematician (Hardy) and geneticist (Weinberg) who published it independently in 1908.

[19] Mayr (1991, 40–42).

[20] See Stigler (2012) for a discussion of Pearson's major contributions to statistics and evolutionary studies.

[21] For a discussion both of this point and of the conflict between biometricians and Mendelians generally, see Provine (2001, Chs. 2–3) and Gigerenzer et al. (1989, 123–161). Presenting the history of this period as a stark contrast between two clearly defined schools of thought is inevitably a bit simplistic. For a more nuanced discussion of this history, see the paper by Shan (2021).

It is an example of a "one locus, two allele" model, meaning that we imagine a single genetic locus that is home to one of two alleles, which we shall denote A_1 and A_2. Suppose that A_1 and A_2 initially appear with frequencies p and $q = 1 - p$. We want to determine the frequencies of these alleles in the next generation, if we assume there is no migration, selection, or genetic drift.

Given our set-up, there are three possible genotypes at this locus, which we can denote by A_1A_1, A_1A_2, and A_2A_2. Let us assume that the population is very large and that mating is random with respect to this allele. The frequencies p and q can be interpreted as the probabilities that a randomly chosen allele from the population will be A_1 or A_2, respectively. In this case, it is easy to see that the frequencies of the three genotypes in the next generation will be p^2, $2pq$, and q^2, respectively. Note that this conclusion depended only on the initial frequencies of the two alleles and that these frequencies did not change from one generation to the next. Consequently, these frequencies will be constant down through the generations, at least until the environment changes so as to make one of our assumptions no longer valid.

Hardy published this in a letter to the editor that opened like this:

> I am reluctant to intrude in a discussion concerning matters of which I have no expert knowledge, and I should have expected the very simple point which I wish to make to have been familiar to biologists. However, some remarks of Mr. Udny Yule, to which Mr. R. C. Punnett has called my attention, suggest that it may still be worth making.[22]

Yule had muffed a question about the frequency of dominant alleles in Mendelian populations, and one can sense Hardy's frustration at having to ride in to set things right. Historically, it tells us something about the novelty of the combinatorial approach at this time that a distinguished geneticist like Yule could have erred on a point that is mathematically elementary.

The combinatorial approach of the Mendelians was initially seen as anti-Darwinian, in the sense that particulate inheritance seemed to support discontinuous evolution through the appearance of sports and not gradual evolution through the accumulation of small variations. However, the work of Ronald Fisher, culminating in his 1930 book *The Genetical Theory of Natural Selection*, reconciled Darwinian gradualism with particulate inheritance. Fisher's mathematical modeling showed that the biometrical emphasis on continuous variation was entirely compatible with Mendelian combinatorics.

Fisher's work was supplemented with additional efforts from Sewall Wright and J. B. S. Haldane, and the modern science of population genetics was born. This body of work was seen as putting natural selection on a firm theoretical footing.[23] Their work made it clear that even small variations conferring tiny reproductive advantages on their possessors would eventually spread through the population, under a variety of plausible assumptions about nature.

[22] Hardy (1908, 49).

[23] See Dobzhansky (1951) and Huxley (2010).

6 Some Equations of Population Genetics

Evolutionary biology addresses big questions about human origins and our place in the natural world. This is why it is sometimes perceived as intruding on religion's territory and why it receives so much attention from the general public.

However, the quotidian concerns of professional biologists are generally more mundane than this. The grand sprawl of natural history does not really lend itself to mathematical modeling, forcing theoreticians to focus instead on the short-term fate of specific alleles over a few generations. It was precisely this sort of modeling that was pioneered by Fisher, Haldane, and Wright, and it remains a focus of attention for researchers today.

Biologist John Gillespie offers the following humorous description of population genetics:

> Population geneticists spend most of their time doing one of two things: describing the genetic structure of populations or theorizing on the evolutionary forces acting on populations. On a good day, these two activities mesh and true insights emerge.[24]

On the theoretical side, the goal is to develop models reflecting various assumptions about the environment and mating habits of a hypothetical species. We then use these models to prove theorems about the likely fate of a given allele, or perhaps a collection of alleles. In these models, there is always a huge trade-off between biological plausibility and mathematical tractability. Evolutionary change in gene pools is governed by so many factors that our models can easily come to look like alphabet soup. That notwithstanding, it is impressive how often biologically simplistic models still turn out to be useful. To experience some of the flavor of this field, let us derive a standard result.[25]

For our starting point, we consider a single genetic locus at which two alleles, still denoted A_1 and A_2, reside. We assume that A_1 initially appears with frequency p and that A_2 appears with frequency $q = 1 - p$. Our interest is in the frequency of A_1 as the generations go forth.

Now, there are four factors that might affect the frequency of A_1 through the generations: migration, mutation, genetic drift, and natural selection. A proper model would contain variables for all of them, but to keep things tractable we will focus primarily on natural selection. We will assume that mating is random with respect to A_1 and A_2, meaning that individuals choose their mates without any regard for their genotype at this locus.

Recall that genetic drift refers to changes in gene frequency that result from stochastic causes. It is intuitively clear, and easy to prove mathematically, that genetic drift will be a more powerful force in small populations than in

[24] Gillespie (1998, 1).

[25] I have based my presentation primarily on the book by Gillespie (1998) and the article by Okasha (2012). The textbook by Hamilton (2021) is also very lucid.

large ones.[26] For this reason, when modeling natural selection, it is customary to assume the population is so large that it is effectively infinite.

Now, with two alleles we have the three genotypes A_1A_1, A_1A_2, and A_2A_2. We assume that initially the three genotypes are at their Hardy-Weinberg frequencies of p^2, $2pq$, and q^2, respectively. The three phenotypes each have a certain probability of surviving long enough to reproduce, and we will denote these probabilities by w_{11}, w_{12}, and w_{22}, respectively. Each probability individually is said to represent the "fitness" of its associated genotype.

We can also define the mean fitness of the population by the formula

$$\overline{w} = p^2 w_{11} + 2pqw_{12} + q^2 w_{22}.$$

In probabilistic terms, this is the expected fitness of a randomly chosen individual in the population.

The number of successful gametes produced by each genotype is proportional to its fitness. If the population size is N, then the total number of successful gametes produced is $N\overline{w}$. If we let p' denote the frequency of A_1 among the offspring, then we can write

$$p' = \frac{Np^2 w_{11} + \frac{1}{2}(2Npqw_{12})}{N\overline{w}} = \frac{p^2 w_{11} + pqw_{12}}{\overline{w}}.$$

If this is unclear, note that the A_1A_1 genotype only produces A_1 gametes, while the A_1A_2 genotype produces A_1 and A_2 gametes in equal number.

Finally, we denote the change in frequency of A_1 due to selection by $\Delta_s p$. We compute

$$\Delta_s p = p' - p = \frac{p^2 w_{11} + pqw_{12} - p\overline{w}}{\overline{w}}.$$

Keeping in mind that $q = 1 - p$, this eventually simplifies to

$$\Delta_s p = \frac{pq[p(w_{11} - w_{12}) + q(w_{12} - w_{22})]}{p^2 w_{11} + 2pqw_{12} + q^2 w_{22}}. \tag{1}$$

We have summarized our findings in Table 1.

Equation (1) becomes easier to study if we employ the following notational convention: Scale the fitnesses of the three genotypes so that A_1A_1 has fitness 1 (implying that A_1A_2 and A_2A_2 have fitnesses w_{12}/w_{11} and w_{22}/w_{11}, respectively). We now let s denote the "selection coefficient" for the relative fitnesses of the homozygous genotypes, and we let h denote the "heterozygous effect," which measures the

[26] This is comparable to saying that in a small number of coin tosses the frequencies of Hs and Ts might differ greatly from the expected 50:50, but in the long run we expect the ratio to be very close to 50:50.

Table 1 Our set-up for analyzing the effect of selection on the frequency of an allele

Genotype	A_1A_1	A_1A_2	A_2A_2
Fitness	w_{11}	w_{12}	w_{22}
Frequency in newborns	p^2	$2pq$	q^2
Frequency after selection	$p^2w_{11}/\overline{w}$	$2pqw_{12}/\overline{w}$	$q^2w_{22}/\overline{w}$

fitness of the heterozygote relative to the selective difference between the homozygotes. With this notation, we can express the fitnesses of A_1A_1, A_1A_2, and A_2A_2 respectively, as 1, $1 - hs$, and $1 - s$. A positive value of s indicates that A_1 has a selective advantage over A_2, while a negative value indicates the reverse.

Equation (1) now simplifies to

$$\Delta_s p = \frac{pqs[ph + q(1-h)]}{\overline{w}}.$$

In this form, it becomes straightforward to analyze the effect of selection on the frequency of p given different assumptions about s and h.

By analyzing the various cases, one is led to the conclusion that natural selection will always alter the frequency of p in a direction that increases the mean fitness of the population. This result is known as the "fundamental theorem of natural selection."[27]

Models of this sort played a central role in the establishment of the so-called synthetic theory of evolution, or simply the "modern synthesis," in the 1940s. Prior to this theoretical modeling, many believed that the small genetic variations on which Darwin put so much emphasis would simply be invisible to natural selection. It was argued that in a typical case, whatever reproductive advantage such variations provided would be too small for selection to overcome other environmental forces. Population genetic models can be used to show that this objection is unfounded.

[27] The fundamental theorem has its own literature attached to it. One can make a stronger statement than mine regarding selection and mean fitness, but doing so requires defining further technical vocabulary, which for the purposes of this brief survey I preferred to avoid. The theorem was first presented by Fisher (1930), but his presentation was a bit muddled and hard to follow. Mathematically rigorous formulations of the theorem can be found in Price (1972) and Edwards (1994). For philosophical analyses of its meaning and significance, see Plutynski (2006) and Okasha (2008). For a more reader-friendly discussion of the theorem, see Grafen (2003). In light of our earlier discussion about statistical mechanics, it is worth noting that Fisher himself saw a connection between his theorem and thermodynamics:

> It will be noted that the fundamental theorem proved above bears some remarkable resemblances to the second law of thermodynamics. Both are properties of populations, or aggregates, true irrespective of the nature of the units which compose them; both are statistical laws; each requires the constant increase of a measurable quantity, in the one case the entropy of a physical system and in the other the fitness … of a biological population. (Fisher 1930, 36)

Even very small selective advantages will be sufficient for natural selection to notice, so to speak.[28]

Models of the sort we have considered in this section are sometimes said to comprise "classical population genetics." Let us now briefly consider a more modern approach to the subject.

7 Coalescent Trees and Phylogenetic Reconstruction

The study of heredity has made a lot of progress since the time of the Mendelians and the biometricians. Today we understand genes down to the molecular level. Researchers no longer have to make do with measurable phenotypic traits of organisms or with hypothetical entities existing only in mathematical equations. Instead, they can study evolution down to the level of individual base pair substitutions in DNA molecules.

The widespread availability of genetic sequence data has opened up new possibilities for population genetics. Whereas the classical approach looked forward in time to model the short-term evolution of allele frequencies, the modern approach looks backward to model the evolutionary history of polymorphism data collected in the present. The result of such an analysis is a rooted, binary tree. The tree's leaves represent the genetic sequence data in the present, while its root represents the most recent common ancestor. If we orient the tree vertically with the root on top, then we can follow time backward from the bottom to the top. We would then see initially disparate nodes gradually uniting, with the number of tree branches decreasing by one with each such unification. In the terminology of population genetics, each such unification is called a "coalescence," and the resulting diagram is called a "coalescent tree."

This is illustrated in Fig. 1, which shows two possible coalescent trees on a starting sample of size four. The two trees have different topologies, corresponding to different possible evolutionary histories for the sample. The subscripts on the time units to the right of the figure indicate the number of branches at that stage of the

[28] There is some debate among evolutionary biologists about the importance of mathematical population genetics. The founding document of the modern synthesis is the book by Huxley (1942), and he devotes considerable space to the work of Fisher, Haldane, and Wright. Nonetheless, Ernst Mayr, another of the major architects of the synthesis, famously derided the combinatorial approach as "beanbag genetics" and believed that the models were too simple to be worth anything (Mayr 1963, 263). This provoked a strong response from Haldane (1964), and the inevitable vast literature ensued. A few of the many references that are worth a look are Lynch (2007), Ewens (2008), Pigliucci (2008), and Crow (2009). For a readable discussion of the differences between Mayr and Haldane, see the book by Dronamraju (2011). Few nowadays would go as far as Mayr, and there is no question that the field of theoretical population genetics has much to be proud of. However, there is a more nuanced objection that the field places too much emphasis on genes and not enough on other factors like embryological development. We will not attempt to adjudicate that dispute here.

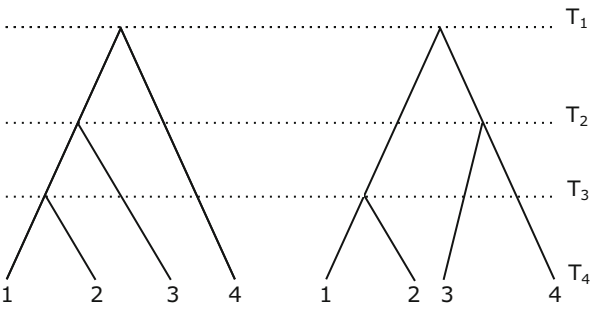


Fig. 1 Two coalescent trees on samples of size four

evolutionary process. We can interpret $T(k)$, where $1 \leq k \leq N$, to be a random variable that records the length of time during which the number of branches is k.

The goal of coalescent analysis is to develop mathematical models for understanding genetic polymorphisms in modern species. The sequence data of the present presumably arose through a lengthy and complex evolutionary process that was influenced by numerous variables. We want to be able to use coalescent analysis to look at the data and say something like, "Data of precisely this sort is what we would expect if the variables X, Y, and Z were at play in the evolutionary process." Clearly, any such model will have to be stochastic. There are so many chance factors in evolution that our conclusions about genealogical relationships, except in the most trivial cases, will have to be probabilistic.

In what follows, to keep our models mathematically tractable, we will assume there is no selection or migration. For our purposes, evolution proceeds solely by genetic drift, with possible mutations along the way.

Probability distributions on evolutionary trees were extensively studied in the 1960s, most notably by Motoo Kimura and Warren Ewens. They based their models on an analogy between genetic drift and Brownian motion. In a stochastic evolutionary process, random interactions among genes in one generation lead to further such interactions in later generations, leading to essentially chaotic changes in gene frequencies further down the line. This is roughly analogous to Brownian motion, in which the chaotic movements of microscopic particles in a medium are explained by a cascading sequence of collisions among other microscopic particles.

There is a rich mathematical theory underlying Brownian motion. This theory is based on diffusion equations, which are a subset of partial differential equations more generally. Even though we generally think of evolution as a discrete time process, it can nonetheless be effectively studied with continuous time models. Kimura and Ewens adapted the theory of diffusion equations to the needs of population genetics, thereby developing probabilistic models for evolutionary trees. These models were then tested against the results of Monte Carlo simulations using computers, which was very pioneering at the time.[29] Modern coalescent theory

[29] Dietrich (1996).

is largely based on adapting their models to look backward in time instead of forward.[30]

Let us see a standard, if somewhat basic, calculation in this field.[31] Our model assumes that all evolution is random, meaning that in any generation, any individual is as likely to reproduce as any other. This implies that as we go backward in time, any two lineages are as likely to coalesce as any other. Likewise, any tree topology will be as probable as any other as well.

If N is the number of endpoints in the present, then the probability that a pair of lineages coalesce one generation back in time is just $1/N$, and the probability that they remain distinct is therefore $1 - (1/N)$. The probability that they remain distinct for at least τ generations is then $(1 - (1/N))^{\tau}$.

If we scale time so that one unit corresponds to N generations, then the probability that two lineages stay distinct for more than t time units is

$$\left(1 - \frac{1}{N}\right)^{\lfloor Nt \rfloor} \to e^{-t},$$

as $N \to \infty$. With this scaling, the coalescence time is exponentially distributed with a mean of 1 in the limit.

We can extend this to k lineages. The probability that none of them coalesce in the previous generation is

$$\prod_{i=0}^{k-1} \frac{N-i}{N} = \prod_{i=1}^{k-1} \left(1 - \frac{i}{N}\right) = 1 - \frac{\binom{k}{2}}{N} + O\left(\frac{1}{N^2}\right),$$

and the probability that more than two do so is $O(1/N^2)$. If N is very large, then this last probability is negligible.

It is illuminating to ask for the expected time to the most recent common ancestor for the entire sample (as opposed to the ancestor for a subset of the initial sample). Our previous work implies that $T(k)$ is exponentially distributed with mean $2/k(k-1)$. Using $E(X)$ to denote the expected value of the random variable X, we now compute

$$E\left(\sum_{k=2}^{N} T(k)\right) = \sum_{k=2}^{N} E(T(k)) = \sum_{k=2}^{N} \frac{2}{k(k-1)} = 2\left(1 - \frac{1}{N}\right).$$

[30] Diffusion equations were first introduced into evolutionary analysis in Fisher (1922), but it was Ewens (1964) and Kimura (1964) who really developed the theory.

[31] The calculations that follow are taken from the video presentation by Nordborg (2018). See also the survey article by Felsenstein (2007) and the textbook by Wakeley (2009) for more about coalescents.

On the other hand, we also have $E(T(2)) = 1$. We conclude that the expected time during which there are only two branches is greater than half the total expected height of the tree.

A useful analogy is to picture a group of cannibalistic beetles in a box. In each beetle encounter, one beetle eats the other. When there are many beetles in the box, collisions will happen frequently and the number of beetles will quickly go down. When there is a small number of beetles in the box, collisions will be far more infrequent.[32]

Let us conclude this section by noting that coalescent theory is a subfield of phylogenetic reconstruction more generally. The grand project of evolutionary biology is to work out the one true tree of life that produced the modern riot of extant species. Trees are fundamentally combinatorial objects, and similar mathematical questions arise whether our data comes from microscopic DNA sequences or macroscopic data of the sort paleontologists might study. The study of Markov models on trees is very highly developed, and this theory is increasingly finding applications in the development of phylogenetic research.[33]

8 The Abuse of Probability by Anti-Evolutionists

With the previous section, we complete our brief survey of the relationships between probability theory and evolutionary biology. We have seen that, historically, the relationship has gone both ways. A desire to test Darwin's ideas led to progress in statistics, and probabilistic modeling was essential to placing evolutionary biology on a solid theoretical footing. The relationship continues into the present, with ever more sophisticated probabilistic methods being applied to the problems of phylogenetic reconstruction.

However, this is not the end of our story. More so than most branches of science, evolutionary biology has a political dimension, especially in the United States. Within certain faith communities, evolution is seen as contrary to religion, and this sometimes leads to conflicts over the science curriculum to be taught in public schools. Given this social relevance, and given that evolution's critics frequently employ probability theory in their discourse, it behooves us to take a look at their arguments, if only to understand why most scientists find them unconvincing.

Leaving aside religious considerations, evolution really does seem to run afoul of a powerful intuition that natural forces do not construct complex, functional systems, no matter how much time you give them. Our everyday experience is that machines break down and fall apart unless energy and intelligence are expended to maintain them, but evolution claims that organisms have become more complex over time. Mathematical anti-evolutionism is about trying to place that intuition on a more rigorous footing. Since probability theory is the branch of mathematics that

[32] This analogy is used by Nordborg (2018) who attributes it to Joseph Felsenstein.

[33] See Semple and Steel (2003) and Steel (2016).

quantifies our notions of what is likely and what is not, it is unsurprising that evolution's critics rely so heavily on it in their writing.

In this regard, a classic anti-evolutionist argument points to the emergence of complex proteins as the thing that is too improbable to countenance. Proteins can be viewed as long chains of amino acids in which each link is drawn from among 20 possibilities. This makes it possible to think of protein formation in combinatorial terms. Here is an especially clear version of the argument, as presented in a mainstream American periodical:

> The specificity of hemoglobin is described by the improbability of the specific amino acid sequence occurring by random chance. Such specificity is capable of exact calculation in the permutation formula
>
> $$P = \frac{N!}{n_1! \times n_2! \times n_3! \ldots \text{etc.}},$$
>
> where N is the total number of amino acids in hemoglobin (574); n_1, etc. are the number of separate kinds of amino acids; . . . In the case of hemoglobin, . . . the specific numerical value of the solution is $P = 10^{654}$. Thus, we can state that the improbability of hemoglobin occurring by random selection can be represented by the infinitely small number 10^{-654}, which means 10 divided by itself 654 times: as near to zero as one could consider.[34]

Arguments of this general sort are ubiquitous in anti-evolutionist literature.[35]

There are many variations on the basic theme, but the logic always conforms to the following general scheme:

1. Identify a complex biological structure, such as a specific gene or protein.
2. Model its evolution as a process of randomly selecting one item from a very large space of equiprobable possibilities.
3. Use elementary combinatorics to determine the size of the space, which we shall call S.
4. Conclude that the probability of the structure having evolved by chance is $1/S$, and assert that this is too small for evolution to be plausible.

I shall refer to this as the Basic Argument from Improbability (BAI).

Since proteins arise ultimately from genes, let us speak, for the moment, of "genotype space" as opposed to "protein space." Mathematically speaking, genotype space is a metric space equipped with a probability distribution. The metric is such that genotypes that are mutationally close to our starting point are assigned a smaller distance than those that are far away. The probability distribution assigns a higher probability to genotypes that are more fit than it does to genotypes that are less fit.

[34] Foster (1999, 80). We should note that 10^{-654} does not really mean to divide 10 by itself 654 times, but this is not the main error with this argument.

[35] Other examples can be found in Roth (1998, 69–70) and Meyer (2013, 173).

The BAI is fallacious because it completely ignores the effect of natural selection on the probabilistic and metric structures of genotype space. Evolution does not navigate genotype space by choosing points at random. Instead, it finds its starting point at the origin of life and then carries out a series of local searches in the neighborhood of wherever it happens to be. The probabilistic and metric structures guarantee that most of the space has a probability close to zero of ever arising, while fit genotypes close to our starting point have a high probability of arising. The probability distribution appropriate to genotype space is highly non-uniform, and this fact shows that the calculation put forth by the BAI is based on a hopelessly unrealistic model.[36]

Moving on, a more sophisticated probabilistic, anti-evolutionary argument has been put forward by mathematician William Dembski in a series of books and articles.[37] He defines a quantity he calls "complex, specified information (CSI)." His basic idea is that it is not improbability by itself that suggests design, but rather improbability coupled with a recognizable pattern. Any sequence of Hs and Ts is as unlikely as any other when we flip a coin 100 times, but 100 Hs, or a perfect alternation of Hs and Ts, show a clear pattern and would therefore suggest trickery of some sort. Any pattern of crags and grooves on a mountain is highly unlikely, but the faces on Mt. Rushmore also fit a pattern and therefore must be explained by intelligent design. Thus, "complex" means low probability, and "specified" means that we find a recognizable pattern. "Information" refers to any event or object whose origin we wish to explain.

As applied to evolutionary questions, the argument might play out like this: Choose a functional structure like the flagellum used for propulsion by some bacteria. It is complex in the sense that it is composed of many individual proteins that must be arranged in a precise way. It is specified in the sense that it functions in a way that is strongly analogous to an outboard motor on a boat. Thus, the flagellum is complex and specified and, therefore, must be designed.

If we are to treat this as a rigorous argument, then within Dembski's framework we must answer two questions: How do we carry out a meaningful probability calculation to establish complexity in the sense he intends? And how do we distinguish the design-suggesting patterns from those we impose on nature through excessive imagination? (For example, it is very easy to look at a fluffy, cumulus cloud and imagine that you see a dragon.) Let us see how Dembski tries to answer these questions.

[36] Since we are only interested in specifically probabilistic arguments, the response we have given is sufficient for our purposes. The question of whether a complex protein really can evolve in a gradual, stepwise manner is separate from the question of whether the BAI is a good argument. However, we should note that protein evolution has been a major area of scientific investigation for many decades. One of the main findings in this area is that standard evolutionary mechanisms are fully capable of exploring large swaths of protein space. See Yeates (2007), Ferrada and Wagner (2010), or Wagner (2014), for example.

[37] See Dembski (2002) for the most formal presentation of his ideas.

To justify the reduction of probability to combinatorics, Dembski relies on the notion of an "irreducibly complex" system. In this context, this idea is the brainchild of biochemist Michael Behe who defined an irreducibly complex system to be one in which several distinct parts work together to carry out some function, such that the removal of any one part causes the system to cease functioning. In his 1996 book *Darwin's Black Box*, Behe argued that such a system could not evolve gradually because the precursor systems would inevitably be missing some of the parts and therefore could not function. Since natural selection has no foresight, it will not preserve a non-functional system in the hope that later developments will somehow render it useful.

Dembski now explains the relevance of irreducible complexity to carrying out probability calculations:

> An irreducibly complex system is a discrete combinatorial object. Probabilities therefore naturally arise and attach to such objects. Such objects are invariably composed of building blocks. Moreover, these building blocks need to converge on some location. Finally, once at this location the building blocks need to be configured to form the object. It follows that the probability of obtaining an irreducibly complex system is the probability of originating the building blocks required for the system, multiplied times the probability of locating them in one place once the building blocks are given, multiplied times the probability of configuring them once the building blocks are given and located in one place.[38]

Dembski's argument was recently endorsed in an article published in the academic *Journal of Theoretical Biology*. Authors Steinar Thorvaldsen and Ola Hössjer present the argument like this:

> [D]embski proposes an equation based on three independent events: A_p: originating the building blocks (protein chains) of the protein complex . . ., A_l: localizing the building blocks in the same place, and A_c: configuring the building blocks correctly to form the complex. Then the probability of a protein complex is the multiplicative product of the probabilities of the origination of its constituent parts, the localization of those parts in one place, and the configuration of those parts into the resulting system (contact topology). This leads to the following estimate for the probability of a protein complex (PC) composed of N independent building blocks:
>
> $$\widehat{P}(A_{PC}) = \prod_{n=1}^{N} \left[P\left(A_p^{(n)} | \widehat{\theta}_p^{(n)}\right) \cdot P\left(A_l^{(n)} | \widehat{\theta}_l^{(n)}\right) \cdot P\left(A_c^{(n)} | \widehat{\theta}_c^{(n)}\right) \right],$$
>
> where $\theta_p^{(n)}, \theta_l^{(n)}$, and $\theta_c^{(n)}$ are the parameters involved in forming the protein chain, the localization, and the configuration of the nth building block.[39]

While this argument is an improvement over the BAI, it is not difficult to find grave errors in its formulation. The first is that it is not legitimate to treat an irreducibly complex system as a discrete, combinatorial object because there are

[38] Dembski (2002, 290–291).

[39] Thorvaldsen and Hössjer (2020, 7).

several plausible scenarios for how such a system could evolve gradually. For example, the removal of redundancy can lead to irreducible complexity, as when a scaffold supports an arch during its construction. When the arch is complete, the scaffold is removed, leaving behind an irreducibly complex structure (the arch cannot support itself without its capstone, and the capstone has nothing to rest on without the arch.) An interdependence of parts can also arise when one or more parts is not required when it is first introduced, but later becomes essential because of changes in the surrounding system.[40]

Moreover, modeling the evolution of a complex structure as a threefold process of origination, localization, and configuration is biologically unrealistic. Instead of combinatorial arrangements of proteins or other sorts of component parts, we should be thinking about the underlying genetic instructions leading to the finished system. However, the same instructions that dictate which proteins get produced also direct those proteins to specific locations and mandate the relationships among them. This shows that the three probabilities in Thorvaldsen's and Hössjer's equation are not independent and therefore cannot be multiplied.

Now let us see how Dembski deals with specification. To address the problem of distinguishing design-suggesting patterns from those we simply impose on nature, he argues that the pattern must be "detachable" from the object itself.

To see the basic idea, we imagine firing an arrow at the wall of a barn. After the arrow embeds itself in the wall, we could paint a small circle around it, thereby making it appear that we hit a small target.[41] This is a non-detachable pattern since

[40] Again, since our interest here is solely in probabilistic arguments, this response is sufficient for our purposes. The anti-evolutionist argument presupposes that an irreducibly complex system cannot evolve in a gradual, stepwise manner, but our hypothetical scenarios show that this is not true. However, we should note that these scenarios are not mere speculation, but seem to have played out in natural history. For example, my scenario about irreducible complexity arising as a result of the removal of redundancy is strongly implicated in the evolution of the bone structure in the mammalian inner ear (Gould 1993). There are numerous other such supposedly irreducibly complex systems for which researchers have found strong evidence for likely evolutionary trajectories. These include the human blood clotting cascade (Miller 1999, 152–158), the avian fight apparatus (Gishlick 2004), and the bacterial flagellum (Musgrave 2004; Beeby et al. 2020), among many others. Interestingly, Darwin himself already noted the possibility that the removal of redundancy could explain the evolution of complex adaptations, in *The Origin of Species*:

> We should be extremely cautious in concluding that an organ could not have formed by transitional gradations of some kind. Numerous cases could be given amongst the lower animals of the same organ performing at the same time wholly distinct functions . . . In such cases natural selection might easily specialise, if any advantage were thus gained, a part or organ, which had performed two functions for one function alone and thus wholly change its nature by insensible steps. Two distinct organs sometimes perform simultaneously the same function in the same individual . . . In these cases, one of the two organs might with ease be modified and perfected so as to perform all the work by itself, being aided during the process of modification by the other organ, and then this other organ might be modified for some other quite distinct purpose, or be quite obliterated. (Darwin 1859, 220)

[41] Mathematicians and statisticians sometimes refer to this as the "Texas sharpshooter fallacy."

we had to see where the arrow landed before identifying the target. A detachable pattern can be described separately from the object or event itself.

Dembski provides a very technical definition of "detachability":

> Given a reference class of possibilities w, a chance hypothesis H, a probability measure induced by H and defined on w (i. e. $P(\cdot \mid H)$), and an event/sample E from w, a rejection function f is *detachable* from E if and only if a subject possesses background knowledge K that is conditionally independent of E (i.e., $P(E \mid H\&K) = P(E|H)$) and such that K explicitly and univocally identifies the function f. Any rejection region R of the form $T^{\gamma} = \{w \in w \mid f(w) \geq \gamma\} T_{\delta} = \{w \in w \mid f(w) \leq \delta\}$ is then said to be *detachable* from E as well. Furthermore, R is then called a *specification* of E and E is said to be *specified*.[42].

Readers with some statistical background will recognize this as standard Fisherian hypothesis testing. In the present context, however, this definition seems unhelpful because we have no way of matching up Dembski's formalism with biologically realistic counterparts in real life. If we are pondering the origins of a bacterial flagellum, for example, then what, exactly, are playing the roles of w, K, E, f, T^{γ}, and T_{δ}?

Indeed, when it comes time to discuss "specification" in the context of particular systems, Dembski does not refer back to this formalism. Instead, he writes:

> Biological specification always refers to function. An organism is a functional system comprising many functional subsystems. In virtue of their function, these systems embody patterns that are objectively given and can be identified independently of the systems that embody them.[43]

This is too casual to serve as an argument that functional systems are specified in the precise technical sense required by his theorizing. Biologists say that evolution by natural selection produces functional systems as a matter of course, and this lends some urgency to the question of whether the function of a system can serve as a proper specification. Does the flagellum instantiate a design-suggesting pattern, or is it just the sort of thing that evolution is able to do?

Our conclusion, then, is that Dembski's theorizing about CSI is not persuasive. We have no way of carrying out meaningful probability calculations, and we have no way of distinguishing design-suggesting patterns in biology from those that can arise through natural causes.

There are many other mathematical arguments in anti-evolutionist discourse, but inevitably they are unsuccessful for the kinds of reasons discussed here.[44] The metaphor of evolution exploring a vast genotype space can be helpful in many contexts, and biologists often employ it in explaining their reasoning. However, the

[42] Dembski (2002, 62–63).

[43] Dembski (2002, 148).

[44] I have written more extensively about mathematical anti-evolutionism elsewhere (Rosenhouse 2016, 2017, 2022). See also Sober (2002) and Elsberry and Shallit (2011), among many other references.

probabilistic and metric structures of the space are too complex to permit broad mathematical conclusions to be drawn regarding what is possible and what is not.

9 Conclusion

Biologist Sergey Gavrilets, himself a major contributor to the mathematical theory of evolution, writes:

> Since the time of the Modern Synthesis, evolutionary biology has arguably remained one of the most mathematized branches of the life sciences, in which mathematical models and methods continuously guide empirical research, provide tools for testing hypotheses, explain complex interactions between multiple evolutionary factors, train biological intuition, identify crucial parameters and factors, evaluate relevant temporal and spatial scales, and point to the gaps in biological knowledge, as well as provide simple and intuitive tools and metaphors for thinking about complex phenomena.[45]

Indeed, among Darwin's most important legacies is the discovery that probability theory is essential to mathematical modeling in biology. With some clear thinking and technical skill, probability can elucidate many formally mysterious aspects of natural history. When its methods are abused by ideologues, it can create obfuscation and confusion. Regardless, explorations of the connections between evolutionary biology and probability have significantly affected the practice of both disciplines, much to their mutual benefit.

References

Beatty J (1984) Chance and natural selection. Philos Sci 51(2):183–211
Beatty J (2006) Chance variation: Darwin on orchids. Philos Sci 73(5):629–641
Beeby M, Ferreira JL, Tripp P, Albers S-V, Mitchell DR (2020) Propulsive nanomachines: the convergent evolution of archaella, flagella, and cilia. FEMS Microbiol Rev 44(3):253–304
Berkman M, Plutzer M (2010) Evolution, creationism, and the battle to control America's classrooms. Cambridge University Press, Cambridge
Bowler PJ (2003) Evolution: the history of an idea, 3rd edn. University of California Press, Berkeley
Charlesworth B, Charlesworth D (2003) Evolution: a very short introduction. Oxford University Press, New York
Crow JF (2009) Mayr, mathematics, and the study of evolution. J Biol 8: Article 13
Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London
Darwin C (1868) The variation of plants and animals under domestication, vol 2. John Murray, London
Dembski WA (2002) No free lunch: why specified complexity cannot be purchased without intelligence. Rowman & Littlefield, Lanham

[45] Gavrilets (2010, 46).

Dietrich MR (1996) Monte Carlo experiments and the defense of diffusion models in molecular population genetics. Biol Philos 11(3):339–356
Dobzhansky T (1951) Genetics and the origin of species, 3rd edn. Columbia University Press, New York
Dronamraju K (2011) Haldane, Mayr, and beanbag genetics. Oxford University Press, New York
Eble G (1999) On the dual nature of chance in evolutionary biology and paleobiology. Paleobiology 25(1):75–87
Edwards AWF (1994) The fundamental theorem of natural selection. Biol Rev 69(4):443–474
Elsberry W, Shallit J (2011) Information theory, evolutionary computation, and Dembski's "complex specified information". Synthese 178(2):237–270
Ewens WJ (1964) Correcting diffusion approximations in finite genetic models, Technical report 4. Department of Mathematics Stanford University
Ewens WJ (2008) On Haldane's 'defense of beanbag genetics'. Int J Epidemiol 37(3):447–451
Felsenstein J (2007) Trees of genes in populations. In: Gascuel O, Steel M (eds) Reconstructing evolution: new mathematical and computational advances. Oxford University Press, New York, pp 3–29
Ferrada E, Wagner A (2010) Evolutionary innovations and the organization of protein functions in genotype space. PLoS One 5(11):1–11
Fisher RA (1922) On the dominance ratio. Proc R Soc Edinb 42:321–341
Fisher RA (1930) The genetical theory of natural selection. Clarendon Press, Oxford
Forrest B, Gross PR (2004) Creationism's Trojan horse: the wedge of intelligent design. Oxford University Press, New York
Foster D (1999) Proving god exists. The Saturday Evening Post November–December:59–61, 78, 80–81, 84
Futuyma DJ, Kirkpatrick M (2017) Evolution, 4th edn. Sinauer, Sunderland
Gavrilets S (2010) High-dimensional fitness landscapes and speciation. In: Pigliucci M, Mueller G (eds) Evolution: the extended synthesis. MIT Press, Cambridge, pp 45–80
Gigerenzer G, Swijtink Z, Porter T, Daston L, Beatty J, Krüger L (1989) The empire of chance: how probability changed science and everyday life. Cambridge University Press, Cambridge
Gillespie JH (1998) Population genetics: a concise guide. The Johns Hopkins University Press, Baltimore
Gishlick A (2004) Evolutionary paths to irreducible complexity: the avian flight apparatus. In: Young M, Edis T (eds) Why intelligent design fails: a scientific critique of the new creationism. Rutgers University Press, New Brunswick, pp 58–71
Gould SJ (1993) Eight little piggies. Norton, New York
Grafen A (2003) Fisher the evolutionary biologist. Statistician 52(3):319–329
Haldane JBS (1964) A defense of beanbag genetics. Perspect Biol Med 7(3):343–360
Hamilton MB (2021) Population genetics, 2nd edn. Wiley-Blackwell, Hoboken
Hardy GH (1908) Mendelian proportions in a mixed population. Science 28:49–50
Herron JC, Freeman S (2014) Evolutionary analysis, 5th edn. Pearson, Glenview
Huxley J (2010) Evolution: the modern synthesis. MIT Press, Cambridge. Reprint of the original 1942 edition
Johnson C (2015) Darwin's Dice: the idea of chance in the thought of Charles Darwin. Oxford University Press, New York
Kardong KV (2007) An introduction to biological evolution, 2nd edn. McGraw Hill, Boston
Kimura M (1964) Diffusion models in population genetics. J Appl Probab 1(2):177–132
Krüger L, Daston L, Heidelberger M (1987) The probabilistic revolution, vols 1 and 2. MIT Press, Cambridge
Laplace PS (1951) A philosophical essay on probabilities. Translated from the sixth French edition, by Truscott FW, Emory FL. Dover, New York
Larson EJ (2004) Evolution: the remarkable history of a scientific theory. The Modern Library, New York

Lynch M (2007) The frailty of adaptive hypotheses for the origin of organismal complexity. Proc Nat Acad Sci 104(Supp. 1):8597–8604
Martin WF (2011) Early evolution without a tree of life. Biol Direct 6(36):25
Maxwell JC (1860a) Illustrations of the dynamical theory of gases. Part 1. On the motions and collisions of perfectly elastic spheres. Philos Mag 19:19–32
Maxwell JC (1860b) Illustrations of the dynamical theory of gases. Part II. On the process of diffusion of two or more kinds of moving particles among one another. Phil Mag 20:21–37
Mayr E (1963) Animal species and evolution. Harvard University Press, Cambridge
Mayr E (1991) One long argument: Charles Darwin and the genesis of modern evolutionary thought. Harvard University Press, Cambridge
Meyer S (2013) Darwin's doubt. HarperOne, New York
Miller KR (1999) Finding Darwin's god: a scientist's search for common ground between god and evolution. Cliff Street Books, New York
Miller KR (2008) Only a theory: evolution and the battle for America's soul. Viking, New York
Millstein R (2000) Chance and macroevolution. Philos Sci 67(4):603–624
Millstein R (2003) Interpretations of probability in evolutionary theory. Philos Sci 70(5):1317–1328
Millstein R (2016) Probability in biology: the case of fitness. In: Hájek A, Hitchcock CR (eds) The Oxford handbook of probability and philosophy. Oxford University Press, Oxford
Morange M (2021) A history of biology. Princeton University Press, Princeton (English translation of the original 2016 French edition)
Musgrave I (2004) Evolution of the bacterial flagellum. In: Young M, Edis T (eds) Why intelligent design fails: a scientific critique of the new creationism. Rutgers University Press, New Brunswick, pp 72–84
Nelkin D (1982) The creation controversy: science or scripture in the schools. Norton, New York
Nordborg M (2018) Introduction to the coalescent theory-lecture 1. Online at: https://www.youtube.com/watch?v=0j0jW0stbB8. Last accessed January 2022
Okasha S (2008) Fisher's fundamental theorem of natural selection: a philosophical analysis. Br J Phil Sci 59(3):319–351
Okasha S (2012) Population genetics. In: Stanford encyclopedia of philosophy. Online at: https://plato.stanford.edu/entries/population-genetics
Pigliucci M (2008) The role of population genetics in modern evolutionary theory. Biol Theory 3(4):316–324
Plutynski A (2006) What was Fisher's fundamental theorem of natural selection and what was it for? Stud Hist Phil Biol Biomed Sci 37(1):59–82
Price GR (1972) Fisher's 'fundamental theorem' made clear. Ann Hum Genet London 36(2): 129–140
Provine WB (2001) The origins of theoretical population genetics. University of Chicago Press, Chicago. Reprint of the original 1971 edition with a new afterward
Ramsey G, Pence CH (eds) (2016) Chance in evolution. University of Chicago Press, Chicago
Ravitch FS (2011) Marketing intelligent design: law and the creationist agenda. Cambridge University Press, Cambridge
Rosenhouse J (2011) Among the creationists: dispatches from the anti- evolutionist front line. Oxford University Press, New York
Rosenhouse J (2016) On mathematical anti-evolutionism. Sci Educ 25(1):95–114
Rosenhouse J (2017) Thermodynamical arguments against evolution. Sci Educ 26(1):3–25
Rosenhouse J (2022) The failures of mathematical anti-evolutionism. Cambridge University Press, Cambridge
Roth AA (1998) Origins: linking science and scripture. Review and Herald Publishing Association, Silver Spring
Ruse M (1979) The Darwinian revolution. University of Chicago Press, Chicago
Scott EC (2009) Evolution vs. creationism: an introduction. Greenwood Press, Westport
Scott EC, Branch G (2006) Not in our schools: why intelligent design is wrong for our schools. Beacon Press, Boston

Semple C, Steel M (2003) Phylogenetics. Oxford University Press, New York
Shan Y (2021) Beyond Mendelism and biometry. Stud Hist Phil Sci Part A 89:155–163
Shanahan T (1991) Chance as an explanatory factor in evolutionary biology. Hist Phil Life Sci 13(2):249–268
Sober E (2002) Intelligent design and probability reasoning. Int J Phil Relig 52(1):65–80
Steel M (2016) Phylogeny: discrete and random processes in evolution. SIAM, Philadelphia
Stigler SM (2012) Studies in the history of probability and statistics L: Karl Pearson and the rule of three. Biometrika 99(1):1–14
Thorvaldsen S, Hössjer O (2020) Using statistical methods to model the fine-tuning of molecular machines and systems. J Theor Biol 507(December 21):1–14
Todhunter I (1865) A history of the mathematical theory of probability from the time of Pascal to that of Laplace. Chelsea, New York
Wagner A (2014) Arrival of the fittest: solving evolution's greatest puzzle. Penguin, New York
Wakeley J (2009) Coalescent theory. Roberts & Co, Greenwood Village
Weldon WFR (1893) On certain correlated variations in *C. maenas*. Proc R Soc Lond 54:318–329
Weldon WFR (1894) An attempt to measure the death-rate due to the selective destruction of *C. maenas* with respect to a particular dimension. Proc R Soc Lond 57:360–379
Weldon WFR (1895) Remarks on variation in animals and plants. Proc R Soc Lond 57:379–382
Yeates TO (2007) Protein structure: evolutionary bridges to new folds. Curr Biol 17(2):R48–R50

Probability Models in Statistical Data Analysis: Uses, Interpretations, Frequentism-as-Model

Christian Hennig

Contents

1 Introduction 1412
2 Basic Concepts 1413
2.1 Interpretations of Probability 1413
2.2 Approaches to Statistical Inference 1419
2.3 Mathematical Models and Reality 1425
3 Frequentism-as-Model 1428
3.1 Frequentism-as-Model as Interpretation of Probability 1430
3.2 The Role of i.i.d. Models and Subject-Matter Knowledge 1432
3.3 How Is Frequentism-as-Model Useful? 1435
3.4 Frequentism-as-Model for Statistical Inference 1439
4 Epistemic Probability-as-Model 1450
4.1 Epistemic Idealization 1450
4.2 Flexible Handling of Priors 1451
4.3 How Different Are Epistemic Probability-as-Model and Frequentism-as-Model? 1453
5 Conclusion 1454
References 1455

Abstract

The application of mathematical probability theory in statistics is quite controversial. Controversies regard both the interpretation of probability and approaches to statistical inference. After having given an overview of the main approaches, I will propose a re-interpretation of frequentist probability. Most statisticians are aware that probability models interpreted in a frequentist manner are not really true in objective reality, but only idealizations. I argue that this is often ignored when actually applying frequentist methods and interpreting the results, and that keeping up the awareness for the essential difference between reality and models can lead to

C. Hennig (✉)
Dipartimento di Scienze Statistiche "Paolo Fortunati", Università di Bologna, Bologna, Italy
e-mail: christian.hennig@unibo.it

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_105

a more appropriate use and interpretation of frequentist models and methods, called "frequentism-as-model." This is elaborated showing connections to existing work, appreciating the special role of independently and identically distributed observations and subject matter knowledge, giving an account of how and under what conditions models that are not true can be useful, giving detailed interpretations of tests and confidence intervals, confronting their implicit compatibility logic with the inverse probability logic of Bayesian inference, re-interpreting the role of model assumptions, and appreciating robustness and the role of "interpretative equivalence" of models. Epistemic probability shares the issue that its models are only idealizations, and an analogous "epistemic-probability-as-model" can also be developed.

Keywords

Foundations of statistics · Probability · Epistemic probability · Aleatory probability · Statistical inference · Constructivism

1 Introduction

Ever since the time of Bernoulli (1713), mathematical probability has been used for statistical inference. Kolmogorov (1933) has given probability a mathematical axiomatic foundation that is the basis of the mathematical theory of probability to this day.

Whereas mathematical probability theory is largely uncontroversial, its application to statistical inference and data analysis is and has been highly contentious throughout the history of probability. This concerns both the very meaning of probability (e.g., Venn (1866); Ramsey (1926); Galavotti (2005)) and how it can be used for learning from data about the world (e.g., Barnard and Cox (1962); Wasserstein et al. (2019); Mayo (2022), the latter even referring to the "statistics wars").

Here I will present a view of some of the central concepts of this controversy that is based on the constructivist philosophy of mathematical modeling elaborated in Hennig (2010), along with some practical consequences. The main idea is that mathematical models are tools for thinking about reality, and they are essentially different from reality. Asking whether a mathematical model is true in reality constitutes a category error; mathematical models have other jobs than being true and need to be assessed based on how well they do these other jobs.

Although many statisticians tend to agree with George Box's famous quote that "all models are wrong but some are useful" (Box 1979) when asked directly, the statistical literature and statistics teaching are full of statements of the kind that "model assumptions have to be fulfilled" and references to "true distributions" and "true parameters" in reality. Many philosophical contributions to the foundations of statistics even more explicitly demand models to be true, or to be "believed." From the perspective of mathematical modeling taken here, such a mindset easily leads to over-interpretation of the results of statistical analysis. It also opens up any position

in the controversy about the foundations of statistics to criticism that could be avoided by a more modest and honest acknowledgement of the role and limitations of the use of probability models.

In Sect. 2, I will give an overview over several key concepts, namely interpretations of probability, approaches to statistical inference, and a constructivist philosophy of mathematical modeling. In Sect. 3, the central part of this chapter introduces "frequentism-as-model," a new re-interpretation of frequentist probability, and inference separating appropriately models and reality and discusses many of its implications. An analogous version of epistemic probability, "epistemic-probability-as-model," is sketched in less detail in Sect. 4 before the chapter is concluded.

2 Basic Concepts

The basic concepts for this chapter are different interpretations of probability, different approaches to statistical inference, and a constructivist philosophy of the relation between mathematical models and reality. Here I give an overview of these. Note that in much literature the distinction between interpretations of probability and approaches to inference is blurred. Often, Bayesian reasoning is identified with a (subjective) epistemic interpretation of probability, but there are non-Bayesian approaches to do inference with epistemic probabilities, and Bayesian reasoning can be combined with frequentist probabilities as well.

2.1 Interpretations of Probability

Interpretations of probability connect probability as mathematical object with the real world by explaining what mathematical probability statements are supposed to mean. There are two major classes of interpretations of probability, aleatory, and epistemic.

2.1.1 Aleatory Probability

Aleatory probability, sometimes referred to as "chance," is a generic term for thinking of a probability as an observer-independent feature of what Hacking (1965) calls "chance setup" or a data generating process in reality. The best-known aleatory interpretation is frequentism. In frequentism, the chance setup is conceived as an experiment or a situation that can be repeated in principle infinitely often, and the probability of an event, $P(A)$, say, is then thought of as the limiting relative frequency of the occurrence of A when the number of random repetitions tends to infinity. I will refer to chance setups as "experiments" in the following, in line with the use of "random experiments" in the literature. This does not necessarily imply controlled experimentation. The axioms of probability are motivated from mathematical properties of relative frequencies.

There have been various attempts to make this notion precise. The most famous of them are probably due to Venn (1866) and von Mises (1939). Both of them

acknowledge that frequentist probabilities are an idealization. Their existence cannot be assured based on finite sequences of observations; in fact, any behavior in any finite part of the sequence is compatible with any limit value from a mathematical point of view. Requiring identical repetitions is an idealization as well. This is bypassed by von Mises (1939) instead requiring the specification of an ideally infinite "Kollektiv" of experiments within which probability is defined rather than relying on "identity." Venn (1866) and many later frequentists take a hypothetical point of view; their probabilities refer to what would happen if the experiment were to be repeated infinitely. Hajek (2009) gives 15 arguments against such a view, many of which have to do with subtleties of mathematical limits, but he also attacks the anti-empirical nature of the idea.

Another key difficulty is the definition of randomness. Randomness can be defined based on algorithmic complexity, i.e., a sequence can count as random if it cannot be generated by certain finitely describable algorithms, see Martin-Löf (1966), in which case he also could prove that limiting relative frequencies exist. However, this cannot be checked given only a finite number of observations. Fine (1973) even showed, given suitable definitions, that "apparently random" finite sequences will automatically look as if their relative frequencies "apparently converge," without any condition regarding what will really happen if that sequence is infinitely continued. Looking at a sequence of relative frequencies of the occurrence of a set A in a "randomly enough" repeated experiment will give the impression that these relative frequencies stabilize. This is a major motivation to think that frequentist probabilities might exist in the real world, but Fine's result shows that the impression is an implication of what kind of sequences an observer attributes randomness to and does not imply anything about the real existence of a limiting frequency.

An alternative aleatory concept is the interpretation of probabilities as "propensities," going back to Popper (1959b). A propensity is conceived as an objective feature of an experiment to have the tendency to produce certain results. If indeed the experiment were repeated infinitely, the idea is that it would produce relative frequencies according to its propensities, but there is no necessity for this to happen for a propensity to be defined. In this way, the propensity interpretation of probability allows to speak of probabilities connected to single realizations of an experiment, and it does not have to be concerned about issues with mathematical limits, although the propensity interpretation does not put the researcher in a better position to make empirical statements about the values or even the very existence of probabilities compared to frequentism.

There are several varieties of propensity interpretations, see Gillies (2000). Gillies' own propensity approach basically postulates that propensities behave according to the probability axioms, and in case repetitions are observed, it can then be tested empirically whether outcomes behave as they should behave given the postulated probability model. Such tests can of course not be very powerful if the number of repetitions is low.

The biggest problems with aleatory probability interpretations stem, in my view, from the idea that their use requires them to exist objectively in reality. Treating

mathematical models as fundamentally different from reality, as I will outline below, an aleatory interpretation of probability should refer to thinking of reality *as if* it behaved in this way, without postulating or believing that it really does, which cannot be ensured, or even very convincingly argued, due to the lack of infinite random repetition in reality. I will call this "frequentism-as-model" (in the following I will mostly refer to "frequentism," as this is in much more frequent use in the literature than "aleatory"). Obviously, there needs to be a justification how models used in this way contribute to science. Many scientists think that objectivity is a main reason to use aleatory probabilities, but they can be called "objective" only in the sense that they *refer* to observer-independent reality, not in the sense that their observer-independent existence could be granted.

2.1.2 Epistemic Probability

An epistemic probability $P(A)$ refers to the belief or knowledge of an agent regarding the event A. Epistemic probabilities do not require any repeatability of experiments; they apply to single situations as well as to experiments that are part of a sequence. The probably simplest and historically original way to explain $P(A)$ in an epistemic way is as a betting rate; the agent would be willing to pay up to $P(A)$ for a gamble in which she receives 1 in case that A obtains and 0 otherwise. The axioms of probability can be motivated from requirements of fair betting prices requiring "coherence," i.e., bets on several events have to be offered in such a way that an opponent cannot put together a system of bets so that the agent incurs sure loss. This was formalized by Ramsey (1926) and de Finetti (1937).

As for aleatory probability, there are various versions of the epistemic interpretation. Ramsey (1926) and de Finetti (1937) held that probabilities are subjective. An agent makes subjective initial prior assignments of probability, and probability theory imposes the coherence of these assignments and how these have to be coherently updated when new evidence occurs, namely following Bayes's Theorem. Different agents can have different probabilities for the same event, even if they hold the same information.

Subjective probabilities may seem unsuitable for science if the aim of science is assumed to be producing objective statements that hold independently of the researcher. Keynes (1921) interpreted probability as an extension of logic to uncertain propositions, the plausibility of which, given the evidence, could be quantified as a probability. Such a probability is epistemic, but there is no subjective agent; the probability is meant to be determined by the evidence without further subjective input. Keynes constrained numerical probabilities to propositions concerning a finite number of cases that are equiprobable, i.e., they have the same probability assigned based on the "Principle of Indifference" in case that for none of them there are reasons to think of them as more likely than the others. Ramsey and de Finetti criticized this as far too restrictive, and there are further issues with the Principle of Indifference; see Gillies (2000). More recent developments of logical and objectivist epistemic interpretations of probabilities, e.g., Jaynes (2003), are less restricted.

Bayesian updating of prior probability assessments ("priors") is central according to many proponents of epistemic probability, which is therefore often identified with

"Bayesianism." At the time of Bayes (1763), there was no explicit distinction between aleatory and epistemic probabilities. Bayes's original paper is somewhat ambiguous on the matter, and his theorem holds as well for aleatory probabilities. Williamson (2010), despite calling his interpretation "Bayesianism," is an example for an epistemic interpretation that does not rely on Bayesian updating.

The major difference between objectivist and subjectivist epistemic interpretations is how the priors should be chosen. The subjective approach may lead to potentially arbitrary choices that may not be an appropriate basis for science, but in very many situations, it is hard if not impossible to impose a unique objective choice. Ideally, at some point there is no evidence. At this point, an "informationless prior" is chosen, and then any incoming evidence can be processed using Bayesian updating. However, general agreement regarding how to choose informationless priors does not exist, and there are issues and paradoxes with them (Kass and Wasserman 1996). Furthermore, much evidence comes in forms that an initially chosen model cannot easily incorporate. Objective probabilities might be seen as desirable, but the subjectivists can state that they are an unachievable ideal in most situations. Berger (2006), in a defense of objective Bayesianism, admits that full objectivity cannot be achieved, particularly when invoking existing information that does not come in an appropriately formalized way, for example, knowledge about logically impossible parameter values or observations. The term "objective Bayesianism" in his view refers mainly to priors that reflect a lack of subjective information. These can be useful also for subjectivists, as a "reference." Howson and Urbach (2006) on the other hand state that the hope for objective priors is misplaced, but that the probability axioms provide an objective basis for logical reasoning with uncertainties. They refer to Cox (1946), who showed that probability calculus including Bayesian updating is implied by a certain set of axioms for such logical reasoning, regardless of the prior. This is largely in line with the subjectivists.

The controversy between subjectivist and objectivist proponents of epistemic probability highlights a more general issue of statistics and even science as a whole, namely the tension between the ideal of objectivity, and the impossibility to achieve it, in its idealized form, in practice. Gelman and Hennig (2017) argue that instead of the ultimately futile demand of objectivity, it is more constructive to aim at a list of virtues that constitute aspects of objectivity, namely transparency, enabling consensus, impartiality, stability, and correspondence with observable reality; acknowledging that there are also virtues connected to subjectivity, namely context-dependence; and acknowledging multiple perspectives. Regarding epistemic probabilities, the objectivist focus is on enabling consensus and impartiality, whereas the subjectivist focus is on context-dependence and potentially multiple perspectives. In order to convince others and potentially a whole scientific community of their analysis, subjectivists need to motivate and justify their prior choices transparently just as well as the objectivists, despite acknowledging that others may legitimately choose another prior.

A general issue with epistemic probabilities is that the prior probability assignments determine the set of potential updates. This means in particular that certain prior choices that are in conflict with later observed data can never be corrected, at least not within the Bayesian framework. A Bayesian agent needs to explicitly assign

nonzero probability to whatever seems possible to them, because otherwise it will remain modeled as "impossible" forever. For example, if in a Bayesian analysis the prior involves the assumption that certain future events are exchangeable, meaning that probabilities do not depend on the order of outcomes, any posterior obtained by Bayesian updating from data will still fulfill this assumption, even if the data indicate that the assumption may not be appropriate. This does not seem to be a reasonable point of view of an epistemic agent. Consider an example regarding binary sequences. Even if initially it may seem an epistemic agent that all possible sequences in the future with, say, 500 ones and 500 zeroes are equally likely, it seems advisable to question her own assumption and potentially to change her mind in some cases. Imagine that 100 ones in a row are observed, then 150 zeroes, and then a few more ones, say 15. This seems to show strong evidence of positive dependence between neighboring outcomes, and most would now think that the outcome "one" should be far more likely in the next go, but having committed themselves to exchangeability, the epistemic agent is forced to ignore this.

It may be demanded, for this reason, that priors should not assume exchangeability. Furthermore, a similar case can be made against most regularity constraints that may be involved in non-exchangeable models. In consequence, the choice of priors would be far more complicated if not practically impossible, and much of the mathematical and algorithmic Bayesian machinery for computing posterior distributions, i.e., the probabilities resulting from updating the prior with the data, could no longer be used. Note also that a famous argument in favor of Bayesian epistemic probability states that the prior choice "wears out," i.e., the posterior distribution will ultimately be dominated by the data and will, in particular, converge toward a "true aleatory probability" of an event in case that this exists. This is often stated to suggest that epistemic probability is just as useful as aleatory probability when it comes to correspondence to observed reality, despite locating probabilities in the agent rather than in reality, but the argument assumes exchangeability, among other things (even though assumptions can be relaxed a bit (see Diaconis and Skyrms 2018, Ch. 6, 9)).

Arguably, this means that pretty much all prior assignments of epistemic probability that are done in practice are "wrong" in the sense that they do not really reflect the full belief or state of information. On top of this, it is well known that human agents are often biased when it comes to specifying probabilities and assessing the implications of their own choices (Kahneman et al. 1982).

Ultimately the same that has been stated for the aleatory interpretation already can also be said for the epistemic interpretation of probability. Epistemic probabilities, be they subjective or supposedly objective, should not be taken as a "really true" reflection of the epistemic state of the agent. They are models, and as such, they idealize and simplify. The later section on "epistemic-probability-as-model" discusses the consequences of explicitly acknowledging the model status of epistemic probabilities.

2.1.3 Postscript on the Controversy About Interpretations of Probability

A lot of work (de Finetti 1974; Jaynes 2003; Howson and Urbach 2006; Diaconis and Skyrms 2018) suggests that epistemic probability should be used universally,

and aleatory probability is flawed and/or superfluous (because of the result mentioned above that under certain prior assignments aleatory probabilities, if they exist, can be recovered by epistemic inference given enough data).

In particular, de Finetti (1974) stated that

> Probability does not exist.

This refers to the existence of objective aleatory probabilities in the world and was meant as an argument that the epistemic interpretation of probability is the only legitimate one. However, I have argued above that the existence of "true" epistemic probabilities is just as dubious. Furthermore, the assumptions that allow aleatory probabilities to be emulated in the epistemic framework may be convenient and in widespread use, but are not necessarily convincing. There seems to be a tendency among authors who favor epistemic probability to elaborate very clearly the deviation from reality of the aleatory concept taken literally, but to deny that the same issue exists for epistemic probability. If the latter is acknowledged, one may wonder whether it is worthwhile to model an epistemic agent rather than reality itself, in cases in which researchers want to make statements about observer-independent reality in the first place.

Senn (2011) raises the issue that Bayesian analyses in practice rarely are in line with how epistemic probability should ideally be handled, particularly when it comes to specifying the prior, and suggests (with arguments similar to mine above) that this may often be impossible. As other authors (e.g., Carnap (1950); Cosmides and Tooby (1996); Gillies (2000)), he acknowledges that both epistemic and aleatory probabilities may be needed.

Defense of aleatory probability and criticism of epistemic probability can also be found in the literature (e.g., Mayo (2018)), but there is more general acknowledgement that aleatory probability is not suitable for all applications (particularly if there is nothing that can be convincingly thought of as repetition) and that there is a role for epistemic probability, if (as some authors think) limited.

I do not think that the adoption of one interpretation of probability in one situation should preclude a statistician from using another interpretation in another situation. A statistician can be a frequentist when analyzing a randomized trial to compare a new drug with a placebo and adopt an epistemic approach when computing a probability that an accused is really the murderer, and in this sense she can be a pluralist. However, if in a single analysis the implied interpretation of probability is not clear, results in terms of probability do not have a clear meaning.

An interesting historical note is that a distinction between aleatory and epistemic probabilities was mentioned for the first time only around 1840 (Hacking 1975), more than a century after the first systematic treatment of probabilities by Bernoulli (1713). Problems associated with games of chance were among the first sources of probability considerations and, originally fair betting, classical probability (i.e., based on counting possibilities that were thought of as equiprobably), and expected frequencies seem to have been thought of as synonymous (in fact the first versions of the law of large numbers were taken as some kind of proof of the

correspondence between classical probability and limiting relative frequencies). There had already been considerations of situations that could hardly be reduced to counting equiprobable possibilities, and Bayes (1763) did not require equiprobability anymore, but his work seems to allude to both epistemic and aleatory ideas. Laplace (1814) defines probabilities more explicitly as epistemic, but his treatment of measurement error distributions looks rather aleatory, and no explicit distinction can be found in his work.

In any case it is not an accident that apparently different concepts such as epistemic and aleatory probabilities come with the same name and can be handled with the same calculus, as they draw from the same sources. This may feed the idea that only one interpretation should be needed, which exists particularly in the epistemic camp.

2.2 Approaches to Statistical Inference

In much literature, it is suggested that adopting either aleatory or epistemic interpretation of probability implies how to do statistical inference. Indeed many (though not all) versions of the epistemic interpretation imply Bayesian inference, whereas classical techniques as hypothesis tests and confidence sets have been devised with an aleatory probability interpretation in mind. There are exceptions to these associations though, and I believe that it clarifies the discussion to distinguish between interpretations of probability and approaches to statistical inference. In particular, there is much criticism of the logic implied by hypothesis tests and confidence intervals that does not concern the aleatory probability interpretation as such, and Bayesian inference is often applied having implicitly or explicitly an aleatory interpretation in mind.

These two approaches are the main focus here, but there are some others, as discussed further.

2.2.1 Compatibility Logic

Hypothesis tests and confidence sets concern parameters of statistical models, or, in the nonparametric case, certain classes of models that are of interest in a real situation. A model parameter can, for example, refer to a physical constant that is assumed to be measured with a random error. Comparing the outcome of a therapy on an ordinal well-being scale to a placebo therapy, a nonparametric set of distributions of interest may model the distribution of results in the therapy group as the same, or alternatively as systematically larger than in the placebo group. In the following, parametric inference will be discussed.

Data are assumed to have been generated according to such a model allowing for an often infinite set of parameters, and it is assumed that there is one true value of it. Inference about this true value is made based on with which parameter values data look incompatible, i.e., are in a suitable region with low probability. This region is usually defined by a statistic that formalizes a "distance" of the data to what is expected given a certain parameter value. If the probability ("p-value") is very low

that a dataset generated from the model with the parameter value in question is as far as or farther away from this expectation than the observed data, then the data are deemed incompatible with the parameter value. In that case, a test will reject a prespecified parameter value (or a set of such values) of interest. A confidence set will collect all compatible parameter values. Test and confidence sets have been criticized for simplifying binary decisions regarding compatibility, and in fact thresholding probabilities to make such decisions can only be arbitrary. *p*-values give continuous information about the degree of incompatibility of the data with the model.

Tests and confidence intervals have been very controversially discussed recently. The centrality of the concept of compatibility can be seen in the 2016 ASA (American Statistical Association) Statement on p-values (Wasserstein and Lazar (2016)), which has a single positive message on p-values, besides a number of negative ones:

> p-values can indicate how incompatible the data are with a specified statistical model.

It is very important to understand that being "compatible" does not imply truth. Particularly, alternative models are not ruled out by any means. There is always an infinite variety of models and usually also parameter values compatible with any dataset. Much misinterpretation of tests and confidence sets ignores this. Confidence sets formalize uncertainty in the sense that they show a range of compatible parameters, but on top of the uncertainty within the assumed model, there is model uncertainty that needs to be addressed in different ways.

Compatibility as defined here is not absolute, but relative to the involved test statistic. A model may be compatible with data in some but not other respects. Data used for a simple linear regression, for example, can be compatible with a model that has a regression slope zero, using the standard t-statistic measuring how far away from zero the estimated slope is, but this may be due to nonlinearity, and a linearity test may reject that same model. Compatibility and incompatibility apply to specific aspects of a model (potentially with a fixed parameter value) rather than to the model as a whole.

It has often been stated that the rejection of a null hypothesis by a statistical test is a "statistical falsification" in the sense of Popper (1959a). Mayo (2018) draws this connection explicitly. Popper's work on falsification does not involve probabilities, and consequently it does not allow for a type I error probability, i.e., a probability that a correct model is rejected. It can be seen as problematic that a probability model can be deemed incompatible based on an event the possibility of which was essential part of the model. This particularly leads to problems with multiple testing; a large enough number of tests will reject a true model with a large probability, even if this probability is small for every single test. However, in case of imprecise measurements modeled as involving random errors, erroneous falsification of theories due to random variation is certainly a possibility. Troubles stemming from the possibility of "falsifying" a true model in the model world reflect a corresponding possibility in the real world and seem as such not inappropriate as extension of Popper's original ideas (in fact early work on statistical tests preceded Popper's work, e.g., Fisher (1922)).

Furthermore, compatibility statements regarding models do not require any "belief" that a true aleatory probability distribution exists or even follows the model assumptions of a test or confidence set. The model is a thought construct, and the thought construct can be seen as reasonable and helpful if the data are compatible (there may be further requirements), even acknowledging that the model is not true (the same point has been made by Greenland and Rafi (2019)). There are pitfalls though; a very large dataset can allow so precise statements that a parameter value might be deemed incompatible even though it is, in terms of its real-world implications, so close to compatible values that it can be seen as more or less equivalent. A parameter might be "rejected" that implies that a certain therapy does not improve on a placebo, even if another parameter is still compatible with the data according to which the improvement is very weak and virtually useless.

Compatibility logic is normally applied together with aleatory probability, but an epistemic agent may as well be interested if her model is compatible with the observed data. This has been operationalized by posterior predictive checks (Box 1980; Gelman et al. 1996) that basically test whether the observed data are in line with the epistemic predictive distribution, once more relative to a chosen test statistic.

2.2.2 Inverse Probability Logic

Bayesian inverse probability logic is the major alternative to compatibility logic regarding the assessment of probability models in the light of the data. To fix terminology, in a standard Bayesian analysis, there is usually a sampling model, i.e., a distribution of the data given a parameter, and a prior distribution of the parameter. Both of these together are chosen prior to the analysis and can therefore be called "prior distribution" (as done by, e.g., de Finetti (1974)), even though in much literature this term is used for what I call "parameter prior." The term "inverse" comes from turning the sampling model of the data given the parameter into a posterior distribution of the parameter given the data with help of the parameter prior. This parameter posterior then also allows to compute predictive distributions for future observations.

Inverse probability logic produces a probability distribution over parameters (or models; in the following I will only mention parameters). Whereas assessing a parameter as compatible with the data does not exclude or reduce the compatibility of other parameters, having a probability distribution implies that parameters compete for probability mass. A larger probability of certain regions of the parameter space implies a lower probability elsewhere. The probability axioms, whether interpreted in an epistemic or in an aleatory manner, imply that there is only one truth (only one of two disjunct events can obtain), even though not known with certainty.

Often posterior distributions are interpreted as distributing probabilities about where to find the true parameter value. The idea of a true parameter value is a traditional frequentist one. De Finetti (1974) argued that posteriors should be interpreted regarding observable quantities such as future observations and not regarding unobservables such as true parameter values that may well not exist. The parameters for him are helpful devices to summarize prior and posterior

assessments, but betting (as crucial for the traditional subjectivist Bayesian approach) should concern observables, so that bets can ultimately be decided. However, according to Diaconis and Skyrms (2018), de Finetti's Theorem means that for a subjectivist, belief in exchangeability or a suitable generalization of it implies belief in the existence of limiting relative frequencies and, therefore, a limiting probability distribution that can be parameterized. Applied Bayesians sometimes argue that even though true aleatory distributions might not exist, in many cases the parameter of interest can be interpreted as a meaningful and potentially objectively existing quantity such as a physical constant that can only be measured with an error treated as random, giving rise to epistemic uncertainty.

2.2.3 Compatibility Logic vs. Inverse Probability Logic

There are advantages and disadvantages of both approaches. Many Bayesians such as Diaconis and Skyrms (2018) have pointed out that p-values and confidence levels are regularly misinterpreted as probabilities regarding the true parameter, because these should be the ultimate quantities of interest in statistical inference, or so it is claimed. Inverse probability logic deals with combining different inferences such as multiple testing, which creates trouble for standard compatibility logic approaches, in a unified and coherent way. Frequentists are not only interested in compatibility, but also in estimation; finding a best model amounts to a competition between models. A Bayesian can argue that in this case a probability distribution over parameters provides more differentiated and appropriate information about how the parameters compare and to what extent one parameter value is more likely than others. Using confidence distributions (Xie and Singh (2013)), which define a probability distribution over the parameters based on the confidence levels at which parameters are still in an appropriately defined confidence set, frequentists involving compatibility logic can give more detailed information about how parameters compare as well, but Bayesians hold that a proper probability distribution is more intuitive and less prone to misinterpretation. Confidence distributions are strongly connected to fiducial probabilities, which Fisher (1936) defined in order to make probability statements about parameters without requiring a prior. This implies some kind of inverse probability logic, even though Fisher distanced himself from the Bayesians. Modern work on fiducial inference (Hannig et al. 2016) uses fiducial distributions of parameters as inferential tool allowing for compatibility statements rather than as actual epistemic distribution alternative to the Bayesian approach.

By allowing many models to be compatible with the data at the same time, compatibility logic is more obviously in line with the attitude that models are idealizations and not really true. Davies (1995) argued that if many models are valid approximations for the same real situation, inverse probability logic is inappropriate, because if, for example, $\mathcal{N}(0,1)$ is a reasonable approximation of the truth ($\mathcal{N}$ denoting the Gaussian or normal distribution; the two numbers in brackets are its mean and standard deviation), $\mathcal{N}(10^{-10}, 1)$ is a reasonable approximation as well, whereas according to inverse probability logic any two models compete for a part of the unit probability mass.

Furthermore, it may seem unfair to criticize tests and confidence intervals based on misinterpretations. Arguably many users do not only want to know the probability for certain parameter values to be true, which indeed tempts them to misinterpret confidence levels and *p*-values. Arguably they also want this probability to be objective and independent of prior probability assessments, which to make up they have a hard time. This combination is not licensed by any properly understood philosophy of statistics, and ultimately statisticians need to accept that their job is often not to give the users what they want, but rather to defy wrong expectations.

The role of the prior distribution in inverse probability logic is a major distinction between the two approaches. Bayesians argue that the prior is a good and very useful vehicle to incorporate prior information. Actually, prior information enters frequentist modeling as well (see the later discussion of frequentism-as-model), but the Bayesian prior is still an additional tool on top of the options that frequentists have to involve information. However, the requirement to set up a prior can also be seen as a major problem with inverse probability logic, given that prior information does not normally come in the form of prior probabilities and that it is actually in most cases very difficult to translate existing information into the required form. It is not an accident that a very large number of applied Bayesian publications come with no or very scarce subject matter justification of the prior, and in most cases the prior information is very clearly compatible with many different potential priors, with comprehensive sensitivity analysis rarely done. If the sample size is large enough for the prior to lose most of its influence, one may wonder why to bother having one. The question whether there is prior information that is meant to have an impact on the analysis and can be encoded convincingly in the form of a prior distribution is a key issue for deciding whether inverse probability or compatibility logic will be more promising in a given application of statistics.

2.2.4 Prediction Quality

Given the problems connecting aleatory and epistemic probabilities to any objectively existing truth, a pragmatic stance may hold that inference about "true parameters" is not of interest in its own right. Rather, probability modeling is worthwhile to the extent that it enables the user to deal with uncertainty about the future and, more precisely, prediction of future observations. This point of view is dominant in the machine learning community and becomes more popular also in the statistical community, following Breiman (2001), who pointed out that traditional statistical modeling (be it aleatory or epistemic) often produces an inferior prediction performance compared to approaches that aim more directly at optimizing prediction quality. Such approaches may not even involve any probability modeling, or only in such a convoluted way that clear interpretations regarding the underlying workings of reality leading to the predictions are hardly possible. Prediction quality can be assessed using data-driven techniques such as cross-validation and bootstrap that apparently do not refer to specific probability models. A worry is that performance achieved on specific training, test, and validation data may not necessarily generalize to future data. Probability models, usually nonparametric and interpreted in an aleatory manner, are used to theoretically underwrite and investigate prediction

algorithms as well as validation techniques (e.g., Schölkopf and Smola (2002); Bates et al. (2021)).

I agree with Breiman that probability modeling is certainly not mandatory, which runs counter to the naive idea that finding the true model is the ultimate aim of statistics. On the other hand, Breiman (2001) and much work from the machine learning community seem to reduce aims of modeling to prediction quality as all-dominating issue, which neglects the role of models for communication and building understanding (see the Sect. 2.3), even though it is probably valid to state that prediction quality is very often at least implied by the aim of analysis. Recently, this issue has been acknowledged also in the machine learning community, and there is a raising number of publications on the interpretability of the outcomes of machine learning algorithms (Rudin et al. 2022).

Probability models allow for a quantification of uncertainty that is often desired in prediction problems. Bayesian approaches deliver such uncertainty quantification automatically, of course relying on the appropriateness of the models. This is harder for machine learning techniques that are not based on probability models, but work exists, e.g., Lakshminarayanan et al. (2017).

Also there are doubts to what extent prediction quality for future events can be reliably assessed. For example, there is the possibility that the process of interest in the future may change, and a transparent model with solid subject matter justification may be more easily adapted, as Hand (2006) argues. Cross-validation and other methods for assessing prediction quality imply independently and identically distributed (i.i.d.) repetition, which may be problematic.

2.2.5 Other Approaches

Many further approaches to statistical inference exist. Technically, parameter estimation, i.e., finding a best guess, does not fall into any of the earlier listed categories. Estimation can be of interest regarding both aleatory and epistemic probabilities; in the latter case particularly where the parameter is identified with a not directly observable quantity that is assumed as existing and meaningful, e.g., in measurement problems with measurement error. Estimators are usually solutions of optimality problems, although there are also qualitative criteria for estimators such as consistency, or robustness against model misspecification. Statisticians often insist that estimators should come with an indication of their uncertainty, which usually then requires either inverse probability or confidence sets (which may involve nonparametric bootstrap techniques).

Some authors (Hacking 1965; Edwards 1972) advocated the use of the likelihood for assessing the support that data lend to certain models, without involving prior distributions. Likelihoods are hard to compare particularly over models with different complexities, but information theoretic approaches such as Akaike's Information Criterion can be used for such comparisons (Burnham and Anderson 2002).

A rich literature exists regarding imprecise probabilities, which combine modeling uncertainty about the occurrence of an event A with the strength of knowledge regarding A. Lower and upper probabilities are defined, with the interpretation that generally weak knowledge is expressed by a big difference between the two. This can be interpreted in an epistemic or aleatory manner, referring either to uncertainty

of an agent, or to upper and lower bounds of probabilities in a set of aleatory probability models (Walley 1991). The latter is related to robust statistics, which aims at good frequentist properties of estimators, tests, and confidence sets given a nominal model, and stability in a suitable neighboring class of models. Dempster-Shafer belief functions (Shafer 1976) are another variant of considering upper and lower plausibility indications. Shafer and Vovk (2019) propose new foundations of probability and inference based on specific betting games.

Less formally, visual and exploratory data analysis also deserves a mention, providing sometimes valuable alternatives to model-based inference. This may be better to appreciate from a philosophical position that does not grant mathematical formalism a generally superior access to reality.

2.3 Mathematical Models and Reality

I wrote Hennig (2010) mainly out of the conviction that the debate about the foundations of statistics often suffers from the lack of a general account of mathematical models and their relation to reality.

On the frequentist side often a rather naive connection between the models and reality is postulated, which is then easily criticized by advocates of epistemic probability, implying that because a naive realist idea of frequentist probabilities is not convincing, probability should better be about human uncertainty rather than directly about the reality of interest. However, modeling human uncertainty is still modeling, and similarly naive ideas of how such models relate to "real" human uncertainty are problematic in similar ways.

Inspired by radical and social constructivist philosophy (von Glasersfeld 1995; Gergen 1999), I have argued that mathematical models are thought constructs and means of communication about reality. I have distinguished "observer-independent (or objective) reality," "personal reality" (the view of reality constructed by an individual observer and basis for her actions), and "social reality" (the view of reality constructed by communication between observers). Taking a constructivist perspective, I have argued that observer-independent reality is not accessible directly to human observers and that, if we want to talk about reality in a meaningful way, it makes sense to refer to personal and social reality, which are accessible by the individual, a social group, respectively, rather than only to observer-independent reality. Observer-independent reality manifests itself primarily in the generally shared experience that neither an individual observer nor a social system can construct their reality however they want; we all face resistance from reality. This resonates with Chang (2012)'s "Active Scientific Realism" (see also Chang 2022):

> I take reality as whatever is not subject to one's will, and knowledge as an ability to act without being frustrated by resistance from reality. This perspective allows an optimistic rendition of the pessimistic induction, which celebrates the fact that we can be successful in science without even knowing the truth. The standard realist argument from success to truth is shown to be ill-defined and flawed.

In Hennig (2010), I interpret mathematical models as a particular form of social reality, based on the idea that mathematics is a particular form of communication that aims at being free of ambiguity and that enables absolute agreement by means of proofs. Science is seen as a social endeavor aiming at general agreement about aspects of the world in free exchange, for which unambiguous mathematical communication should be obviously useful. This comes to the price that mathematics needs to unify potentially diverging personal and social observations and constructions. There is no guarantee that by doing so mathematical models come closer to observer-independent reality; in fact, this view sees the domain of mathematics as distinct from the domain of observer-independent reality. The connection is that the constructs of personal and social reality are affected by "perturbation" (or, as Chang puts it, "resistance") from observer-independent reality, and it can be hoped that such perturbation, if observed in sufficiently related ways by different individuals and social systems, can be represented in mathematical modeling, which then allows to analyze the represented reality using the machinery of mathematics and logic. There is also repercussion of communication, including mathematics, on social and personal constructions of reality. Mathematical modeling has an effect on the world view and dealings with the world of social systems and individuals. We can hope for this effect to be beneficial, but this may not always be the case, because important subtleties and differences between individual and social realities may be lost in mathematics. See, e.g., the criticism of the unreflected use of quantification in psychology (Tafreshi et al. 2016) and a more general exposition of the role and history of quantification (Porter 1996).

According to constructivism, it should not be hoped that mathematical models represent reality "correctly," but they can have many uses anyway. They can represent important aspects of the modeler's perception of reality. Accordingly, a major benefit that they can have is to communicate the modeler's view unambiguously and in this way to support agreement (or potentially "agreement to disagree"). They can explore different scenarios, reduce complexity, help decision making, and enable predictions (that can be checked against observed reality). Sometimes the best way to learn from a model is to understand in which way it fails.

Another important benefit is the stimulation of creativity and method development. A standard way to solve a data analytic problem is to set up a probability model and to derive inference methods that are in some formal sense justified or even optimal given that model. This does not rely on the truth of the model, but of course if the model is inappropriate, results can be misleading. Constructivism does not advocate that "anything goes"; the modeler needs to convince themselves and her community that the model is a good basis for inference, which can involve both checking against observations and explaining how existing knowledge and thought is related to the model. Rather than licensing an arbitrary use of models, the focus is on the responsibility of the modeler to explain and defend her modeling choices. Widely accepted mathematical models in science have been discussed, questioned, defended, and possibly updated, and have been found to be in agreement with observations in lots of situations.

The constructivism I adhere to may be these days rather unpopular in philosophy, and what Chang refers to as "standard realism" in his quote above has won some ground. However, given the rather tedious controversy that philosophers have about

the objective "existence" of long run frequentist or single-event chances and propensities (see Eagle (2019)), it seems as attractive as ever to treat them as a thought construct and to look at how they are used and how they relate to observations in order to circumvent the enormous amount of problems that come with postulating the objective "existence" of such a thing.

I acknowledge that there is a lot of philosophical work on models in science that is not naïvely realistic, and concerned in more detail with how modeling actually works, see Frigg and Nguyen (2020). Much of this is compatible with the view taken here, as also a constructivist view implies that models represent reality, reality referring here to the perceived reality of the modeler and possibly her social system.

2.3.1 Reality, Observations, and Operations

In order to make a convincing case for a proposed model, it needs to be connected with reality or, more precisely, with the reality perceptions of the people in the community at which the model is addressed, at an individual and social level. Von Mises, Reichenbach (another contributor to the foundations of frequentism), de Finetti, and Ramsey were all influenced by logical empiricist ideas and by the operationalist view of measurement (Bridgman 1927; Creath 2021; Gillies 2000; Galavotti 2005) when proposing their foundations of aleatory and subjectivist epistemic probability, respectively. They built their concepts on well-defined evaluation operations referring to observations and tried to avoid metaphysical concepts (even though von Mises' limiting relative frequency has been charged as metaphysical). Von Mises defined probabilities as what is measured, with increasing precision, by observed relative frequencies; de Finetti detailed the betting operations necessary for eliciting prior belief and declared that the resulting bets should concern future observations rather than unobservable parameters.

Measurement is central to the philosophy of mathematical models outlined above, as measurement is the transformation of informal perceptions to mathematical objects that can then be handled by mathematical models and methodology.

Logical empiricism and operationalism have been criticized in various ways and are no longer very popular (Creath 2021; Chang 2021), but Chang (2021) argues that operationalism is still relevant regarding central issues in the philosophy of science, and Van Fraassen (1980) proposed a more modern version of empiricism, sharing its anti-metaphysical (sometimes branded "anti-realist") commitment. From my constructivist perspective, early empiricists and operationalists took a too extreme and reductionist point of view by denying the meaning of thought constructs that could not directly be tied to operations and observations. Operations and observations rely themselves on theories. It is important to acknowledge that measurement operations and outcomes are not an objective "given" and that the distinction between a measurement and what is measured is an elementary part of scientific reasoning. Scientists need to argue the validity of their measurement operations regarding what they want to achieve, and observations can be criticized meaningfully as potentially erroneous. This does in my view not refer to the inaccessible observer-independent reality, but rather to the thought constructs of the scientists and their community.

However, I agree with Chang and Van Fraassen that the focus on observations and operations is generally helpful for clarifying meaning, validity, and implications of

scientific concepts. Observations and operations do not have absolute authority and need to be negotiated and agreed, but they seem much better suited to enable clear negotiations and agreement (or clear disagreement) than reference to an unobservable "truth," Particularly regarding epistemic probability, betting schemes give a much clearer idea what it means to state that "the probability of an event is 0.4" than referring to the rather abstract setup of Cox's Theorem as motivation why reasoning about plausibilities can be formalized by probabilities. For frequentism, von Mises's concept of the "Kollektiv" clarifies that aleatory probabilities need to be defined relative to a "reference class" of experiments that qualify to be counted, which is often ignored in practical data analysis but can be a source of trouble (Hajek 2009).

In this vein, the contributions of von Mises, de Finetti, and others can be re-appreciated despite the criticism that they attracted when interpreting their work as attempt to convince the community of the "truth" or "objectivity" of these concepts. Without this ambition, what they achieved is a thorough elaboration of the (partly problematic) implications and the operational and empirical meaning of the corresponding probability interpretations. My position is not that models can be arbitrarily chosen, but rather that the modeler needs to be clear about how the involved probabilities are to be understood and to what extent they can be empirically tested, or used in empirical work. This will ultimately contribute to transparent communication and to stability and reliability of the outcomes.

3 Frequentism-as-Model

The frequentist (aleatory) interpretation of probability and frequentist inference such as hypothesis tests and confidence intervals have been strongly criticized recently (e.g., Hajek (2009); Diaconis and Skyrms (2018); Wasserstein et al. (2019)). In applied statistics, they are still in very widespread use, and in theoretical statistics, the number of still appearing new works based on frequentist principles is probably not lower than that of any competing approach to statistics.

Even defenders of frequentist statistics such as Mayo (2018) admit that frequentist inference is often misinterpreted, while arguing that such misinterpretation is a major source of unfair criticism of frequentism. By and large I agree with this. I will argue that what I call "frequentism-as-model" avoids many issues that critics have raised and allows for a better understanding of how results of frequentist inference can be interpreted.

Here are some influences and important ideas from the literature. I already mentioned the "long run propensity" interpretation of probability by Gillies (2000, Ch. 7). Although his outlook is more traditionally realist than mine, I see this as an insightful interpretation of how frequentist statisticians actually use their models.

Tukey (1997) re-interpreted the role of frequentist models in statistical inference as "challenges" that procedures have to face in order to show that they are worthwhile in a given situation. He recommended using

> bouquets of alternative challenges

to investigate how well a procedure may work in a real situation, including worst cases, rather than finding an optimal procedure for a single model. He wrote

> (more honest foundations of data analysis) have to include not assuming that we always know what in fact we never know – the exact probability structure involved (. . .) almost nothing is harder than verifying a stochastic assumption, something I find hard to believe has ever been done in even a single instance.

Many different models can be compatible with the same situation, and it is often misleading to base an analysis on just one model.

Davies (2014)'s view of statistical models is similar:

> The problem is how to relate a given data set and a given model. Models in statistics are, with some exceptions, ad hoc, although not arbitrary, and so should be treated consistently as approximations and not as truth. I believe there is a true but unknown amount of copper in a sample of water, but I do not believe that there is a true but unknown parameter value. Nor am I prepared to behave as if there were a 'true but unknown parameter value', here put in inverted commas as statisticians often do to demonstrate that they do not actually believe that there is such a thing. Moreover the model has to be an approximation to the data and not to some 'true but unknown generating mechanism'. There may well be some physical truth behind the data but this will be many orders of magnitude more complex than anything that can be reasonably modelled by a probability measure. Furthermore there will be a different truth behind every single observation. (. . .) The definition of approximation used in this book is that a model is an adequate approximation to a data set if 'typical' data generated under the model 'look like' the real data.

Again this implies that many different models can be adequate approximations of the same data and that the job of a model is not to be true.

A difference between Davies's and my own view is that I think that "treating reality *as if* it behaves as the model says" (temporarily, and with skepticism) is a valid and helpful description of the handling of models, whereas Davies thinks it is enough to define the probability models' ability to approximate data mathematically, and no further interpretation is required. Davies and Dümbgen (2022) give an example for a nonstandard use of probability models that uses them as calibration device rather than "assuming" them to be real in any way.

An aspect of the work of both Tukey and Davies is an emphasis on procedures for data analysis rather than models. The latter may be used to assess and potentially define or inspire procedures, but play just an auxiliary role, and are often not required. Tukey (1993) wrote

> (. . .) once we have ceased to give a model's truth a special role, we cannot allow it to "prescribe" a procedure. What we really need to do is to choose a procedure, something we can be helped in by a knowledge of the behavior of alternative procedures when various models (i.e., various challenges) apply (. . .)

I sympathize with this attitude, even though the present chapter on models is not the right place to discuss it in detail. In my impression, Tukey and Davies treat the models as data generating processes, and as such they are interpreted as aleatory.

I will first define what I mean by frequentism-as-model. Then, in order to understand its implications, the role of i.i.d. models and subject matter information is discussed. Reasons why frequentism-as-model can be useful despite the separation between reality and model are given. I will then treat statistical inference based on frequentism-as-model, including the role of model assumptions, robustness and stability, and the key concept of "interpretative equivalence." The last subsection concerns the use of inverse probability logic with aleatory probabilities in the spirit of frequentism-as-model.

3.1 Frequentism-as-Model as Interpretation of Probability

Frequentism-as-model is a version of the aleatory interpretation of probability and can be seen as a flavor of frequentism in the sense that a probability of an event A refers to the limiting relative frequency for observing A imagining an idealized infinite repetition of a random experiment with possible outcome A. This comes with problems, as discussed earlier.

Von Mises and most other frequentists claimed that their interpretation is applicable to experiments allowing for lots of (idealized infinitely many, idealized identical) repetitions, but in fact in standard frequentist probability modeling, such repetitions are modeled as i.i.d., meaning that a whole sequence of repetitions is modeled by a single probability model, modeling probabilities not only for the outcomes of a single replicate but for combinations of outcomes of some or all replicates. Central results of probability theory such as the laws of large numbers and the central limit theorem apply to such models. Applying the traditional frequentist interpretation of probability to them would require whole independent and potentially infinite sequences to be repeated, which of course is impossible in reality.

Frequentism-as-model deals with the issues in two ways. Firstly, it emphasizes that probability is a model and as such fundamentally different from observed reality. Adopting frequentism-as-model means to think about an experiment or a situation *as if* it were generated by a frequentist process, temporarily, without implying that this really is the case. This means that it is not required to make a case that the experiment is really infinitely or even very often repeatable and will really lead to the implied relative frequencies. Particularly, this means that all available observations can be modeled as i.i.d. without having to postulate that the whole sequence of observations can be infinitely or even just once repeated. Frequentism-as-model does not imply a belief in objective randomness and is also compatible with determinism. Obviously, insisting on the fundamental difference between model and reality in this way raises the question how we can use such a model to learn something useful about reality, which I will address further. Von Mises and other frequentists already acknowledged that frequentist probabilities are idealizations. To me it seems, however, that in traditional frequentism this is only a response to critics and that otherwise it does not have consequences regarding data analyses and their interpretations.

Secondly, whereas it is not clear how to connect the limits of traditional frequentist infinite sequences to an observed finite amount of real data as any limit

is invariant against arbitrary changes of any finite beginning of a sequence, i.i.d. models, as well as models with other fully specified dependence and non-identity structures, make probability statements about finite sequences of observations, and these can be checked against the actual observations. It can be found out if and in which sense the observed sequence would have been "typical" according to the specified model. In fact, whereas we may not be able to observe another real-data sequence of the same length, and surely not arbitrarily many, we are able to generate an almost arbitrary number of repetitions of sequences from the model involving just the light idealization that we would need a perfect random number generator. Even without a random number generator, in many situations, probability theory can tell us what to expect if the model holds. This resonates with Gillies's long run propensity interpretation of probability, where probabilities are not defined referring to infinite sequences, but rather as tendencies to observe in finite experiments a "typical" (i.e., large probability) outcome as specified by the model.

An issue with this is the definition of "typical." For example, assuming an i.i.d Bernoulli model, a sequence that has 50 zeroes, then 50 ones, then 50 zeroes, than 50 ones again, and so on (called 50-50-sequence in the following), has the same probability as a randomly looking sequence of zeroes and ones that gives us no intuitive reason to doubt the i.i.d. assumption. On what basis can we claim that the i.i.d. model is adequate (to use Davies's terminology) for the latter sequence but not for the 50-50-sequence? This requires subjective researcher input. The researcher needs to decide about the specific way, or ways, in which a sequence has to be typical according to the model in order for the model to be "adequate." For example, the researcher may decide that the model is only useful for data in which the lengths of "runs" of zeroes or ones in a row are not longer to what is expected under the model, in which case the model will be ruled out for the 50-50-sequence (which can be formally done using the runs test, Lehmann (1986), p. 176), whereas the irregular sequence will count as typical. A reason for such a decision may be that the researcher thinks that sequences with long runs of the same outcome are better modeled by models involving positive dependence between the individual binary experiments and that, under such models, it can be shown that the method of analysis the researcher would want to apply to an i.i.d. Bernoulli sequence will be misleading. This could, for example, be a confidence interval for the probability of observing a one. Keeping in mind the fundamental difference between model and reality, the researcher in this situation does not make a decision about which model she believes is in fact true. For such a belief, the subjective decision to distinguish between two sequences of the same individual probability may look questionable. The actual decision is rather about case-dependent criteria regarding in terms of what models they want to think when analyzing the data, informed by the aim of data analysis as well as the characteristics of potential alternative models, and knowledge about the nature of the data and potential reasons for dependence. Note though that if the researcher uses the runs test as decision rule, she will never be able to observe a 50–50-sequence whenever she uses the i.i.d. Bernoulli model, under which such sequences should actually be possible. This is an instance of the "misspecification paradox"; see Sect. 3.4.3.

Probability theory offers ways to help the researcher in making such decisions, by, for example, addressing questions of the form "If indeed the model is as assumed, how likely is it to decide against it; or if a different model is true, how likely is it to not detect it and wrongly stick to the original one; and how bad would the consequences of such an erroneous decision be in terms of the probability of getting a misleading result of the final analysis?" The researcher may also decide to adopt a more robust final method of analysis in order to deal appropriately with a larger number of models between which the data cannot distinguish.

By always keeping in mind that reality is only treated temporarily *as if* generated by a certain model, not ruling out the possibility that it could be different (for which we may consider different models), and even taking this possibility into account when making further data analytic decisions, frequentism-as-model takes the fundamental difference between model and reality much more seriously into account than a standard frequentist analysis, in which a model is used and normally no longer questioned after some potential initial checks. The latter leads to interpretations of the final results that often seem to naively imply that the model is true, even if the researcher using such an approach may well admit, when asked explicitly, that reality actually differs from the model.

3.2 The Role of i.i.d. Models and Subject-Matter Knowledge

Thinking about reality *as if* it were generated by a certain probability model has implications. A critical discussion of these implications is a way to decide about whether or not to adopt a certain model, temporarily, and on top of this, it bears the potential to learn about the real situation. If a sequence is thought of as generated by an i.i.d. process, it means that the individual experiments are treated as identical and independent. De Finetti made the point that whatever can be distinguished is not identical. Treating experiments as identical in frequentism-as-model does not mean that the researcher believes that they are really identical, and neither do they have to be really identical in order to license the application of such a model. Rather it implies that the differences between the individual experiments are temporarily ignored or, equivalently, treated as irrelevant. Analogously, using an independence model does not mean that the researcher believes that experiments are really independent in reality, but rather that she assesses potential sources of dependence as irrelevant to what she wants to do and how she wants to interpret the result. In this way, i.i.d. repetitions are constructed by the researcher. Of course, depending on what kind and strength of dependence can be found in reality, this may be inappropriate. On such grounds, the model can be criticized based on subject-matter knowledge and may be revised. This allows for a discussion about whether the model is appropriate, on top of possible checks against the data, which may have a low power detecting certain deviations from an i.i.d. model.

I.i.d. models, and their Bayesian counterpart, exchangeability models, play an important role in frequentism-as-model as well as in applied statistics in general. Of course, many non-i.i.d. models are used such as regression where the response

distribution is non-identical depending on the explanatory variables, and models for time series or spatial dependence. Even those models will normally have an i.i.d. element, such as i.i.d. residuals or innovations, be it conditional in a Bayesian setup. This is not an accident. Treating experiments as identical and independent (or having regular and fully specified dependence or non-identity structures that will usually involve an i.i.d. element) specifies how an observation from one experiment can be used for learning about another, which is the core aim of statistics. Statistics relies on models that allow us to use different observations to accumulate information. This particularly means that the statistician will need to make decisions to ignore certain potential irregular differences or dependencies in order to be able to do her job to aggregate information, regardless of whether her position is frequentism-as-model, traditional frequentism, or an epistemic Bayesian approach. Obviously, i.i.d. models do not necessarily lead to convincing results, and ultimately, it is a matter of experience and practical success in what situations the use of these models is appropriate. Note in particular that it is possible with a sufficiently large amount of data to distinguish, with large probability, i.i.d. data from data with a regular dependence structure such as ARIMA. It is however not possible to distinguish i.i.d. data from arbitrarily irregular structures of dependence or non-identity. A model that states that the first observation is random but conditionally on it all further observations are fixed with probability one can never be ruled out based on the data alone, and subject-matter arguments are required to make a reasonably regular model plausible.

At the stage of analysis, both checking of adequacy based on the data and understanding and discussion of the subject matter background are of crucial importance. The charge of an ignorance of subject-matter information that Bayesians often raise against frequentists does not affect frequentism-as-model. On the flipside, frequentism-as-model obviously relies on subjective decisions of the researchers, so that a statistician holding it cannot proudly claim that analyses are fully objective, as some frequentists and objective Bayesians like to do, which is questionable anyway (Gelman and Hennig (2017)).

As an example for subject-matter arguments, in a much-cited and influential study that prompted a Lancet editorial (Lancet (2005)) with the title "The end of homeopathy," Shang et al. (2005) carried out a meta-analysis of eight studies comparing homeopathy with a placebo using a standard meta-analysis model with a random study effect assumed i.i.d. (note that this implies dependence between the observations within the same study). They found that there is no evidence that the odds ratio between "homeopathy works" and "placebo works" is different from one and concluded (not inappropriately) that

> this finding is compatible with the notion that the clinical effects of homeopathy are placebo effects.

Here are two implications of the i.i.d. model for the random study effect. Firstly, modeling it as identically distributed implies that the differences between studies are attributed to random variation, which basically means that knowledge of known

differences between the studies is ignored for modeling. This particularly includes differences between different modes of applying homeopathy. For example, there were only two studies used in the meta-analysis in which homeopathy was applied in the classical way with individual repertorization, which is seen as the only proper way of applying homeopathy by many homeopaths. Modeling the study effect of these studies as coming from the same distribution as all the other study effects may be justifiable if there is an a priori belief that homeopathy basically is a placebo and the differences between application modes are irrelevant, but it is not suitable for convincing skeptical supporters of classical homeopathy who would be willing to accept a fair study.

Secondly, an expected value of 0 of the random effect is not assumed for every single study, but overall studies, with a potentially arbitrary large between-studies variance. This implies that the effective sample size for estimating the odds ratio is not the number of individual patients treated in all studies (several thousands), but the number of studies, $n = 8$, and $n = 2$ for classical homeopathy only. The authors may use this to justify their decision to not make a modeling difference between different modes of applying homeopathy. The low number of studies was selected by Shang et al. (2005) by a study quality criterion. The power of the resulting test is therefore very low, and consequently, the confidence interval that the authors give for the odds ratio is very wide. This argument applies to meta-analyses with a random effect generally; in the literature one can sometimes find the informal remark that such analyses require a large number of studies; see Kulinskaya et al. (2008). On top of that, it could be discussed whether there may be reasons for assuming systematic bias over all studies, which in the model would amount to a nonzero expectation of the random effect.

Obviously, $n = 8$ is not sufficient to check the i.i.d. assumption for the random study effect at any reasonable power. Rather than having any connection to an observable "truth," the random effect is a convenient modeling device allowing for potential individual study effects to be captured by a single parameter. A low number of parameters is, according to statistical theory, useful with a small n. This comes at the price of precluding the possibility to involve specific information about differences between the individual trials. This implication is usually ignored.

Given this, it looks very dubious that the Lancet (2005)'s issue editorial concluded:

> Surely the time has passed for selective analyses, biased reports, or further investment in research to perpetuate the homeopathy versus allopathy debate.

Regarding the purely statistical evidence, the meta-analysis could as well be used to argue that there are not yet enough high-quality studies.

In principle such discussions can also take place regarding the traditional frequentist use of models. However, frequentism-as-model reframes these discussions in a helpful way. It makes us aware of the implicit meaning of the model assumptions, which can be used to discuss them. Moreover, it emphasizes that what has to be decided is not the truth of the model, but rather a balance between the capacity of the model to enable learning from the existing observations about future

observations or underlying mechanisms on one hand and on the other hand taking into account all kinds of peculiarities that may be present in the real situation of interest, but that may be hard to incorporate in such a model despite potentially having a more or less strong impact on the results.

A key aspect is that rather than thinking in binary terms about whether a model is true or not, frequentism-as-model implies that the connection between observed reality and models can be assessed gradually as more or less close. It is neither impossible nor inadmissible to conceive a situation as a random experiment that cannot or only a very limited number of times be repeated and to use a model in the sense of frequentism-as-model for it. This would imply to *think of* the situation as hypothetically repeatable, and of certain outcomes having a certain tendency, or, in standard philosophical terms, propensity to happen, which would materialize in case of repetition, but not in reality.

Therefore, frequentism-as-model does not bar single-case probabilities as von Mises's flavor of frequentism does. However, it is more difficult for the person who puts up such a model to convincingly justify it compared to a situation in which there are what is interpreted as "replicates." Convincing justification is central due to the central role of communication and agreement for science in the philosophy outlined above. Whatever the model is used for, it must be taken into account that hypothesized values of the single-case probability cannot be checked against data. This changes if the experiment is embedded in a set of experiments that are not directly seen as repetitions of each other, but for which an assumption is made that their results are put together from systematic components and an error term, and the latter is interpreted as i.i.d. repetition. This is actually done in standard frequentist regression analyses, in which the explanatory variable or vector $\mathbf{x}$ is often assumed fixed. For a given value or vector of values of the explanatory variable(s) $\mathbf{x}$, a distribution for the response y is implied, despite the fact that the random experiment can in reality not be repeated for any given $\mathbf{x}$ in situations in which the researcher cannot control the $\mathbf{x}$. Insisting on repeatability for the existence of traditional frequentist probabilities would imply that nothing real corresponds to the distribution of y for fixed $\mathbf{x}$, and there is no way to check any probability model. The unobservable i.i.d. error term constructs repeatability artificially, as in this way observations with different $\mathbf{x}$ can be interpreted as involving repeated realizations of the error term. I have not seen this mentioned anywhere. From a frequentism-as-model perspective, this needs to be acknowledged, and dependent on the situation it may be seen as appropriate or inappropriate, once more using information from the data and about the subject matter background, but there is nothing essentially wrong or particularly suspicious about it, apart from the fact that it turns out that our possibilities to test model assumptions are quite generally more limited than many might think (see Sect. 3.4.3).

3.3 How Is Frequentism-as-Model Useful?

From the discussion of mathematical models and reality, it follows that the separation between reality and model is a general feature of mathematical models, not an

exclusive one of frequentism-as-model, although frequentism-as-model acknowledges it explicitly. The question "how is frequentism-as-model useful, and how is it connected to reality?" is therefore connected with the general question how mathematical models can be useful, even if "wrong" in Box's terms. The way to use mathematical models is obviously to take the mathematical objects that are involved in the model as representations of either perceived or theoretically implied aspects of reality and then to use this to interpret the results of using the models.

In order for the model to be useful, does not reality have to be "somehow" like the model? According to the constructivist view presented earlier, reality is accessible only through personal perception and communication, so the best we can hope for is that the model can correspond to how we perceive, communicate, and think about reality. Regarding frequentism-as-model, this concerns in the first place the concept of i.i.d. repetition as already discussed above. I.i.d. repetition is a thought construct and not an objective reality; however, even apart from probability modeling, it is easy to see how much of our worldview even before any science relies on the idea of repetition, be it the cycles of day and night and the seasons, or be it the expectation, out of experience, of roughly the same behavior or even roughly the same distribution of behaviors when observing any kind of recurring process as diverse as what a baby needs to do to attract the attention of the parents, look, size, and edibility of plants of the same species, or any specific industrial production process. Objections against the truth of "i.i.d" can be easily constructed for observations in all of these examples, yet analyzing them in terms of i.i.d. models often gives us insight, new ideas, and practical competence to deal with these processes, be it with additional ingredients such as explanatory variables, regular dependence structures, or seasonal components. Ultimately, the success of probability modeling has to be evaluated empirically, relative to researchers' aims, and obviously the result will sometimes, but not always, be positive. I remain agnostic regarding to what extent positive results confirm that "objective reality really is like that," to which there ultimately is no answer beyond what occurs in personal and social reality.

A major general reason for using probability models is that they allow us to quantify random variation. It is not an accident that probability theory is a relatively young branch of mathematics. Unlike most historically older mathematical modeling, probability modeling is essentially not only about what was, is, or will be observed, but also about what could have been observed but was not and/or will not be observed. The latter does not and will not exist as observation. Connecting this to observations is tricky and, in my view, a major reason why probability models are so controversial and problematic. However, they allow us to deal with a concept that was hardly acknowledged when probability theory emerged and became hugely influential, namely, the distinction between a meaningful pattern and meaningless variation.

What can be done in order to support useful and avoid misleading results of probability modeling interpreted as frequentism-as-model? Below are some questions that need to be addressed. These questions are not very surprising, and statisticians, be they traditional frequentists, Bayesians, or something else, may think that they address these anyway. What makes frequentism-as-model different is that a standard approach would be to take these issues into account when setting up and justifying the model, and then all further analysis and interpretation is done

conditionally on the model being true. In frequentism-as-model, there is no such thing as a true model, meaning that the issues below need to be kept in mind, and they still play a role when running model-based analyses and interpreting the results.

1. What is the aim of modeling? Probability models are used in different ways, and this is important for how to set them up. Hand (1994) emphasized the importance of clarifying the research question for statistical analysis, with several examples illustrating how much of a difference this can make, regarding modeling and analysis, but also in earlier stages such as the planning of experiments. Although the aim of modeling is not always ignored in traditional frequentist or Bayesian analysis, there is a general tendency to think that the ultimate aim is always to find the true model, and if this (or at least a model as "close" as possible to it) is found, everything "relevant" that can be known is known. This is fundamentally different from the attitude implied by frequentism-as-model.

   Cox (1990) lists three different classes of modeling aims, namely, "substantive models," which model (often causal) mechanisms that the researcher is interested in, "empirical models," which try to aid inference about relations between variables such as effect sizes, without making an attempt at reconstructing how the relations work in any detail, and "indirect models," where models are not directly connected with a real situation. This encompasses the use of models for deriving or comparing methodology, but also, in a real situation but only indirectly serving the main aim of analysis, issues such as the imputation of missing values. The classes may be mixed or intermediate in some specific situations.

   This has methodological consequences. Substantive models may have aspects that are there for representing the researcher's view but may refer to unobservables, in which case they cannot or in the best case very indirectly be tested against the data. A substantive model may even be there to communicate a researcher's view without the ambition to fit any data. It may at times represent scenarios that are seen as best or worst cases, intentionally not modeling what is seen as most realistic. Parsimony may be a strong concern if the model is used to communicate an idea, a weaker concern if the model is used for prediction, and no concern at all if the model is used to simulate data for exploring potential future variation in a complex system. In some cases, the major use of the model is learning from discovering its lack of fit and the exploration of reasons.

   At first sight, it seems central for empirical models to fit the data. However, for such models, the specific modeling aim is important, because deviations of the data from what would be typical for the model are only a problem to the extent that they raise the danger of misleading results, which depends on what kind of result the researcher is after, e.g., modeling integer number data by a Gaussian distribution is not normally a problem if only inference about means is required, at least not as long as the i.i.d. assumption is implied and the distributional shape is not strongly skew or with large outliers, both of which can happen with supposedly continuous data as well.

   A major indirect use of models is the investigation of properties of inference methods by theory or simulation. For example, maximum likelihood (ML)

estimators can be seen as inspired by probability models that are interpreted in a frequentism-as-model sense: "Imagine a real situation as modelled, then the ML estimator makes the data most likely." The major aim of modeling there is to find a good method for estimation. Even many Bayesians use models with an implicit frequentist interpretation in methodological work in order to investigate the quality of their methods in a situation with a constructed known truth. It could be argued that in order for such work to be applicable to objective reality, objective reality has to be as the model specifies, but if we accept that "all models are wrong" in objective reality, there is no better alternative than studying the behavior of method under a variety of such models. Simulations and theory will seem more relevant if the used models seem "realistic," but as long as the aim is not to address a very specific application, models used in this way are not checked against data. This only can happen when the resulting methods are applied.

Furthermore, for some modeling aims, there are methods that can more or less directly measure how well the aim is fulfilled. A major example is cross-validation for prediction quality. If prediction is the major modeling aim, good cross-validation results can be seen as dominating other concerns such as fit of the existing data or correspondence with subject matter knowledge.

Another example for direct measurement of what is of interest is the objective function of *k*-means clustering (for details, see Hennig et al. (2015)), which directly measures the quality of the approximation of any data point by the closest cluster centroid (more generally, often loss functions can be constructed that formalize the aim of analysis). If achieving a good performance in this respect is the aim of analysis (e.g., if clustering is used for data set size reduction before some other analysis, and all clustered objects are replaced by the cluster centroid in a final analysis), *k*-means clustering is a suitable method even in situations that are very different from the model for which *k*-means is ML, namely Gaussian distributions with same spherical covariance matrix in all clusters but different cluster means. Statisticians have argued that *k*-means often does not work well if this assumption is violated (e.g., Vermunt (2011) in a discussion of Steinley and Brusco (2011)), but this relies on the view that reconstructing the true model is required, which may deviate from what is needed in practice.

2. Is the model compatible with the data? I will discuss this in more detail in the subsection on statistical inference.
3. How does the model relate to subject matter knowledge? The model should be informed by perceptions and ideas (subject matter knowledge, which is not necessarily "objective") regarding the real process to be modeled, such as specific reasons for dependence, non-identity, or also for or against particular distributional shapes such as symmetry. Just to give an example, dependence by having the same teacher and communication between students makes the use of an i.i.d. model for analyzing exam results for students from the same class or even school suspicious, and it can be controversially discussed to what extent introducing dependence only in form of an additive random effect addresses this appropriately.
4. How stable are conclusions against modeling differently? The ultimate answer to the question how we can be sure that a model is appropriate for the situation we

want to analyze is that we cannot be. Breiman (2001) mentioned the "multiplicity of good models" in line with Tukey (1997) and Davies (2014), and different "good" models may lead to contradicting conclusions. On top of that, research hypotheses may be operationalized in different ways, using different measurements, different data collection, and so on. There is no way to be sure of a scientific claim backed up by a statistical result based on the naive idea that the model is true and the method is optimal. In order to arrive at reliable conclusions, science will need to establish the stability of the conclusions against different ways of operationalizing the problem. This is basically the same way in which we as individual human beings arrive at a stable concept of the world, as far as it concerns us; after lots of experiences, looking at something from different angles in different situations, using different senses, communicating with others about it, etc., we start to rely on our "understanding" of something. In the same way, every single analysis only gives us a very restricted view of a problem and does not suffice to secure stability; see the discussion of "interpretative equivalence" in Sect. 3.4.1. Currently, there is talk of a "reproducibility crisis" in a number of disciplines (Fidler and Wilcox (2018)). I think that a major reason for this is that no single analysis can establish stability, which is all too conveniently ignored, given that researchers and their funders like to tout big meaningful results with limited effort. Even if researchers who try to reproduce other researchers' work have the intention to take the very same steps of analysis described in the original work that they try to reproduce, more often than not there are subtle differences that were not reported, for example, data-dependent selection of a methodology that the reproducer then uses unconditionally. Any single analysis will depend on researchers' decisions that are rarely fully documented, could often be replaced by other decisions that are not in any way obviously worse, and may have a strong impact on the result. Having something confirmed by as large as possible number of analyses investigating the same research hypothesis of interest in different ways is the royal road to increase the reliability of scientific results. In this I agree with the spirit of Mayo (2018)'s "severity," and it also resonates with her "piecemeal testing"; a scientific claim is the more reliable; the harder researchers have tried to falsify it and failed. Frequentism-as-model surely does not license claims that anything substantial about the world can been proved by rejecting a single straw man null hypothesis.

A further "use" of the consciousness and acknowledgement of the essential difference between model and reality may be a better awareness that probability modeling is not mandatory; data can be often analyzed descriptively or visually without the need to compute p-values or posterior probabilities, and in some situations, quantification itself can be regarded with more skepticism (Tafreshi et al. (2016) and Chang (2022)).

3.4 Frequentism-as-Model for Statistical Inference

Adopting frequentism-as-model as interpretation of probability does not imply that classical methods of frequentist inference, tests, confidence intervals, or estimators

have to be used. It may also be combined with Bayesian inference as discussed later. However, the classical methods of statistical inference have valid frequentism-as-model interpretations, and I believe that understanding these interpretations properly will help to apply the methods in a meaningful and useful way. Therefore, I disagree with recent calls to abandon significance testing or even frequentist inference as a whole (Wasserstein et al. (2019)).

3.4.1 Interpretative Equivalence, Stability, and Robustness

The question how stable our conclusions are as raised above does not only depend on models, methods, and the data, but also on the conclusions that are drawn from them, i.e., the interpretation of the results by the researcher. A helpful concept to clarify things is "interpretative equivalence." I call two models "interpretatively equivalent" with respect to a researcher's aim if their subject-matter interpretation of what they are interested in is the same regardless of which of the two models is true. For example, if a researcher is interested in testing whether a certain treatment improves blood pressure based on paired data before and after treatment, she will in all likelihood draw the same conclusion, namely that the treatment overall does not change the blood pressure, if the distribution of differences between after and before treatment is symmetric about zero, be it a Gaussian or a t_2-distribution, for example. For her research aim, the precise shape of the distribution is irrelevant. Comparing two models with expectation zero, one of which is symmetric whereas the other one is not, this is not so clear; the treatment changes the shape of the distribution, and whether this can be seen as an improvement may depend on the precise nature of the change. For example, consider the model

$$0.99\mathcal{N}\left(\mu^*,\ \sigma^2\right) + 0.01^*\delta_{100}, \tag{1}$$

δ_x being the one-point distribution in x. The researcher may think of δ_{100}-distribution as modeling erroneous observations, in which case $\mu^* = 0$ makes the model interpretatively equivalent to $\mathcal{N}(0,\ \sigma^2)$, whereas for $\mu^* = -\frac{1}{0.99}$, despite expectation zero, the researcher will interpret an average effect of lowering the blood pressure, which the model implies for a subpopulation of 99% of people.

In fact, the term "interpretative equivalence" implies an oversimplification, as models can be more or less interpretatively similar, and it is not always possible or helpful to make a binary distinction between "interpretative equivalence" and "interpretative difference," but I will stick to this distinction here for ease of discussion.

The importance of interpretative equivalence is that it allows to discuss what deviations of a temporarily assumed model should be handled in which way. If conclusions are drawn from a method involving certain model assumptions, it would be desirable, and conclusions can be seen as stable, if the probability would be roughly the same that the researcher arrives at the same conclusion from the same data if data are actually generated by an interpretatively equivalent model other than the assumed one. If this is not the case, it would be of interest if the models can be told apart with good probability by the data, in which case the data can be used to

decide on which one of these models to base a data analysis. This also applies to the decision between a specific parametric and a more general but potentially less precise nonparametric approach.

The concept also clarifies the issue that tests on large samples reject point null hypotheses too easily. Considering a $\mu = 0$ null hypothesis, a model with $\mu \neq 0$ but a very small absolute value will lead to a rejection of $\mu = 0$ with a large probability for large n. This is a problem if and only if models with, say, $|\mu| \leq \varepsilon, \varepsilon > 0$ are considered interpretatively equivalent to the model with $\mu = 0$, in other words, if a difference as small as ε is considered irrelevant from a subject matter point of view. If this can be specified, the question whether values of μ with $|\mu| \leq \varepsilon$ are in the corresponding confidence interval, or what severity the test achieves ruling out $|\mu| \leq \varepsilon$, see Mayo (2018), is more relevant than whether a test of $\mu = 0$ is significant. Note that in a real situation it will hardly be possible to specify a precise borderline ε, and one could think that things are rather continuous than binary; however, due to the discreteness of language, in many situations at least certain nonzero differences will be called irrelevant.

Being concerned about interpretative equivalence of models is an entry point for robust methods, which are designed for dealing with the possibility that a nominal model is violated in ways that are hard to diagnose but may make a difference regarding analysis. Robustness theory (e.g., Hampel et al. (1986); Huber and Ronchetti (2009)) is about limiting changes in results under small changes of the modeled processes. An issue with standard robustness theory is that it is often implied that "contamination" of a distribution should not have an influence of the results, whereas in practice the contamination may actually be meaningful. To decide this is a matter of assessing interpretative equivalence. Frequentism-as-model acknowledges that desirable behavior of a statistical method is not just something "objective" but depends on how we interpret and assign meaning to differences between distributions. Robustness is helpful where the reduced sensitivity of robust methods to certain changes in the data or model takes effect if the changes are indeed interpreted as meaningless relative to the aim of analysis; for example, there needs to be a decision whether in model (1) what orresponds to the quantity of interest is rather the mean of the distribution, or rather the μ^*. In some applications, it is inappropriate to use a method of which the results may not change much when replacing up to half of the data, namely, where the resulting processes are considered as interpretatively very different, but in any case comparing the results of such a method with a more "sensitive" non-robust one may allow for a more differentiated perception of what is going on than does every single method on its own.

3.4.2 Significance Tests and Frequentist Inference

Statistical hypothesis tests have become very controversial (e.g., Wasserstein et al. (2019)), but the interpretation of tests from the position of frequentism-as-model is rather straightforward. They address to what extent models are compatible with the data in the sense formalized by the test statistic. For simplicity, I mostly call models "compatible" or "incompatible" with data, but in fact compatibility is gradual, as expressed by p-values.

Obviously, a non-rejection cannot be an indication that the model is in fact true, but it means the absence of evidence in the data that reality is different in the way implied by the test statistic, making it impossible to claim such evidence. Sometimes tests, particularly two-sided tests, are criticized because "point null hypotheses are usually scientifically implausible and hence only a straw man" (e.g., Rice et al. (2020)), but then a parametric model allowing for any parameter value cannot be "really true" anyway, so that for the same reason for which the point null hypothesis (H_0) is "implausible," the whole parametric family is implausible as well, but that does not necessarily stop inference about it from being informative and useful. What a test can do is neither to confirm H_0 as true, nor to allow to infer any specific alternative in case of rejection. A two-sided test should be run if compatibility of the data with H_0 is a possibility of interest; even then it may not be the method of choice if the sample size is too large, see the earlier discussion. Recall that tests are often chosen according to Neyman-Pearson logic (Lehmann 1986). This means that they are meant to maximize the power, i.e., to minimize the probability for non-rejection given a specific alternative hypothesis and given a fixed value for the probability to reject a true H_0. This is often read as being only meaningful if either the nominal H_0 or the specific alternative is true. Instead, I see this reasoning as a rather clever idea to motivate test construction in a real situation, using mathematical modeling. A reasonable formal criterion is optimized in the model world, leading to a procedure that can be applied in the real world, which differs from the model world. It is important to ask to what extent the differences between the real world and the model world have the potential to mislead the procedure, but this can only be assessed case by case. For sure it cannot be required that the real situation satisfies the formal model optimality. Demanding this would be a category error.

Often significance tests are presented implying that a rejection of the H_0 is meaningful whereas a non-rejection is not. Above I have explained the meaning of a non-rejection in frequentism-as-model; the meaning of rejection is a more complex issue. Obviously, rejection of the H_0 does not imply that there is evidence in favor of any particular model that is not part of the H_0. On the positive side, a rejection of the H_0 gives information about the "direction" of deviation from the H_0, which can and mostly should be substantiated using confidence sets of compatible parameter values, and particularly data visualization for exploring how this plays out without relying on the specified model.

Consider a one-sample t-test of the H_0: $\mu = 0$ regarding $\mathcal{N}(\mu,\ \sigma^2)$ against $\mu > 0$. The test statistic T is the difference between the sample mean and $\mu = 0$ standardized by the sample standard deviation. Sticking to model-based thinking while not implying that $\mathcal{N}(\mu,\ \sigma^2)$ is true, rejection of the H_0 can be interpreted as providing evidence in favor of any distribution of the two samples for which a larger value of T is observed with larger probability than under H_0, compared with the H_0. We do not only have evidence against H_0; we also learn that the problem with H_0 is that $\mu = 0$ is most likely too low, which is informative.

In many situations in which the one-sample t-test is used, the user is interested in testing $\mu = 0$ or rather its interpretative meaning, but not in the distributional shape $\mathcal{N}(\mu,\ \sigma^2)$, which just enables them to run a test of $\mu = 0$. If in fact there were a true

distribution, which of course we can consider, in the "as if"-model world, which is not $\mathcal{N}(\mu, \sigma^2)$ with any specific μ or σ^2, it is not clear in a straightforward manner what would correspond to μ. It may be the expected value, but in a model such as (1) it may well be seen as the μ^* that governs 99% of the observations. In any case, there are distributions that could be considered as interpretatively equivalent to the actual H_0. Any distribution symmetric about zero is in most applications of t-tests interpretatively equivalent to the H_0, meaning that if indeed a distribution different from $\mathcal{N}(\mu, \sigma^2)$ but still symmetric about $\mu = 0$ were true, it would be desirable that the probability to reject H_0 would be as low as if H_0 were true.

One can then ask how likely it is to reject the H_0 under a model that is not formally part of the H_0, but interpretatively equivalent. For the one-sample t-test, assuming existing variances, many distributions interpretatively equivalent to $\mathcal{N}(\mu, \sigma^2)$ are asymptotically equivalent as well, and results in Cressie (1980) and some own simulations indicate that it is very hard if not impossible even in finite samples to generate an error probability of more than 7% for rejecting an interpretatively true H_0 with a nominal level of 5%, even if the underlying distribution is skew and μ is taken to be the expected value, meaning that testing at a slightly smaller nominal level than the maximum error probability that we want to achieve will still allow for the intended interpretation.

However, the situation becomes much worse under some other deviations from the model assumptions, particularly positive dependence between observations, which increases the variation of T and may lead to significant T and increased type I error probability very easily, also in situations that are interpretatively equivalent to the H_0. Consider an example in which the expected change of the turnover generated by a salesperson after attending a sales seminar is of interest in order to evaluate the quality of the instructor. If all salespersons in the sample attend the sales seminar together, they may learn something useful from talking with each other about sales strategies, even if what the instructor does itself is useless.

What is required in order to achieve a reliable conclusion is just what was discussed in the section on the use of models. The general question to address is whether a rejection of H_0 could have been caused by something interpretatively equivalent to the H_0, in which case the conclusion could not be relied upon. The questions to ask are as follows: (1) is there any evidence in the data that such a thing may have happened, (2) is there any subject matter knowledge that suggests this (e.g., salespersons may have learnt from each other rather than from the seminar), (3) does the meaning of the test statistic correspond to what is of interest, and (4) is it feasible to run further analyses, with the same or other data, that can confirm the conclusion? (4) is important because even the best efforts to use (1) and (2) will not be able to remove all doubt. Models are always conceivable that can cause trouble but cannot be distinguished from the nominal model based on the data, and thinking about the subject matter may miss something important. (3) is in my view a very important issue that is not usually appreciated (but see Hand (1994)). Tests are usually derived using optimality considerations under the nominal model, but this does not imply that they are optimal for distinguishing the "interpretative H_0" from the "interpretative alternative." In many cases at least, they make some good sense,

but a model like (1) may lead them astray. The full interpretative H_0 and the interpretative alternative will normally be too complex to derive any optimal test from them, but the form of the test statistic itself suggests what kinds of distributions the test actually distinguishes. This should be in line with what is interpreted as the difference of interest between H_0 and the alternative. At least under a point H_0, the distribution of any test statistic can be simulated (if the H_0 is a set of distributions, parametric bootstrap can be used, if potentially involving a certain bias). Assuming that rejection happens for large values of the test statistic, it implicitly defines the interpretative H_0 and the interpretative alternative distinguished by the test as the set of distributions for which the test statistic is expected low and the set of distributions for which it is expected larger, respectively. In many situations, a nonstandard test statistic may be better suited than what is optimal under a simple nominal model (see Hennig and Lin (2015) for an example).

Bayesians often argue against frequentist inference by stating that the probabilities that characterize the performance of frequentist inference methods such as test error and confidence set coverage probabilities are pre-data probabilities and that they do not tell the researcher about the probability for their inferences to be true after having seen the data. One example is that a 95% confidence interval for the mean of a Gaussian distribution in a situation in which the mean is constrained to be larger than zero may consist of negative values only, meaning that after having seen the data the researcher can be sure that the true value is not in the confidence interval. This is a problem if the coverage probability of the confidence interval is indeed interpreted as a probability for the true parameter value to be in the confidence interval, which is a misinterpretation; the coverage probability is a performance characteristic of the confidence interval pre-data, assuming the model. From the point of view of frequentism-as-model, this is not a big problem, because there is no such thing as a true model or a true parameter value, and therefore, a probability for any model to be true is misleading anyway, although it can be given a meaning within a model in which the prior distribution is also interpreted in a frequentism-as-model sense (see Sect. 3.4.4). In practice, if in fact a confidence interval is found that does not contain any admissible value, this means that either the restriction of the parameter space that makes all the values in the confidence interval impossible can be questioned or that a truly atypical sample was observed. Davies (1995) defined "adequacy regions" that are basically confidence sets defined based on several statistics together (such as the mean or median, an extreme value index, and a discrepancy between distributional shapes; appropriately adjusting confidence levels) that can by definition in principle rule out all distributions of an "assumed" parametric family, meaning that no member of the family is compatible with the data given the combination of chosen statistics. Normally confidence intervals are interpreted as giving a set of truth candidate values assuming that the parametric model is true; but if that assumption is dropped from the interpretation and it is just a set of models compatible with the data defined by parameter values within a certain parametric family, it is possible that the whole family is not compatible with the data.

Overall, I suspect that the biggest problem with interpreting standard frequentist inference is that many of the people who use it want to make stronger claims, and

want to have bigger certainty, than the probability setup that they use allows. An example for this is the ubiquity of implicit or explicit claims that the null hypothesis is true in case that it was not rejected by a test. Interpreting the results of classical frequentist inference assuming the truth of the model will generally lead to over-interpretation. According to frequentism-as-model, classical frequentist inference allows compatibility and incompatibility statements of models with data. As far as the models are used to make substantive statements, such compatibility assessment is of course of interest. What exactly can be learnt from this depends on the analyzed situation as well as on what the researcher sees as interpretatively equivalent in that situation. From the point of view of frequentism-as-model, research about the expected variation of results under interpretatively equivalent models would be desirable, acknowledging that interpretative equivalence may mean different things in different applications. Deciding on a parametric model and from then on ignoring that other models can be compatible with the data as well will easily lead to overinterpretation.

3.4.3 Model Assumptions

A standard statement regarding statistical methods is that these are based on model assumptions and that the model assumptions have to be fulfilled for a method to be applied. This is misleading. Model assumptions cannot be fulfilled in reality, because models are thought constructs and operate on a domain different from observer-independent reality. For this reason, they cannot even be approximately true, in a well-defined sense, because no distance between the unformalized "underlying real truth" and the model can be defined, although, following Davies, the approximation notion can be well defined comparing observed data to a model.

What is the role of model assumptions then? There are theorems that grant a certain performance of the method, sometimes optimal, sometimes just good in a certain sense, under the model assumptions. The model assumptions are not required for applying the method, but for securing the performance achieved in theory. The theory can be helpful to choose a method, and some theory leads to the development of methods, but the theoretical performance can never be granted in reality. This does not mean that the performance will be bad whenever the model assumptions are not fulfilled. In fact, some aspects of the model assumptions are usually almost totally irrelevant for the performance, such as the application of methods derived from models for continuous data to data that is rounded to, say, two decimal places. Interpretative equivalence is an important concept also in this respect, because optimally methods would not only distinguish what is deemed relevant to distinguish by the researcher under the model assumptions, e.g., a certain H_0 and its nominal alternative, but they would lead to largely the same results for interpretatively equivalent distributions that do not fulfill the model assumptions. As far as this is the case, the corresponding model assumption is irrelevant.

It makes sense, at first sight, to think that the method will be appropriate if the assumed model is a good model, i.e., if reality looks very much like it. This itself can be modeled, meaning that one can look at the performance of a method in a situation in which the assumed method does not hold, but a somewhat similar model, which

can be formalized using dissimilarity measures between distributions. This has been considered in robust statistics (Hampel et al. (1986); Huber and Ronchetti (2009), "qualitative robustness" in particular) with the unsettling result that the performance of some classical statistical methods can decline strongly in arbitrarily small neighborhoods of the assumed model. As a consequence, more robust methods have been developed, i.e., methods with a more stable performance in close neighborhoods. This is surely valuable progress, but on one hand, it does not change the fact that models can be set up that cannot be rejected by any amount of data and that can annihilate the performance of any method, e.g., using irregular dependence and/or non-identity structures, so full safety cannot even be had in the modeled world, let alone the real one. On the other hand, the situation is not quite as hopeless as this may suggest. Usually, bad results about standard methods are worst case results, very often regarding extreme outliers, and it is often plausible, from checks, data visualization, or subject matter considerations, that the worst case does not occur. To argue against detrimental dependence or non-identity structures is much harder, and the best that can be done in this respect is looking for conceivable reasons for trouble, using a suitable method if no such reason comes to mind, and then hope that the world will behave when acting based on the result. Science makes mistakes and is hopefully open-minded enough to correct itself in case of bad outcomes.

Mayo (2018, Sec. 4.8–4.11) and others argue that model assumptions can and should be tested. Whereas the truth of the model cannot be secured, much can indeed be found out about whether a model is adequate for a purpose. However, model checking has a number of issues. The first one is that if a certain model assumption turns out to be compatible with the data, it does not mean that other models that may lead to different results are ruled out. Misspecification testing cannot secure robustness of results against models that cannot be excluded.

A second issue is that if a method is used conditionally on passing a model misspecification test for checking a model assumption, the distribution of the data that eventually go into the method becomes conditional on passing, and this will normally violate the model assumptions, even if they were not violated before. Particularly, it will make observations dependent even if they were not dependent before, because if, for example, $n - 1$ observations are on the borderline for passing the misspecification test, the nth observation has to fit well for passing. I called this "goodness-of-fit or misspecification paradox" in Hennig (2007). In many situations (but not always), this will not change error probabilities associated with the model-based method strongly, meaning that in case the model was true before misspecification testing, not much harm is done.

On the other hand, in case that the model assumptions are violated in a problematic way, the distribution conditionally under passing a misspecification test will often not make the method work better, and sometimes even worse than before testing; keep in mind that just because the misspecification test does not reject the model assumption, it does not mean that it is fulfilled. Shamsudheen and Hennig (2019) reviewed work investigating the actual performance of procedures that involve a misspecification test for one or more model assumptions before running a method that is based on these model assumptions, looking at data that originally

fulfilled or did not fulfill the model assumption. Occasionally, a combined procedure was investigated in which, in case of the violation of a model assumption, a method with lighter assumptions is used. The philosophy of such work, which is in line with frequentism-as-model, is that the key question is not whether the model assumptions are really fulfilled, but what the performance of such combined procedures is in several situations, compared to both the model-based and the less model-based (often nonparametric) test. The results are mixed, depending on the specific combinations of tests, and surprisingly many authors advise against misspecification testing based on their results, e.g., Fay and Proschan (2010):

> The choice between t- and Wilcoxon-Mann-Whitney decision rules should not be based on a test of normality.

Whereas formally requiring lighter model assumptions, the use of nonparametric methods only pays off if the probability to arrive at a conclusion that is interpretatively in line with a modeled truth is better than for a competing parametric method for most models compatible with data and existing knowledge and traded off potentially against the expected sizes of "interpretative differences," as far as these can be specified. This is not always the case.

Shamsudheen and Hennig (2019) showed a theoretical result that presents combined procedures in a somewhat more positive light. They looked at the overall performance of combined methods in a setup where datasets could be, with a certain probability λ, generated by the assumed model, and with probability $1 - \lambda$ by a distribution that could cause trouble for the model-based method (this probability could be taken as rather idealized aleatory one considering many studies of a similar kind). Under certain assumptions for the involved methods, they showed that for a range of values of λ, the combined procedure beats both involved tests, the model-based one and the one not requiring the specific model assumption, regarding power, even if not winning for $\lambda = 0$ and $\lambda = 1$, which is what authors of previous work had investigated.

The surprisingly pessimistic assessment of misspecification testing by authors who investigated its effect may be due to the fact that most available misspecification tests test the model assumption against alternatives that are either easy to handle or very general, whereas little effort has been spent on developing tests that rule out specific violations of the model assumptions that are known to affect the performance of the model-based method strongly, i.e., leading to different conclusions for interpretatively equivalent models, or same conclusions for models that are interpretatively very different. Such tests would be specifically connected to the model-based method with which they are meant to be combined and to the assessment of interpretative equivalence.

Furthermore, as already mentioned, not everything can be tested on data. For example, many conceivable dependence structures do not lead to patterns that can be used to reject independence. For example, it may occasionally but not regularly happen that one observation determines or changes the distribution of the next one. Think of psychological tests in which sometimes a test person discusses the test with

another participant who has not yet been tested and where such communication can have a strong influence on the result.

Frequentism-as-model could inspire work that looks at performance of methods involving model assumptions and misspecification testing under all kinds of models that seem realistic, including Bayesian combinations of different models with different probabilities. Generally, regarding methodological research, a researcher adhering to frequentism-as-model will always be interested in the performance of methods that are supposedly model-based under situations in which the model assumptions are violated, knowing that model assumptions can never be relied upon and that looking at other models is the only way to investigate what happens then. This is in the tradition of robust statistics, but with less focus on worst cases and more focus on interpretative equivalence for deciding what outcome would actually be desirable. The cases that are really the worst ones are hopeless under any approach, and even the best statistics cannot always save the day.

3.4.4 Frequentism-as-Model and Falsificationist Bayes

Epistemic probabilities model a personal or "objective" degree of belief, not the data generating process as such, and therefore, they cannot be checked against and falsified by the data. See Dawid (1982b) for a discussion of "calibration," i.e., agreement or potential mismatch between predictions based on epistemic Bayesian probabilities and what is actually observed. Gelman and Shalizi (2013) argued that Bayesian statistics should allow for checking the model against the data and interpret Bayesian models as modeling data generating processes rather than epistemic uncertainty. This is in line with a lot of applied Bayesian work in which the posterior distribution of the parameter is interpreted as encoding probabilities for a certain parameter being true descriptors of the underlying data generating process. This allows to check the model against the data and potentially to revise it, even though this violates the coherence requirement, which is taken as essential by most advocates of epistemic Bayesian interpretations. Gelman and Hennig (2017) called the connection of inverse probability logic and model checks "falsificationist Bayes."

My interpretation of it is that probabilities in falsificationist Bayes are interpreted in the same manner as in frequentism-as-model. The parametric model is handled *as if* it describes the data-generating process but can be dropped or modified if falsified by the data. The involved interpretation of probabilities is consistently aleatory if the parameter prior is interpreted as a model of a parameter-generating process in the same way. Gelman and Shalizi (2013) stated that the prior distribution may encode "*a priori* knowledge" or a "subjective degree of belief." This seems to mix up an epistemic interpretation of the parameter prior with an aleatory interpretation of the parametric model, and it is hard to justify using them in the same calculus. I believe that it would be better to refer to the parameter prior as an idealistic model of a process that generates parameters for different situations that are based on the same information, i.e., to interpret it in a frequentism-as-model way. This is in line with the fact that Gelman in presentations sometimes informally refers to the parameter prior as a distribution over parameters realistic in a distribution of different situations of similar kind in which datasets can be drawn. This is a very idealistic concept, and it is

probably hard to connect the setup of parameter generation precisely to real observations. The modeling will normally indeed rely more on belief and informal knowledge than on observation of replicates of what is supposed to be parameter generation, but an arbitrary amount of data can be generated from the fully specified model and can be compared with the observed data. Testing the parameter prior is hard. In a standard simple Bayesian setup, it is assumed that only one (possibly multidimensional) parameter value generated all observed data, so the effective sample size for checking the parameter prior is smaller than one, because the single parameter is not even precisely observed. Therefore, sensitivity against prior specification will always be a concern, but see the earlier discussion of single-case probabilities. In any case, the parametric model can be tested in frequentist ways. In case the parameter prior encodes valuable information about the parameter that can most suitably be encoded in this way, Bayesian reasoning based on such a model is clearly useful and open to self-correction by falsificationist compatibility logic. As for model assumptions testing followed by traditional frequentism methods (Shamsudheen and Hennig (2019)), it may be of interest to analyze the behavior of Bayesian reasoning conditionally on model checking in case of fulfilled and not fulfilled model assumptions.

A benefit of such an interpretation could be that the prior no longer either has to be claimed to be objective or to model a specific person. It is a not necessarily unique researcher's suggestion how to imagine the parameter-generating process based on a certain amount of information and can as such be compared with alternatives and potentially rejected, if not by the data, then by open discussion. As earlier, stability against different prior choices compatible with the same information can be investigated. Models may be set up to formalize not only most realistic processes but also, depending on the application, worst or best case scenarios in order to explore what range of possibilities this generates.

Alternatively, the sampling model may be interpreted in a frequentist manner, and the parameter prior may not be of interest in its own right, but rather be used as a tool to achieve inference with good frequentist properties such as confidence coverage or small estimator mean squared error (Bayarri and Berger 2004; Wasserman 2006). In this case, the posterior probabilities cannot be interpreted, as the posterior inherits meaning from the prior, and the prior has not been chosen in order to be meaningful. Often, Bayesian methods have better frequentist properties for certain parameter values and worse for others, so that the prior can also be used to improve matters for certain parameter values that are seen as more relevant, to the price of having a worse result in some other cases, being potentially in line with the idea that prior information suggests that certain parameter values are more likely than others.

Falsificationist Bayes combines the inverse probability logic and compatibility logic by applying the former within a Bayesian model, of which the compatibility with the data should also be investigated. Gelman et al. (1996) emphasize that the posterior distribution of the parameter is conditional on the truth of the model, and according to frequentism-as-model, a researcher can interpret results temporarily *as if* this were the case, without ignoring that this is just a thought construct and that many other models are compatible as well. This differs from the epistemic

probability interpretation, where the "truth of the model" is not a matter of mechanisms in the real world, but rather of the degrees of belief of the analyst, and the possibility that the analyst's degrees of belief are not properly reflected by the model is rarely discussed.

Frequentism-as-model allows to interpret the parameter prior in falsificationist Bayes in a way that requires neither to mix epistemic and frequentist meanings of probability nor to demand that the prior corresponds to an in principle infinitely repeatable data-generating process. The Achilles heel of this is that in a standard situation, based on a single realization that is not even directly observable, the potential to check the parameter prior against the data is very weak. The prior still needs to be convincingly defended in other ways, and sensitivity analysis is certainly desirable.

4 Epistemic Probability-as-Model

As for aleatory probability, it is also possible to use epistemic probability in a way that explicitly takes into account that models are tools for thought rather than referring to any trut and be it a truth concerning a subjective agent.

This chapter had the main focus on frequentism-as-model, and I will only make short remarks on epistemic probability-as-model. Some of the discussion regarding frequentism-as-model also applies to epistemic probability-as-model, such as regarding the role of the exchangeability assumption, which is analogous to the role of i.i.d. in frequentism as discussed earlier. Considerations regarding interpretative equivalence and robustness are also relevant.

4.1 Epistemic Idealization

As the idea of limiting relative frequencies, the axioms of epistemic probability are idealizations, and as such epistemic probability models are, as models, essentially different from reality, which here is the reality of agents reasoning under uncertainty. Adopting such a model amounts to adopting an idealization, and the "as-model" approach taken here states that this should not be forgotten. In fact, idealization is conceded in many places (e.g., Howson and Urbach (2006)), but as in frequentist analyses, results in practice are often interpreted as if idealization has not happened (e.g., probabilities that a certain parameter lies in a suitably chosen set may be portrayed without stating relative to which choices this has been computed, or without much justification of these choices).

De Finetti's (and others') arguments regarding coherent betting require linear utility of money. Leaving this issue aside, the involved betting regarding completely specified distributions is hardly ever done in practice; if this happens at all (e.g., in some financial applications), human agents with specific points of view bet against other human agents, and taking into account knowledge about the opponent or more generally human behavior may influence the offered bets. Several authors (Jaynes 2003;

Howson and Urbach 2006) prefer the derivation of the probability axioms from desiderata for rational reasoning about plausibility under uncertainty as in Cox (1946)'s Theorem to the betting approach, but also this involves idealization (i.e., the assumption that such plausibilities always can be expressed by bounded real numbers; any two events can be compared on a one-dimensional scale, and a number of other not necessarily intuitive requirements; see Van Horn (2003)), and it does not give an operational meaning to the probabilities chosen in any specific situation.

There are two levels at which idealization takes place. The first one is the choice of priors. It has already been mentioned that prior choices in practice will necessarily be simplified by not taking everything into account that could possibly happen. Furthermore, there are different principles for designing informationless priors that in a given situation may lead to different choices and different results (Kass and Wasserman 1996). The second level is the justification of Bayesian (or other) updating as a supposedly objective way to process probability statements given new data, whether or not the priors are taken as objective. However, this second level can hardly be considered independently of the prior choice, as modifying an ill-chosen prior may require a deviation from Bayesian updating and coherence. See Van Horn (2003) for a comprehensive discussion of other possible objections to the desiderata of Cox's Theorem.

4.2 Flexible Handling of Priors

Many pragmatic Bayesians have been well aware for a long time of the problems of all too dogmatically following the axioms of epistemic probability, and many recommendations exist that deviate from these, so that many practical implications of epistemic probability-as-model are not new. For example, Bernardo and Smith (2000) wrote:

> In this work, we shall proceed on the basis of a prescriptive theory which assumes precise quantification, but then pragmatically acknowledges that, in practice, all this should be taken with a large pinch of salt and a great deal of systematic sensitivity analysis.

More suggestions and references can be found in Gelman et al. (2013), particularly in Chap. 6.

Dropping the requirement to find a "true" prior, it is possible to explore a range of priors. For example, Spiegelhalter et al. (1994) propose to look at a "skeptical," an informationless, and an "enthusiastic" prior in randomized trials, in order to see clearly how posterior results are affected by different attitudes expressed in the prior. Sensitivity analysis looking at different priors that are seen as encoding (more or less) the same prior information is sometimes used. It is also possible to choose a prior that intentionally only uses certain prior information (which may be uncontroversial) while not accounting for other information that may be controversial among stakeholders, to allow for better agreement. See Dawid (1982a) on intersubjective choices of epistemic probabilities.

Maybe the biggest problem with the coherence requirement for epistemic probabilities is that it does not allow the user to deviate from the prior choices other than by Bayesian updating, as abovementioned. Pragmatic Bayesians know that it is practically impossible to specify prior probabilities for everything that can possibly happen. In case of a conflict between prior and data (to be discovered using compatibility logic, temporarily), it is often seen as more appropriate to change the model (see, e.g., Gelman et al. 2013, Ch. 6) even if this violates the axioms of epistemic probability. According to the betting logic proposed by Ramsey and de Finetti, this will allow an opponent to choose bets that incur sure loss for the user who chooses to change their model. However of course, the betting setup is itself an idealization, and in much applied data analysis, bets based on epistemic probabilities do not really happen. It may therefore seem harmless to "reject" earlier prior choices (Bernardo and Smith 2000, Ch. 6) back this up by utility considerations (incurring sure loss may be milder than exposing oneself to higher losses due to an inappropriate model choice), but they admit that rejecting a model cannot be fit into the Bayesian framework unless alternative models are already specified. Another reason to change priors is miscalibration, i.e., bad prediction results (Dawid 1982b).

It should not be ignored that if the data analyst is willing to violate the axioms for epistemic probability, these axioms can no longer be used to sell the Bayesian approach as coherent and unified. All Bayesian posterior probabilities are functions of the prior. If the prior is changed conditionally on the data, the interpretation of the posterior is questionable. A way to still justify the resulting posterior probabilities is to argue that a prior may have been specified in advance that involves both the original and the changed model, and given the incompatible data, the posterior probability for the original model should be very small, so that posterior probabilities from the new changed prior are very close to the posterior probabilities had the analyst specified a prior comprising all possibilities from the outset. It could be objected, though, that also even further models could have been considered (which could have been chosen in case the original model would have been rejected for different reasons), and as this was in fact not done, it cannot be known what implications this would have had. This can be taken as a reason to not change the prior based on the data too lightly.

The topic here is the use of epistemic probabilities in statistical data analysis. Philosophers interested in a general formal logic of plausibility will use epistemic probabilities somewhat differently, even if they follow the same Bayesian formalism with reference to Cox's Theorem. Jaynes (2003) motivates his introduction to Bayesian probability theory with equipping a robot with the formal tools to do automated reasoning. A robot cannot be expected to take axioms and their implications "with a large pinch of salt," as a human statistician can, and such an attitude is probably not to most philosophers' tastes either. A robot can be taught handling of epistemic probabilities, but not (at least not currently) epistemic probabilities-as-model, flexibly distancing itself from the absoluteness of a formal calculus and violating (or re-adjusting) it for a variety of good reasons that may not all be obvious in advance.

In Gelman and Hennig (2017), we have argued that it is worthwhile in science to aim at aspects of objectivity such as transparency, consensus, and impartiality.

Following a consistent formal system is surely attractive in these respects, but there is always an irreducible tension between these aims and the context-dependence of reasoning and the recognition of multiple perspectives. We should be prepared to question and at least temporarily suspend any consistent formal system if required, acknowledging it as a model, an idealization with limitations. A major issue with artificial intelligence (AI) (like Jaynes's robot) is that AI has to rely on formal models and cannot question them, or only in the framework of a bigger formal model.

4.3 How Different Are Epistemic Probability-as-Model and Frequentism-as-Model?

In frequentism-as-model, an agent adopts a probability model temporarily, acknowledging its essential difference from reality, which means that she considers the modeled experiment as if it has a tendency to give rise to the corresponding relative frequencies under repetition. In epistemic probability-as-model, an agent adopts a probability model temporarily, acknowledging its essential difference from reality, which means that she considers the uncertainties in the modeled experiment as if, where she bets on future observations, she would do this according to the model.

Involving exchangeability (which of course is questionable) in the epistemic model actually implies that relative frequencies are expected to converge, and using a frequentist model could then be justified just as well, although a Bayesian may want to keep the prior, at least if it encodes relevant information that would otherwise be lost. It has also been recommended in various places (Gelman et al. 2013, Ch. 1) to think about a situation as repeatable and to consider expected relative frequencies when setting up Bayesian priors. In fact, Bayesians often think of the sampling model part of the prior as modeling an aleatory mechanism (possibly idealized, in an "as-model" manner), and often posterior probabilities regarding a parameter are interpreted as if this were about a "true" parameter value defined within an aleatory sampling model.

Taking also into account that epistemic and aleatory probabilities were historically not distinguished, frequentism-as-model and epistemic probability-as-model are indeed strongly related. It will however contribute to the clarity of communication to specify, when doing analyses, explicitly what the involved probabilities are supposed to mean. Does the modeler model a data generating mechanism in the world, and do probabilities refer to the outcomes of it? Or do probabilities refer to the agent's uncertainty? I have elaborated that this is not an issue of belief in the existence of objective aleatory probabilities, which is not more or less problematic than belief in a general epistemic calculus formalizing rational handling of uncertainty and in specific prior choices, even if the latter only refer to the agent herself. Modeling does not require belief in the truth of a model; a model has to be justified from its use.

Aleatory models can in principle, by bold idealization, even be applied to situations that are essentially non-repeatable and may not even allow for repeatable residual or innovation terms and the like. However, anchoring such applications in

observations and defending them convincingly may be hard if not impossible, and a modeler has an easier task for sure justifying the use of epistemic probabilities.

On the other hand, many sequences of experiments are treated in probability modeling as if they behave like aleatory processes, and results are interpreted in this way, even if used as sampling model in a Bayesian prior. It seems to be more direct to me in such cases to talk about models for data-generating processes in reality, rather than about models for how an agent thinks about data-generating processes in reality, and therefore I often prefer the aleatory interpretation, not ruling out using Bayesian calculus.

5 Conclusion

It may seem to be my core message that models are models and as such different from reality. This is of course commonplace, and agreed by many if not all statisticians, although it rarely influences applied statistical analyses or even discussions about the foundations of statistics.

Here are some less obvious implications:

- The usual way of talking about model assumptions, namely that they "have to be fulfilled," is misleading. The aim of model assumption checking is not to make sure that they are fulfilled, but rather to rule out issues that misguide the interpretation of the results. Combining model assumption checking and analyses chosen conditionally on the model checking results can itself be modeled and analyzed, and depending on what exactly is done, it may or may not turn out to work well.
- There are always lots of models compatible with the data. Some of these can be favored or excluded by plausibility considerations, prior information, or the data, but some irregular ones have to be excluded simply because inference would be hopeless if they were correct.
- Model assessment based on the data involves decisions about "in what way to look," i.e., what model deviations are relevant. Whether a model is compatible with the data cannot be decided independently of such considerations.
- Choosing a model implies decisions to ignore certain aspects of reality, e.g., differences between the conditions under which observations modeled as i.i.d. were gathered. These decisions should be transparent and open for discussion.
- Consideration of interpretative equivalence, i.e., what different models would be interpreted in the same or different way regarding the subject matter, is important in order to investigate robustness and stability, i.e., to what extent different models compatible with the data would lead to results on the same data that have a different meaning.
- Frequentism-as-model is compatible with both compatibility and inverse probability logic. A key to decide which one to prefer is to ask whether the parameter prior distribution required for inverse probability logic can be used to add

valuable information in a convincing way. The parameter prior itself cannot normally be checked against the data with satisfactory power. Compatibility logic can also be used for epistemic-probability-as-model; a model may be dropped that turns out to be incompatible with the observations, even though this violates coherence.

Acknowledgments I thank Laurie Davies, Andrew Gelman, Donald Gillies, Sander Greenland, David Hand, Johannes Hüsing, and Keith O'Rourke for helpful comments on an earlier version of this chapter.

References

Barnard GA, Cox DR (eds) (1962) The foundations of statistical inference: a discussion. Methuen, London
Bates S, Hastie T, Tibshirani R (2021) Cross-validation: what does it estimate and how well does it do it? https://arxiv.org/abs/2104.00673
Bayarri MJ, Berger JO (2004) The interplay of Bayesian and frequentist analysis. Stat Sci 19:58–80. https://doi.org/10.1214/088342304000000116
Bayes T (1763) An essay toward solving a problem in the doctrine of chances. Philos Trans R Soc Lond 53:370–418
Berger JO (2006) The case for objective Bayesian analysis. Bayesian Anal 1:385–402
Bernardo JM, Smith AFM (2000) Bayesian theory. Wiley, Chichester
Bernoulli J (1713) Ars Conjectandi. Thurnisiorum Fratum, Basel
Box GEP (1979) Robustness in the strategy of scientific model building. In: Launer RL, Wilkinson GN (eds) Robustness in statistics. Academic Press, Cambridge, MA, pp 201–236
Box GEP (1980) Sampling and Bayes' inference in scientific modelling and robustness. J R Stat Soc Ser A General 143:383–430
Breiman L (2001) Statistical modeling: the two cultures. Stat Sci 16:199–215
Bridgman PW (1927) The logic of modern physics. Macmillan, New York
Burnham KP, Anderson DR (2002) Model selection and multimodel inference. Springer, New York
Carnap R (1950) Logical foundations of probability. Chicago University Press, Chicago
Chang H (2012) Is water H_2O? Evidence, realism and pluralism. Springer, Dordrecht
Chang H (2021) Operationalism. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Fall 2021 edn. Metaphysics Research Lab, Stanford University
Chang H (2022) Realism for realistic people: a new pragmatist philosophy of science. Cambridge University Press, Cambridge
Cosmides L, Tooby J (1996) Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition 58:1–73
Cox RT (1946) Probability, frequency, and reasonable expectation. Am J Phys 14:1–13
Cox DR (1990) Role of models in statistical analysis. Stat Sci 5:169–174
Creath R (2021) Logical empiricism. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Winter 2021 edn. Metaphysics Research Lab, Stanford University
Cressie N (1980) Relaxing assumptions in the one-sample t-test. Aust J Stat 22:143–153
Davies PL (1995) Data features. Statistica Neerlandica 49:185–245
Davies PL (2014) Data analysis and approximate models. Chapman and Hall/CRC, New York
Davies L, Dümbgen L (2022) Covariate selection based on a model-free approach to linear regression with exact probabilities. https://arxiv.org/abs/2202.01553
Dawid AP (1982a) Intersubjective statistical models. In: Koch G, Spizichino F (eds) Exchangeability in probability and statistics. North-Holland, Amsterdam, pp 217–232
Dawid AP (1982b) The well-calibrated Bayesian (with discussion). J Am Stat Assoc 77:605–613

de Finetti B (1937) Foresight: its logical laws, its subjective sources (English translation). In: Kyburg HE, Smokler HE (eds) Studies in subjective probability (1964). Krieger, Huntington
de Finetti B (1974) Theory of probability, vol 1. Wiley, New York
Diaconis P, Skyrms B (2018) Ten great ideas about chance. Princeton University Press, Princeton, NJ
Eagle A (2019) Chance versus randomness. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Spring 2019 edn. Metaphysics Research Lab, Stanford University
Edwards AWF (1972) Likelihood. Cambridge University Press, Cambridge
Fay MP, Proschan MA (2010) Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. Stat Surv 4:1–39
Fidler F, Wilcox J (2018) Reproducibility of scientific results. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Winter 2018 edn. Metaphysics Research Lab, Stanford University
Fine TL (1973) Theories of probability. Academic Press, New York
Fisher RA (1922) On the mathematical foundation of theoretical statistics. Philos Trans R Soc Lond A 22:309–368
Fisher RA (1936) Uncertain inference. Proc Am Acad Arts Sci 71:248–258
Frigg R, Nguyen J (2020) Modelling nature: an opinionated introduction to scientific representation. Springer, Cham
Galavotti MC (2005) Philosophical introduction to probability. CSLI Publications, Stanford, CA
Gelman A, Hennig C (2017) Beyond objective and subjective in statistics (with discussion). J R Stat Soc Ser A 180:967–1033
Gelman A, Shalizi C (2013) Philosophy and the practice of Bayesian statistics (with discussion). Br J Math Stat Psychol 66:8–80
Gelman A, Meng XL, Stern H (1996) Posterior predictive assessment of model fitness via realized discrepancies (with discussion). Stat Sin 6(4):733–808
Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian Data Analysis, 3rd edn. Chapman and Hall/CRC, New York
Gergen KJ (1999) An invitation to social construction. Sage, Thousand Oaks, CA
Gillies D (2000) Philosophical theories of probability. Routledge, London
Greenland S, Rafi Z (2019) To aid scientific inference, emphasize unconditional descriptions of statistics. https://arxiv.org/abs/1909.08583
Hacking I (1965) Logic of statistical inference. Cambridge University Press, Cambridge, NY
Hacking I (1975) A philosophical study of early ideas about probability, induction and statistical inference. Cambridge University Press, Cambridge, NY
Hajek A (2009) Fifteen arguments against hypothetical frequentism. Erkenntnis 70(2):211–235. http://www.jstor.org/stable/40267419
Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986) Robust statistics: the approach based on influence functions. Wiley, New York
Hand DJ (1994) Deconstructing statistical questions. J R Stat Soc Ser A Stat Soc 157:317–356
Hand DJ (2006) Classifier technology and the illusion of progress. Stat Sci 21:1–14
Hannig J, Iyer H, Lai RCS, Lee TCM (2016) Generalized fiducial inference: a review and new results. J Am Stat Assoc 111:1346–1361
Hennig C (2007) Falsification of propensity models by statistical tests and the goodness-of-fit paradox. Philos Math 15(2):166–192
Hennig C (2010) Mathematical models and reality: a constructivist perspective. Found Sci 15: 29–48
Hennig C, Lin CJ (2015) Flexible parametric bootstrap for testing homogeneity against clustering and assessing the number of clusters. Stat Comput 25:821–833
Hennig C, Meila M, Murtagh F, Rocci R (eds) (2015) Handbook of cluster analysis. Chapman & Hall/CRC, Boca Raton, FL
Howson C, Urbach P (2006) Scientific reasoning: the Bayesian approach, 3rd edn. Open Court, Chicago
Huber PJ, Ronchetti EM (2009) Robust statistics, 2nd edn. Wiley, Hoboken, NJ

Jaynes ET (2003) Probability theory: the logic of science. Cambridge University Press, Cambridge
Kahneman D, Slovic P, Tversky A (1982) Judgment under uncertainty: heuristics and biases. Cambridge University Press, Cambridge
Kass RE, Wasserman L (1996) The selection of prior distributions by formal rules. J Am Stat Soc 91:1343–1370
Keynes JM (1921) A treatise on probability. Macmillan, London
Kolmogorov A (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin/Heidelberg
Kulinskaya E, Morgenthaler S, Staudte RG (2008) Meta analysis. Wiley, New York
Lakshminarayanan B, Pritzel A, Blundell C (2017) Simple and scalable predictive uncertainty estimation using deep ensembles. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (eds) Advances in neural information processing systems, vol 30. Curran Associates
Lancet (2005) The end of homoeopathy (editorial). Lancet 366:690
Laplace PS (1814) Essai philosophique sur les probabilités. Courcier, Paris
Lehmann EL (1986) Testing statistical hypotheses, 2nd edn. Wiley, New York
Martin-Löf P (1966) The definition of random sequences. Inf Control 9:602–619
Mayo DG (2018) Statistical inference as severe testing: how to get beyond the statistics wars. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781107286184
Mayo DG (2022) The statistics wars and intellectual conflicts of interest. Conserv Biol 36(1):e13861
Popper KR (1959a) The logic of scientific discovery. Routledge, London
Popper KR (1959b) The propensity interpretation of probability. Br J Philos Sci 10:25–42
Porter TM (1996) Trust in numbers: the pursuit of objectivity in science and public life. Princeton University Press, Princeton NJ
Ramsey FP (1926) Truth and probability. In: Braithwaite R (ed) Foundations of mathematics and other logical essays (1931). Kegan, Paul, Trench, Trubner, London
Rice K, Bonnett T, Krakauer C (2020) Knowing the signs: a direct and generalizable motivation of two-sided tests. J R Stat Soc A Stat Soc 183(2):411–430
Rudin C, Chen C, Chen Z, Huang H, Semenova L, Zhong C (2022) Interpretable machine learning: fundamental principles and 10 grand challenges. Stat Surv 16:1–85
Schölkopf B, Smola AJ (2002) Learning with kernels. MIT Press, Cambridge, MA
Senn S (2011) You may believe you are a bayesian but you are probably wrong. RMM 2(42):48–66
Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton, NJ
Shafer G, Vovk V (2019) Game-theoretic foundations for probability and finance. Wiley, Hoboken, NJ
Shamsudheen MI, Hennig C (2019) Should we test the model assumptions before running a model-based test? https://arxiv.org/abs/1908.02218. Accepted for publication in Journal of Data Science, Statistics, and Visualisation
Shang A, Huwiler-Müntener K, Nartey L, Jüni P, Dörig S, Sterne JAC, Pewsner D, Egger M (2005) Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. Lancet 366:726–732
Spiegelhalter DJ, Freedman LS, Parmar MKB (1994) Bayesian approaches to randomized trials. J R Stat Soc A Stat Soc 157(3):357–387
Steinley D, Brusco MJ (2011) Choosing the number of clusters in k-means clustering. Psychol Models 16:285–297
Tafreshi D, Slaney KL, Neufeld SD (2016) Quantification in psychology: critical analysis of an unreflective practice. J Theor Philos Psychol 36:233–249
Tukey JW (1993) Issues relevant to an honest account of data-based inference, partially in the light of Laurie Davies' paper
Tukey JW (1997) More honest foundations for data analysis. J Stat Plan Inference 57:21–28
Van Fraassen B (1980) The scientific image. Clarendon Pressn, Oxford
Van Horn KS (2003) Constructing a logic of plausible inference: a guide to cox's theorem. Int J Approx Reason 34(1):3–24

Venn J (1866) The logic of chance. Macmillan, London
Vermunt JK (2011) K-means may perform as well as mixture model clustering but may also be much worse: comment on Steinley and Brusco (2011). Psychol Models 16:82–88
von Glasersfeld E (1995) Radical constructivism: a way of knowing and learning. The Falmer Press, London
von Mises R (1939) Probability, statistics, and truth, 2nd edn. Macmillan, London
Walley P (1991) Statistical reasoning with imprecise probabilities. Chapman and Hall, London
Wasserman L (2006) Frequentist Bayes is objective (comment on articles by Berger and by Goldstein). Bayesian Anal 1:451–456
Wasserstein RL, Lazar NA (2016) The ASA statement on p-values: context, process, and purpose. Am Stat 70(2):129–133
Wasserstein RL, Schirm AL, Lazar NA (2019) Moving to a world beyond $p < 0:05$. Am Stat 73: 1–19
Williamson J (2010) In defence of objective Bayesianism. Oxford University Press, Oxford
Xie M, Singh K (2013) Confidence distribution, the frequentist distribution estimator of a parameter: a review. Int Stat Rev 81(1):3–39

Section V

Constructing the Reals: Paradoxes and Foundations

Paul M. Livingston and Bharath Sriraman

Introduction to the History and Philosophy of Mathematical Practice in Constructing the Reals

Paul M. Livingston

Contents

1 Introduction 1462
2 Summary 1471
3 Cross-References 1472
References 1472

Abstract

The ancient problem of the relationship of the continuous to the discrete, since its discovery by the Greeks, has posed a range of immensely fruitful challenges to both philosophical and mathematical thought, leading to a variety of mathematical and conceptual innovations whose positive development actively continues today. In this brief section introduction, I selectively outline some significant moments at which this problem has provided important historical occasions for concrete mathematical innovation as well as closely linked philosophical insights, before introducing the particular contributions of the section authors to historical and contemporary research and practice into the nature and structure of the continuum, the construction of real numbers, the foundations of analysis, and the logical, set-theoretical, and computational underpinnings of the structure of continuously valued numbers in the most general sense, in both constructive and nonconstructivist contexts.

Keywords

Continuum · Discrete · Real numbers · Incommensurability · Intuitionism · Constructivism · Plato · Leibniz · Weyl · Russell · Poincare · Dedekind

P. M. Livingston (✉)
Philosophy Department, University of New Mexico, Albuquerque, NM, USA
e-mail: pmliving@unm.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_129

1 Introduction

The ancient problem of the relationship of the continuous to the discrete, since its discovery by the Greeks, has posed a range of immensely fruitful challenges to both philosophical and mathematical thought, leading to a variety of mathematical and conceptual innovations whose positive development actively continues today. In this brief section introduction, I selectively outline some significant moments at which this problem has provided important historical occasions for concrete mathematical innovation as well as closely linked philosophical insights, before introducing the particular contributions of the section authors to historical and contemporary research and practice into the nature and structure of the continuum, the construction of real numbers, the foundations of analysis, and the logical, set-theoretical, and computational underpinnings of the structure of continuously valued numbers in the most general sense, in both constructive and nonconstructivist contexts.

Ancient sources credit the discovery of what we would today call the "irrationality" of the square roots of some integers to a Pythagorean of the early fifth century BC, perhaps Hippasus of Metapontum, and also relate the (probably apocryphal) story of his having been drowned at sea as punishment for making public secret Pythagorean doctrines. From the beginning, what dominates the attention of the Greeks in this discovery is not simply the numerical fact of the nonexistence of any ratio of integers a/b corresponding to the irrational quantities, but (rather more basically) the geometric *incommensurabilty* that is exhibited by (for example) the diagonal of the unit square in relation to its sides (Heath 1921). That the diagonal is incommensurable with the side in this sense means that there can be no common unit length l, no matter how small, that can "measure" both in the sense that both lengths can be expressed as integer multiples of it. There is some evidence (von Fritz 1945) that the first discovery of the mutual incommensurability of lengths was not that of the incommensurability of the diagonal of the square, but rather that of the incommensurability of the pentagon with its side, for which there is a directly geometrical demonstration. However this may be, for the case of the diagonal of the unit square (which was independently known to have the length root $\sqrt{2}$, either by the Pythagorean theorem or by the independent demonstration, given by Plato in the *Meno*, of the possibility of constructing a square of area 2 on the diagonal of the unit square), a primarily numerical proof by contradiction is recorded in an appendix to book 10 of Euclid's *Elements* (although the proof is probably a later addition, not included by Euclid himself, and is omitted from Heath's translation) (Euclid 1908).

Several historians of Greek mathematics, especially in the early twentieth century, have seen this discovery of the incommensurable and hence of the impossibility of understanding all magnitudes in terms of ratios of whole numbers as representing a fundamental crisis in the history of Greek mathematics and as thereby leading to increasingly radical attempts to remedy it (see Hesse and Scholz 1928; Becker 1933; Saito 2004). Whether or not this is the case, the rapid efforts of Greek mathematicians to classify the irrational magnitudes evidently figure prominently in Plato's own thinking about the nature of knowledge and (in particular) the possibility of clarifying the reality of infinite continuous variation through the discovery of

discrete *forms* or *eide* associated with specific discursive definitions. Here, Plato appears to make reference to the project of determining types of quantities capable of classifying this reality of infinite variation, most evidently shown in the infinite multiplicity of types of number or quantity, into discrete identities or groups. The early development of these methods of classification is due primarily to the mathematicians Theaetetus and Theodorus, both associates of Socrates and Plato and connected to his school, and book ten of Euclid's *Elements,* which is devoted to incommensurable and commensurable magnitudes, appears largely to summarize their results (the connection is attested by Pappus: see Sayre 1983, p. 108). Near the beginning of Plato's dialogue *Theaetetus* (147d, ff.) which is primarily concerned with the definition and nature of knowledge, a young Theaetetus is presented as in the middle of classifying the initial integers, in part by demonstrating the irrationality of the roots of the nonsquares from 3 through 17. Just as the integers can be classified into different groups according to whether they are "square" (perfect squares) or "oblongs," so the infinite range of variation of a concept such as that of knowledge might be tamed, Socrates suggests, by means of a specification of the *eide* or types among which this infinity is distributed under a single overarching concept (148d).

The suggestion that one might in this way "tame the infinite" through the specification of determinate limits by using the best *mathematical* techniques available at the time also appears to play a role in the later Plato's conception of how rational thought can comprehend the discrete ideal forms of quantities and magnitudes that are naturally in continuous variation. A concern with the problem of continuous variation appears in many different places in Plato's corpus, perhaps most centrally in the late dialogue *Philebus* in connection with the project of making a "division" of whatever exists into the two categories of the "limit" and "unlimited" (16c–17a; 24c–25a). Here, the suggested "method" of rational inquiry is that of looking, in each case of continuous variation (for example, between the hotter and the colder or between the higher and lower pitch), for the unity of forms which operate as limits and assign discrete measure and proportion between the extremes.

Noting the evident connection that this bears to Theaetetus' problem of the classification of the irrational (i.e., incommensurable) magnitudes, Sayre (1983) suggests that Plato's description of the "method" of determining the ideal forms of these magnitudes may draw on a method for approximating irrationals to an arbitrary degree of precision due to Eudoxus of Cnidus, himself a student and associate of Plato's, which several modern writers have compared to the central innovation of the modern project of "constructing" the reals, Dedekind's so-called "cut." The method is implicit in Euclid's definition 5 in book 5 of the *Elements*, which specifies that two magnitudes a/b and c/d (whether rational or irrational) are in the same ratio if and only if, for any m and n, either (i) $ma > nb$ and $mc > nd$; (ii) $ma = nb$ and $mc = nd$, or (iii) $ma < nb$ and $mc < nd$. Given this, the value of a magnitude incommensurable to the unit length can be approximated in rational terms to an arbitrary degree of accuracy. For example, in a geometric illustration that captures the method as it would have been known to Eudoxus and Plato, the magnitude $\sqrt{2}$ can be approximated by imagining two series of points placed at intervals next to each other. The first series are placed at intervals of the unit length, while the second are placed at

intervals of (the length) $\sqrt{2}$. Then it will be the case that the first point of the second column will fall between the first and second points of the first column: the 10th between the 14th and the 15th, the 100th between the 141st and 142nd, and so on. In this way, the decimal expansion of $\sqrt{2}$ is approximated as falling between two ordered sets of quantities expressed in terms of unit measures, essentially dividing the rationals into those above and those below the relevant irrational magnitude.

With respect to devices such as these, it is clear that the mathematics available to Plato at the time of his writing had already progressed very significantly in the direction of what much later would be understood as the "construction" of the real continuum and even the substantive determination of its defining properties (Eudoxus is also responsible, for instance, for the introduction of the "Archemidean" property of the continuum (*Elements* book V, def. 4) whereby for any two magnitudes m, n, there is an integer p such that $pm > n$, according to Archimedes' own testimony). Moreover, it is clear that what underlies the actual philosophical fruitfulness of these constructions, for example, in the case of Plato's own philosophy of forms and their participants, is the way that they offer to address the underlying problem of incommensurability which threatens to cleave the discrete from the continuous, thereby express the most central kinds of access of rational, discursive, and conceptual thought to what Plato envisioned as its taming of a world of infinite continuous variation. In this light, the philosophical relevance of the mathematical techniques for comprehending and constructing the continuous in terms of the discrete do not come so much from their overcoming or solution of this problem as from the ways in which they offer the problem itself to be reinscribed in more logically and philosophically revealing terms. Much the same is the case, in fact, for much of the later history of the development of the differential and integral calculus, the set-theoretical foundation of the real numbers suggested by Weierstrass, Cantor, and Dedekind, and even the foundational work that is still being done today.

Another opportunity for reflection on the way that the problem of the relationship between the discrete and the continuous has simultaneously led to significant innovations of both a philosophical and a mathematical character is afforded by reflection on the fundamental contributions of G. W. F. Leibniz to both domains. By 1686, Leibniz had arrived at what are essentially the fundamental methods of both the integral and differential calculi, appealing (without using the term) to what are essentially infinitesimal quantities including infinitely small but nonzero distances, and conceiving of curves as essentially infilateral polygons (Bell 2022). His work on the calculus together with his reflections on what he saw as the impossibility of understanding the continuous line as an aggregate of points appears decisive in producing his overarching metaphysical picture of *monadism*, according to which all real unities are simple substances amounting to "formal atoms" (Leibniz 1695). This metaphysics of the composition of all real entities, including matter, as aggregates of monads led Leibniz to his conviction of the ideality of space and time: Since (Leibniz supposed) all real unities were either pointlike monads or aggregates thereof, and since space and time, being continua, could not be understood as such aggregates, they must not be realities at all. However, Leibniz holds that bodies are not only divisible but actually divided into real parts, ad infinitum (Leibniz 1686);

more generally the idea of the actual composition of perceived phenomena by parts that are separate to infinity provides support for what has been called his "principle of continuity," which holds (in various different formulations) that there are nowhere any discontinuities or "gaps" in natural phenomena.

Although Leibniz thus denies the ultimate reality of the continuum along with extension in time and space themselves, these claims also lead him positively to understand the distinction between the necessity and contingency of propositions as a result of the mutual *incommensurability* of continuous magnitudes and the specific kind of irreducibility this may be thought to imply in the logical analysis of subject/predicate propositions. Leibniz saw the truth of any such proposition as following from the "containment" of the predicate term within the subject, whether or not we are in a position to perceive the particular containment by means of any finite analysis. In particular, Leibniz argues in a 1689 fragment (Leibniz 1689) that what we understand as "necessary" truths may be seen as those in which the containment of the predicate in the subject concept may be seen as following from basic truths, or statements of identity, by means of at most *finitely* many deductive steps. For contingent truths, by contrast, the containment of the predicate in the subject cannot be shown in any finite number of steps, although Leibniz suggests it may nevertheless be known to God, who can see the *whole* of the infinite series of steps which could alone resolve it. In this respect, Leibniz argues, the analysis (whether finite or infinite) of containment is analogous to determining the ratios of rational as well as irrational numbers. Just as in every truth the predicate is contained in the subject, so in every such ratio a smaller number is "in" a larger one (or an "equal in an equal"). But whereas in certain ratios (i.e., those corresponding to rational numbers) one can analyze the larger without residue into the smaller, or into a common measure of both, in ratios such as that of the side to the diagonal of the square, the incommensurability of the ratio ensures that there can be no common finite measure and the analysis is hence continued "into infinity" (Leibniz 1689, p. 97). Thus, in the case of contingent propositions whose truth we can see only by means of sensory perception, it is impossible for our reasoning to reduce these to identities by a proof of any (finite) number of steps or to see the necessity of the containment of the predicate in the subject. Nevertheless, just as we do nevertheless have some knowledge through geometry of incommensurable magnitudes and even have proofs (Leibniz is, doubtless, thinking of his the methods of his own infinitesimal calculus) about infinite series, so we may suppose that God has (to an even greater extent) knowledge of the necessary truths through an infallible kind of intellectual intuition of the whole that here precedes the infinite series of parts.

For both Plato and Leibniz, the fact of actual incommensurability and the inherent involvement of the problematic of the infinite to which it leads played an essential role in articulating both the possibility and the inherent limits of any project of conceptual or metaphysical discursive analysis of the phenomena of motion, change, and continuous variation that appear to be given to us primarily rather in sensory, intuitive, or temporal forms. Here, the problem for the logic of the analysis of concepts and propositions that is posed by the continuous is not so much that of applying a pregiven logical framework to its structure as that of comprehending how

conceptual terms can grasp or even approximate its structure at all, given the demands that the phenomenon of incommensurability already itself places on the projects of exact measurement, demarcation, and analysis. In these projects as Plato and Leibniz understood them, there is thus nothing like a full-fledged "construction" of the continuum or of the reals in general yet in view. But with the range of developments in the eighteenth, nineteenth, and twentieth centuries in the foundations of analysis which accomplished the "arithmetization" of the reals, the culmination of which was the foundation of the concepts of the calculus by means of the Cauchy-Weierstrass conception of the limit, and the set-theoretical construction of the reals in general by means of Dedekind's "cut," the logical and conceptual problems of foundation surrounding the grasp of linguistic or discursive thought on the structure of the continuum are not only articulated in a new way, but also in a certain sense intensified and displaced onto a new and more radical logical ground.

One underlying form that these problems do take in the altered context provided by these nineteenth- and early twentieth-century developments is evident in Hermann Weyl's 1918 *The Continuum,* which is simultaneously a devastating conceptual critique of the contemporary set-theoretical "foundations" of real analysis and an attempt to build up these foundations positively on a different and more limited ground, more closely related to the basic form in which the continuous is plausibly given in our *intuitive* experience of space and time. The central issue underlying Weyl's critique is that of the apparent *impredicativity* of Dedekind's attempt at set-theoretical definition of real numbers and of the real number continuum. The terminology and issue of predicativity traces originally to Russell, who used "non-predicative" to describe any propositional function which fails to non-circularly determine a class (Feferman 2005). Such a failure generally may occur when the attempted definition of a class fails to exclude a "vicious circle," for example, by including in the purported definition the notion of definability itself. More generally, Russell held that the definition of a class or totality is nonpredicative whenever the class or totality so defined would have to include members that are definable only on the basis of that totality itself. Russell saw impredicativity in this sense as the underlying origin of a number of set-theoretical and semantic paradoxes, including his own (and Zermelo's) paradox of the set of all sets not elements of themselves, and therefore sought to prohibit impredicativity in general by imposing a system of (simple or, later, "ramified") types ensuring that sets and propositional functions can only be formed from those of lower type (Russell 1908).

As Poincaré had already pointed out in 1906, another paradox plausibly having the same form and origin is Richard's paradox concerning the definition of real numbers generally (Poincaré 1906; cited in Feferman 2005, pp. 1–2). Briefly, since it follows from Cantor's proof of the uncountability of the reals that there are more real numbers than there are possible definitions, it is possible to define (by diagonalization) a real number different from any that is definable. The basis of Weyl's critique of the by-then standard definition of the reals by means of Dedekind's cut is, then, that (as Poincaré in fact also points out) this definition renders familiar properties of the real numbers impredicative. For example, Dedekind's device of defining the real numbers as (generally infinite) sets of

rationals renders it impossible to appeal to the greatest lower bound (g.l.b.) and least upper bound properties of bounded sets of real numbers (i.e., the properties that every bounded subset S of **R** has a greatest lower bound and a least upper bound) in general without essential impredicativity (Feferman 2005). For the identification of the g.l.b. and l.u.b. for any such set is possible, given Dedekind's technique, only by presupposing the *completed* totality of Dedekind cuts itself, the vast majority of which are not definable by any definition of finite length.

For Weyl in 1918, this implies that the Dedekind (or more broadly, any simple set-theoretical) device for defining the reals in general is hopelessly flawed from a logical standpoint: in particular, that it assumes without reason that the concept of a "real number" is (what Weyl calls) an "extensionally determinate" one. Such a concept, in Weyl's terminology, is one that determines a set of *existent* objects falling under it and also determines unequivocally, for any given property, *whether* there is or is not any object falling under the concept that has that property (Weyl 1919, pp. 109–110). This kind of determinacy fails, Weyl argues, for the real numbers exactly in the case of determining (for example) the l.u.b. of a set of reals. If a real number is understood as a definite set of rationals, and thus as corresponding to a single "definite property" of rationals, then a set of real numbers corresponds to a property A *of properties* of rational numbers. The least upper bound of a set of reals is itself a real number, and thus (on Dedekind's definition) a set of rational numbers x which possess a certain property P_A, that is, just the property *that there is* a property (of properties) of the kind A belonging to x. But this kind of expression, which requires of the definition of the property P_A that it be formulated in a way that ranges over the *totality* of properties of (i.e., of sets of) rational numbers, is (Weyl says) "*clearly meaningless*" (Weyl 1919, p. 111). For, as the expression "property of rational numbers" is itself not extensionally determinate, we are in no position to use phrases definitively which require us to reason about what exists or does not within the (supposed) totality of all such properties. Alternatively, though, if we do restrict the relevant kind of property to ones falling under an extensionally determinate concept, then P_A will itself be one which is, as a matter of its *sense,* almost certainly outside the range of such extensionally constrained properties. It follows from this, Weyl argues, that if the continuum and its most characteristic properties are to be founded at all, its foundations must be a different kind of predicative one, in particular one that makes no essential appeal to the (fictitious) totality of *all* properties of rational numbers. Such an alternative can be found, Weyl argues, only if one takes the continuum to have its own, distinct formulation in the *intuitive givenness* of continuous phenomena and quantities, and most basically in the phenomenon of time (for more on the positive form that this development takes in *The Continuum* and in Weyl's later work, see John L. Bell's article in this section).

After *The Continuum*, Weyl was increasingly led by these suggestions toward the philosophical conceptions of Husserl and Brouwer, according to which the "givenness" of the continuum to any project of conceptual analysis or foundation must depend on the specific and temporal form of that very givenness in living intuition prior to, and as an essential ongoing constraint on, its conceptual or logical analysis. That this conception of the actual foundation of the continuum in intuition

does not preclude the significant and ongoing development of the implications of *logic* for its (properly mathematical) structure is shown sufficiently by the developments of intutionist logics that modify the principles of classical logic such as those of excluded middle and double negation elimination, and by the vast range of the implications of these which continue to follow on them today. I will not go into these implications in this introduction, though several of the individual section contributions (discussed below) do so. But it is nevertheless worth noting how the problems of the relationship of what is plausibly the structure of conceptual or discursive thought and the logical dictate of the avoidance of vicious paradox there which yields the issue of predicativity and Weyl's critique of the set-theoretical foundations of the real, on these grounds, appear to manifest once more (as in the cases of Plato and Leibniz as well) the deep and practically insoluble implications of the basic and underlying problem of the irreducibility of the scarcely delimitable variability of the continuous as such to the determinacy and discreteness of conceptual or predicative thought itself.

The various historical and topical contributions to the section further develop the implications of these problems of the relationship of the structure of the continuum and its relationship to the discrete, including logical, constructive, abstractive, and computational aspects, for important historical moments in the history of their development as well as cutting-edge developments today.

First, John L. Bell's comprehensive "The Continuum and the Evolution of the Concept of Real Number" offers a wide-ranging survey of the historical development and contemporary specifications of the concepts of real number and continuum. Beginning with ancient mathematics and the Greek discovery of the incommensurable, Bell treats the first introduction of the general concept of a continuous arithmetic (by Stevin in 1585) and then develops the controversial issues historically surrounding the nature and existence of infinitesimals in the nineteenth century. Following this, Bell describes the motivations and achievement of the project of Weierstrass, Cantor, and Dedekind in "arithmetizing" the continuum by analyzing it in terms of the discrete, and the subsequent set-theoretical identification of real numbers as sets of rationals or sets of points, closely accompanied by the extirpation of infinitesimals that was achieved by means of the Cauchy-Weierstrass redefinition of the basic concepts of the calculus in terms of limits. In the third section of his article, Bell discusses the positions of several nineteenth- and twentieth-century dissidents to the emerging consensus around this set-theoretical analysis of the continuum in discrete terms. For these mathematicians and philosophers (including du Bois-Reymond, Poincaré, Brouwer, Weyl, Brentano, and Peirce), the underlying mysteries of the continuum or the ineliminability of its actual givenness in intuition as continuously or temporally real verified that there could be, in principle, no secure foundation of the continuous in the discrete and led to various attempts (drawing on mathematical, philosophical, and phenomenological sources) to base the reality of the continuous (and, in some cases, the infinitesimal) on alternative grounds.

For Brouwer in particular, the continuous was always a matter of the intuitive givenness of continuous magnitudes actually given in a temporally real intuition, and

all mathematical activity involves essentially such "construction" in intuition. As Bell documents in section 4 of his paper, a development of Brouwer's suggestion outside the specific context of the psychology or phenomenology of intuitive givenness – but preserving the importantly "constructivist" idea that a mathematical problem is to be regarded as solved only if an explicit solution can be produced – leads to the influential project of "constructive" mathematics and, with it, to a completely different way of logically founding the continuum. Here, in particular, certain typical regularities of classical logic are abandoned in favor of what has been called "intuitionistic" logic: In particular, existential statements require the production of an instance, and disjunctions can only be affirmed only when one or the other of the disjuncts (or both) is explicitly demonstrated. For this reason, both the classical law of excluded middle and the classical procedure of "strong" *reductio* (relying on double negation elimination to establish positive existential claims) are not assumed to hold in the intuitionist context. This modification has important consequences, some negative and others positive, for the construction of the reals themselves and the conditions of their identification. On the constructivist conception, real numbers themselves are then not arbitrary sets of rationals, but rather definite sequences determined either by an effective rule or indeed by successive "free choices" conceived as, in principle, able to be made by an actual mathematical consciousness existing in time, or an idealization thereof.

In the last two sections of his contribution, Bell discusses some of the radical consequences and new developments that have followed the reintroduction, primarily by Abraham Robinson, of a mathematics of infinitesimal quantities and the attendant creation of the new project of Nonstandard Analysis. On one development of the project, the standard classical real line **R** is supplemented by a new and additional structure called a *nonstandard universe*: Defining properties of the original continuum are preserved while at the same time the new universe admits a wide variety and richness of new objects, including both infinite and infinitesimal numbers. On another and more directly geometrically founded approach, that of *Smooth Infinitesimal Analysis* (SIA), a restriction of functions to those that are "smooth" or arbitrarily many times differentiable allows for the postulation of a set of infinitesimal lengths within the real continuum. The requirement of smoothness and continuity of functions has the consequence that the underlying logic of SIA does not affirm the law of excluded middle, and so is intuitionist, and as a further consequence allows the refutation of the classical set-theoretical Axiom of Choice. Moreover, and even more radically, the identity relation here fails even to be decidable since it is not universally the case for all x, y that either $x = y$ or $x \neq y$. However, as Bell discusses, models of SIA can also be understood category-theoretically as certain topoi, leading to a fascinating connection to the topos-theoretical modeling of set theories in general (including, in the general case, intuitionistic set theory). In the more generalized context of a topos-theoretic universe, a consequence is that there are now various alternative constructions which may serve the role of "the" mathematical continuum: In Bell's own words, here, "the classical view that the linear continuum is a uniquely determined entity gives way to a pluralistic conception under which the continuum has a number of embodiments with essentially different properties" (p. 81, below). Bell's

article concludes with a brief and illuminating contrast of the logical structure and structural consequences of Nonstandard Analysis with those of SIA, and a suggestive consideration of the plausibility of considering SIA in particular as a potent model for real relations and results in classical and quantum physics.

Following Bell's article, the next three contributions to the section develop historically important conceptions and results from the nineteenth century concerning the structure of reals and the achieved mathematics of real analysis. First, Piotr Blasczyk and Marlena Fila's "On Bolzano and Greek Concepts of Continuity" traces the development of the problem and conception of the continuum from its ancient Greek sources, including their treatment in Euclid, through the more mechanical treatment of curves given by Pappus and then by Descartes. In the last part of their contribution, Blasczyk and Fila treat the conceptual innovations and philosophical implications of the 1817 work of the philosopher and mathematician Bernard Bolzano, *Rein analytischer Beweis*. . ., which culminates in his successful proof of (a formulation of) the Intermediate Value Theorem. Elías Fuentes Guillén's contribution, "Bolzano's Theory of Meβbare Zahlen: Insights and Uncertainties Regarding the Number Continuum," then considers the theory of "measurable numbers" introduced by Bolzano in the 1830s. As Guillén shows, Bolzano's thinking about these quantities was closely connected with his thinking about the general concept of numbers, including imaginary and complex ones, and of the infinite. Despite his initial rejection of the reality of infinitesimals, Bolzano's thinking is, Guillén shows, equally determined by his consideration of rational numbers that are capable of increasing or decreasing to an infinite or an infinitely small quantity. Following on these, John K. Dagsvik's article "Abel's approach to Elliptic Integrals: the Addition Theorems" contextualizes and explains the remarkable contributions of the nineteenth-century mathematician Niels Henrik Abel to the theory and analysis of elliptic integrals and functions, in particular developing his so-called "addition theorems" and bringing out the intuitions behind his distinctive approach to these functions as well as those underlying the previous approaches of others.

The last two papers of the section discuss contemporary developments of the continuum and its potential generalizations, including the dramatic logical, proof-theoretical, and computational implications of these recent innovations around the problem of continuous structure. In "The Absolute Arithmetic Continuum and its Geometric Counterpart," Philip Ehrlich employs the surreal numbers discovered by John Horton Conway in the 1970s as a maximally general field (**No**) of "all numbers great and small," including both reals and ordinals as well as infinites and infinitesimals. Ehrlich describes the development of what is essentially a tree of all the elements of **No** ordered recursively by the increasing lexicographic complexity of their canonical "proper names." Ehrlich suggests that this ordered field of surreal numbers may be considered to define a kind of "absolute" arithmetic continuum, above and beyond the ordered field of real numbers (but without the "Archimedean" property that is definitive in **R** itself). In the last two sections of the paper, Ehrlich fascinatingly suggests that a fundamental kind of bridging between the continuity characteristic of geometrical relationships and the discreteness characteristic of numbers may be provided by considering the Cartesian space over the ordered

field of surreal numbers thus structured. Here, indeed, Ehrlich suggests that it is just the properties that the surreal numbers inherit from their recursive unfolding in the simplicity-hierarchical tree that qualifies their elementary continua to define an absolute continuum suitable for use across a broad variety of specific developments of (generally non-Archimedean) continua in diverse fields.

The final chapter of the section, John Stillwell's "Reverse Mathematics," develops the related consequences for the proof-theoretical and computational bases of real analysis of the project that has been called that of "reverse mathematics," that is, the project of (rather than deriving theorems from "given" axioms) considering *which* axioms are needed to prove certain results. As Stillwell discusses, the first historical applications of this kind of project were marked in the consideration of the specific implications of, and consequences of omitting, Euclid's fifth (or "parallel") postulate for geometry; another important example is provided by set-theoretical reflection about the dispensability of the Axiom of Choice. After discussing the proof-theoretic foundations of the important developments of analysis in the nineteenth century as well as Dedekind's construction of the reals, Stillwell describes how the consideration of the proof-theoretic strength of various axioms systems may be related to considerations of the recursive computability of certain proposed derivations by means of certain determinate rules. Following the precise definition of computability achieved by Turing, Gödel, and Church in the 1930s, it is then possible to distinguish rigorously between recursively enumerable and nonenumerable sets of consequences and thus, in a natural way motivated by the computability considerations themselves, to separate "constructive" from "non-constructive" theorems of analysis and related fields. In particular, the addition of a single "recursive comprehension axiom" to the original Peano axioms for natural numbers leads to the extension of arithmetic to what may be considered a wholly constructive development of the reals. Beyond this, however, it is also possible to define further nonconstructive axiom systems of increasing degrees of complexity and to reflect on the extension which thereby becomes possible into more far-flung results of analysis, including those that essentially involve impredicative "constructions" in the sense discussed above.

2 Summary

With these considerations and constructions, the basic and underlying "philosophical" problem of the extent and limits of the actual positive construction of the reals by the finite conceptual methods of discursive thought, as against the apparently much vaster "reality" of the continuum understood as comprising the vast multiplicity of nonpredicative and nonconstructible elements, is brought to a revealing new level of formal and deductive reflective self-consciousness. At the same time, however, the underlying gulf between the powers of rational discursive thought and the generality of the continuum and its structure (or structures) seems to remain unabated and indeed appears likely to continue to provide a decisive and fruitful stimulus simultaneously for both mathematical innovation and philosophical reflection long into the future.

3 Cross-References

- ▶ Abel's Approach to Elliptic Integrals
- ▶ Bolzano's Theory of *meßbare Zahlen*: Insights and Uncertainties Regarding the Number Continuum
- ▶ On Bolzano and Greek Concepts of Continuity
- ▶ Reverse Mathematics
- ▶ The Absolute Arithmetic Continuum and its Geometric Counterpart
- ▶ The Continuum and the Evolution of the Concept of Real Number

References

Becker O (1933) Eudoxus-Studien: I: Eine voreudoxische Proportionenlehre und ihre Spuren bei Aristoteles und Euklid. Quellen und Studien zur Geschichte der Mathematik, Astronomie und Phyik B. II (1933), 311–330. [Reprinted in Jean Christianidis (ed) Classics in the history of Greek mathematics. Boston studies in the philosophie of science, vol 240, Dordrecht/Boston: 2004, 191–209, with intro. by Ken Saito, 188–9]

Bell JL (2022) Continuity and infinitesimals. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Spring 2022 edn. https://plato.stanford.edu/archives/spr2022/entries/continuity/

Euclid (1908) The thirteen books of Euclid's elements, 3 vols (trans: Heath TL). Cambridge University Press, Cambridge. Second edition in 1926

Feferman S (2005) Predicativity. In: Shapiro S (ed) Oxford handbook of philosophy of mathematics and logic. Oxford University Press, Oxford, pp 590–624

Heath TA (1921) A history of Greek mathematics, 2 vols. Clarendon, Oxford

Hesse H, Scholz H (1928) Die Grundlagenkrisis der grieschen Mathematik. Karl Metzner, Charlottenburg

Leibniz GWF (1686) Primary truths. In: Leibniz: philosophical essays. Translated by Roger Ariew and Daniel Garber. Hackett, Indianapolis, 1989

Leibniz GWF (1689) On freedom. In: Leibniz: philosophical essays. Translated by Roger Ariew and Daniel Garber. Hackett, Indianapolis, 1989

Leibniz GWF (1695) A new system of the nature and communication of substances, and of the union of soul and body. In: Leibniz: philosophical essays. Translated by Roger Ariew and Daniel Garber. Hackett, Indianapolis, 1989

Russell B (1908) Mathematical logic as based on the theory of types. Am J Math 30(3):222–262. https://doi.org/10.2307/2369948

Saito K (2004) "Introduction" to studies on the proportion theory and incommensurability. In: Christianids J (ed) Classics in the history of Greek mathematics. Kluwer, Dordrecht

Sayre KM (1983) Plato's late ontology: a riddle resolved. Princeton University Press, Princeton. Republished with a new introduction and the essay, "Excess and deficiency at statesman 283c–285d", Parmenides Publishing, 2005

von Fritz K (1945) The discovery of incommensurability by Hippasus of Metapontum. Ann Math 46(2):242–264. https://doi.org/10.2307/1969021

Weyl H (1919) The continuum: a critical examination of the foundation of analysis. Translated by Stephen Pollard and Thomas Bole. Thomas Jefferson University Press, Lanham, 1987

The Continuum and the Evolution of the Concept of Real Number

John L. Bell

Contents

1 The Origins of the Real Number Concept 1474
2 The Cantor-Dedekind Theory of Real Numbers and the Arithmetization of the Continuum 1481
3 Dissenting Voices: Divergent Conceptions of the Continuum in the Nineteenth and Early Twentieth Centuries 1496
4 The Continuum in Constructive and Intuitionistic Mathematics 1521
5 Nonstandard Analysis and the Hyperreal Line 1533
6 The Continuum in Smooth Infinitesimal Analysis 1536
References 1560

Abstract

This chapter traces the historical and conceptual development of the idea of the continuum and the allied concept of real number. Particular attention is paid to the idea of infinitesimal, which played a key role in the development of the calculus during the seventeenth and eighteenth centuries, and which has undergone a revival in the later twentieth century.

Keywords

Continuum · Real number · Arithmetization · Infinitesimal · Constructive mathematics · Intuitionistic mathematics · Nonstandard analysis · Smooth infinitesimal analysis

J. L. Bell (✉)
Department of Philosophy, The University of Western Ontario, London, ON, Canada
e-mail: jbell@uwo.ca

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_76

1 The Origins of the Real Number Concept

The concept of number has undergone a long evolution, leading to the emergence of several types of "number." The fundamental numbers are, of course, the *natural or whole numbers* 1, 2, 3, ... associated with the process, temporal in nature, of *counting.* Counting is applied to *discrete* collections of well-distinguished objects such as flocks of sheep, piles of stones, etc. Discrete collections are not always exactly divisible into numerically equal parts of a prescribed size. Thus, for example, a flock of 31 sheep, not being divisible into an exact number of pairs or triplets, cannot be exactly "counted by twos," or "counted by threes." Once the unit, in this case "individual sheep" is fixed, it cannot be arbitrarily changed and used to provide an exact count of the flock. This "rigidity of the unit" is a characteristic feature of the use of natural numbers in counting discrete collections. In this sense counting is *absolute.*

Natural numbers can be *compared.* The basic standard of comparison of numbers is, of course, *greater* ($>$) and *less* ($<$). A subtler standard of comparison is *proportionality* or *ratio.* Intuitively, given two numbers $m < n$, the ratio or proportion $m: n$ represents the *relative size* of m to n. In cases where m divides n exactly, $m: n$ can be regarded as a natural number, the number of times m divides n. For example 4: 24 may be identified with the number 6. But when m is not a divisor of n, the ratio $m: n$ cannot be represented by a natural number in this way. Such ratios came to be known as *fractions,* and later as *rational numbers.*

Such is the case for counting. In the case of *measuring,* additional problems arise from the fact that measuring is concerned not with discrete collections ultimately composed of distinct indivisible "atoms," but with *infinitely divisible, continuous entities,* in a word, with *continua.* In the case of discrete collections, like flocks of sheep, we are automatically provided with a natural "intrinsic" unit for counting purposes, so generating a natural number. Frege's brilliant idea of identifying the unit with a property or concept is illuminating here. In the case of a flock of 31 sheep, the number "31" is based, according to Frege, on the "concept-unit" "individual sheep." The "rigidity of the unit" in this case is illustrated by the fact that changing the concept-unit to "pair of sheep" renders it incapable of associating a natural number with the flock. Now because of the unlimited divisibility of continua, no such intrinsic units are, or can be provided. In measuring the length of a straight line, for example, we say that it is so many inches, or centimeters, long. Apart from the fact that these inches or centimeters represent lengths of straight lines, their lengths are entirely arbitrary "units," not determined in any further way by the extent of the given line. Measurement arises from a *comparison* of the length of the given line with the length of the linear unit, inch or centimeter as it may be. Measurement is thus *relative:* in measuring the length of a straight line an *additional* straight line must be specified as a unit of measurement. This fact further complicates the issue of defining the "ratio" of the lengths of straight lines, and divisible magnitudes generally. In the case of discrete collections, ratios are given directly in terms of the numbers of units constituting the collections: thus the ratio of a collection of 31 sheep to 68 sheep is 31: 68. But in the case of divisible magnitudes such as straight

lines, it is not always immediately clear how the "ratio" between them should be defined. Of course, they are straightforward cases in the divisible case such as the following: suppose we are given a straight line L and we construct two straight lines L' and L'' by laying down the given line L 31 and 68 times, respectively. Then the ratio of (the lengths of) L' to L'' is 31: 68. In effect, the situation has been reduced to the discrete case by taking (the length of) L as a common unit. But when the lines L' and L'' are *arbitrary*, what guarantee is there that they are *commensurable,* that is, can a line L be found to play the role of common unit of measure for L' and L'' so enabling the ratio or proportion to be presented numerically? Let us call this the problem of *commensurability.*

Certain ancient cultures such as the Egyptians and Babylonians handled ratios in a pragmatic manner. Treating them as "fractions," which later became known as *rational numbers,* they formulated rules of calculation for them analogous to, if more complicated than, the familiar rules of calculation for natural numbers. While we do not know whether they regarded fractions as being true "numbers," for all practical purposes they treated them as such. And it seems likely that they treated the issue of commensurability, when it arose in practice, in a rough-and-ready way.

As far as we know, in antiquity, only the mathematicians of the Greek school were exercised by the problem of commensurability, which had naturally arisen in their development of geometry – itself, as its name shows, a by-product of the practice of measurement. The Greek mathematicians had a strict conception of number, *arithmos,* defined by Euclid as a multitude composed of units, which resisted expansion to embrace proportions, ratios, or fractions. They accordingly regarded numerical proportions or ratios– that is, proportions or ratios *between* numbers – as being quite distinct from numbers *themselves.* By 500 B.C.E. they, more precisely the Pythagorean School, had developed a theory of numerical proportions.

It was a fundamental principle of the Pythagoreans that the world is explicable in terms of properties of, and relations between, whole numbers – that number, in fact, forms the very essence of the real. In geometry this doctrine led to the sweeping aside of the problem of commensurability by the affirmation of *universal commensurability*. In particular, given any pair of lines, they took it for granted that it is possible to choose a unit of length sufficiently small so as to enable the proportion of the lengths of the lines to be presented as a numerical ratio.

It must then have come as a great shock to the Pythagoreans to find, as they did, that universal commensurability *cannot be upheld*. This followed upon the shattering discovery, probably made by the later Pythagoreans before 410 B.C., that ratios of whole numbers do not suffice to enable the diagonal of a square or a pentagon to be compared in length with its side. They showed that these line segments are *incommensurable*, that is, it is not possible to choose a unit of length sufficiently small so as to enable both lines to be measured by an integral number of the chosen units. Pythagorean philosophy was dealt a devastating blow by this discovery. Their geometry, too, was affected, since the demonstrations of certain basic propositions employed the principle of universal commensurability.

The discovery of incommensurability (which can be seen as an early instance of a failure to reduce the continuous to the discrete) had made it impossible for

proportions between continuous magnitudes to be generally representable as numerical proportions. The response was to detach the general concept of proportion from the idea of numerical proportion and provide it with an axiomatic development independent both of numerical proportion and of commensurability. This is the theory presented in Book V of Euclid's *Elements.* The commentator Proclus attributes the general theory of proportion to Eudoxus of Cnidus (c.400–350 B.C.E.)

In Eudoxus's theory, we are given a collection of *similar magnitudes,* e.g., line segments, or planes, or volumes, or angles, etc., which can be compared in size, together with the notion of *ratio* of similar magnitudes satisfying certain *postulates,* of which the following, expressed in modern terms, are the most important:

P_1. Given any two similar magnitudes, there is an integral multiple of the one which exceeds the other.

This postulate, which is also known as *Archimedes' principle*, has the effect of excluding infinitely small or infinitely large quantities.

The second postulate is in effect a *definition of equality* for ratios:

P_2. The ratio *A:B* of two magnitudes *A, B* is equal to the ratio *C:D* of two other magnitudes *C, D* if and only if for any natural numbers *m, n, mA* is greater than, equal to, or less than *nB* according as *mC* is greater than, equal to, or less than *nD.*

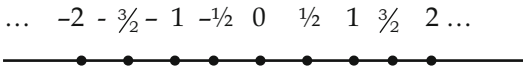
This postulate covers both the cases when A and B are commensurable, and when they are incommensurable. In the former case, numbers p, q can be found to satisfy $pA = qB$. It follows that $pC = qD$ and it is then easily shown that the remaining conditions of the definition are satisfied. In this case equality between the ratios $A{:}B$ and $C{:}D$ is established in a finite number of steps, so the ratios are numerical, and correspond to *rational numbers*.

In the case of *incommensurable* A and B, the content of **P_2** has been interpreted by historians of mathematics as an anticipation of the concept of *irrational number*, and so, by extension, of the general concept of *real number.* There is indeed a strikingly close relationship (Fine 1917) between the idea of the ratio of two incommensurables suggested in **P_2** and that of an irrational number as defined by Dedekind in the nineteenth century (discussed in the next section). The key element in **P_2** is the fact that when A and B are incommensurable the definition of the ratio $A : B$ demands comparing mA and nB for *all* pairs (m, n) of all natural numbers m and $n,$ and thus, implicitly at least, the assembling of these pairs into two separate classes, the class $\{(m, n)\colon mA > nB\}$ and the class $\{(m, n)\colon mA < nB\}$. This separation of pairs of numbers into two classes is, remarkably, identical with the Dedekind cut defining the irrational number which expresses the ratio $A : B$. It follows that the collection of ratios, commensurable and incommensurable, as defined in Euclid Book V, is actually a *subset* of the set $\mathbb{R}$ of real numbers as defined by Dedekind. For this subset to be identifiable with the whole of $\mathbb{R}$ it is necessary to postulate, as Dedekind (and Cantor) did, the existence of a one-to-one correspondence between real numbers and points on a line. This amounts, in the present case, to the assertion that for every Dedekind cut there exists a pair of magnitudes A and B which will yield this cut in the manner just described.

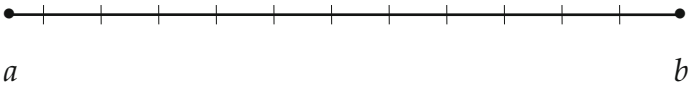
This is not, of course, to claim that Eudoxus or Euclid had literally "anticipated" Dedekiind's definition of irrational or real number. As already observed, they would have rejected the very idea that a numerical ratio could be identified as an actual number. They avoid the fraction and the irrational by basing the theory of proportion upon the equality of ratios instead of ratio itself. Nevertheless, there is no question that the basis of Dedekind's sound definition of the concept of real number millennia later was implicit in notions familiar to Eudoxus and Euclid.

Had incommensurability never been discovered, there would have been no necessity to extend the number system beyond the rational numbers.

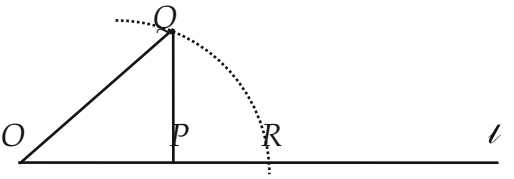
It was only very much later that fractions (along with negative numbers and 0, which, together with the natural numbers, constitute the *integers*) came to be regarded as genuine numbers. This essentially came about in the sixteenth and seventeenth centuries through an expansion of the Greek concept of number as discrete collections of units to numbers conceived as measuring arbitrary quantities such as geometric magnitudes. Numbers came to be associated with *points on a (continuous) line.* In particular the positive rationals (i.e., those of the form m/n with m and n positive integers) appear to the right of the origin 0 and the negative rationals (i.e., those of the form p/q with p negative and q positive) to the left of 0.



Regarding the rationals as points on a line enables the line to be divided as finely as one pleases. For example, to represent all rationals of the form $m/10^9$ as points on the line, we divide the interval (0, 1) of the line between 0 and 1 into a billion equal pieces; similarly for all other intervals (1, 2), (2, 3),. . . and the points of subdivision then correspond to fractions of the form $m/10^9$. Since the denominator of these fractions can be made arbitrarily large, 10^{10}, 10^{100}, or whatever, thereby producing subdivisions of unlimited fineness, it would be natural to suppose that in this way one would capture as points of subdivision *all* the points on the line – in other words, that every point is represented by a *rational* number. Now it is certainly true that rational numbers suffice for all practical purposes of measuring; indeed, the Pythagorean doctrine of universal commensurability essentially elevated this practical fact to a general principle. Moreover, the rational points are *dense* on the line in the sense that they may be found in any interval (a, b), however small, with rational endpoints $a < b$ – we have only to observe that the rational number $(a + b)/2$ lies between a and b. It may be inferred from this fact that each such interval contains *infinitely* many rational points. For if the interval (a, b) contained only finitely many, n say, then we could mark them off as shown below and then any interval between two adjacent points would be free of rational points, contradicting what we have already established.



All this would seem to lend support to the idea that every point on the line is represented by a rational number. But, of course, the Pythagorean discovery of the incommensurability of the side and diagonal of a square shows that this idea is incorrect. If we take two perpendicular lines *OP* and *PQ* of length 1 and use compasses to mark out on the line ℓ a line *OR* of the same length as *OQ* then the incommensurability of *OP* and *OQ*, and hence also *OR,* just means that *R* is not a rational point. In fact, if we designate



the "number" associated with the point *R,* then we have, by the Pythagorean theorem, $r^2 = 1^2 + 1^2 = 2$, so that, replacing *r* by the customary symbol $\sqrt{2}$, we conclude that *no rational number is equal to* $\sqrt{2}$ or that $\sqrt{2}$ is *irrational.* Here the term "irrational" is to be understood in the sense of "that which cannot be expressed as a ratio" as opposed to its more usual (but related) meaning "contrary to reason."

The fact that not every point on the line corresponds to a rational number means that, if the correspondence between points on a continuous line and "numbers" is to be maintained, the system of rational numbers has to be *enlarged still further.* It was essentially this observation, first made by the Dutch mathematician Simon Stevin (1548–1620), that led to the development of the system of *real numbers.* In his *L'Arithmétique* of 1585 Stevin explicitly introduces the idea of extending the discrete arithmetic of the Greeks to a continuous arithmetic. Stevin thinks of numbers as measures, and measures of continuous magnitudes are by their very nature continuous:

> *as well as to continuous water corresponds a continuous wetness*, *so to a continuous magnitude corresponds a continuous number.*

The term "continuous number" would accordingly have been more apposite than the term "real number." But history decreeds otherwise. The term "real number" was introduced in the seventeenth century by Descartes to distinguish them from what had come to be termed "imaginary" numbers – that is, square roots of negative numbers (such as $i = \sqrt{-1}$) arising formally as roots to algebraic equations. These latter were regarded as fictitious, the mere product of imagination. In this respect the term "imaginary" is to be contrasted with the term "irrational."

Stevin's most important innovation was the *decimal representation of real numbers,* which appears (although not of course in modern notation) in his 1585 work *De Thiende* ("the art of tenths"). This still today provides the most familiar way of presenting the continuum arithmetically.

In a major departure from the rational numbers which are given simply by pairs of integers, under the decimal representation each real number takes the form of an *infinite* sequence of integers

$$m.n_0\,n_1\,n_2\ldots.$$

A real number presented in this way is then a rational number precisely when it is *periodic,* i.e., displays a repeating pattern indefinitely after a certain stage. For example,

$$6\,{}^{5}\!/_{8} = 6.6250000\ldots \qquad {}^{3}\!/_{7} = 0.428571\,428571\ldots.$$

It follows that just a *finite* amount of information is needed to represent a rational number as a decimal fraction.

Any *nonperiodic* decimal must then represent an *irrational* number; it is easy to furnish examples of these, for instance, the following decimal containing an increasing number of zeros:

$$0.1010010001000010000001\ldots.$$

Surprisingly, perhaps, no simple rule of this kind exists for constructing the decimal representation of familiar irrational numbers such as $\sqrt{2}$.

The system of *real numbers* can be formally defined to be the set of all finite or infinite (positive or negative) decimals: considered geometrically the real numbers constitute the *geometric continuum* or *real line.* The operations of addition and multiplication can be naturally extended to the real numbers so that, like the rational numbers, they constitute a *field,* which we shall denote by the symbol $\mathbb{R}$. Thus the real numbers resemble the rational numbers insofar as they are subject to the same operational laws. On the other hand, if we regard the integers as the basic ingredients from which the other numbers are constructed, then, as pointed out above, while each rational number can be defined in terms of just two integers, in general a real number requires *infinitely many* integers to define it. The fact that infinite processes play an essential role in the construction of the real numbers places them in sharp contrast with the rationals.

The definition of real numbers as infinite decimals is not entirely satisfactory since, for one thing, there is no compelling mathematical reason to choose the number 10 as a base for them. Moreover, the real numbers are supposed to correspond exactly to points on a line; but how do we know that every such point corresponds to a real number thus defined as a decimal? To establish this it is necessary to show that there are no "gaps" in our set of real numbers, and to define with precision what is to be understood by this assertion. This was carried out in the latter half of the nineteenth century and resulted in the modern theory of real numbers.

The Vexed Issue of Infinitesimals. Closely associated with the concept of an infinitely divisible continuum is that of an *infinitesimal.* According to the *Oxford English Dictionary* the term *infinitesimal* was originally

> *an ordinal, viz. the "infinitieth" in order; but, like other ordinals, also used to name fractions, thus infinitesimal part or infinitesimal came to mean unity divided by infinity* $\left(\frac{1}{\infty}\right)$*, and thus an infinitely small part or quantity.*

An *infinitesimal magnitude* has been, traditionally, somewhat hazily conceived as a continuum "viewed in the small," an "ultimate part" of a continuum. In something like the same sense as a discrete entity is made up of its individual units, its "indivisibles," so, it was maintained, a continuum is "composed" of infinitesimal magnitudes, its ultimate parts. (It is in this sense, for example, that mathematicians of the seventeenth century held that continuous curves are "composed" of infinitesimal straight lines.) Now the "coherence" of a continuum entails that each of its (connected) parts is also a continuum, and, accordingly, divisible. Since individual points are indivisible, it follows that no point can be part of a continuum. Points are, in fact, just *locations* in a continuum. Infinitesimal magnitudes, as parts of continua, cannot, of necessity, be points; they are, in a word, *nonpunctiform.*

Magnitudes are normally taken as being *extensive* quantities, like mass or volume, which are defined over extended regions of space. By contrast, infinitesimal magnitudes have been conceived as *intensive* magnitudes resembling locally defined intensive quantities such as temperature or density. The effect of "distributing" or "integrating" an intensive quantity over an infinitesimal magnitude is to convert the former into an infinitesimal extensive quantity; thus temperature is transformed into infinitesimal heat and density into infinitesimal mass. When the continuum is the trace of a motion, the associated infinitesimal/intensive magnitudes have been identified as *potential* magnitudes – entities which, while not possessing true magnitude themselves, embody a *tendency* to generate magnitude through motion, so manifesting "becoming" as opposed to "being."

An infinitesimal *number* has been conceived as a number so small that, while not coinciding with zero, is in some sense smaller than any finite positive number. An infinitesimal number is, so to speak, "greater than nothing but less than anything" (Pyle 1997, p. 208). An infinitesimal number has been construed as a "number' which fails to satisfy *Archimedes' Principle*, that is, as a nonzero "number" a such that, for any integer n, $n.a$ is less than any finite nonzero number. We have already pointed out that Eudoxus and Euclid explicitly exclude the existence of such "numbers."

The concept of infinitesimal was beset by controversy from its beginnings. The idea makes an early appearance in the mathematics of the Greek atomist philosopher Democritus c. 450 B.C., only to be banished c. 350 B.C. by Eudoxus in what was to become official "Euclidean" mathematics. Taking the form of "indivisibles," infinitesimals resurfaced in the sixteenth and seventeenth centuries and were systematically employed by Kepler, Galileo's student Cavalieri, the Bernoulli clan, and a number of other mathematicians. In the guise of the delightfully named "linelets" and "timelets," infinitesimals played an essential role in Barrow's "method for finding tangents by calculation," which appears in his *Lectiones Geometricae* of 1670. As "evanescent quantities" infinitesimals were instrumental (although later abandoned) in Newton's development of the calculus, and, as "inassignable quantities," in Leibniz's. The Marquis de l'Hôpital, who in 1696 published the first treatise on the differential calculus (entitled *Analyse des Infiniments Petits pour l'Intelligence des Lignes Courbes*), invokes the concept in postulating that "a curved line may be regarded as being made up of infinitely small straight line segments,"

and that "one can take as equal two quantities differing by an infinitely small quantity."

However useful infinitesimals may have been in practice, they could scarcely withstand logical scrutiny. Derided by Berkeley in the eighteenth century as "ghosts of departed quantities," in the nineteenth century execrated by Cantor as "cholera-bacilli" infecting mathematics, and in the twentieth roundly condemned by Bertrand Russell as "unnecessary, erroneous, and self-contradictory," the use of infinitesimals in the calculus and mathematical analysis was believed to have been finally supplanted by the limit concept which took rigorous and final form in the latter half of the nineteenth century. By the beginning of the twentieth century, the concept of infinitesimal had become, in analysis at least, essentially an "unconcept."

Nevertheless, the proscription of infinitesimals did not succeed in extirpating them; they were, rather, driven further underground. Physicists and engineers, for example, never abandoned their use as a heuristic device for the derivation of correct results in the application of the calculus to physical problems. Differential geometers of the stature of Lie and Cartan relied on their use in the formulation of concepts which would later be put on a "rigorous" footing. Even in mathematical analysis they survived in Du Bois-Reymond's "orders of infinity," and, in a mathematically rigorous sense, they lived on in the algebraists' investigations of nonarchimedean fields.

A new phase in the saga of infinitesimals has opened in the past few decades with the refounding on a solid basis of the concept of the infinitesimal in analysis and differential geometry. This has been achieved in two essentially different ways, *Nonstandard Analysis* and *Smooth Infinitesimal Analysis.* These will be discussed in the concluding sections of this survey.

2 The Cantor-Dedekind Theory of Real Numbers and the Arithmetization of the Continuum

By the beginning of the nineteenth century the conception of real numbers as corresponding to points on a line, and so forming a continuum, was firmly established in mathematical practice. But solid definitions of the concept of continuum, and indeed the concept of real number itself, were lacking. Mathematical arguments involving real numbers frequently made appeal to geometric – even spatiotemporal – intuition and the intuitive idea of continuity. (Respectability on the use of spatiotemporal intuition in mathematical arguments had been conferred by the central role such intuition played in Kant's philosophy of mathematics. Kant had claimed that arithmetic and geometry were grounded in the intuition of time and space, respectively. The influence of Kant's philosophy on mathematicians waned after the emergence of non-Euclidean of geometry early on in the nineteenth century.) In addition, the concept of infinitesimal, which had played such an important role in the development of the calculus over the previous two centuries, but whose logical foundations were worryingly shaky, had begun to come under scrutiny.

Cauchy, Bolzano, and Hamilton had all attempted to clarify the concepts of real number, but it was **Karl Weierstrass** (1815–97) who made the first systematic attack on the problem. He was determined to expel spatiotemporal intuition, and the infinitesimal, from the foundations of analysis. To instill complete logical rigor Weierstrass proposed to establish mathematical analysis on the basis of natural number alone – to "arithmetize" it. According to Hobson (1957, p. 22), "the term 'arithmetization' is used to denote the movement which has resulted in placing analysis on a basis free from the idea of measurable quantity, the fractional, negative, and irrational numbers being so defined that they depend ultimately upon the conception of integral number."

In pursuit of this goal Weierstrass had first to formulate a rigorous "arithmetical" definition of real number. He did this by defining a (positive) real number to be a countable set of positive rational numbers for which the sum of any finite subset always remains below some preassigned bound, and then specifying the conditions under which two such "real numbers" are to be considered equal, or strictly less than, one another.

Weierstrass was also concerned to make precise the concept of *continuous function.* The concept of function had by this time been greatly broadened: in 1837 Dirichlet suggested that a variable y should be regarded as a function of the independent variable x if a rule exists according to which, whenever a numerical value of x is given, a unique value of y is determined. (This idea was later to evolve into the set-theoretic definition of function as a set of ordered pairs.) Dirichlet's definition of function as a correspondence, from which all traces of continuity had been purged, made necessary Weierstrass's independent definition of continuous function.

Weierstrass formulated the familiar (ε, δ) definition of continuous function: a function $f(x)$ is continuous at a if for any $\varepsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ for all x with $|x - a| < \delta$. He also proved his famous *approximation theorem for continuous functions:* any continuous function defined on a closed interval of real numbers can be uniformly approximated to by a sequence of polynomials.

The notion of *uniform continuity* for functions was later introduced (in 1870) by Heine: a real valued function f is uniformly continuous if for any $\varepsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all x and y in the domain of f with $|x - y| < \delta$. In 1872 Heine proved the important theorem that any continuous real-valued function defined on a closed bounded interval of real numbers is uniformly continuous.

Following Weierstrass's efforts, another attack on the problem of formulating rigorous definitions of continuity and the real numbers was mounted by **Richard Dedekind** (1831–1916). We learn from the introductory remarks to *Continuity and Irrational Numbers* (1872) that Dedekind was inspired to embark on his investigations by his belief that, in presenting the differential calculus, "geometric intuition," while "exceedingly useful from a didactic standpoint," can "make no claim to being scientific." Accordingly he made "the fixed resolve to keep meditating on the question till I should find a purely arithmetic and perfectly rigorous foundation for the principles of infinitesimal analysis." He went on to observe

> *The statement is so frequently made that that the differential calculus deals with continuous magnitude, and yet an explanation of this continuity is nowhere given; even the most rigorous expositions of the differential calculus do not base their proofs upon continuity but, with more or less consciousness of the fact, they either appeal to geometric notions or those suggested by geometry, or upon theorems which are never established in a purely arithmetic manner.*

Dedekind considered the concept of number in general to be a part of *logic,* and not, as Kant had claimed, dependent on spatiotemporal intuition. In the Preface to his later book "*The Nature and Meaning of Numbers*" (*1888*) *he says*

> *In speaking of arithmetic (algebra, analysis) as a part of logic I mean to imply that I consider the number-concept entirely independent of the notions or intuitions of space and time, that I consider it an immediate result from the laws of thought.*

In his investigations of continuity Dedekind focused attention on the question: exactly what is it that distinguishes a continuous domain from a discontinuous one? He seems to have been the first to recognize that the property of density, possessed by the ordered set of rational numbers, is insufficient to guarantee continuity. In *Continuity and Irrational Numbers* he remarks that when the rational numbers are associated to points on a straight line, "there are infinitely many points [on the line] to which no rational number corresponds" (Ewald 1999, p. 770 *et seq*.), so that the rational numbers manifest "a gappiness, incompleteness, discontinuity," in contrast with the straight line's "absence of gaps, completeness, continuity." He goes on:

> *In what then does this continuity consist? Everything must depend on the answer to this question, and only through it shall we obtain a scientific basis for the investigations of all continuous domains. By vague remarks upon the unbroken connection in the smallest parts obviously nothing is gained; the problem is to indicate a precise characteristic of continuity that can serve as the basis for valid deductions. For a long time I pondered over this in vain, but finally I found what I was seeking. This discovery will, perhaps, be differently estimated by different people; but I believe the majority will find its content quite trivial. It consists of the following. In the preceding Section attention was called to the fact that every point p of the straight line produces a separation of the same into two portions such that every point of one portion lies to the left of every point of the other. I find the essence of continuity in the converse, i.e., in the following principle:*
>
> *'If all points of the straight line fall into two classes such that every point of the first class lies to the left of every point of the second class, then there exists one and only one point which produces this division of all points into two classes, this severing of the straight line into two portions.'*

Dedekind regards this principle as being essentially indemonstrable; he ascribes to it, rather, the status of an axiom "by which we attribute to the line its continuity, by which we think continuity into the line." It is not, Dedekind stresses, necessary for *space* to be continuous in this sense, for "many of its properties would remain the same even if it were discontinuous." And in any case, he continues,

> *if we knew for certain that space were discontinuous there would be nothing to prevent us . . . from filling up its gaps in thought and thus making it continuous; this filling up would consist in a creation of new point-individuals and would have to be carried out in accordance with the above principle.*

The filling-up of gaps in the rational numbers through the "creation of new point-individuals" is the key idea underlying Dedekind's construction of the domain of real numbers. He first defines a *cut* to be a partition (A_1, A_2) of the rational numbers such that every member of A_1 is less than every member of A_2. After noting that each rational number corresponds, in an evident way, to a cut, he observes that infinitely many cuts fail to be engendered by rational numbers. The discontinuity or incompleteness of the domain of rational numbers consists precisely in this latter fact. That being the case, he continues,

> *whenever we have a cut (A_1, A_2) produced by no rational number, we create a new number, an irrational number α, which we regard as completely defined by this cut (A_1, A_2); we shall say that the number α corresponds to this cut, or that it produces this cut. From now on, therefore, to every definite cut there corresponds a definite rational or irrational number, and we regard two numbers as different or unequal if and only if they correspond to essentially different cuts.*

It is to be noted that Dedekind does not identify irrational numbers with cuts; rather, each irrational number is newly "created" by a mental act and remains quite distinct from its associated cut.

Dedekind goes on to show how the domain of cuts, and thereby the associated domain of real numbers, can be ordered in such a way as to possess the property of *continuity*, viz.

> *if the system R of all real numbers divides into two classes A_1, A_2 such that every number a_1 of the class A_1 is less than every number a_2 of the class A_2, then there exists one and only one number α by which this separation is produced.*

Dedekind notes that this property of continuity is actually equivalent to two principles basic to the theory of limits; these he states as:

> *If a magnitude grows continually but not beyond all limits it approaches a limiting value*

and

> *if in the variation of a magnitude x we can, for every given positive magnitude δ, assign a corresponding interval within which x changes by less than δ, then x approaches a limiting value.*

Dedekind's definition of real numbers as cuts was to become an essential part of the rigorous analysis of the continuum.

The most visionary "arithmetizer" of all was **Georg Cantor** (1845–1918). Cantor's analysis of the continuum in terms of infinite point sets led to his theory of

transfinite numbers and to the eventual freeing of the concept of set from its geometric origins as a collection of points, so paving the way for the emergence of the concept of general abstract set central to today's mathematics. It is of interest to note that neither Dedekind nor Cantor took real numbers to be actual *sets*. This step was in fact first taken by Bertrand Russell in 1903 in *The Principles of Mathematics* (Russell 1964).

At about the same time that Dedekind published his researches into the nature of the continuous, Cantor formulated his theory of the real numbers. This was presented in the first section of a paper (Cantor 1872) on trigonometric series. Like Weierstrass and Dedekind, Cantor aimed to formulate an adequate definition of the irrational numbers which avoided the presupposition of their prior existence, and he follows them in basing his definition on the rational numbers. Following Cauchy, Cantor calls a sequence $a_1, a_2, \ldots, a_n, \ldots$ of rational numbers a *fundamental sequence* (nowadays called a *Cauchy sequence*) if, for any positive rational ε, there exists an integer N such that, $|a_{n+m} - a_n| < \varepsilon$ for all m and all $n > N$. Any sequence $<a_n>$ satisfying this condition is said to have a *definite limit b*. Dedekind had taken irrational numbers to be "mental objects" associated with cuts, so, analogously, Cantor regards these definite limits as nothing more than *formal symbols* associated with fundamental sequences (Dauben 1979, p. 38).

The domain B of such symbols may be considered an enlargement of the domain A of rational numbers, since each rational number r may be identified with the formal symbol associated with the fundamental sequence $r, r, \ldots, r, \ldots$. Order relations and arithmetical operations are then defined on B: for example, given three such symbols b, b', b'' associated with the fundamental sequences $<a_n>$, $<a'_n>$, $<a''_n>$, the inequality $b < b'$ is taken to signify that, for some $\varepsilon > 0$ and N, $a_n - a'_n > \varepsilon$ for all $n > N$, while the equality $b + b' = b''$ is taken to express the relation $\lim (a_n + a'_n - a''_n) = 0$.

Having imposed an arithmetical structure on the domain B, Cantor is emboldened to refer to its elements as (real) *numbers*. Nevertheless, he still insists that these "numbers" have no existence except as representatives of fundamental sequences: in his theory,

> *the numbers* (*above all lacking general objectivity in themselves*) *appear only as components of theorems which have objectivity, for example, the theorem that the corresponding sequence has the number as limit*. (Dauben 1979, p. 39).

Cantor next considers how real numbers are to be associated with points on the linear continuum. If a given point on the line lies at a distance from the origin O bearing a rational relation to the point at unit distance from that origin, then it can be represented by an element of A. Otherwise, it can be approached by a sequence $a_1, a_2, \ldots a_n, \ldots$ of points each of which corresponds to an element of A. Moreover, the sequence $<a_n>$ can be taken to be a fundamental sequence; Cantor writes:

> *The distance of the point to be determined from the point O* (*the origin*) *is equal to b*, *where b is the number corresponding to the sequence* (Dauben 1979, p. 40).

In this way Cantor shows that each point on the line corresponds to a definite element of *B*. Conversely, each element of *B* should determine a definite point on the line. Realizing that the intuitive nature of the linear continuum precludes a rigorous proof of this property, Cantor simply assumes it as an axiom, just as Dedekind had done in regard to his principle of continuity:

> *Also conversely, to every number there corresponds a definite point of the line, whose coordinate is equal to that number* (Dauben 1979, p. 40).

For Cantor, who began as a number-theorist, and throughout his career cleaved to the discrete, it was numbers, rather than geometric points, that possessed objective significance. Indeed, the isomorphism between the discrete numerical domain *B* and the linear continuum was regarded by Cantor essentially as a device for facilitating the manipulation of numbers.

Cantor and Dedekind had offered different ways of "constructing" the arithmetical continuum of real numbers: the former, using Cauchy sequences of rational numbers, the latter, "cuts" in the rationals. Since each claimed that his system of real numbers corresponded exactly to the points on a geometric line, it would seem to follow that Cantor's and Dedekind's systems of real numbers were essentially "the same," i.e., isomorphic. However, neither provided an explicit proof of this assertion. That the two systems are indeed isomorphic later came to be rigorously proved by showing that they are both *complete ordered fields* and that any two complete ordered fields are isomorphic. An *ordered field* is an algebraic field on which a linear ordering is defined which is compatible with the algebraic operations. *Completeness* is the condition that any bounded subset has a greatest lower bound and a least upper bound (McShane and Botts 1959, Ch. 1). It can be shown that every complete ordered field has the *Archimedean property:* for any $a, b > 0$, there is a natural number n such that $na > b$. This is essentially the condition $\mathbf{P_1}$ laid down in Eudoxus' theory of proportions.

Cantor's arithmetization of the continuum had another important consequence. It had long been recognized that the sets of points of any pair of line segments, even if one of them is infinite in length, can be placed in one-one correspondence. This fact was taken to show that such sets of points have no well-defined "size." But Cantor's identification of the set of points on a linear continuum with a domain of numbers enabled the *sizes* of point sets to be compared in a definite way, using the well-grounded idea of *one-one correspondence between sets of numbers*. Thus, in a letter to Dedekind written in November 1873 Cantor notes that the totality of natural numbers can be put into one-one correspondence with the totality of positive rational numbers, and, more generally, with the totality of finite sequences of natural numbers. It follows that these totalities have the same "size"; they are all *denumerable.* Cantor next raises the question of whether the natural numbers can be placed in one-one correspondence with the totality of all positive *real* numbers (Ewald 1999, p. 844).

He quickly answers his own question in the negative. In letters to Dedekind written during December 1873 Cantor shows that, for any sequence of real numbers, one can define numbers in every interval that are not in the sequence. It follows in

particular that the whole set of real numbers is nondenumerable. Another important consequence concerns the existence of *transcendental numbers*, that is, numbers which are not algebraic in the sense of being the root of an algebraic equation with rational coefficients. In 1844 Liouville had established the transcendentality of any number of the form

$$\frac{a_1}{10} + \frac{a_2}{10^{2!}} + \frac{a_3}{10^{3!}} + \cdots$$

where the a_i are arbitrary integers from 0 to 9 (Kline 1972, p. 981).

In his reply to Cantor's letter of November 1873, Dedekind had observed that the set of algebraic numbers is denumerable; it followed from the nondenumerability of the real numbers that *there must be many transcendental numbers* (in fact nondenumerably many, although Cantor did not make this fact explicit until sometime later).

By this time Cantor had come to regard nondenumerability as a necessary condition for the continuity of a point set, for in a paper of 1874 he asserts:

> *Moreover, the theorem. . .represents the reason why aggregates of real numbers which constitute a so-called continuum (say the totality of real numbers which are* > 0 *and* ≤ 1*), cannot be uniquely correlated with the aggregate (v); thus I found the clear difference between a so-called continuum and an aggregate like the entirety of all real algebraic numbers* (Dauben 1979, p. 53).

Cantor next became concerned with the question of whether the points of spaces of different dimensions – for instance a line and a plane – can be put into one-one correspondence. In a letter to Dedekind of January 1874 he remarks:

> *It still seems to me at the moment that the answer to this question is very difficult – although here too one is so impelled to say no that one would like to hold the proof to be almost superfluous* (Ewald 1999, p. 850).

Nevertheless, 3 years later Cantor, in a dramatic *volte-face*, established the existence of such correspondences between spaces of different dimensions. He showed, in fact, that (the points of) a space of any dimension whatsoever can be put into one-one correspondence with (the points of) a line. This result so startled him that, in a letter to Dedekind of June 1877 he was moved to exclaim: *Je le vois, mais je ne le crois pas* ("I see it, but I don't believe it." (Ewald 1999, p. 860).

Cantor's discovery caused him to question the adequacy of the customary definition of the dimension of a continuum. For it had always been assumed that the determination of a point in an n-dimensional continuous manifold requires *n* independent coordinates, but now Cantor had shown that, in principle at least, the job could be done with just a *single* coordinate. For Cantor this fact was sufficient to justify the claim that

> *. . . all philosophical or mathematical deductions that use that erroneous presupposition are inadmissible. Rather the difference that obtains between structures of different dimensionnumber must be sought in quite other terms than in the number of independent coordinates – the number that was hitherto held to be characteristic* (Ewald 1999, p. 860).

In his reply to Cantor, Dedekind conceded the correctness of Cantor's result, but balked at Cantor's radical inferences therefrom. Dedekind maintained that the dimension-number of a continuous manifold was its "first and most important invariant" and emphasized the issue of continuity:

> *For all authors have clearly made the tacit, completely natural presupposition that in a new determination of the points of a continuous manifold by new coordinates, these coordinates should also (in general) be continuous functions of the old coordinates, so that whatever appears as continuously connected under the first set of coordinates remains continuously connected under the second* (Ewald 1999, p. 863).

Dedekind also noted the extreme discontinuity of the correspondence Cantor had set up between higher dimensional spaces and the line:

> *. . . it seems to me that in your present proof the initial correspondence between the points of the ρ-interval (whose coordinates are all irrational) and the points of the unit interval (also with irrational coordinates) is, in a certain sense (smallness of the alteration), as continuous as possible; but to fill up the gaps, you are compelled to admit, a frightful, dizzying discontinuity in the correspondence, which dissolves everything to atoms, so that every continuously connected part of the one domain appears in its image as thoroughly decomposed and discontinuous* (Ewald 1999, pp. 863–4).

Dedekind avows his belief that no one-one correspondence between spaces of different dimensions can be continuous:

> *If it is possible to establish a reciprocal, one-to-one, and complete correspondence between the points of a continuous manifold A of a dimensions and the points of a continuous manifold B of b dimensions, then this correspondence itself, if a and b are unequal, is necessarily utterly discontinuous* (Ewald 1999, p. 863).

In his reply to Dedekind of July 1877 Cantor clarifies his remarks concerning the dimension of a continuous manifold:

> *. . .I unintentionally gave the appearance of wishing by my proof to oppose altogether the concept of a ρ-fold extended continuous manifold, whereas all my efforts have rather been intended to clarify it and to put it on the correct footing. When I said: "Now it seems to me that all philosophical and mathematical deductions which use that erroneous presupposition –" I meant by this presupposition not "the determinateness of the dimension-number" but rather the determinates of the independent coordinates, whose number is assumed by certain authors to be in all circumstances equal to the number of dimensions. But if one takes the concept of coordinate generally, with no presuppositions about the nature of the intermediate functions, then the number of independent, one-to-one, complete coordinates, as I showed, can be set to any number* (Ewald 1999, p. 863).

Nevertheless he agrees with Dedekind that if "we require that the correspondence be continuous, then only structures with the same number of dimensions can be related to each other one-to-one." (Ewald 1999, p. 863). In that case, an invariant *can* be found in the number of independent coordinates, "which ought to lead to a definition of the dimension-number of a continuous structure" (Ewald 1999, p. 863).

The problem is to correlate that dimension-number, a perfectly definite mathematical object, with something as elusive as an arbitrary continuous correspondence. Cantor writes:

> *However, I do not yet know how difficult this path (to the concept of dimension-number) will prove, because I do not know whether one is able to limit the concept of continuous correspondence in general. But everything in this direction seems to me to depend on the possibility of such a limiting.*
>
> *I believe I see a further difficulty in the fact that this path will probably fail if the structure ceases to be thoroughly continuous; but even in this case one wants to have something corresponding to the dimension-number – all the more so, given how difficult it is to prove that the manifolds that occur in nature are thoroughly continuous* (Ewald 1999, p. 863).

In rendering the continuous discrete, and thereby admitting *arbitrary* correspondences "of [a] frightful, dizzying discontinuity" between geometric objects "dissolved to atoms," Cantor grasps at the same time that he has rendered the intuitive concept of spatial dimension a hostage to fortune.

In 1878 Cantor published a fuller account of his ideas (Cantor 1878). Here he explicitly introduces the concept of the *power (Mächtigkeit)* of a set of points: two sets are said to be of equal power if there exists a one-one correspondence between them. Cantor presents demonstrations of the denumerability of the rationals and the algebraic numbers, remarking that "the sequence of positive whole numbers constitutes. . .the least of all powers which occur among infinite aggregates." (Dauben 1979, p. 59).

The central theme of Cantor's 1878 paper is the study of the powers of continuous n-dimensional spaces. He raises the issue of invariance of dimension and its connection with continuity:

> *Apart from making the assumption, most are silent about how it follows from the course of this research that the correspondence between the elements of the space and the system of values $x_1, x_2, \ldots, x_n$ is a continuous one, so that any infinitely small change of the system $x_1, x_2, \ldots, x_n$ corresponds to an infinitely small change of the corresponding element, and conversely, to every infinitely small change of the element a similar change in the coordinates corresponds. It may be left undecided whether these assumptions are to be considered as sufficient, or whether they are to be extended by more specialized conditions in order to consider the intended conceptual construction of n-dimensional continuous spaces as one ensured against any contradictions, sound in itself* (Dauben 1979, p. 60).

The remarkable result obtained when one no longer insists on continuity in the correspondence between the spatial elements and the system of coordinates is described by Cantor in the following terms:

> *As our research will show, it is even possible to determine uniquely and completely the elements of an n-dimensional continuous space by a single real coordinate t. If no assumptions are made about the kind of correspondence, it then follows that the number of independent, continuous, real coordinates which are used for the unique and complete determination of the elements of an n-dimensional continuous space can be brought to*

any arbitrary number, and thus is not to be regarded as a unique feature of the space (Dauben 1979, p. 60).

Cantor shows how this result can be deduced from the existence of a one-one correspondence between the set of reals and the set of irrationals, and then, by means of an involved argument, constructs such a correspondence.

Cantor seems to have become convinced by this time that the essential nature of a continuum was fully reflected in the properties of sets of points – a conviction which was later to give birth to abstract set theory. In particular a continuum's key properties, Cantor believed, resided in the range of powers of its subsets of points. Since the power of a continuum of any number of dimensions is the same as that of a linear continuum, the essential properties of arbitrary continua were thereby reduced to those of a line. In his investigations of the linear continuum Cantor had found its infinite subsets to possess just two powers, that of the natural numbers and that of the linear continuum itself. This led him to the conviction that these were the only possible powers of such subsets – a thesis later to be enunciated as the *continuum hypothesis*. The question of the truth or falsity of the continuum hypothesis was to become one of the most celebrated problems of set theory. In the twentieth century its independence of the axioms of set theory was established by Kurt Gödel and Paul Cohen.

The problem of establishing the invariance of dimension of spaces under continuous correspondences remained a pressing issue. Soon after the publication of Cantor's 1878 paper, a number of mathematicians, for example, Lüroth, Thomae, Jürgens, and Netto attempted proofs, but all of these suffered from shortcomings which did not escape notice (Dauben 1979, pp. 70–72). In 1879 Cantor himself published a proof which seems to have passed muster at the time, but which also contained flaws that were not detected for another 20 years (Dauben 1979, pp. 72–76).

Satisfied that he had resolved the question of invariance of dimension, Cantor returned to his investigation of the properties of subsets of the linear continuum. The results of his labors are presented in six masterly papers published during 1879–84, *Über unendliche lineare Punktmannigfaltigkeiten* ("On infinite, linear point manifolds"). Remarkable in their richness of ideas, these papers contain the first accounts of Cantor's revolutionary theory of infinite sets and its application to the classification of subsets of the linear continuum. In the third and fifth of these are to be found Cantor's observations on the nature of the continuum.

In the third article, that of 1882, which is concerned with multidimensional spaces, Cantor applies his result on the nondenumerability of the continuum to prove the startling result that continuous motion is possible in discontinuous spaces. To be precise, he shows that, if M is any countable dense subset of the Euclidean plane R^2, (for example, the set of points with both coordinates algebraic real numbers), then any pair of points of the discontinuous space $\mathrm{A} = \mathrm{R}^2 - M$ can be joined by a continuous arc lying entirely within A. In fact, Cantor claims even more:

After all, with the same resources, it would be possible to connect the points…by a continuously running line given by a unique analytic rule and completely contained within the domain A.

(Spaces like A are today called *arcwise connected*.) Cantor points out that the belief in the continuity of space is traditionally based on the evidence of continuous motion, but now it has been shown that continuous motion is possible even in discontinuous spaces. That being the case, the presumed continuity of space is no more than a hypothesis. Indeed, it cannot necessarily be assumed that physical space contains every point given by three real number coordinates. This assumption, he urges,

> *. . .must be regarded as a free act of our constructive mental activity. The hypothesis of the continuity of space is therefore nothing but the assumption, arbitrary in itself, of the complete, one-to-one correspondence between the 3-dimensional purely arithmetic continuum* (*x, y, z*) *and the space underlying the world of phenomena* (Dauben 1979, pp. 72–76).

These facts, so much at variance with received views, confirmed for Cantor once again that geometric intuition was a poor guide to the understanding of the continuum. For such understanding to be attained reliance must instead be placed on arithmetical analysis.

In the fifth paper in the series, the *Grundlagen* of 1883, is to be found a forthright declaration of Cantor's philosophical principles, which leads on to an extensive discussion of the concept of the continuum. Cantor distinguishes between the *intrasubjective* or *immanent* reality and *transsubjective* or *transient* reality of concepts or ideas. The first type of reality, he says, is ascribable to a concept which may be regarded

> *as actual in so far as*, *on the basis of definitions*, *[it] is well distinguished from other parts of our thought and stand[s] to them in determinate relationships*, *and thus modifies the substance of our mind in a determinate way* (Ewald 1999, p. 895).

The second type, transient or transsubjective reality, is ascribable to a concept when it can, or must, be taken

> *as an expression or copy of the events and relationships in the external world which confronts the intellect.*

Cantor now presents the principal tenet of his philosophy, to wit, that the two sorts of reality he has identified invariably occur together,

> *in the sense that a concept designated in the first respect as existent always also possesses*, *in certain*, *even infinitely many ways*, *a transient reality.*

Cantor's thesis is tantamount to the principle that correct thinking is, in its essence, a reflection of the order of Nature (or, perhaps, vice-versa. In a footnote Cantor places his thesis in the context of the history of philosophy. He claims that "it agrees essentially both with the principles of the Platonic system and with an essential tendency of the Spinozistic system" and that it can be found also in Leibniz's philosophy. But philosophy since that time has come, in Cantor's eyes, to deviate from this cardinal principle:

> *Only since the growth of modern empiricism, sensualism, and scepticism, as well as of the Kantian criticism that grows out of them, have people believed that the source of knowledge and certainty is to be found in the senses or in the so-called pure form of intuition of the world of appearances, and that they must confine themselves to these. But in my opinion these elements do not furnish us with any secure knowledge. For this can be obtained only from concepts and ideas that are stimulated by external experience and are essentially formed by inner induction and deduction as something that, as it were, was already in us and is merely awakened and brought to consciousness.*

This linkage between the immanent and transient reality of mathematical concepts – the fact that correct mathematical thinking reflects objective reality – has, in Cantor's view, the important consequence that

> *mathematics, in the development of its ideas has only to take account of the immanent reality of its concepts and has absolutely no obligation to examine their transient reality.*

It follows that

> *mathematics in its development is entirely free and is only bound in the self-evident respect that its concepts must both be consistent with each other and also stand in exact relationships, ordered by definitions, to those concepts which have previously been introduced and are already at hand and established. In particular, in the introduction of new numbers it is only obliged to give definitions of them which will bestow such a determinacy and, in certain circumstances, such a relationship to older numbers that they can in any given instance be precisely distinguished.*

For these reasons Cantor bestows his blessing on rational, irrational, and complex numbers, which "one must regard as being every bit as existent as the finite positive integers"; even Kummer's introduction of "ideal" numbers into number theory meets with his approval. But not infinitesimal numbers, as we shall see.

Cantor begins his examination of the continuum with a tart summary of the controversies that have traditionally surrounded the notion:

> *The concept of the 'continuum' has not only played an important role everywhere in the development of the sciences but has also evoked the greatest differences of opinion and event vehement quarrels. This lies perhaps in the fact that, because the exact and complete definition of the concept has not been bequeathed to the dissentients, the underlying idea has taken on different meanings; but it must also be (and this seems to me the most probable) that the idea of the continuum had not been thought out by the Greeks (who may have been the first to conceive it) with the clarity and completeness which would have been required to exclude the possibility of different opinions among their posterity. Thus we see that Leucippus, Democritus, and Aristotle consider the continuum as a composite which consists ex partibus sine fine divisilibus, but Epicurus and Lucretius construct it out of their atoms considered as finite things. Out of this a great quarrel arose among the philosophers, of whom some followed Aristotle, others Epicurus; still others, in order to remain aloof from this quarrel, declared with Thomas Aquinas that the continuum consisted neither of infinitely many nor of a finite number of parts, but of absolutely no parts. This last opinion seems to me to contain less an explanation of the facts than a tacit confession that one has not got to the bottom of the matter and prefers to get genteely out of its way. Here we see the medieval-scholastic origin of a point of view which we still find represented today, in which the*

> *continuum is thought to be an unanalysable concept, or, as others express themselves, a pure a priori intuition which is scarcely susceptible to a determination through concepts. Every arithmetical attempt at determination of this mysterium is looked on as a forbidden encroachment and repulsed with due vigour. Timid natures thereby get the impression that with the 'continuum' it is not a matter of a mathematically logical concept but rather of religious dogma* (Ewald 1999, p. 903).

It is not Cantor's intention to "conjure up these controversial questions again." Rather, he is concerned to "develop the concept of the continuum as soberly and briefly as possible, and only with regard to the *mathematical* theory of sets." This opens the way, he believes, to the formulation of an exact concept of the continuum – nothing less than a demystification of the *mysterium.* Cantor points out that the idea of the continuum has heretofore merely been presupposed by mathematicians concerned with the analysis of continuous functions and the like and has "not been subjected to any more thorough inspection."

Cantor next repudiates any use of temporal intuition in an exact determination of the continuum:

> *. . .I must explain that in my opinion to bring in the concept of time or the intuition of time in discussing the much more fundamental and more general concept of the continuum is not the correct way to proceed; time is in my opinion a representation, and its clear explanation presupposes the concept of continuity upon which it depends and without whose assistance it cannot be conceived either objectively (as a substance) or subjectively (as the form of an a priori intuition), but is nothing other than a helping and linking concept, through which one ascertains the relation between various different motions that occur in nature and are perceived by us. Such a thing as objective or absolute time never occurs in nature, and therefore time cannot be regarded as the measure of motion; far rather motion as the measure of time – were it not that time, even in the modest role of a subjective necessary a priori form of intuition, has not been able to produce any fruitful, incontestable success, although since Kant the time for this has not been lacking.*

These strictures apply, *pari passu,* to spatial intuition:

> *It is likewise my conviction that with the so-called form of intuition of space one cannot even begin to acquire knowledge of the continuum. For only with the help of a conceptually already completed continuum do space and the structure thought into it receive that content with which they can become the object, not merely of aesthetic contemplation or philosophical cleverness or imprecise comparisons, but of sober and exact mathematical investigations.*

Cantor now embarks on the formulation of a precise arithmetical definition of a continuum. Making reference to the definition of real number he has already provided (i.e., in terms of fundamental sequences), he introduces the n-dimensional arithmetical space G_n as the set of all n-tuples of real numbers $(x_1 \, |x_2| \, \ldots \, |x_n)$, calling each such an *arithmetical point* of G_n. The distance between two such points is given by

$$\sqrt{\left(x_1' - x_1\right)^2 + \left(x_2' - x_2\right)^2 + \ldots \left(x_n' - x_n\right)^2}.$$

Cantor defines an *arithmetical* point-set in G_n to be any "aggregate of points of the points of the space G_n that is given in a lawlike way."

After remarking that he has previously shown that all spaces G_n have the same power as the set of real numbers in the interval (0, 1), and reiterating his conviction that any infinite point set has either the power of the set of natural numbers or that of (0, 1) – the continuum hypothesis – Cantor turns to the definition of the general concept of a continuum within *Gn.* For this he employs the concept of *derivative* or *derived set* of a point set introduced in his 1872 paper on trigonometric series. Cantor had defined the derived set of a point set *P* to be the set of *limit points* of *P,* where a limit point of *P* is a point of *P* with infinitely many points of *P* arbitrarily close to it. A point set is called *perfect* if it coincides with its derived set (in the terminology of general topology, a set is perfect if it is closed and has no isolated points). Cantor observes that this condition does not suffice to characterize a continuum, since perfect sets can be constructed in the linear continuum which are dense in no interval, however small: as an example of such a set he offers the set consisting of all real numbers in (0, 1) whose ternary expansion does not contain a "1." This set later became known as the *Cantor ternary set* or the *Cantor discontinuum.*

Accordingly, an additional condition is needed to define a continuum. Cantor supplies this by introducing the concept of a *connected* set. A point set *T* is connected in Cantor's sense if for any pair of its points *t, t'* and any arbitrarily small number ε there is a finite sequence of points $t_1, t_2, \ldots, t_n$ of *T* for which the distances $\overline{tt_1}, \overline{t_1t_2}, \overline{t_2t_3}, \ldots, \overline{t_nt'}$ are all less than ε. Cantor now observes:

> *all the geometric point-continua known to us fall under this concept of connected pointset, as it easy to see; I believe that in these two predicates 'perfect" and 'connected' I have discovered the necessary and sufficient properties of a point-continuum. I therefore define a point-continuum inside G_n as a perfect-connected set. Here 'perfect' and 'connected' are not merely words but completely general predicates of the continuum; they have been conceptually characterized in the sharpest way by the foregoing definitions* (Ewald 1999, p. 906).

Cantor points out the shortcomings of previous definitions of continuum such as those of Bolzano and Dedekind, and in a note dilates on the merits of his own definition:

> *Observe that this definition of a continuum is free from every reference to that which is called the dimension of a continuous structure; the definition includes also continua that are composed of connected pieces of different dimensions, such as lines, surfaces, solids, etc. . ..I know very well that the word 'continuum' has previously not had a precise meaning in mathematics; so my definition will be judged by some as too narrow, by others as too broad. I trust that that I have succeeded in finding a proper mean between the two.*
>
> *In my opinion, a continuum can only be a perfect and connected structure. So, for example, a straight line segment lacking one or both of its end-points, or a disc whose boundary is excluded, are not complete continua; I call such point-sets semi-continua.*

It will be seen that Cantor has advanced beyond his predecessors in formulating what is in essence a *topological* definition of continuum, one that, while still dependent on metric notions, does not involve an order relation. It is interesting to

compare Cantor's definition with that given in modern general topology. There we find a *continuum* defined as a compact connected subset of a topological space. Now within any *bounded* region of Euclidean space it can be shown that Cantor's continua coincide with continua in the sense of the modern definition. While Cantor lacked the definition of compactness, his requirement that continua be "complete" (which led to his rejecting as continua such noncompact sets as open intervals or discs) is not far away from the idea.

Cantor's analysis of infinite point sets had led him to introduce transfinite numbers (Hallett 1984) and he had come to accept their objective existence as being beyond doubt. But throughout his mathematical career he maintained an unwavering, even dogmatic opposition to infinitesimals, attacking the efforts of mathematicians such as du Bois-Reymond and Veronese to formulate rigorous theories of actual infinitesimals. As far as Cantor was concerned, the infinitesimal was beyond the realm of the possible; infiinitesimals were no more than "castles in the air, or rather just nonsense," to be classed "with circular squares and square circles" (Fisher 1981). His abhorrence of infinitesimals went so deep as to move him to outright vilification, branding them as "cholera-bacilli of mathematics" (Dauben 1979, p. 233).

Cantor believed that the theory of transfinite numbers could be employed to explode the concept of infinitesimal once and for all. Cantor's specific aim was to refute all attempts at introducing infinitesimals through the abandoning of the *Archimedean principle* – i.e., the assertion that for any positive real numbers $a < b$, there is a sufficiently large natural number n such that $na > b$. (Domains in which this principle fails to hold are called *nonarchimedean*). In a paper of 1887, Cantor attempted to demonstrate that the Archimedean property was a necessary consequence of the "concept of linear quantity" and "certain theorems of transfinite number theory," so that the linear continuum could contain no infinitesimals. He concludes that "the so-called Archimedean axiom is not an axiom at all, but a theorem which follows with logical necessity from the concept of linear quantity (Fisher 1981). Cantor's argument (in amended form) seems to have convinced a number of influential mathematicians, including Peano and Russell, of the untenability of infinitesimals.

Cantor's argument relied on the claim that the product of a positive infinitesimal, should such exist, with one of his transfinite numbers could never b finite. But since a proof of this claim was not supplied, Cantor's alleged demonstration that infinitesimals are impossible must be regarded as inconclusive (Fisher 1981). In any case consistent theories of infinitesimals were later constructed (see Sects. 5 and 6 below).

Cantor's rejection of infinitesimals stemmed from his conviction that his own theory of transfinite ordinal and cardinal numbers exhausted the realm of the numerable, so that no further generalization of the concept of number, in particular any which embraced infinitesimals, was admissible. For Cantor, transfinite numbers were grounded in *transient reality*, while infinitesimals and similar chimeras (in his estimation) could not be accorded such a status. Recall Cantor's assertion:

> *In particular, in the introduction of new numbers it is only obliged to give definitions of them which will bestow such a determinacy and, in certain circumstances, such a relationship to older numbers that they can in any given instance be precisely distinguished.*

Accordingly Cantor could not, or would not, grant the infinitesimal an immanent reality which was compatible with "older" numbers – among which he of course included his transfinite numbers – for had he done so he would perforce have had, in accordance with his own principles, to grant the infinitesimal transient reality. This seems to be the reason for Cantor's determination to demonstrate the inconsistency of the infinitesimal with his concept of transfinite number.

It has to be said (Fisher 1981) that concerning infinitesimals Cantor displayed a dogmatic attitude, arguing, in effect, that the admission of infinitesimals was self-contradictory. While Cantor's intolerant attitude toward the infinitesimal is not, strictly speaking, inconsistent with the "freedom" of mathematics he proclaimed, it does seem to reflect his deep-seated conviction (Dauben 1979, pp. 288–91) that his transfinite set theory was the product of "divine inspiration," so that anything in conflict with it must be anathematized.

3 Dissenting Voices: Divergent Conceptions of the Continuum in the Nineteenth and Early Twentieth Centuries

Despite the remarkable success of Weierstrass, Dedekind, and Cantor in constructing the continuum from arithmetical materials, a number of thinkers of the late nineteenth and early twentieth centuries remained opposed, in varying degrees, to the idea of explicating the continuum concept entirely in discrete terms, whether arithmetical or set-theoretic. These include the mathematicians *du Bois-Reymond, Poincaré, Brouwer,* and *Weyl,* and the philosophers *Brentano* and *Peirce.*

Paul du Bois-Reymond (1831–1889), against whose theory of infinities and infinitesimals Cantor fought so hard, was a prominent mathematician of the later nineteenth century who made significant contributions to real analysis, differential equations, mathematical physics, and the foundations of mathematics. While accepting many of the methods of the Dedekind-Cantor school, and indeed embracing the idea of the actual infinite, he rejected its associated philosophy of the continuum on the grounds that it was committed to the reduction of the continuous to the discrete. In 1882 he writes:

> *The conception of space as static and unchanging can never generate the notion of a sharply defined, uniform line from a series of points however dense, for, after all, points are devoid of size, and hence no matter how dense a series of points may be, it can never become an interval, which must always be regarded as the sum of intervals between points* (Ehrlich 1994, p. x).

Du Bois-Reymond took a somewhat mystical view of the continuum, asserting that its true nature, being beyond the limits of human cognition, would forever elude the understanding of mathematicians. This echoes his older brother Emile's famous *ignorabimus* declaration in 1880 concerning "world riddles," certain of which, such as the ultimate nature of matter and force, and the origin of sensations, would always elude explanation.

Nevertheless, Du Bois-Reymond's view of the continuum as a *mysterium* did not prevent him from developing his own theory of the mathematical continuum, a *continuum of functions,* during the 1870s and 1880s. This was introduced in an article of 1870–1871 as the *calculus of infinities.* Here du Bois-Reymond considers "functions ordered according to the limit of their quotients" (this and all subsequent quotations of Du Bois-Reymond are taken from Fisher 1981).

The orderings of functions, in du Bois-Reymond's notation,

$$f(x) \succ \varphi(x), f(x) \approx \varphi(x), f(x) \prec \varphi(x),$$

are defined, respectively, by

$$\lim_{x\to\infty} f(x)/\varphi(x) = \infty, \lim_{x\to\infty} f(x)/\varphi(x) \text{ is finite and } \neq 0, \lim_{x\to\infty} f(x)/\varphi(x) = 0.$$

Thus, for example, $e^x \succ x \succ \log(x)$, $x^p \succ x$ for any $p > 1$, while $cx^r \approx x^r$ for any c and r. When $f(x) \succ \varphi(x)$, $f(x)$ is said to have an "infinity greater than $\varphi(x)$"; when $f(x) \prec \varphi(x)$, $f(x)$ may be thought of (although du Bois-Reymond does not say this explicitly) as being infinitesimal in comparison with $\varphi(x)$. Du Bois-Reymond considers sequences of functions linearly ordered under $\prec$ or $\succ$. Such "scales of infinity" can be caused to become arbitrarily complex by the continued interpolation of new such sequences between terms. Du Bois-Reymond draws an analogy with the ordered set of real numbers:

> *Just as between two functions two functions as close with respect to their infinities as one may want, one can imagine an infinity of others forming a kind of passage from the first function to the second, one can compare the sequence F [a scale of infinity] to the sequence of real numbers, in which one can also pass from one number to a number very little different from it by an infinity of other ones.*

While du Bois-Reymond uses the term "infinities" in connection with his classification of functions, he does not at this point speak of infinite numbers or actual infinities. But in an article of 1875 he drops his reservations on the matter, and boldly begins by asserting:

> *I decided to publish this continuation of my research on functions becoming infinite in German after I overcame my aversion to using the word 'infinite* (unendlich)' *as a substantive, like the French their 'infini'. I even flatter myself that, by this 'infinite* (unendlich)', *I have enriched our mathematical vocabulary in a noteworthy way.*

He goes on to say:

> *In earlier articles I have distinguished the different infinities of functions by their different magnitudes so that they form a domain of quantities (the infinitary) with the stipulation that the infinity of $\varphi(x)$ is to be regarded as larger than that of $\psi(x)$ or equal to it. . .according as the quotient $\varphi(x)/\psi(x)$ is infinite or finite. Thus in the infinitary domain of quantities the quotient enters in place of the difference in the ordinary domain of numbers. Between the two domains there are many analogies. . . I can add further that the most complete symmetry*

exists between functions becoming zero and becoming infinity, in such a way that everywhere the positive numbers correspond in the most striking way to becoming infinity, the negative numbers to becoming zero, zero to remaining finite. Instead of numbers as fixed signs in the domain of numbers, one has in the infinitary domain of quantities an unlimited number of simple functions; the exponential functions, the powers, the logarithmic functions, that likewise form fixed points of comparisons, and between whose arbitrarily close infinities a limitless number of infinities different from each other can be inserted.

In a paper of 1877 du Bois-Reymond compares his system of "infinities" and that of "ordinary" numbers. He introduces the concept of "numerical continuity," an idea which he suggests underlies the introduction of irrational numbers. To illustrate the idea, du Bois-Reymond offers as a metaphor the distribution of the stars on a great circle in the sky. The readily identified brighter stars he compares to rational numbers with small numerators and denominators. Use of telescopes reveals the presence of new stars in any region, however small, but patches of darkness are always found between them. And then

our imagination, or speculation, peoples this as it were asymptotically uniform nothingness which always remains, with matter whose radiation or our observation can no longer make accessible. In our thought, we may believe there is no end, and we admit no empty spot in the sky.

This is analogous to the generation of rational and irrational numbers:

Thus through more precise consideration the rational numbers always approach more closely to one another, yet in our minds gaps are always left between them, which mathematical speculation then fills with the irrationals.

According to du Bois-Reymond this is essentially the way in which "numerical continuity" has arisen. He sees mathematical intuition as assigning equal authenticity to geometric and numerical quantity, but the attainment of complete equality between the two can only be attained through the use of the limit concept in introducing the irrationals. And the insertion of the irrationals between the rationals is an extension of the primitive concept of number to an equally primitive, but more comprehensive, concept of continuous quantity. That being the case, the comparison between numerical and geometric quantities may conceal further subtleties.

One such subtlety is brought to light in connection with continuous families of curves. When these are allowed to increase with growing rapidity their approximative behavior is quite different from that associated with ordinary spatial continuity. Du Bois-Reymond writes:

If we think of two different quickly increasing functions, then all the transitions from one to the other are spatially conceivable and present in our minds. We cannot conceive anywhere a gap between two curves increasing to infinity or in the neighbourhood of one such curve, which could not be filled with curves; on the contrary, each curve is accompanied by curves which proceed arbitrarily close to it, to infinity.

Now, unlike the points on a line segment, the curves which run between two such curves do not form a "simple infinity," that is, they do not depend on just a single parameter. Du Bois-Reymond shows that this infinity is "unlimited" in the sense that it is not n-fold for any finite n. He continues:

> *. . .just as in the ordinary domain of quantities we can only express quantities numerically exactly by means of rational numbers, since the other numbers are not actual numbers but only limits of such numbers: so we can only express infinities with well-defined functions, of which we only have at our disposal up to now those belonging to the family of logarithms, powers, exponential functions.*

Du Bois-Reymond next notes the difference between the approximative behavior of real numbers and that of "infinities" associated with functions. While one can approximate a number, say ½, by many sequences in such a way that any number, however close to ½, will fall between two members of any such sequence, the situation is quite different for the functions associated with infinities. For example, consider the sequence of functions

$$x^{1/2}, x^{2/3}, \ldots, x^{p/p+1}, \ldots.$$

The exponents of the members of this sequence approach 1, but it is not hard to establish the existence of functions whose infinities fall between *all* of the infinities of the members of this sequence and the function x to which the sequence converges in an appropriate sense. For example, $x^{\log\log x/\log\log x+1}$ is readily shown to be such a function.

Du Bois-Reymond next proceeds to demonstrate the generality of this phenomenon:

> *One cannot approximate a given infinity $\lambda(x)$ with any sequence of functions $\varphi_p(x)$, $p = 1, 2, \ldots$ in such a way that one could not always specify a function $\psi(x)$ which satisfies for arbitrarily large values of p $\lambda(x) \gtrless \psi(x) \gtrless \varphi_p(x)$.*

He continues:

> *Now the fact that we can with no conceivable sequence of functions approach without limit a given infinity, certainly has something strange about it. For it would be . . . completely counter to our intuition to suppose that there is necessarily a gap, for example around the line $y = x$. We can always fill this gap in our thoughts with curves which accompany the line $y = x$ to infinity.*

However, he finds here

> *no irreconcilable conflict of the results of different forms of thought, but only one of the idea of a perhaps not very familiar but still not inaccessible spatial behaviour.*

For du Bois-Reymond this only indicates the presence of "a gap between in the analogy between ordinary and infinitary quantities," the manifestation of "a behaviour peculiar to the infinitary domain."

In his book *Die allgemeine Functionentheorie* of 1882 du Bois-Reymond presents his views on the nature and existence of infinitesimals. He begins by stating that in the analysis of "continuous mathematical quantities," one begins with a "geometric quantity" and tries to relate other quantities to it. So the finite decimals are assigned correlates on a segment, that is, "points." This correlation between finite decimals and points is then extended to infinite decimals by a limit process. But the totality of such points can never form a complete segment, since

> *points are just dimensionless, and therefore an arbitrarily dense sequence of points can never become a distance.*

Here we see once again a rejection of the idea that the continuous is reducible to the discrete. Consequently, a geometric segment must contain something other than finite and infinite decimals. These "others," according to du Bois-Reymond, are infinitesimal segments: there are infinitely many of these in any line segment, however short.

Du Bois-Reymond provides just a few rules of calculation for infinitesimal segments, reminiscent of those used by seventeenth and eighteenth mathematicians. To wit:

> *A finite number of infinitely small segments joined to one another do not form a finite segment, but again an infinitely small segment . . . no upper bound can be specified either for the finite or for the infinitely small.*

> *I say two finite segments are equal when there is no* finite *difference between them . . . Two finite quantities whose difference is infinitely small are equal to one another . . .* A *finite quantity does not change if an infinitely small quantity is added to it or taken away from it.*

While we may be incapable of forming a mental image of the *relation* of the infinitesimal to the finite, according to du Bois-Reymond we can visualize the infinitesimal *in itself*, and when we do so we find that it behaves just like the finite:

> *The infinitely small is a mathematical quantity and has all its properties in common with the finite.*

Du Bois-Reymond took a dim view of some mathematicians' conception of infinitesimals as being ordinary magnitudes continually in a state of flux toward zero, remarking sarcastically

> *As long as the book is closed there is perfect repose, but as soon as I open it there commences a race of all the magnitudes which are provided with the letter d towards the zero limit.* (Ehrlich 1994, pp. 9–10.)

The admission of the infinitesimal in relation to the *finite* opens the way to the infinitesimal in relation to the *infinitesimal* (so entailing, reciprocally, the presence of the infinitely large):

> *In this way, there arises a series of types of quantities, whose successive relation always is that a finite number of quantities of one kind never yields a quantity of the preceding kinds.*

Such quantities accordingly form a nonarchimedean domain. Moreover,

> *If within one and the same of these types of quantities, the properties of ordinary mathematical quantities hold, hence the same types of calculation as in the finite, then the comparison of the different types of quantities with each other is the object of the so-called infinitary* calculus. *This calculus reckons with the relations of the infinitely large or infinitely small from type to type, and these types show connections with each other that do not fall under the ordinary concept of equality. The passages of one type into another do not show, for example, the continuity of change of mathematical quantities, although no jump changes result.*

Du Bois-Reymond concludes his musings on the infinitesimal with the observation that there is an imbalance between belief in the infinitely large and belief in the infinitely small. A majority of educated people, he says, will admit an "infinite" (i.e., actual infinite) in space and time, and not just an "unboundedly large" (i.e., potential infinite). But only with difficulty will they accept the infinitely small, despite the fact that it has the same "right to existence" as the infinitely large. In sum,

> *A belief in the infinitely small does not triumph easily. Yet when one thinks boldly and freely, the initial mistrust will soon mellow into a pleasant certainty. . .. Were the sight of the starry sky lacking to mankind; had the race arisen and developed troglodytically in enclosed spaces; had its scholars, instead of wandering through the distant places of the universe telescopically, only looked for the smallest constituents of form and so were used in their thoughts to advancing into the boundless in the direction of the unmeasurably small: who would doubt that then the infinitely small would take the same place in our system of concepts that the infinitely large does now? Moreover, hasn't the attempt in mechanics to go back down to the smallest active elements long ago introduced into science the atom, the embodiment of the infinitely small? And don't as always skilful attempts to make it superfluous for physics face with certainty the same fate as Lagrange's battle against the differential?*

Du Bois-Reymond was, indeed, a doughty champion of the infinitesimal.

In his later years the Austrian philosopher **Franz Brentano** (1838–1917) became preoccupied with the nature of the continuous. Much of Brentano's philosophy has its starting point in Aristotelian doctrine, and his conception of the continuum constitutes no exception. Aristotle's theory of the continuum, it will be recalled, rests upon the assumption that all change is continuous and that continuous variation of quality, of quantity, and of position are inherent features of perception and intuition. Aristotle considered it self-evident that a continuum cannot consist of points. Any pair of unextended points, he observes, are such that they either touch or are totally separated: in the first case, they yield just a single unextended point, in the second there is a definite gap between the points. Aristotle held that any continuum – a continuous path, say, or a temporal duration, or a motion – may be divided ad infinitum into other continua but not into what might be called "discreta" – parts that cannot themselves be further subdivided. Accordingly, paths may be divided into shorter

paths, but not into unextended points; durations into briefer durations but not into unextended instants; motions into smaller motions but not into unextended "stations." Nevertheless, this does not prevent a continuous line from being divided at a point constituting the common border of the line segments it divides. But such points are, according to Aristotle, just *boundaries,* and not to be regarded as actual *parts* of the continuum from which they spring. If two continua have a common boundary, that common border unites them into a single continuum. Such boundaries exist only *potentially,* since they come into being when they are, so to speak, marked out as connecting parts of a continuum; and the parts in their turn are similarly dependent as parts upon the existence of the continuum.

In its fundamentals Brentano's account of the continuous is akin to Aristotle's. Brentano regards continuity as something given in perception, primordial in nature, rather than a mathematical construction. He held that the idea of the continuous is a fundamental notion abstracted from sensible intuition:

> *Thus I affirm that. . . the concept of the continuous is acquired not through combinations of marks taken from different intuitions and experiences, but through abstraction from unitary intuitions. . .Every single one of our intuitions – both those of outer perception as also their accompaniments in inner perception, and therefore also those of memory – bring to appearance what is continuous* (Brentano 1988, p. 6).

Brentano suggests that the continuous is brought to appearance by sensible intuition in three stages. First, sensation presents us with objects having parts that coincide. From such objects the concept of *boundary* is abstracted in turn, and then one grasps that these objects actually *contain* coincident boundaries. Finally, it becomes clear that this is all that is required in order to understand the concept of a continuum.

Continuity is manifested in sensation in a variety of ways. In *visual* sensation, we are presented with *extension*, something possessing length and breadth, and hence with something such that between any two of its parts, provided these are separated, there is a third part. Every sensation possesses a certain *qualitative* continuity in that the object presented in the sensation could have a given manifested quality (color, for example) in a greater or less degree, and between any two degrees of that quality lies still another degree of that quality. Finally, each sensation manifests *temporal* continuity: this is most evident when we perceive something as moving or at rest.

Brentano recognizes that continua have *qualities* which cause them to possess *multiplicity*—a continuum may manifest continuity in several ways simultaneously. This led him to classify continua into *primary* and *secondary:* a secondary continuum being one whose manifestation is dependent upon another continuum. Here is Brentano himself on the matter:

> *Imagine, for example, a coloured surface. Its colour is something from which the geometer abstracts. For him there comes into consideration only the constantly changing manifold of spatial differences. But the colour, too, appears extended with the spatial surface, whether it manifests no specific colour-differences of its own – as in the case of a red colour which fills out a surface uniformly – or whether it varies in its colouring – perhaps in the manner of a*

> *rectangle which begins on one side with red and ends on the other side with blue, progressing uniformly through all colour-differences from violet to pure blue in between. In both cases we have to do with a multiple continuum, and it is the spatial continuum which appears thereby as primary, the colour-continuum as secondary.* A *similar double continuum can also be established in the case of a motion from place to place or of a rest, in which case it is a temporal continuum as such that is primary, the temporally constant or varying place that is the secondary continuum. Even when one considers a boundary of a mathematical body as such, for example a curved or straight line, a double continuity can be distinguished. The one presents itself in the totality of the differences of place that are given in the line, which always grows uniformly, whether in the case of straight, bent, or curved lines, and is that which determines the length of the line. The other resides in the direction of the line, and is either constant or alternating, and may vary continuously, or now more strongly, now less. It is constant in the case of the straight line, changing in the case of the broken line, and continuously varying in every line that is more or less curved. The direction-continuum here is to be compared with the colour-continuum discussed earlier and with the continuum of place in the case of rest or motion of a corporeal point in time. In the double continuum that presents itself to us in the line it is this continuum of directions that is to be referred to as the secondary, the manifold of differences of place as such as the primary continuum* (Brentano 1988, p. 21f).

It is worth mentioning that Brentano's distinction of primary and secondary continua can be neatly represented within *category theory*: to put it succinctly, a *primary continuum* is a *domain,* a *secondary continuum* a *codomain.* We form a category C – the *category of continua* – by taking continua as objects and *correlations* between continua as arrows. Then, given any arrow $f: A \rightarrow B$ in C, the domain A of f may be taken as a "primary" continuum and its codomain B as a "secondary" continuum. In Brentano's example of a colored surface, for instance, the primary continuum A is the given spatial surface, the secondary continuum B is the color spectrum, and the correlation f assigns to each place in A its color as a position in B. In the case of a corporeal point moving in space, the primary continuum A is an interval of time, the secondary continuum B a region of space, and the correlation f assigns to each instant in A the position in B occupied by the corporeal point. Finally, in the case of the varying direction of a curve the primary continuum A is the curve itself, the secondary continuum is the continuum of measures of angles, and the correlation f assigns to each point on the curve the slope of the tangent there: thus f is nothing other than the *first derivative* of the function associated with the curve.

For Brentano the essential feature of a continuum is its inherent capacity to engender boundaries, and the fact that such boundaries can be grasped as coincident. Boundaries themselves possess a quality which Brentano calls *plerosis* ("fullness"). Plerosis is the measure of the number of directions in which the given boundary actually bounds. Thus, for example, within a temporal continuum the endpoint of a past episode or the starting point of a future one bounds in a single direction, while the point marking the end of one episode and the beginning of another may be said to bound doubly. In the case of a spatial continuum there are numerous additional possibilities: here a boundary may bound in all the directions of which it is capable of bounding, or it may bound in only some of these directions. In the former case, the boundary is said to exist in *full plerosis;* in the latter, in *partial plerosis.* Brentano writes:

> *. . .the spatial nature of a point differs according to whether it serves as a limit in all or only in some directions. Thus a point located inside a physical thing serves as a limit in all directions, but a point on a surface or an edge or a vertex serves as a limit in only some direction. And the point in a vertex will differ in accordance with the directions of the edges that meet at the vertex. . . I call these specific distinctions differences of plerosis. Like any manifold variation, plerosis admits of a more and a less. The plerosis of the centre of a cone is more complete than that of a point on its surface; the plerosis of a point on its surface is more complete than that of a point on its edge, or that of its vertex. Even the plerosis of the vertex is the more complete the less the cone is pointed.* (Brentano 1988, p. xvii).

Brentano believed that the concept of plerosis enabled sense to be made of the idea that a boundary possesses "parts," even when the boundary lacks dimensions altogether, as in the case of a point. Thus, while the present or "now" is, according to Brentano, temporally unextended and exists only as a boundary between past and future, it still possesses two "parts" or aspects: it is both the end of the past and the beginning of the future. It is worth mentioning that for Brentano it was not just the "now" that existed only as a boundary; since, like Aristotle he held that "existence" in the strict sense means "existence now," it necessarily followed that existing things exist only as boundaries of what has existed or of what will exist, or both.

Brentano ascribes particular importance to the fact that points in a continuum can *coincide.* On this matter he writes:

> *Various other thorough studies could be* made [on the continuum concept] *such as a study of the impossibility of adjacent points and the* possibility of coincident points, *which have, despite their coincidence, distinctness and full relative independence.* [This] *has been and is misunderstood in many ways. It is commonly believed that if four different-coloured quadrants of a circular area touch each other at its centre, the centre belongs to only one of the coloured surfaces and must be that colour only. Galileo's judgment on the matter was more correct; he expressed his interpretation by saying paradoxically that the centre of the circle has as many parts as its periphery. Here we will only give some indication of these studies by commenting that everything which arises in this connection* follows from the point's relativity as involves a continuum and the fact that it is essential for it to belong to a continuum. *Just as the possibility of the coincidence of different points is connected with that fact, so is the existence of a point in diverse or more or less perfect* plerosis. *All of this is overlooked even today by those who understand the continuum to be an actual infinite multiplicity and who believe that we get the concept not by abstraction from spatial and temporal intuitions but from the combination of fractions between numbers, such as between 0 and* 1 (Brentano 1974, p. 357).

Brentano's doctrines of plerosis and coincidence of points are well illustrated by applying them to the traditional philosophical problem of the initiation of motion: if a thing begins to move, is there a last moment of its being at rest or a first moment of its being in motion? The usual objection to the claim that both moments exist is that, if they did, there would be a time between the two moments, and at that time the thing could be said neither to be at rest nor to be in motion – in violation of the law of excluded middle. Brentano's response would be to say that both moments do exist, but that they *coincide*, so that there are no times between them; the violation of the law of excluded middle is thereby avoided. More exactly, Brentano would assert that

the temporal boundary of the thing's being at rest – the end of its being at rest – is the same as the temporal boundary of the thing's being in motion – the beginning of its being in motion – but the boundary is *twofold* in respect of its plerosis. The boundary is, in fact, in half plerosis at rest and in half plerosis in motion.

Brentano took a dim view of the efforts of mathematicians to construct the continuum from numbers. His attitude varied from rejecting such attempts as inadequate to according them the status of "fictions." In a letter to Husserl drafted in 1905, Brentano asserts that "I regard it as absurd to interpret a continuum as a set of points." This is not surprising given his Aristotelian inclination to take mathematical and physical theories to be genuine descriptions of empirical phenomena rather than idealizations: in his view, if such theories were to be taken as literal descriptions of experience, they would amount to nothing better than "misrepresentations." Indeed, Brentano writes:

> *We must ask those who say that the continuum ultimately consists of points what they mean by a point. Many reply that a point is a cut which divides the continuum into two parts. The answer to this is that a cut cannot be called a thing and therefore cannot be a presentation in the strict and proper sense at all. We have, rather, only presentations of contiguous parts. . . . The spatial point cannot exist or* be conceived of *in isolation. It is just as necessary for it to belong to a spatial continuum as for the moment of time to belong to a temporal continuum* (Brentano 1974, p. 354).

Concerning Poincaré's approach to the continuum (see below) Brentano has this to say:

> *Poincaré . . .follows extreme empiricists in the in the area of sensory psychology and therefore does not believe that there is granted to us an intuition of a continuous space. Poincaré's entire mode of procedure reveals that he also denies that we are in possession of an intuition of a continuous time. We saw how first of all he inserted between 0 and 1 fractions having a whole number as numerator and a whole power of 2 as denominator. In similar fashion, he then inserted all proper fractions whose denominator is a whole power of 3, and then also all those whose denominators are powers of every other whole number. He obtained thereby a series containing all rational fractions which, as he said, already has a certain continuity about it. He then inserted . . . a series of irrational fractions. To these one now adds the series of fractions involving transcendental ratios. . .. Poincaré was prepared to admit that this process will never come to an end. . .. But he believed that he could be satisfied with the insertions already made. And nothing is more self-evident than that we have here a confession that the attempt to obtain a true continuum in this way has broken down* (Brentano 1988, p. 39).

Dedekind's account of the continuum does not fare much better (Brentano 1988, pp. 40–41):

> *Dedekind differs from Poincaré already in the fact that he does not wish to deny that we have an intuition of a continuum – he simply does not want to make any use thereof. . . . Dedekind's and Poincaré's constructions share in common that they fail to recognise the essential character of the continuum, namely that it allows the distinguishing of boundaries, which are nothing in themselves, but yet in conjunction make a contribution to the continuum. Dedekind believes that either the number ½ forms the beginning of the series ½ to 1, so*

> *that the series 0 to ½ would thereby be spared a final member, i.e. an end point which would belong to it, or conversely. But this is not how things are in the case of a true continuum. Rather it is the case that, when one divides a line, every part has a starting point, but in half plerosis. . . .*

Here Brentano appears to be saying that when one divides a closed interval $[a, b]$ at an intermediate point c, one necessarily obtains the closed intervals $[a, c]$, $[c, b]$, with the common point c (in half plerosis). In that case, Brentano probably has regarded a continuous line as indecomposable into disjoint intervals.

He goes on:

> *If a red and a blue surface are in contact with each other then a red and a blue line coincide, each with different plerosis. And if a circular area is made up of three sectors, a red, a blue and a yellow, then the mid-point is a whole which consists to an equal extent of a red, a blue and a yellow part. According to Dedekind this point would belong to just one of the three colour-segments, and we should have to say that it could be separated from this while the segment in question remained otherwise unchanged. Indeed, the whole circular surface would then be conceivable as having been deprived of its mid-point, like Dedekind's number-series from which only the number ½ has fallen away. One sees immediately that this is absurd if one keeps in mind that the true concept of the continuum is obtained through abstraction from an intuition, and thus also that the entire conception has missed its target.*

That Brentano considered "absurd" the idea of removing a single point from a continuum seems to indicate that his continuum has the same "syrupy" property as those of intuitionistic and smooth infinitesimal analysis: see Sects. 4 and 6 below.

In conclusion,

> *One sees that in this entire putative construction of the concept of what is continuous the goal has been entirely missed; for that which is above all else characteristic of a continuum, namely the idea of a boundary in the strict sense* (*to which belongs the possibility of a coincidence of boundaries*), *will be sought after entirely in vain. Thus also the attempt to have the concept of what is continuous spring forth out of the combination of individual marks distilled from intuition is to be rejected as entirely mistaken, and this implies further that what is continuous must be given to us in individual intuition and must therefore have been extracted therefrom* (Brentano 1988, pp. 4f).

Brentano's analysis of the continuum centered on its phenomenological and qualitative aspects, which are by their very nature incapable of reduction to the discrete. Brentano's rejection of the mathematicians' attempts to construct it in discrete terms is thus hardly surprising.

The American philosopher-mathematician **Charles Sanders Peirce**'s (1839–1914) view of the continuum was, in a sense, intermediate between that of Brentano and the arithmetizers. Like Brentano, he held that the cohesiveness of a continuum rules out the possibility of it being a mere collection of discrete individuals, or points, in the usual sense:

> *The very word* continuity *implies that the instants of time or the points of a line are everywhere welded together.*

> [*The*] *continuum does not consist of indivisibles, or points, or instants, and does not contain any except insofar as its continuity is ruptured* (Peirce 1976, p. 925).

And even before Brouwer (see below) Peirce seems to have been aware that a faithful account of the continuum will involve questioning the law of excluded middle:

> *Now if we are to accept the common idea of continuity . . . we must either say that a continuous line contains no points or . . . that the principle of excluded middle does not hold of these points. The principle of excluded middle applies only to an individual . . . but places being mere possibilities without actual existence are not individuals.* (Peirce 1976, p. xvi: the quotation is from a note written in 1903).

But Peirce also held that any continuum harbors an *unboundedly large* collection of points – in his colorful terminology, a *supermultitudinous* collection – what we would today call a *proper class*. Peirce maintained that if "enough" points were to be crowded together by carrying insertion of new points between old to its ultimate limit, they would – through a *logical* "transformation of quantity into quality" – lose their individual identity and become fused into a true continuum. Here are his observations on the matter:

It is substantially proved by Euclid that there is but one assignable quantity which is the limit of a convergent series. That is, if there is an increasing convergent series, A *say, and a decreasing convergent series,* B *say, of which every approximation exceeds every approximation of* A*, and if there is no rational quantity which is at once greater than every approximation of* A *and less than every approximation of* B*, then there is but one surd quantity so intermediate. . .There is one surd quantity and only one for each convergent series, calling two series the same if their approximations all agree after a sufficient number of terms, or if their difference approximates toward zero. But this is only to say that the multitude of surds equals the multitude of denumerable sets of rational numbers which is. . . the primipostnumeral multitude.* [Peirce assumed what amounts to the generalized continuum hypothesis in supposing that each possible infinite set has one of the cardinalities $\aleph_0, 2^{\aleph_0}, 2^{2^{\aleph_0}}, \ldots$. These he termed denumerable, primipostnumeral, secundipostnumeral, etc.]

> *. . .We remark that there is plenty of room to insert a* secundipostnumeral *multitude of quantities between* [*a*] *convergent series and its limit. Any one of those quantities may likewise be separated from its neighbours, and we thus see that between it and its nearest neighbours there is ample room for a* tertiopostnumeral *multitude of other quantities, and so on through the whole denumerable series of postnumeral quantities.*

> *But if we suppose that all such orders of systems of quantities have been inserted, there is no longer any room for inserting any more. For to do so we must select some quantity to be thus isolated in our representation. Now whatever one we take, there will always be quantities of higher order filling up the spaces on the two sides.*

> *We therefore see that such a supermultitudinous collection sticks together by logical necessity. Its constituent individuals are no longer distinct and independent subjects. They*

have no existence – no hypothetical existence – except in their relations to one another. They are not subjects, but phrases expressive of the properties of the continuum.

. . . Supposing a line to be a supermultitudinous collection of points, . . . to sever a line in the middle is to disrupt the logical identity of the point there and make it two points. It is impossible to sever a continuum by separating the connections of the points, for the points only exist by virtue of those connections. The only way to sever a continuum is to burst it, that is, to convert what was one into two (Peirce 1976, p. 95).

There is some resemblance between Peirce's conception of the continuum and John Conway's system of *surreal numbers* (Ehrlich 1994a). Conway's system may be characterized as being an η_α-field for every ordinal α, that is, a real-closed ordered field S which satisfies the condition that, for any pair of subsets X, Y for which every member of X is less than every member of Y, there is an element of S strictly between X and Y.

In their Introduction to Peirce (1992), Ketner and Putnam characterize Peirce's conception of the continuum as "a possibility of repeated division which can never be exhausted in any possible world, not even in a possible world in which one can complete [nondenumerably] infinite processes." (This description would seem to apply equally well to Conway's conception.) It is not hard to show that, between any pair of members of S there is a *proper class* of members of S – in Peirce's terminology, a supermultitudinous collection. Nevertheless, S is still discrete: its elements, while supermultitudinous, remain distinct and unfused (were it not for this fact, Conway would scarcely be justified in calling the members of S "numbers"). On the face of it the discreteness of S would seem to imply that the presence of superabundant quantity in Peirce's sense is not enough to ensure continuity. Of course, Brentano would have dismissed this idea altogether, in view of his critical attitude toward any construction of the continuum by repeated insertion of points.

Peirce's conception of the number continuum is also notable for the presence in it of an abundance of *infinitesimals*, a feature it shares with du Bois-Reymond's nonarchimedean number systems. In defending infinitesimals, Peirce remarks that

It is singular that nobody objects to $\sqrt{-1}$ *as involving any contradiction, nor, since Cantor, are infinitely great quantities much objected to, but still the antique prejudice against infinitely small quantities remains* (Peirce 1976, p. 123).

Peirce held the view that the conception of infinitesimal is suggested by introspection – that the *specious present* is in fact an infinitesimal:

It is difficult to explain the fact of memory and our apparently perceiving the flow of time, unless we suppose immediate consciousness to extend beyond a single instant. Yet if we make such a supposition we fall into grave difficulties, unless we suppose the time of which we are immediately conscious to be strictly infinitesimal (Peirce 1976, p. 124).

We are conscious of the present time, which is an instant, if there be any such thing as an instant. But in the present we are conscious of the flow of time. There is no flow in an instant. Hence, the present is not an instant. (Peirce 1976, p. 925).

Peirce also championed the retention of the infinitesimal concept in the foundations of the calculus, both because of what he saw as the efficiency of infinitesimal methods, and because he regarded infinitesimals as constituting the "glue" causing points on a continuous line to lose their individual identity.

The idea of continuity played a central role in the thought of the great French mathematician **Henri Poincaré** (1854–1912). But sorting out his views on the continuum, concerning which he made numerous scattered remarks, is by no means an easy task. Indeed, there seems to be an inconsistency in his attitude toward the set-theoretical, or arithmetized, continuum. On the one hand, he rejected actual infinity and impredicative definition – both cornerstones of the Cantorian theory of sets which underpins the construction of the arithmetized continuum. And yet in his mathematical work he employs variables ranging over all the points of an interval of the set-theoretical continuum, and he "accepts the standard account of the least upper bound, which is impredicative" (Folina 1992, p. xv). (Impredicativity is a form of circularity: a definition of a term is impredicative if it contains a reference to a totality to which the term under definition belongs. See, e.g., Fraenkel et al. 1973, pp. 193–200.)

Beneath this apparent inconsistency lies his belief that what ultimately underpins mathematics, creating its linkage with objective reality, is *intuition – that* "intuition is what bridges the gap between symbol and reality" (Folina 1992, p. 113).

His view of the continuum, in particular, is informed by this credo. For Poincaré the continuum and the range of points on it is grasped in intuition in something like the Kantian sense, and yet the continuum cannot be treated as a completed mathematical object, as a "mere set." (Folina 1992, p. xvi).

Of the arithmetical continuum Poincaré remarks:

> *The continuum so conceived is only a collection of individuals ranged in a certain order, infinite to one another, it is true, but* exterior *to one another. This is not the ordinary conception, wherein is supposed between the elements of the continuum a sort of intimate bond which makes of them a whole, where the point does not exist before the line, but the line before the point. Of the celebrated formula "the continuum is unity in multiplicity", only the multiplicity remains, the unity has disappeared. The analysts are none the less right in defining the continuum as they do, for they always reason on just this as soon as they pique themselves on their rigor. But this is enough to apprise us that the veritable mathematical continuum is a very different thing from that of the physicists and the metaphysicians.* (Poincaré 1946, pp. 43–44).

But despite Poincaré's apparent acceptance of the arithmetic definition of the continuum, he questions the fact that (as with Dedekind and Cantor's formulations) the (irrational) numbers so produced are mere symbols, detached from their origins in intuition:

> *But to be content with this [fact] would be to forget too far the origin of these symbols; it remains to explain how we have been led to attribute to them a sort of concrete existence, and, besides, does not the difficulty begin even for the fractional numbers themselves? Should we have the notion of these numbers if we had not known a matter that we conceive as infinitely divisible, that is to say, a continuum?* (Poincaré 1946, pp. 45–46).

That being the case, Poincaré asks whether the notion of the mathematical continuum is "simply drawn from experience." To this he responds in the negative, for the reason that our sensations, the "raw data of experience," cannot be brought under an acceptable scheme of measurement:

> *It has been observed, for example, that a weight* A *of 10 grams and a weight B of 11 grams produce identical sensations, that the weight B is just as indistinguishable from a weight C of 12 grams, but that the weight* A *is easily distinguished from the weight C. Thus the raw results of experience may be expressed by the following relations:*

$$A = B, B = C, A < C,$$

which may be regarded as the formula of the physical continuum.

According to Poincaré it is the "intolerable discord with the principle of contradiction" of this formula which has forced the invention of the mathematical continuum. (Actually, the formula ceases to be contradictory if the identity relation $=$ is replaced by a symmetric, reflexive, but *nontransitive* relation $\approx$: here $x \approx y$ is taken to assert that the sensations or perceptions x and y are indistinguishable.) The continuum is then obtained in two stages. First, formerly indistinguishable terms are distinguished and a new term, indistinguishable from both, inserted between them. Repeating this procedure indefinitely gives rise to what Poincaré calls a *first-order continuum*, in essence the rational number line. A second stage now becomes necessary because two first-order continua, for example, the diagonal of a square and its inscribed circle, need not intersect. This second stage, in which are added all possible "boundary" points between first-order continua leads to the second-order or mathematical continuum. Here is how Poincaré describes the process:

> *But conceive of a straight line divided into two rays. Each of these rays will appear to our imagination as a band of a certain breadth; these bands moreover will encroach one on the other, since there must be no interval between them. The common part will appear to us as a point which will always remain when we try to imagine our bands narrower and narrower, so that we admit as an intuitive truth that if a straight line is cut into two rays their common boundary is a point; we recognize here the conception of Dedekind, in which an incommensurable number was regarded as the common boundary of two classes of rational numbers.*
>
> *Such is the origin of the continuum of second order, which is the mathematical continuum so called.*

Poincaré goes on to discuss continua of higher dimensions. To obtain these he considers *aggregates* of sensations. As with single sensations, any given pair of these aggregates may or may not be distinguishable. He remarks that, while these aggregates, which he terms *elements*, are analogous to mathematical points, they are not in fact quite the same thing, for

> *we cannot say that our element is without extension, since we cannot distinguish it from neighbouring elements and it is thus surrounded by a sort of haze. If the astronomical*

> *comparison may be allowed, our 'elements' would be like nebulae, whereas the mathematical points would be like stars* (Poincaré 1946, p. 49).

This leads to a definition of a physical continuum:

> *a system of elements will form a* continuum *if we can pass from any one of them to any other, by a series of consecutive elements such that each is indistinguishable from the preceding. This* linear series *is to the* line *of the mathematician what an isolated* element *was to the point.*

Poincaré defines a *cut* in a physical continuum *C* to be a set of elements removed from it "which for an instant we shall regard as no longer belonging to this continuum." Such a cut may happen to subdivide *C* into several distinct continua, in which case *C* will contain two distinct elements *A* and *B* that must be regarded as belonging to two distinct continua. This becomes necessary

> *because it will be impossible to find a linear series of consecutive elements of C, each of these elements indistinguishable from the preceding, the first being A and the last B,* without one of the elements of this series being indistinguishable from one of the elements of the cut.

On the other hand, it may happen that the cut fails to subdivide the continuum C, in which case it becomes necessary to determine precisely which cuts will subdivide it. Poincaré calls a continuum *one-dimensional* if it can be subdivided by a cut which reduces to a finite number of elements all distinguishable from one another (and so forming neither a continuum nor several continua). When *C* can be subdivided only by cuts which are themselves continua, *C* is said to possess several dimensions:

> *If cuts which are continua of one dimension suffice, we shall say that C has two dimensions; if cuts of two dimensions suffice, we shall say that C has three dimensions, and so on.*

Thus is defined the concept of a multidimensional physical continuum, based on "the very simple fact that two aggregates of sensations are distinguishable or indistinguishable."

Unlike Cantor, Poincaré accepted the infinitesimal, even if he did not regard all of the concept's manifestations as useful. This emerges from his answer to the question: "Is the creative power of the mind exhausted by the creation of the mathematical continuum?" He responds:

> *No; the works of Du Bois-Reymond demonstrate it in a striking way. We know the mathematicians distinguish between infinitesimals and that those of second order are infinitesimal not only in an absolute way, but also in relation to those of first order. It is not difficult to imagine infinitesimals of fractional and even irrational order, and thus we find again that scale of the mathematical continuum which has been dealt with in the preceding pages.*
>
> *Further, there are infinitesimals which are infinitely small in relation to those of the first order, and, on the contrary, infinitely great in relation to those of order 1 + ε, and that*

> *however small ε may be. Here, then, are new terms intercalated in our series . . . I shall say that thus has been created a sort of continuum of the third order.*
>
> *It would be easy to go further, but that would be idle; one would only be imagining symbols without possible application, and no one would think of doing that. The continuum of the third order, to which the consideration of the different orders of infinitesimals leads, is itself not useful enough to have won citizenship, and geometers regard it as a mere curiosity. The mind uses its creative faculty only when experience requires it.*

Poincaré's attitude toward the continuum resembles in certain respects that of the intuitionists (see below): while the continuum exists, and is knowable intuitively, it is not a "completed" set-theoretical object. It is geometric intuition, not set theory, upon which the totality of real numbers is ultimately grounded.

The Dutch mathematician *L. E. J. Brouwer* (1881–1966) is best known as the founder of the philosophy of *(neo)intuitionism*. Brouwer's highly idealist views on mathematics bore some resemblance to Kant's. For Brouwer, mathematical concepts are admissible only if they are adequately grounded in intuition, mathematical theories are significant only if they concern entities which are constructed out of something given immediately in intuition, and mathematical demonstration is a form of construction in intuition. Brouwer's insistence that mathematical proof be constructive in this sense required the jettisoning of certain received principles of classical logic, notably the *law of excluded middle*: the assertion that, for any proposition *p,* either *p* or not *p*. Brouwer maintained, in fact, that the applicability of the law of excluded middle to mathematics

> *was caused historically by the fact that, first, classical logic was abstracted from the mathematics of the subsets of a definite finite set, that, secondly, an a priori existence independent of mathematics was ascribed to the logic, and that, finally, on the basis of this supposed apriority it was unjustifiably applied to the mathematics of infinite sets* (Kneebone 1963, p. 246).

Brouwer held that much of modern mathematics is based on an illicit extension of procedures valid only in the restricted domain of the finite. He therefore embarked on the radical course of jettisoning virtually all of the mathematics of his day – in particular the set-theoretical construction of the continuum – and starting anew, using only concepts and modes of inference that could be given clear intuitive justification. In the process it would become clear precisely what the logical laws are that intuitive, or constructive, mathematical reasoning actually obeys, making possible a comparison of the resulting *intuitionistic,* or *constructive logic* with classical logic. This is not to say that Brouwer was primarily interested in *logic,* far from it: indeed, his distaste for formalization caused him to be quite dismissive of subsequent codifications of intuitionistic logic.

While admitting that the emergence of noneuclidean geometry had discredited Kant's view of space, Brouwer maintained, in opposition to the logicists (whom he called "formalists") that arithmetic, and so all mathematics, must derive from *temporal intuition.* In his own words:

> *Neointuitionism considers the falling apart of moments of life into qualitatively different parts, to be reunited only while remaining separated by time, as the fundamental phenomenon of the human intellect, passing by abstracting from its emotional content into the fundamental phenomenon of mathematical thinking, the intuition of the bare two-oneness. This intuition of two-oneness, the basal intuition of mathematics, creates not only the numbers one and two, but also all finite ordinal numbers, inasmuch as one of the elements of the two-oneness may be thought of as a new two-oneness, which process may be repeated indefinitely; this gives rise still further to the smallest infinite ordinal ω. Finally this basal intuition of mathematics, in which the connected and the separate, the continuous and the discrete are united, gives rise immediately to the intuition of the linear continuum, i.e., of the "between", which is not exhaustible by the interposition of new units and which can therefore never be thought of as a mere collection of units. In this way the apriority of time does not only qualify the properties of arithmetic as synthetic a priori judgments, but it does the same for those of geometry, and not only for elementary two- and three-dimensional geometry, but for non-euclidean and n-dimensional geometries as well. For since Descartes we have learned to reduce all these geometries to arithmetic by means of coordinates.* (Benacerraf and Putnam 1977, p. 80).

Brouwer maintained that it is the awakening of awareness of the temporal continuum in the subject, an event termed by him "The Primordial Happening" or "The Primordial Intuition of Time," that engenders the fundamental concepts and methods of mathematics. In "Mathematics, Science and Language" (1929), he describes how the notion of number – the discrete – emerges from the awareness of the continuous:

> Mathematical Attention *as an act of the will serves the instinct for self-preservation of individual man; it comes into being in two phases;* time awareness *and* causal attention. *The first phase is nothing but the fundamental intellectual phenomenon of the falling apart of a moment of life into two qualitatively different things of which one is experienced as giving away to the other and yet is retained by an act of memory. At the same time this split moment of life is separated from the Ego and moved into a world of its own, the world of perception. Temporal twoity, born from this time awareness, or the two-membered sequence of time phenomena, can itself again be taken as one of the elements of a new twoity, so creating temporal threeity, and so on. In this way, by means of the self-unfolding of the fundamental phenomenon of the intellect, a* time sequence of phenomena is created of arbitrary multiplicity (Mancosu 1998, p. 45)

Brouwer took the elucidation of the continuum as his *raison d'être.* In his early thought he rejected out of hand the set-theoretic account of the continuum in terms of discrete points as put forward by Cantor and Dedekind, asserting that

> *. . .The* continuum as a whole [is] *given to us by intuition; a construction for it, an action which would create from the mathematical intuition 'all' its points as individuals, is inconceivable and impossible.* (Mancosu 1998, p.45)

Later Brouwer was to modify this doctrine. In his mature thought, he radically transformed the concept of point, endowing points with sufficient fluidity to enable them to serve as generators of a "true" continuum. This fluidity was achieved by admitting as "points," not only fully defined discrete numbers such as $\sqrt{2}$, π, *e,* and the like – which have, so to speak, already achieved "being" – but also "numbers"

which are in a perpetual state of becoming in that their the entries in their decimal (or dyadic) expansions are the result of free acts of choice by a subject operating throughout an indefinitely extended time. The resulting *choice sequences* cannot be conceived as finished, completed objects: at any moment only an initial segment of each is known (Fraenkel et al. 1973, pp. 255–261). In this way Brouwer obtained the mathematical continuum in a way compatible with his belief in the primordial intuition of time – that is, as an unfinished, indeed unfinishable entity in a perpetual state of growth, a "medium of free development." In this conception, the mathematical continuum is indeed "constructed," not, however, by initially shattering, as did Cantor and Dedekind, an intuitive continuum into isolated points, but rather by assembling it from a complex of continually changing overlapping parts.

The mathematical continuum as conceived by Brouwer displays a number of features that seem bizarre to the classical eye. For example, in the Brouwerian continuum the usual law of comparability, namely, that for any real numbers *a, b* either $a < b$ or $a = b$ or $a > b$, fails. Even more fundamental is the failure of the law of excluded middle in the form that for any real numbers *a, b,* either $a = b$ or $a \neq b$. The failure of these seemingly unquestionable principles in turn vitiates the proofs of a number of basic results of classical analysis, for example, the Bolzano–Weierstrass theorem, as well as the theorems of monotone convergence, intermediate value, least upper bound, and maximum value for continuous functions. The failure of these important results of classical analysis has caused most mathematicians of the day to shun intuitionistic, and even constructive mathematics. It was not until the 1960s that adequate constructive versions were worked out (see Sect. 4 below).

While the Brouwerian continuum may possess a number of negative features from the standpoint of the classical mathematician, it has the merit of corresponding more closely to the continuum of intuition than does its classical counterpart. Hermann Weyl pointed out a number of respects in which this is so:

> *In accordance with intuition, Brouwer sees the essential character of the continuum, not in the relation between element and set, but in that between part and whole. The continuum falls under the notion of the 'extensive whole', which Husserl characterizes as that "which permits a dismemberment of such a kind that the pieces are by their very nature of the same lowest species as is determined by the undivided whole* (Weyl 1949, p. 52).

Far from being bizarre, the failure of the law of excluded middle for points in the intuitionistic continuum is seen by Weyl as "fitting in well with the character of the intuitive continuum":

> *For there the separateness of two places, upon moving them toward each other, slowly and in vague gradations passes over into indiscernibility. In a continuum, according to Brouwer, there can be only continuous functions.* The continuum is not composed of parts (Weyl 1949, p. 54).

For Brouwer had indeed shown, in 1924, that every function defined on a closed interval of the continuum as he conceived it is uniformly continuous. Now one might be inclined to regard this claim as impossible: is not a counterexample provided by, for example, the function f given by $f(0) = 0, f(x) = \mid x \mid /x$ otherwise? No, because

from the intuitionistic standpoint, this function is not everywhere defined on the interval $[-1, 1]$, being undefined at those arguments x for which it is unknown whether $x = 0$ or $x \neq 0$.

As a consequence, the intuitionistic continuum is *indecomposable* or *cohesive,* that is, *it cannot be split into two nonempty disjoint parts in any way whatsoever.* In contrast with a discrete entity, the cohesive Brouwerian continuum cannot be composed of its parts.

Hermann Weyl (1885–1955), one of most important and versatile mathematicians of the twentieth century, devoted a great deal of thought to the nature of the continuum. In his *Philosophy of Mathematics and Natural Science* he reflects on what he calls the "inwardly infinite" nature of a continuum:

> *The essential character of the continuum is clearly described in this fragment of Anaxagoras: "Among the small there is no smallest, but always something smaller. For what is cannot cease to be no matter how small it is being subdivided." The continuum is not composed of discrete elements which are "separated from one another as though chopped off by a hatchet." Space is infinite not only in the sense that it never comes to an end; but at every place it is, so to speak, inwardly infinite, inasmuch as a point can only be fixed as step-by-step by a process of subdivision which progresses* ad infinitum. *This is in contrast with the resting and complete existence that intuition ascribes to space. The "open" character is communicated by the continuous space and the continuously graded qualities to the things of the external world. A real thing can never be given adequately, its "inner horizon" is unfolded by an infinitely continued process of ever new and more exact experiences; it is, as emphasized by Husserl, a limiting idea in the Kantian sense. For this reason it is impossible to posit the real thing as existing, closed and complete in itself. The continuum problem thus drives one to epistemological idealism. Leibniz, among others, testifies that it was the search for a way out of the "labyrinth of the continuum" which first suggested to him the conception of space and time as orders of phenomena* (Weyl 1949, p. 41).

Weyl identifies three attempts in the history of thought "to conceive of the continuum as Being in itself." These are, respectively, *atomism, the infinitely small,* and *set theory.* In Weyl's view, despite atomism's brilliant success in unraveling the structure of matter, it had failed in that regard as to space, time, and mathematical extension because it "never achieved sufficient contact with reality." As for the infinitely small, it was not so much supplanted as rendered superfluous by the limit concept. Weyl saw the limit concept as providing the necessary link between the microcosm of the infinitely small and the realm of macroscopic objects. Without that link, the fact that the microcosm is governed by "elementary laws" making for ease of calculation, would remain entirely useless in drawing conclusions about the macrocosm (but in this connection see the remarks on the Constancy Principle in Sect. 6).

In the case of set theory Weyl saw the danger in employing it a basis for explicating the continuum as lying in the fact that set theory bestows a finished, completed character on "the places in the continuum, i.e. to the possible sequences or sets of natural numbers." In Weyl's view, this was a double error for, neither the aggregate of sets of natural numbers nor (in general) individual such sets can be considered finished entities. Rather the continuum should be considered as an

essentially incompletable "field of constructive possibilities." To suppose otherwise is to risk running up against set-theoretic paradoxes such as Russell's.

During the period 1918–1921 Weyl wrestled with the problem of providing the continuum with an exact mathematical formulation free of objectionable set-theoretic assumptions. As he saw it in 1918, there is an unbridgeable gap between intuitively given continua (e.g., those of space, time, and motion) on the one hand, and the discrete exact concepts of mathematics (e.g., that of real number) on the other. For Weyl the presence of this gap meant that the construction of the mathematical continuum could not simply be "read off" from intuition. Rather, he believed at this time that the mathematical continuum must be treated as if it were an element of the realm immediately accessible to intuition, and so, in the end, justified in the same way as a physical theory. In Weyl's view, it was not enough that the mathematical theory of the continuum be *consistent*; it must also be *reasonable.*

Das Kontinuum (1918) embodies Weyl's attempt at formulating a theory of the continuum which satisfies the first, and, as far as possible, the second, of these requirements. In the following passages from this work he acknowledges the difficulty of the task:

> *. . . the conceptual world of mathematics is so foreign to what the intuitive continuum presents to us that the demand for coincidence between the two must be dismissed as absurd.*

> *. . . the continuity given to us immediately by intuition* (*in the flow of time and of motion*) *has yet to be grasped mathematically as a totality of discrete "stages" in accordance with that part of its content which can be conceptualized in an exact way.*

> *Exact time- or space-points are not the ultimate, underlying atomic elements of the duration or extension given to us in experience. On the contrary, only reason, which thoroughly penetrates what is experientially given, is able to grasp these exact ideas. And only in the arithmetico-analytic concept of the real number belonging to the purely formal sphere do these ideas crystallize into full definiteness.*

> *When our experience has turned into a real process in a real world and our phenomenal time has spread itself out over this world and assumed a cosmic dimension, we are not satisfied with replacing the continuum by the exact concept of the real number, in spite of the essential and undeniable inexactness arising from what is given.*

However much he may have wished it, in *Das Kontinuum* Weyl did not aim to provide a mathematical formulation of the continuum as it is presented to intuition, which, as the quotations above show, he regarded as an impossibility (at that time at least). Rather, his goal was first to achieve *consistency* by putting the *arithmetical* notion of real number on a firm logical basis, and then to show that the resulting theory is *reasonable* by employing it as the foundation for a plausible account of continuous process in the objective physical world.

Weyl had come to believe that mathematical analysis at the beginning of the twentieth century would not bear logical scrutiny, for its essential concepts and procedures involved vicious circles to such an extent that, as he says, "every cell (so to speak) of this mighty organism is permeated by contradiction." In *Das Kontinuum*

he tries to overcome this by providing analysis with a *predicative* formulation – not, as Russell and Whitehead had attempted in their *Principia Mathematica,* by introducing a hierarchy of logically ramified types, which Weyl seems to have regarded as too complicated – but rather by confining the basic principle of set formation to formulas whose bound variables range over just the initial given entities (numbers). Thus, he restricts analysis to what can be done in terms of natural numbers with the aid of three basic logical operations, together with the operation of substitution and the process of "iteration," i.e., primitive recursion. Weyl recognized that the effect of this restriction would be to render unprovable many of the central results of classical analysis – e.g., Dirichlet's principle that any bounded set of real numbers has a least upper bound – but he was prepared to accept this as part of the price that must be paid for the security of mathematics.

In Sect. 6 of *Das Kontinuum* Weyl presents his conclusions as to the relationship between the intuitive and mathematical continua. He poses the question: Does the mathematical framework he has erected provide an adequate representation of physical or temporal continuity as it is *actually experienced*? He begins his investigation by noting that, according to his theory, if one asks whether a given function is continuous, the answer is not fixed once and for all, but is, rather, dependent on the extent of the domain of real numbers which have been defined up to the point at which the question is posed. Thus the continuity of a function must always remain *provisional*; the possibility always exists that a function deemed continuous *now* may, with the emergence of "new" real numbers, turn out to be discontinuous *in the future.*

To reveal the discrepancy between this formal account of continuity based on real numbers and the properties of an intuitively given continuum, Weyl next considers the experience of seeing a pencil lying on a table before him throughout a certain time interval. The position of the pencil during this interval may be taken as a function of the time, and Weyl takes it as a fact of observation that during the time interval in question this function is continuous and that its values fall within a definite range. And so, he says,

> *This observation entitles me to assert that during a certain period this pencil was on the table; and even if my right to do so is not absolute, it is nevertheless reasonable and well-grounded. It is obviously absurd to suppose that this right can be undermined by "an expansion of our principles of definition" – as if new moments of time, overlooked by my intuition could be added to this interval, moments in which the pencil was, perhaps, in the vicinity of Sirius or who knows where. If the temporal continuum can be represented by a variable which "ranges over" the real numbers, then it appears to be determined thereby how narrowly or widely we must understand the concept "real number" and the decision about this must not be entrusted to* logical *deliberations over principles of definition and the like.*

To drive the point home, Weyl focuses attention on the fundamental continuum of *immediately given phenomenal time*, that is, as he characterizes it,

> *. . . to that constant form of my experiences of consciousness by virtue of which they appear to me to flow by successively.* (*By "experiences" I mean what I experience, exactly as I*

experience it. I do not mean real psychical or even physical processes which occur in a definite psychicsomatic individual, belong to a real world, and, perhaps, correspond to the direct experiences.)

In order to correlate mathematical concepts with phenomenal time in this sense Weyl grants the possibility of introducing a rigidly punctate "now" and of identifying and exhibiting the resulting temporal points. On the collection of these temporal points is defined the relation of *earlier than* as well as a congruence relation of *equality of temporal intervals,* the basic constituents of a simple mathematical theory of time. Now Weyl observes that the discrepancy between phenomenal time and the concept of real number would vanish if the following pair of conditions could be shown to be satisfied:

1. *The immediate expression of the intuitive finding that during a certain period I saw the pencil lying there were construed in such a way that the phrase "during a certain period" was replaced by "in every temporal point which falls within a certain time span OE.* [Weyl goes on to say parenthetically here that he admits "that this no longer reproduces what is intuitively present, but one will have to let it pass, *if it is really legitimate to dissolve a period into temporal points."*]
2. If P is a temporal point, then the domain of rational numbers to which l belongs if and only if there is a time point L earlier than P such that OL = l.OE *can be constructed arithmetically in pure number theory on the basis of our principles of definition, and is therefore a real number in our sense.*

Condition 3 means that, if we take the time span *OE* as a unit, then each temporal point *P* is correlated with a definite real number. In an addendum Weyl also stipulates the converse.

But can temporal intuition itself provide evidence for the truth or falsity of these two conditions? Weyl thinks not. In fact, he states unequivocally that

. . . everything we are demanding here is obvious nonsense: *to* these *questions, the intuition of time provides no answer – just as a man makes no reply to questions which clearly are addressed to him by mistake and, therefore, are unintelligible when addressed to him.*

The grounds for this assertion are by no means immediately evident, but one gathers from the passages following it that Weyl regards the experienced *continuous flow* of phenomenal time as constituting an insuperable barrier to the whole enterprise of representing this continuum in terms of individual points, and even to the characterization of "individual temporal point" itself. As he says,

The view of a flow consisting of points and, therefore, also dissolving into points turns out to be mistaken: precisely what eludes us is the nature of the continuity, the flowing from point to point; in other words, the secret of how the continually enduring present can continually slip away into the receding past. Each one of us, at every moment, directly experiences the true character of this temporal continuity. But, because of the genuine primitiveness of phenomenal time, we cannot put our experiences into words. So we shall content ourselves with the following description. What I am conscious of is for me both a being-now and, in its essence,

> *something which, with its temporal position, slips away. In this way there arises the persisting factual extent, something ever new which endures and changes in consciousness.*

Weyl sums up what he thinks can be affirmed about "objectively presented time" – by which I take it is meant "phenomenal time described in an objective manner" – in the following two assertions, which he claims apply equally, mutatis mutandis, to every intuitively given continuum, in particular, to the continuum of spatial extension:

> 1. *An individual point in it is non-independent,* i.e., *is pure nothingness when taken by itself, and exists only as a "point of transition" (which, of course, can in no way be understood mathematically);*
> 2. *it is due to the essence of time (and not to contingent imperfections in our medium) that a fixed temporal point cannot be exhibited in any way, that always only an approximate, never an exact determination is possible.*

The fact that single points in a true continuum "cannot be exhibited" arises, Weyl continues, from the fact that they are not genuine individuals and so cannot be characterized by their properties. In the physical world they are never defined absolutely, but only in terms of a coordinate system, which, in an arresting metaphor, Weyl describes as "the unavoidable residue of the eradication of the ego." This metaphor, which Weyl was to employ more than once (Weyl 1950, p. 8 and 1949, p. 123) reflects the continuing influence of Husserlian phenomenological doctrine: in this case, the thesis that the existent is given in the first instance as the contents of a consciousness.

By 1919 Weyl had come to embrace Brouwer's views on the intuitive continuum. The latter's influence looms large in Weyl's next paper on the subject, *On the New Foundational Crisis of Mathematics,* written in 1920. Here Weyl identifies two distinct views of the continuum: "atomistic" or "discrete"; and "continuous." In the first of these the continuum is composed of individual real numbers which are well-defined and can be sharply distinguished. Weyl describes his earlier attempt at reconstructing analysis in Das Kontinuum as atomistic in this sense:

> *Existential questions concerning real numbers only become meaningful if we analyze the concept of real number in this extensionally determining and delimiting manner. Through this conceptual restriction, an ensemble of individual points is, so to speak, picked out from the fluid paste of the continuum. The continuum is broken up into isolated elements, and the flowing-into-each other of its parts is replaced by certain conceptual relations between these elements, based on the "larger-smaller" relationship. This is why I speak of the* atomistic *conception of the continuum* (Weyl 1998, p. 91).

Weyl now repudiated atomistic theories of the continuum, including that of *Das Kontinuum*. He writes:

> *In traditional analysis, the continuum appeared as the set of its points; it was considered merely as a special case of the basic logical relationship of* element and set. *Who would not have already noticed that, up to now, there was no place in mathematics for the equally fundamental relationship of* part and whole? *The fact, however,* that it has parts, *is a fundamental property of the continuum; and so* (*in harmony with intuition, so drastically*

> *offended against by today's "atomism"*) *this relationship is taken as the mathematical basis for the continuum by Brouwer's theory. This is the real reason why the method used in delimiting subcontinua and in forming continuous functions starts out from* intervals *and not* points *as the primary elements of construction. Admittedly a set also has parts. Yet what distinguishes the parts of sets in the realm of the "divisible" is the existence of "elements" in the set-theoretical sense, that is, the existence of* parts that themselves do not contain any further parts. *And indeed, every part contains at least one "element". In contrast, it is inherent in the nature of the continuum that every part of it can be further divided without limitation. The concept of a point must be seen as an idea of a limit, "point" is the idea of a* limit *of a division extending* in infinitum. *To represent the continuous connection of the points, traditional analysis, given its shattering of the continuum into isolated points, had to have recourse to the concept of a* neighbourhood. *Yet, because the concept of continuous function remained mathematically sterile in the resulting generality, it became necessary to introduce the possibility of "triangulation" as a restrictive condition* (Weyl 1998, p. 115).

Like Brentano, Weyl knew that to "shatter a continuum into isolated points" would be to eradicate the very feature which characterizes a continuum – the fact that its cohesiveness is inherited by every one of its (connected) parts.

Weyl welcomed Brouwer's construction of the continuum by means of sequences generated by free acts of choice, thus identifying it as a "medium of free Becoming" which "does not dissolve into a set of real numbers as finished entities." Weyl felt that Brouwer's intuitionistic approach had brought him closer than anyone else to bridging that "unbridgeable chasm" between the intuitive and mathematical continua. In particular, he found compelling the fact that the Brouwerian continuum is not the union of two disjoint nonempty parts – that it is *indecomposable* or *cohesive.* "A genuine continuum," Weyl says, "cannot be divided into separate fragments." In 1921 Weyl observed:

> *. . .if we pick out a specific point, say, $x = 0$, on the number line C* (*i.e., on the variable range of a real variable x*), *then one cannot, under any circumstance, claim either coincides with it or is disjoint from it. The point $x = 0$ thus does not at all split the continuum C into two parts $C^-: x < 0$ and $C^+: x > 0$, in the sense that C would consist of the union of C^-, C^+ and the one point 0 . . . If this appears offensive to present-day mathematicians with their atomistic thought habits, it was in earlier times a self-evident view held by everyone: Within a continuum, one can very well generate subcontinua by introducing boundaries; yet it is irrational to claim that the total continuum is made up of the boundaries and the subcontinua. The point is, a genuine continuum is something connected in itself, and it cannot be divided into separate fragments; this conflicts with its nature* (Weyl 1921, p. 111).

In later publications he expresses this more colorfully by quoting Anaxagoras to the effect that a continuum "defies the chopping off of its parts with a hatchet."

There being only minor differences between Weyl's and Brouwer's accounts of the continuum, Weyl abandoned his earlier attempt at the reconstruction of analysis and "joined Brouwer." At the same time, however, Weyl recognized that the resulting gain in intuitive clarity had been bought at a considerable price:

> *Mathematics with Brouwer gains its highest intuitive clarity. He succeeds in developing the beginnings of analysis in a natural manner, all the time preserving the contact with intuition much more closely than had been done before. It cannot be denied, however, that in*

> *advancing to higher and more general theories the inapplicability of the simple laws of classical logic eventually results in an almost unbearable awkwardness. And the mathematician watches with pain the greater part of his towering edifice which he believed to be built of concrete blocks dissolve into mist before his eyes* (Weyl 1949, p. 54).

Although he later practiced intuitionistic mathematics very rarely, Weyl remained an admirer of intuitionism. And the "riddle of the continuum" retained its fascination for him, as is attested to by the observation he made one of his last papers, *Axiomatic and Constructive Procedures in Mathematics*, written in 1954,

> *. . . the constructive transition to the continuum of real numbers is a serious affair. . . and I am bold enough to say that not even to this day are the logical issues involved in that constructive concept completely clarified and settled* (Weyl 1985, p. 17).

4 The Continuum in Constructive and Intuitionistic Mathematics

As we have observed, Brouwer maintained that mathematical activity is a form of construction in intuition. If the philosophically loaded term "in intuition" is removed, we are left simply with the term "construction," a much more concrete term on whose meaning mathematicians have achieved some measure of agreement. This leads to the idea of *constructive mathematics*.

The central principle of constructive mathematics is that a problem is to be regarded as solved only if an explicit solution can, in principle at least, be produced. Thus, for example, "There is an x such that $P(x)$" *means* that, in principle at least, we can explicitly produce an x such that $P(x)$. If the solution to the problem involves parameters, we must be able to present the solution explicitly by means of some *algorithm* or *rule* when given values of the parameters. That is, "for every x there is a y such that $P(x, y)$ *means* that, we possess an explicit method of determining, for any given x, a y for which $P(x, y)$.

The practice of constructive mathematics requires, as Brouwer recognized, the abandoning of certain laws of classical logic. In constructive mathematical reasoning *an existential statement can be considered affirmed only when an instance is produced*, and *a disjunction can be considered affirmed only when an explicit one of the disjuncts is demonstrated*. Consequently, neither the classical law of excluded middle nor the law of strong reductio ad absurdum can be constructively admissible. Hermann Weyl said of nonconstructive existence proofs that "they inform the world that a treasure exists without disclosing its location."

Consider the existential statement *there exists an odd perfect number* (i.e., an odd number equal to the sum of its proper divisors) which we shall write as $\exists nP(n)$. Its contradictory is the statement $\forall n\neg P(n)$. Classically, the law of excluded middle then allows us to affirm the disjunction

$$\exists nP(n) \vee \forall n\neg P(n) \tag{1}$$

Constructively, however, in order to affirm this disjunction we must *either* be in a position to affirm the first disjunct $\exists n P(n)$, i.e., to possess, or have the means of obtaining, an odd perfect number, *or* to affirm the second disjunct $\forall n \neg P(n)$, i.e., to possess a demonstration that no odd number is perfect. Since at the present time mathematicians have neither of these, the disjunction (1), and *a fortiori* the law of excluded middle is not constructively admissible.

The logical principles underlying constructive reasoning have been codified into what is known as *constructiive* or *intuitionistic* logic. Roughly speaking, constructive/intuitionistic logic is the body of arguments of classical logic, which can be established without the use, explicitly or implicitly, of the law of excluded middle.

Intuitionistic logic guides how one should *reason* in constructive mathematics. But how are mathematical objects to be given, or, as we shall say, *specified*, in constructive mathematics? To begin with, everybody knows what it means to specify an *integer*. For example, $7 \bullet 10^4$ is specified, while the number n defined to be 0 if an odd perfect number exists, and 1 if an odd perfect number does not exist, is not specified. The number of primes less than, say, $10^{1000000}$ is specified, in the sense intended here, since we could, *in principle at least*, calculate this number. Constructive mathematics as we shall understand it is not concerned with questions of feasibility, nor in particular with what can actually be computed in real time by actual computers.

In *constructive real analysis* (Bishop and Bridges 1985) a *rational number* may be defined as a pair of integers (a, b) without a common divisor (where $b > 0$ and a may be positive or negative, or a is 0 and b is 1). The usual arithmetic operations on the rationals, together with the operation of taking the absolute value, are then easily supplied with explicit definitions. Accordingly it is clear what it means to specify a rational number.

To specify exactly what is meant by giving a *real number* explicitly is not quite so simple. For a real number is by its nature an infinite object, but one normally regards only finite objects as capable of being given explicitly. This difficulty may be overcome by stipulating that, to be given a real number, we must be given a (finite) *rule* or *explicit procedure* for calculating it to any desired degree of accuracy. Intuitively speaking, to be given a real number r is to be given a method of computing, for each positive integer n, a rational number r_n such that

$$|r - r_n| < \frac{1}{n}.$$

These r_n will then obey the law

$$|r_m - r_n| \leq \frac{1}{m} + \frac{1}{n}.$$

So, given any numbers k, p, we have, setting $n = 2\,k$,

$$\left|r_{n+p} - r_n\right| \leq \frac{1}{n+p} + \frac{1}{n} \leq \frac{2}{n} = \frac{1}{k}.$$

One is thus led to *define* a (constructive) real number to be a sequence of rationals $(r_n) = r_1, r_2, \ldots$ such that, for any k, a number n can be found such that, for all p $\left|r_{n+p} - r_n\right| \leq \frac{1}{k}$.

Here we understand that to be given a *sequence* we must be in possession of a *rule* or explicit method for generating its members. Each rational number α may be regarded as a real number by identifying it with the real number (α, α, …). The set of all real numbers – the *constructive real line* or the *constructive continuum* – will be denoted, as usual, by $\mathbb{R}$. Now of course, for any "given" real number there are a variety of ways of giving explicit approximating sequences for it. Thus it is necessary to define an equivalence relation, "equality on the reals." The correct definition here is: $r =_{\mathbb{R}} s$ if and only if for any k, a number n can be found so that

$$\left|r_{n+p} - s_{n+p}\right| \leq \frac{1}{k} \text{ for all } p.$$

When we say that two real numbers are equal we shall mean that they are equivalent in this sense, and so write simply "=" for "$=_{\mathbb{R}}$". To assert the inequality $r \neq s$ of two real numbers r, s is to assert that the equality $r = s$ leads to a contradiction. Inequality in this sense is constructively weak. In constructive mathematics a stronger notion of inequality, that of *apartness*, is normally used instead. We say that r and s are *apart*, or *distinguishable*, written $r \# s$, if n and k can actually be found so that $\left|r_{n+p} - s_{n+p}\right| > \frac{1}{k}$ for all p. Clearly $r \# s$ implies $r \neq s$, but the converse cannot be affirmed constructively. (In fact the converse is equivalent to *Markov's Principle*, which asserts that, if, for each n, $x_n = 0$ or 1, and if it is contradictory that $x_n = 0$ for all n, then there exists n for which $x_n = 1$. This thesis is accepted by some, but not all schools of constructivism.)

Is it constructively the case that for any real numbers x and y, we have $x = y \vee x \neq y$? The answer is no. For if this assertion were constructively true, then, in particular, we would have a method of deciding whether, for any given rational number r, whether $r = \pi^{\sqrt{2}}$ or not. But at present no such method is known — it is not known, in fact, whether $\pi^{\sqrt{2}}$ is rational or irrational. We can, of course, calculate $\pi^{\sqrt{2}}$ to as many decimal places as we please, and if in actuality it is unequal to a given rational number r, we shall discover this fact after a sufficient amount of calculation. If, however, $\pi^{\sqrt{2}}$ is *equal* to r, even several centuries of computation cannot make this fact certain; we can be sure only that is very close to r. We have no method which will tell us, in finite time, whether $\pi^{\sqrt{2}}$ exactly coincides with r or not. This situation may be summarized by saying that equality on the reals is *not decidable*. (By contrast, equality on the integers or rational numbers is decidable.) Observe that this does *not* mean $\neg(x = y \vee x \neq y)$. We have not actually derived a *contradiction* from the assumption $x = y \vee x \neq y$; we have only given an example showing its implausibility. It is natural to ask whether it can actually be *refuted*. For this it would be necessary to make some assumption concerning the real numbers which contradicts classical mathematics. Certain schools of constructive mathematics are willing to make such assumptions; but the majority of constructivists confine themselves to methods which are also classically correct.

Despite the fact that constructive equality of real numbers is not a decidable relation, it is *balanced* in the sense of satisfying the *law of double negation* $\neg(r \neq s) \Rightarrow r = s$. In fact, we can prove the stronger assertion that $\neg(r \# s) \Rightarrow r = s$. For, given k, we may choose n so that $|r_{n+p} - r_n| \leq \frac{1}{4k}$ and $|s_{n+p} - s_n| \leq \frac{1}{4k}$ for all p. If $|r_n - s_n| \geq \frac{1}{k}$, then we would have $|r_{n+p} - s_{n+p}| \geq \frac{1}{2k}$ for all p, which entails $r \neq s$. If $\neg(r \# s)$, it follows that $|r_n - s_n| < \frac{1}{k}$ and $|r_{n+p} - s_{n+p}| \leq \frac{2}{k}$ for every p. Since for every k we can find n so that this inequality holds for every p, it follows that $r = s$. (From the facts that $\neg(r \# s) \Rightarrow r = s$ it follows easily that $r \neq s \Leftrightarrow \neg\neg(r \# s)$).

One should not, however, conclude from the stability of equality that the law of double negation $\neg\neg A \rightarrow A$ is generally affirmable. That it is not can be seen from the following example. Write the decimal expansion of π and below the decimal expansion $\rho = 0.333\ldots$, terminating it as soon as a sequence of digits 0123456789 has appeared in π. Then if the 9 of the first sequence 0123456789 in π is the k^{th} digit after the decimal point, $\rho = \frac{10^k - 1}{3 \cdot 10^k}$. Now suppose that ρ were not rational; then $\rho = \frac{10^k - 1}{3 \cdot 10^k}$ would be impossible and no sequence 0123456789 could appear in π, so that $\rho = \frac{1}{3}$, which is also impossible. Thus the assumption that ρ is not rational leads to a contradiction; yet we are not warranted to assert that r is rational, for this would mean that we could calculate integers m and n for which $\rho = \frac{m}{n}$. But this evidently requires that we can produce a sequence 0123456789 in π or demonstrate that no such sequence can appear, and at present we can do neither.

Order on the Constructive Real Line. The *order relation* on the reals is given constructively by stipulating that $r < s$ is to mean that we have an explicit lower bound on the distance between r and s. That is,

$$r < s \Leftrightarrow n \text{ and } k \text{ can be found so that } s_{n+p} - r_{n+p} > 1/k \text{ for all } p.$$

It can readily be shown that, for any real numbers x, y such that $x < y$, there is a rational number α such that $x < \alpha < y$.

We observe that $r \# s \Leftrightarrow r < s \vee s < r$. The implication from right to left is clear. Conversely, suppose that $r \# s$. Find n and k so that $|r_{n+p} - s_{n+p}| > \frac{1}{k}$ for every p, and determine $m > n$ so that $|r_m - r_{m+p}| < \frac{1}{4k}$ and $|s_m - s_{m+p}| < \frac{1}{4k}$ for every p. Either $r_m - s_m > \frac{1}{k}$ or $s_m - r_m > \frac{1}{k}$; in the first case $r_{m+p} - s_{m+p} > \frac{1}{2k}$ for every p, whence $s < r$; similarly, in the second case, we obtain $r < s$.

We define $r \leq s$ to mean that $s < r$ is false. Notice that $r \leq s$ is not the same as $r < s$ or $r = s$: in the case of the real number ρ defined above, for instance, clearly $\rho \leq \frac{1}{3}$; but we do not know whether $\rho < \frac{1}{3}$ or $\rho = \frac{1}{3}$. Still, it is true that $r \leq s \wedge s \leq r \Rightarrow r = s$. For the premise is the negation of $r < s \vee s < r$, which, by the above, is equivalent to $\neg r \# s$. But we have already seen that this last implies $r = s$.

There are several common properties of the order relation on real numbers which hold classically but which cannot be established constructively. Consider, for example, the trichotomy law $x < y \vee x = y \vee y < x$. Suppose we had a method enabling us to decide which of the three alternatives holds. Applying it to the case $y = 0$, $x = \pi^{\sqrt{2}} - r$ for rational r would yield an algorithm for determining whether $\pi^{\sqrt{2}} = r$ or not, which we have already observed is an open problem. One can also

demonstrate the failure of the trichotomy law (as well as other classical laws) by the use of "fugitive sequences." Here one picks an unsolved problem of the form $\forall nP(n)$, where P is a decidable property of integers – for example, Goldbach's conjecture that every even number ≥ 4 is the sum of two odd primes. Now one defines a sequence – a "fugitive" sequence – of integers (n_k) by $n_k = 0$ if $2\,k$ is the sum of two primes and $n_k = 1$ otherwise. Let r be the real number defined by $r_k = 0$ if $n_k = 0$ for all $j \leq k$, and $r_k = 1/m$ otherwise, where m is the least positive integer such that $n_m = 1$. It is then easy to check that $r \geq 0$ and $r = 0$ if and only if Goldbach's conjecture holds. Accordingly the correctness of the trichotomy law would imply that we could resolve Goldbach's conjecture. Of course, Goldbach's conjecture might be resolved in the future, in which case we would merely choose another unsolved problem of a similar form to define our fugitive sequence.

A similar argument shows that the law $r \leq s \vee s \leq r$ also fails constructively: define the real number s by $s_k = 0$ if $n_k = 0$ for all $j \leq k$; $s_k = 1/m$ if m is the least positive integer such that $n_m = 1$, and m is even; $s_k = -1/m$ if m is the least positive integer such that $n_m = 1$, and m is odd. Then $s \leq 0$ (resp. $0 \leq s$) would mean that there is no number of the form $2 \cdot 2\,k$ (resp. $2 \cdot (2\,k + 1)$) which is not the sum of two primes. Since neither claim is at present known to be correct, we cannot assert the disjunction $s \leq 0 \vee 0 \leq s$.

In constructive analysis there is a convenient substitute for trichotomy known as the *comparison principle*. This is the assertion

$$r < t \Rightarrow r < s \vee s < t.$$

Its validity can be established in a manner similar to the foregoing.

Albebraic Operations on the Constructive Reals. The fundamental operations $+$, $-$, $\cdot$, $^{-1}$ and $|\cdot|$ are defined for real numbers as one would expect, viz.

- $r + s$ is the sequence $(r_n + s_n)$
- $r - s$ is the sequence $(r_n - s_n)$
- $r \cdot s$ or rs is the sequence $(r_n s_n)$
- if $r \# 0$, r^{-1} is the sequence (t_n), where $t_n = {r_n}^{-1}$ if $t_n \neq 0$ and $t_n = 0$ if $r_n = 0$
- $|\,r\,|$ is the sequence $(|r_n|)$

It is then easily shown that $rs \# 0 \Leftrightarrow r \# 0 \wedge s \# 0$. For if $r \# 0 \wedge s \# 0$, we can find k and n such that $|r_{n+p}| > \frac{1}{k}$ and $|s_{n+p}| > \frac{1}{k}$ for every p, so that $|r_{n+p}s_{n+p}| > \frac{1}{k^2}$ for every p, and $rs \# 0$. Conversely, if $rs \# 0$, then we can find k and n so that

$$|r_{n+p}s_{n+p}| > \frac{1}{k}, |r_{n+p} - r_n| < 1, |s_{n+p} - s_n| < 1$$

for every p. It follows that

$$|r_{n+p}| > \frac{1}{k}(|s_n| + 1) \text{ and } |s_{n+p}| > \frac{1}{k}(|r_n| + 1)$$

for every p, whence $r \# 0 \wedge s \# 0$.

But it is not constructively true that, if $rs = 0$, then $r = 0$ or $s = 0$! To see this, use the following prescription to define two real numbers r and s. If in the first n decimals of π no sequence 0123456789 occurs, put $r_n = s_n = 2^{-n}$; if a sequence of this kind does occur in the first n decimals, suppose the 9 in the first such sequence is the k^{th} digit. If k is odd, put $r_n = 2^{-k}$, $s_n = 2^{-n}$; if k is even, put $r_n = 2^{-n}$, $s_n = 2^{-k}$. Then we are unable to decide whether $r = 0$ or $s = 0$. But $rs = 0$. For in the first case above r_n $s_n = 2^{-2n}$; in the second r_n $s_n = 2^{-k-n}$. In either case $|r_n s_n| < \frac{1}{m}$ for $n > m$, so that $rs = 0$.

Convergence of Sequences and Completeness of the Constructive Reals. As usual, a sequence (a_n) of real numbers is said to *converge* to a real number b, or to have *limit* s if, given any natural number k, a natural number n can be found so that for every natural number p,

$$\left|b - a_{n+p}\right| < 2^{-k}.$$

As in classical analysis, a constructive necessary and condition that a sequence (a_n) of real numbers be convergent is that it be a *Cauchy* sequence, that is, if, given any natural number k, a natural number n can be found so that for every natural number p,

$$\left|a_{n+p} - a_n\right| < 2^{-k}.$$

However, some classical theorems concerning convergent sequences are no longer valid constructively. For example, a bounded momotone sequence need no longer be convergent. A simple counterexample is provided by the sequence (a_n) defined as follows: $a_n = 1–2^{-n}$ if among the first n digits in the decimal expansion of π no sequence 0123456789 occurs, while $a_n = 2–2^{-n}$ if among these n digits such a sequence does occur. Since it is not known whether the limit of this sequence, if it exists, is 1 or 2, we cannot claim that that this limit exists as a well-defined real number.

In classical analysis $\mathbb{R}$ is *complete* in the sense that every nonempty set of real numbers that is bounded above has a supremum. As it stands, this assertion is constructively incorrect. For consider the set A of members $\{x_1, x_2, \ldots\}$ of any fugitive sequence of 0 s and 1 s. Clearly A is bounded above, and its supremum would be either 0 or 1. If we knew which, we would also know whether $x_n = 0$ for all n, and the sequence would no longer be fugitive.

Nevertheless, the completeness of $\mathbb{R}$ can be salvaged by defining suprema and infima somewhat more delicately than is customary in classical mathematics. A nonempty set A of real numbers is *bounded above* if there exists a real number b, called an *upper bound* for A, such that $x \leq b$ for all $x \in A$. A real number b is called a *supremum*, or *least upper bound*, of A if it is an upper bound for A and if for each $\varepsilon > 0$ there exists $x \in A$ with $x > b - \varepsilon$. We say that A is *bounded below* if there exists a real number b, called a *lower bound* for A, such that $b \leq x$ for all $x \in A$. A real number b is called an *infimum*, or *greatest lower bound*, of A if it is a lower bound for

A and if for each $\varepsilon > 0$ there exists $x \in A$ with $x < b + \varepsilon$. The supremum (respectively, infimum) of A, is unique if it exists and is written sup A (respectively, inf A).

Let us prove the *constructive least upper bound principle*.

Theorem Let A be a nonempty set of real numbers that is bounded above. Then sup A exists if and only if for all $x, y \in \mathbb{R}$ with $x < y$, either y is an upper bound for A or there exists $a \in A$ with $x < a$.

Proof If sup A exists and $x < y$, then either sup $A < y$ or $x <$ sup A; in the latter case we can find $a \in A$ with sup A – (sup $A - x$) $< a$, and hence $x < a$. Thus the stated condition is necessary.

Conversely, suppose the stated condition holds. Let a_1 be an element of A and choose an upper bound b_1 for A with $b_1 > a_1$. We construct recursively a sequence (a_n) in A and (b_n) of upper bounds for A such that, for each $n \geq 0$,

(i) $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$

and

(ii) $b_{n+1} - a_{n+1} \leq \frac{3}{4}(b_n - a_n)$.

Having found $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, if $a_n + \frac{3}{4}(b_n - a_n)$ is an upper bound for A, put $b_{n+1} = a_n + \frac{3}{4}(b_n - a_n)$ and $a_{n+1} = a_n$; while if there exists $a \in A$ with $a > a_n + \frac{3}{4}(b_n - a_n)$, we set $a_{n+1} = a$ and $b_{n+1} = b_n$. This completes the recursive construction.

From (i) and (ii) we have

$$0 \leq b_n - a_n \leq (\tfrac{3}{4})^{n-1}(b_1 - a_1).$$

It follows that the sequences (a_n) and (b_n) converge to a common limit 1 with $a_n \leq 1 \leq b_n$ for $n \geq 1$. Since each b_n is an upper bound for A, so is 1. On the other hand, given $\varepsilon > 0$, we can choose n so that $1 \geq a_n > 1 - \varepsilon$, where $a_n \in A$. Hence $1 = \sup A$.

An analogous result for infima can be stated and proved in a similar way.

Functions on the ConstructiveReals. Considered constructively, a *function* from $\mathbb{R}$ to $\mathbb{R}$ is a rule F which enables us, when given a real number x, to compute another real number $F(x)$ in such a way that, if $x = y$, then $F(x) = F(y)$. It is easy to check that every polynomial is a function in this sense, and that various power series and integrals, for example, those defining tan x and e^x, also determine functions. Viewed constructively, some classically defined "functions" on $\mathbb{R}$ can no longer be considered to be defined on the whole of $\mathbb{R}$. Consider, for example, the "blip" function B defined by $B(x) = 0$ if $x \neq 0$ and $B(0) = 0$. Here the domain of the function is $\{x \in \mathbb{R}: x = 0 \vee x \neq 0\}$. But we have seen that we cannot assert dom$(B) = \mathbb{R}$. Accordingly,

the blip function is not well defined as a function from $\mathbb{R}$ to $\mathbb{R}$. Of course, classically, B is the simplest *discontinuous* function defined on $\mathbb{R}$. The fact that the simplest possible discontinuous function fails to be defined on the whole of $\mathbb{R}$ gives grounds for the suspicion that *no* function defined on $\mathbb{R}$ can be discontinuous; in other words, that, constructively speaking, *all functions defined on* $\mathbb{R}$ *are continuous* (as remarked in the previous section, this was a central tenet of intuitionism's founder, Brouwer). The claim is plausible. For if a function F is well defined on all reals x, it must be possible to compute the value for all rules x determining real numbers, that is, determining their sequences of rational approximations $x_1, x_2, \ldots$. Now $F(x)$ must be computed to accuracy ε in a finite number of steps – the number of steps depending on ε. This means that only finitely many approximations can be used, i.e., $F(x)$ can be computed to within ε only when x is known within δ for some δ. Thus F should indeed be continuous. In fact all known examples of constructive functions defined on $\mathbb{R}$ are continuous. Constructively, a real-valued function f is *continuous* if for each $\varepsilon > 0$ we can find a specific $\delta > 0$ such that $|f(x) - (y)| \leq \varepsilon$ whenever $|x - y| < \delta$. If all functions on $\mathbb{R}$ are continuous, then a subset A of $\mathbb{R}$ may fail to be genuinely *complemented*: that is, there may be no subset B of $\mathbb{R}$ disjoint from A such that $\mathbb{R} = A \cup B$. To see this, we first suppose that A, B are disjoint subsets of $\mathbb{R}$ and that there is a point $a \in A$ which can be approached arbitrarily closely by points of B (or vice versa). Then, assuming all functions on $\mathbb{R}$ are continuous, it cannot be the case that $\mathbb{R} = A \cup B$. For if so, we may define the function f on $\mathbb{R}$ by $f(x) = 0$ if $x \in A$, $f(x) = 1$ if $x \in B$. Then for all $\delta > 0$ there is $b \in B$ for which $|b - a| < \delta$, but $|f(b) - f(a)| = 1$. Thus f fails to be continuous at a, and we conclude that $\mathbb{R} \neq A \cup B$. Now if we take A to be any finite set of real numbers, any union of open or closed intervals, or the set $\mathbb{Q}$ of rational numbers, then in each case the set B of points "outside" A satisfies the above condition. Accordingly, for each such subset A, $\mathbb{R}$ is not "decomposable" into A and the set of points "outside" A, in the sense that these two sets of points together do not exhaust $\mathbb{R}$. This fact indicates that the constructive continuum is a great deal more "cohesive" than its classical counterpart. For classically, the continuum is merely *connected* in the sense that it is not (nontrivially) decomposable into two open (or closed) subsets. Constructively, however, $\mathbb{R}$ is indecomposable into subsets which are neither open nor closed. In this sense the constructive real line can be brought close to the ideal of a true continuum.

Certain well-known theorems of classical analysis concerning continuous functions fail in constructive analysis. One such is the *theorem of the maximum*: a uniformly continuous function on a closed interval assumes its maximum at some point. For consider, as Fig. 1 below, a function f: $[0, 1] \rightarrow \mathbb{R}$ with two relative

Fig. 1

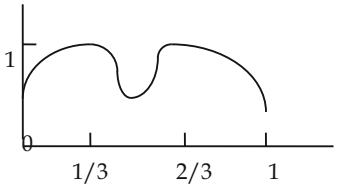
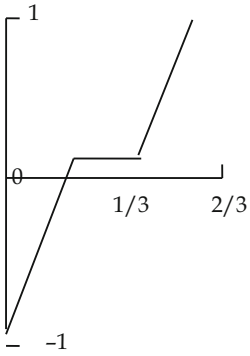


Fig. 2



maxima, one at $x = 1/3$ and the other at $x = 2/3$ and of approximately the same value. Now arrange things so that $f\left(\frac{1}{3}\right) = 1$ and $f\left(\frac{2}{3}\right) = 1 + t$, where t is some small parameter. If we could tell where f assumes its absolute maximum, clearly we could also determine whether $t \leq 0$ or $t \geq 0$, which, as we have seen, is not, in general, possible. Nevertheless, it can be shown that from f we can in fact calculate the maximum value itself, so that at least one can assert the existence of that maximum, even if one cannot tell exactly where it is assumed.

Another classical result that fails to hold constructively in its usual form is the well-known *intermediate value theorem*. This is the assertion that, for any continuous function f from the unit interval [0, 1] to $\mathbb{R}$, such that $f(0) = -1$ and $f(1) = 1$, there exists a real number $a \in [0, 1]$ for which $f(a) = 0$. To see that this fails constructively, consider the function f depicted in Fig. 2 above: here f is piecewise linear, taking the value t (a small parameter) between $x = 1/3$ and $x = 2/3$. If the intermediate value theorem held, we could determine a for which $f(a) = 0$. Then either $a < 2/3$ or $a > 1/3$; in the former case $t \geq 0$; in the latter $t \leq 0$. Thus we would be able to decide whether $t \geq 0$ or $t \leq 0$; but we have seen that this is not constructively possible in general.

Nevertheless, it can be shown that, constructively, the intermediate value theorem is "almost" true in the sense that

$$\forall f \forall \varepsilon > 0 \exists a \, (|f(a)| < \varepsilon)$$

and also in the sense that, if we write $P(f)$ for

$$\forall b \forall a < b \exists c \, (a < c < b \wedge f(c) \neq\neq 0),$$

then

$$\forall f [P(f) \rightarrow \exists x \, (f(x) = 0)].$$

This example illustrates how a single classical theorem "refracts" into several constructive theorems.

Axiomatizing the Constructive Reals. The constructive reals can be furnished with an axiomatic description (Bridges 1999). We begin by assuming the existence of a set $\boldsymbol{R}$ with

- A binary relation $>$ (*greater than*)
- A corresponding *apartness relation* # defined by $x \# y \Leftrightarrow x > y$ or $y > x$
- A unary operation $x \mapsto -x$
- Binary operations $(x, y) \mapsto x + y$ (*addition*) and $(x, y) \mapsto xy$ (*multiplication*)
- Distinguished elements 0 (*zero*) and 1 (*one*) with $0 \neq 1$
- A unary operation $x \mapsto x^{-1}$ on the set of elements $\neq 0$

The elements of $\boldsymbol{R}$ are called *real numbers*. *A* real number x is *positive* if $x > 0$ and *negative* if $-x > 0$. The relation $\geq$ (*greater than or equal to*) is defined by

$$x \geq y \Leftrightarrow \forall z(y > z \Rightarrow x > z).$$

The relations $<$ and $\leq$ are defined in the usual way; x is *nonnegative* if $0 \leq x$.

The sets $\boldsymbol{N}$ of natural numbers, $\boldsymbol{N}^+$ of positive integers, $\mathbf{Z}$ of integers and $\boldsymbol{Q}$ of rational numbers are identified with the usual subsets of $\boldsymbol{R}$; for instance, $\boldsymbol{N}^+$ is identified with the set of elements of R of the form $1 + 1 + \cdots + 1$.

These relations and operations are subject to the following three groups of axioms, which, taken together, form the system **CA** of axioms for *constructive analysis*, or the *constructive real numbers.*

Field Axioms

$$x + y = y + x \quad (x + y) + z = x + (y + z) \quad 0 + x = x \quad x + (-x) = 0 \quad xy = yx$$
$$(xy)z = x(yz) \quad 1x = x \quad x \# 0 \rightarrow xx^{-1} = 1 \quad x(y + z) = xy + xz$$

Order Axioms

$$(x \geq y \wedge y \geq x) \leftrightarrow x = y.$$
$$\neg(x > y \wedge y > x) \quad x > y \rightarrow \forall z(x > z \vee z > y) \quad \neg(x \# y) \rightarrow x = y$$
$$x > y \rightarrow \forall z(x + z > y + z) \quad (x > 0 \wedge y > 0) \rightarrow xy > 0.$$

Notice that from the fourth of the order axioms it follows that $\neg(x \neq y) \rightarrow x = y$, that is, the equality relation is *balanced*. (Here and in the sequel $x \neq y$ is an abbreviation for $\neg(x = y)$).

The final two axioms introduce special properties of $>$ and $\geq$. In the second of these the notions *bounded above*, *bounded below*, *and bounded* are defined as in classical mathematics, and the *least upper bound*, if it exists, of a nonempty set S of real numbers is the unique real number b such that

- b is an upper bound for S.
- For each $c < b$ there exists $s \in S$ with $s > c$.

Here a "nonempty" has the stronger constructive meaning of being inhabited, to wit, that an element of the set in question can actually be constructed.

Special Properties of >.

Archimedean axiom. For each $x \in \boldsymbol{R}$ such that $x \geq 0$ there exists $n \in \boldsymbol{N}$ such that $x < n$.

The least upper bound principle. Let S be a nonempty subset of R that is bounded above, such that for all real numbers a, b with $a < b$, either b is an upper bound for S or else there exists $s \in S$ with $s > a$. Then S has a least upper bound.

The following basic properties of $>$ and $\geq$ can then be established.

$$\neg(x > x)\ x \geq x\ (x > y \wedge y > z) \rightarrow x > z\ \neg(x > y \wedge y \geq x)$$
$$(x > y \geq z) \rightarrow x > z \neg(x > y) \leftrightarrow y \geq x\ \neg(x \geq y) \leftrightarrow \neg$$
$$\neg(y > x)\ (x \geq y \geq z) \rightarrow x \geq z\ \neg(x > y \wedge x = y)$$
$$x \geq 0 \rightarrow (x = 0 \leftrightarrow \forall \varepsilon > 0(x < \varepsilon))x + y > 0 \rightarrow (x > 0 \vee y > 0)$$
$$x > 0 \rightarrow -x < 0\ (x > y \wedge z < 0) \rightarrow yz > xz\ x0 \leftrightarrow x^2 > 0\ 1 > 0$$
$$0 < x < 1 \rightarrow x > x^{-2}\ n \in N^+ \rightarrow n^{-1} > 0$$
$$\textit{if}\ x > 0 \text{ and } y \geq 0, \text{then there exists } n \in \mathbf{Z} \text{ such that } nx > y$$
$$x > 0 \rightarrow x^{-1} > 0\ xy > 0 \rightarrow (x \neq 0 \vee y \neq 0)$$
$$\text{if } a < b, \text{then there exists } r \in \mathbf{Q} \text{ such that } a < r < b$$

The constructive real line $\mathbb{R}$ as introduced above is a model of **CA**. Are there any other models, that is, models not isomorphic to $\mathbb{R}$? If classical logic is assumed, **CA** is a categorical theory and so the answer is no. But this is not the case within intuitionistic logic, for it can be shown that, in intuitionistic set theory, both the Dedekind and Cantor reals are models of **CA**, while these may fail to be isomorphic.

The Intuitionistic Continuum. In constructive analysis, a real number is an infinite (convergent) sequence of rational numbers generated by an effective rule, so that the constructive real line is essentially just a restriction of its classical counterpart. Brouwerian intuitionism takes a more liberal view of the matter, resulting in a considerable enrichment of the arithmetical continuum over the version offered by strict constructivism. As conceived by intutionism, the arithmetical continuum admits as real numbers "not only infinite sequences determined in advance by an effective rule for computing their terms, but also ones in whose generation free selection plays a part." (Dummett 1977, p. 62). The latter are called (*free*) *choice sequences*. Without loss of generality we may and shall assume that the entries in choice sequences are natural numbers.

Hermann Weyl describes Brouwer's conception of choice sequences in the following terms:

> *In Brouwer's analysis*, *the individual place in the* continuum, *the real number*, *is to be defined not by a set but by a* sequence *of natural numbers*, *namely*, *by a law which correlates with every natural number n a natural number* $\varphi(n)$. . . *How then do assertions arise which concern*. . . *all real numbers*, *i.e.*, *all values of a real variable? Brouwer shows that frequently statements of this form in traditional analysis*, *when correctly interpreted*, *simply concern the totality of natural numbers*. *In cases where they do not*, *the notion of sequence changes its meaning*: *it no longer signifies a sequence determined by some law or other*, *but rather one that is created* step by step by free acts of choice, *and thus necessarily remains* in

> statu nascendi. *This* "*becoming*" selective sequence (werdende Wahlfolge) *represents the continuum*, *or the variable*, *while the sequence determined* ad infinitum *by a law represents the individual real number in the continuum*. *The continuum no longer appears*, *to use Leibniz's language*, *as an aggregate of fixed elements but as a medium of free* "*becoming*". *Of a selective sequence* in statu nascendi, *naturally only those properties can be meaningfully asserted which already admit of a yes-or-no decision* (*as to whether or not the property applies to the sequence*) *when the sequence has been carried to a certain point*; *while the continuation of the sequence beyond this point*, *no matter how it turns out*, *is incapable of overthrowing that decision*. (Weyl 1949, p. 52).

While constructive analysis does not formally contradict classical analysis and may in fact be regarded as a subtheory of the latter, a number of intuitionistically plausible principles have been proposed for the theory of choice sequences, which render intuitionistic analysis divergent from its classical counterpart. (Intuitionistic analysis, nevertheless, an extension of CA.)

One such principle is *Brouwer's Continuity Principle*. This asserts that, given a relation $R(\alpha, n)$ between choice sequences α and numbers n, if for each α a number n may be determined for which $R(\alpha, n)$ holds, then n can already be determined on the basis of the knowledge of a finite number of terms of α. This may be seen to be plausible if one considers that according to Brouwer the construction of a choice sequence is incompletable; at any given moment we can know nothing about it outside the identities of a finite number of its entries. Brouwer's Continuity Principle amounts to the assertion that every function from $\mathbb{N}^{\mathbb{N}}$ to $\mathbb{N}$ is continuous.

From this one can prove a weak version of the Continuity Theorem, namely, that every function from $\mathbb{R}$ to $\mathbb{R}$ is continuous (Bridges and Richman 1987, p. 109). Another such principle is *Bar Induction*, a certain form of induction for well-founded sets of finite sequences (Dummett 1977; Bridges and Richman 1987). Brouwer used Bar Induction and the Continuity Principle in proving his Continuity Theorem that every real-valued function defined on a closed interval is uniformly continuous, from which it follows that the intuitionistic continuum, and all of its closed interval, are cohesive.

Brouwer gave the intuitionistic conception of mathematics an explicitly subjective twist by introducing the *creative subject*. The creative subject was conceived as a kind of idealized mathematician for whom time is divided into discrete sequential stages, during each of which he may test various propositions, attempt to construct proofs, and so on. In particular, it can always be determined whether or not at stage n the creative subject has a proof of a particular mathematical proposition p. While the theory of the creative subject remains controversial, its purely mathematical consequences can be obtained by a simple postulate, which is entirely free of subjective and temporal elements. The creative subject allows us to define, for a given proposition p, a binary sequence $<a_n>$ by $a_n = 1$ if the creative subject has a proof of p at stage n; $a_n = 0$ otherwise.

Now if the construction of these sequences is the only use made of the creative subject, then references to the latter may be avoided by postulating the principle known as *Kripke's Scheme:*

For each proposition p there exists an increasing binary sequence $<a_n>$ such that p holds if and only if $a_n = 1$ for some n.

Taken together, these principles have been shown to have remarkable consequences for the cohesiveness of subsets of the continuum. Not only is the intuitionistic continuum cohesive, but, assuming Brouwer's Continuity Principle and Kripke's Scheme, it remains cohesive even if one pricks it with a pin "The [intuitionistic] continuum has, as it were, a syrupy nature, one cannot simply take away one point." If in addition Bar Induction is assumed, then, even more surprisingly, cohesiveness is maintained even when all the rational points are removed from the continuum. Thus it is appropriate to describe the classical continuum as the "frozen intuitionistic continuum."

5 Nonstandard Analysis and the Hyperreal Line

Once the continuum had been provided with a set-theoretic foundation, the use of the infinitesimal in mathematical analysis was largely abandoned. And so the situation remained for a number of years. The first signs of a revival of the infinitesimal approach to analysis surfaced in the 1950s with a paper by A. H. Laugwitz and C. Schmieden (Laugwitz and Schmieden 1958). But the major breakthrough came in 1960 when it occurred to the mathematical logician **Abraham Robinson** (1918–1974) that "the concepts and methods of contemporary Mathematical Logic are capable of providing a suitable framework for the development of the Differential and Integral Calculus by means of infinitely small and infinitely large numbers." (Robinson 1996, p. xiii). This insight led to the creation of *Nonstandard Analysis* (**NSA**), which Robinson regarded as realizing Leibniz's conception of infinitesimals and infinities as ideal numbers possessing the same properties as ordinary real numbers. In the introduction to his book on the subject he writes:

It is shown in this book that Leibniz's ideas can be fully vindicated and that they lead to a novel and fruitful approach to classical Analysis and to many other branches of mathematics. The key to our method is provided by the detailed analysis of the relation between mathematical languages and mathematical structures which lies at the bottom of contemporary model theory.

After Robinson's initial insight, a number of ways of presenting Nonstandard Analysis were developed. Here is a sketch of one of them (Bell and Machover 1977; Keisler 1994).

Starting with the classical real line $\mathbb{R}$, a set-theoretic universe – the *standard universe* – is first constructed over it: here by such a universe is meant a set U containing $\mathbb{R}$ which is closed under the usual set-theoretic operations of union, power set, Cartesian products, and subsets. Now write $\mathbf{U}$ for the structure $(U, \in)$, where $\in$ is the usual membership relation on U: associated with this is the extension $\mathcal{L}(U)$ of the first-order language of set theory to include a name u for each element u

of U. Now, using the well-known compactness theorem for first-order logic, U is extended to a new structure $^*\mathbf{U} = (^*U, {}^*\in)$, called a *nonstandard universe*, satisfying the following key principle:

> ***Saturation Principle.*** *Let* $\mathbf{\Phi}$ *be a collection of* $\mathcal{L}(U)$*-formulas with exactly one free variable. If* $\mathbf{\Phi}$ *is* finitely satisfiable *in* $\mathbf{U}$, *that is, if for any finite subset* $\mathbf{\Phi}'$ *of* $\mathbf{\Phi}$ *there is an element of* U *which satisfies all the formulas of* $\mathbf{\Phi}'$ in $\mathbf{U}$, *then there is an element of* $^*U$ *which satisfies all the formulas of* $\mathbf{\Phi}$ *in* $^*\mathbf{U}$.

The saturation property expresses the intuitive idea that the nonstandard universe is very rich in comparison to the standard one. Indeed, while there may exist, for each finite subcollection F of a given collection P of properties, an element of U satisfying the members of F in U, there may not necessarily be an element of U satisfying *all* the members of P. The saturation of $^*\mathbf{U}$ guarantees the existence of an element of $^*U$ which satisfies, in $^*\mathbf{U}$, all the members of P. For example, suppose the set $\mathbb{N}$ of natural numbers is a member of U; for each $n \in \mathbb{N}$ let $P_n(x)$ be the property $x \in \mathbb{N}$ & $n < x$. Then clearly, while each finite subcollection of the collection $\mathrm{P} = \{P_n: n \in \mathbb{N}\}$ is satisfiable in U, the whole collection is not. An element of $^*U$ satisfying all the members of P in $^*\mathbf{U}$ will then be a "natural number" greater than every member of $\mathbb{N}$, that is, an *infinite* number.

From the saturation property it follows that $^*\mathbf{U}$ satisfies the important

Transfer Principle. *If* σ *is any sentence of* $\mathcal{L}$(U), *then* σ *holds in* $\mathbf{U}$ *if and only if it holds in* $^*\mathbf{U}$.

The transfer principle may be seen as a version of Leibniz's continuity principle: it asserts that all first-order properties are preserved in the passage to or "transfer" from the standard to the nonstandard universe.

The members of U are called *standard sets*, or *standard objects*; those in $^*U - U$ *nonstandard sets* or *nonstandard objects*: $^*U$ thus consists of both standard and nonstandard objects. The members of $^*U$ will also be referred to as **-sets* or **-objects.* Since $U \subseteq {}^*U$, under this convention every set (object) is also a *-set (object). The **-members* of a *-set A are the *-objects x for which $x \,{}^*\in A$

If A is a standard set, we may consider the collection $\mathrm{A}^\star$ – the *inflate* of A – consisting of all the *-members of A: this is not necessarily a standard set nor even a *-set. The inflate of a standard set may be regarded as the same set viewed from a nonstandard vantage point. While clearly $A \subseteq A^\star$, $A^\star$ may contain "nonstandard" elements not in A. It can in fact be shown that *infinite* standard sets always get "inflated" in this way. Using the transfer principle, any function f between standard sets automatically extends to a function – also written f – between their inflates.

Each mathematical structure $\mathbf{A} = (A, R, \ldots)$ has an inflate $\mathbf{A}^\star = (A^\star, R^\star)$. From the transfer principle it follows that $\mathbf{A}$ and $\mathbf{A}^\star$ have precisely the same first-order properties.

Now suppose that the set $\mathbb{N}$ of natural numbers is a member of U. Then so is the set $\mathbb{R}$ of real numbers, since each real number may be identified with a set of natural numbers. $\mathbb{R}$ may be regarded as an ordered field, and the same is therefore true of its inflate $\mathbb{R}^\star$. This, the *hyperreal line*, has precisely the same first-order properties as $\mathbb{R}$.

The members of $\mathbb{R}^\star$ are called *hyperreals*. A standard hyperreal is then just a real, to which we shall refer for emphasis as a *standard real*. Since $\mathbb{R}$ is infinite, nonstandard hyperreals must exist. The saturation principle implies that there must be an *infinite* (nonstandard) hyperreal, that is, a hyperreal a such that $a > n$ for every $n \in \mathbb{N}$. In that case its reciprocal $\frac{1}{a}$ is infinitesimal in the sense of exceeding 0 and yet being smaller than $\frac{1}{n+1}$ for every $n \in \mathbb{N}$. In general, we call a hyperreal *a infinitesimal* if its absolute value $| a |$ is $< \frac{1}{n+1}$ for every $n \in \mathbb{N}$. In that case the set I of infinitesimals contains not just 0 but a substantial number (in fact, infinitely many) of other elements. Clearly I is an additive subgroup of $\mathbb{R}$, that is, if $a, b \in I$, then $a - b \in I$.

Note that $\mathbb{R}^\star$ is thus a nonarchimedean ordered field. One might question whether this is compatible with the facts that $\mathbb{R}^\star$ and $\mathbb{R}$ share the same first-order properties, but the latter is archimedean. These data are consistent because the archimedean property is not first-order. However, while $\mathbb{R}^\star$ is nonarchimedean, it is **-archimedean* in the sense that, for any $a \in \mathbb{R}^\star$ there is $n \in \mathbb{R}^\star$ for which $a < n$.

The members of the inflate $\mathbb{N}^\star$ of $\mathbb{N}$ are called *hypernatural numbers*. As for the hyperreals, it can be shown that $\mathbb{N}^\star$ also contains nonstandard elements which must exceed every member of $\mathbb{N}$: these are called *infinite* hypernatural numbers.

For hyperreals a, b we define $a \approx b$ and say that a and b are *infinitesimally close* if $a - b \in I$. This is an equivalence relation on the hyperreal line: for each hyperrreal a we write $\mu(a)$ for the equivalence class of a under this relation and call it the *monad* of a. The monad of a hyperreal a thus consists of all the hyperreals that are infinitesimally close to a: it may be thought of as a small cloud centered at a. Note also that $\mu(0) = I$.

A hyperreal a is *finite* if it is not infinite; this means that $| a | < n$ for some $n \in \mathbb{N}$. It is not difficult to show that finiteness is equivalent to the condition of *near-standardness*: here a hyperreal a is *near-standard* if $a \approx r$ for some standard real r.

Much of the usefulness of Nonstandard Analysis stems from the fact that statements of classical analysis involving limits or the (ε, δ) criterion admit succinct, intuitive translations into statements involving infinitesimals or infinite numbers, in turn enabling comparatively straightforward proofs to be given of classical theorems. Here are some examples of such translations:

- *Let* $< s_n >$ be *a standard infinite sequence of real numbers and let s be a standard real number. Then s is the limit of* $< s_n >$ *within* $\mathbb{R}$, $\lim_{n\to\infty} s_n = s$ *in the classical sense, if and only if* $s_n \approx s$ *for all* infinite *subscripts n*.
- A *standard sequence* $< s_n >$ *converges if and only if* $s_n \approx s_m$ *for all infinite n and m*. (Cauchy's criterion for convergence.)

Now suppose that f is a real-valued function defined on some open interval (a, b). We have remarked above that f automatically extends to a function – also written f – on (a, b).

- *In order that the standard real number c be the limit of* $f(x)$ *as x approaches* x_0, $\lim_{x\to x_0} f(x) = c$, *with* x_0 *a standard real number in* (a, b), *it is necessary and sufficient that* $f(x) \approx f(x_0)$ *for all* $x \approx x_0$.

- *The function f is continuous at a standard real number x_0 in (a, b) if and only if $f(x) \approx f(x_0)$ for all $x \approx x_0$.* (This is *equivalent to saying that f maps the monad of x_0 into the monad of $f(x_0)$.*
- *In order that the standard number c be the derivative of f ay x_0 it is necessary and sufficient that*

$$\frac{f(x) - f(x_0)}{x - x_0} \approx c$$

for all $x \neq x_0$ in the monad of x_0.

Many other branches of mathematics admit elegant and fruitful nonstandard formulations.

Finally, it should be pointed out that while the usual models of Nonstandard Analysis are obtained using highly nonconstructive tools, other methods have been developed for producing such models which are constructively acceptable. (Moerdijk 1995; Palmgren 1998, 2001).

6 The Continuum in Smooth Infinitesimal Analysis

Smooth Infinitesimal Analysis. In differential geometry it is customary to confine attention to functions between spaces which are *smooth*, that is, arbitrarily many times differentiable. *Smooth Infinitesimal Analysis* (**SIA**) provides a framework for mathematical analysis which embodies the practice of differential geometry in a remarkably direct and powerful way, namely, by stipulating that *all* functions – however defined – between spaces be smooth. In particular all functions defined on the real line **R** to itself are required to be smooth. In **SIA** this requirement is satisfied through the following considerations. Starting with a differentiable function f: $\mathbf{R} \to \mathbf{R}$, draw the graph of f, so generating a curve in the Cartesian plane. Now pick any point on the curve and consider a short length L of the curve around the point. If L is taken to be sufficiently short, it will appear to be *approximately* a straight line. In **SIA** it is required that, if L is taken to be of *infinitesimal* length, then it will be a straight line *exactly*. This idea in its turn is realized by postulating the existence in **R** of a set Δ of infinitesimals with the property that all functions on Δ are linear.

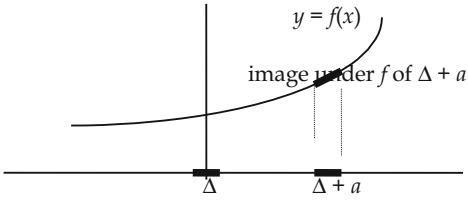
Thus in **SIA** we are given:

- An object **R** called the *real* line (or *smooth line*), on which the usual operations of addition and multiplication are defined and satisfy the usual algebraic laws governing real numbers (in particular 0 and 1 are present and are neutral elements under + and $\times$, respectively.
- A subset $\boldsymbol{\Delta}$ of the set **R** of real numbers called the *domain of infinitesimals*. $\boldsymbol{\Delta}$ is assumed to satisfy the following conditions:
  (1) $0 \in \boldsymbol{\Delta}$
  (2) *Any multiple of an infinitesimal is an infinitesimal*, that is, for any $a \in \mathbf{R}$, $x \in \boldsymbol{\Delta} \Rightarrow ax \in \boldsymbol{\Delta}$. It follows immediately from this that $\boldsymbol{\Delta}$ is *symmetric around* 0, that is, $\forall x\, (x \in \boldsymbol{\Delta} \Leftrightarrow -x \in \boldsymbol{\Delta})$.

(3) The ***Infinitesimal Linearity Principle*** or the ***Kock-Lawvere axiom.*** This asserts that all maps $\mathbf{\Delta}\to\mathbf{R}$ are *linear* (strictly speaking, *affine*, but no matter) in the following strong sense: for any f: $\mathbf{\Delta}\to\mathbf{R}$ there is a *unique* $a \in \mathbf{R}$ such that, for all $\varepsilon \in \mathbf{\Delta}$, $f(\varepsilon) = f(0) + \varepsilon a$. (Here and in the sequel we shall use symbols ε, η to denote arbitrary elements of $\mathbf{\Delta}$.) This real number a is called the *slope* of f, and is written $\mathbf{slp}(f)$. Thus, for any f: $\mathbf{\Delta}\to\mathbf{R}$ and all $\varepsilon \in \mathbf{\Delta}$, $f(\varepsilon) = f(0) + \varepsilon\, \mathbf{slp}(f)$.

Given f: $\mathbf{\Delta}\to\mathbf{R}$, let f^*: $\mathbf{\Delta}\to\mathbf{R}\times\mathbf{R}$ defined by $f^*(\varepsilon) = (\varepsilon, f(\varepsilon))$. The map f^* may be thought of as the graph of the map f n the plane. The Infinitesimal Linearity Principle can be seen as asserting that this graph is a straight line with slope $\mathbf{slp}(f)$ passing through $(0, f(0))$. Thus the effect of any map on $\mathbf{\Delta}$ is to translate and rotate it; in effect $\mathbf{\Delta}$ behaves like a short "rigid rod," just long enough to have a slope, but too short to bend. It is, as it were, a geometric object possessing location and direction but lacking extension.

The Infinitesimal Linearity Principle can also be presented in the following way. Think of a real function f: $\mathbf{R}\to\mathbf{R}$ as defining a curve C in the plane. Then the Infinitesimal Linearity Principle amounts to the assertion that, for any such f and for any point $a \in \mathbf{R}$, the image under f of the "infinitesimal interval" $\mathbf{\Delta} + a$ obtained by translating $\mathbf{\Delta}$ to a is a straight line and coincides with the tangent to C at $x = a$.



Thus each real function has the effect of "bringing" $\mathbf{\Delta}$ into coincidence" with the tangent vector to the curve associated with the function at any point on it. In this sense, then, $\mathbf{\Delta}$ plays the role of a *generic tangent vector.* Also, since the image of $\mathbf{\Delta}$ under a function is necessarily a straight line and a part of the associated curve, it follows that each point on a curve is contained in a nondegenerate infinitesimal straight segment of the curve. Accordingly in **SIA** *curves are infinitesimally straight*. This is called the **Infinitesimal Straightness Principle**.

The Infinitesimal Linearity Principle has two immediate formal consequences:

- **The nondegeneracy of Δ:** $\Delta \neq \{0\}$. For suppose $\Delta = \{0\}$, and let a, b be any unequal real numbers. Then the two maps $\varepsilon \mapsto \varepsilon a$ and $\varepsilon \mapsto \varepsilon b$ would both be identically 0, contradicting the Linearity Principle.
- **The Cancellation Principle:** for any a, $b \in \mathbf{R}$, if $\varepsilon a = \varepsilon b$ for all $\varepsilon \in \Delta$, then $a = b$. This follows immediately from the uniqueness condition in the Linearity Principle.

We can now show that Δ consists of *nilpotent*, in fact *nilsquare* quantities, that is,

$$\forall x\big(x \in \Delta \Rightarrow x^2 = 0\big).$$

To prove this, write s: $\mathbf{\Delta}\rightarrow\mathbf{R}$ for the map $\varepsilon \mapsto \varepsilon^2$. Then we have, for $\varepsilon \in \Delta$,

$$\varepsilon^2 = \mathrm{s}(\varepsilon) = s(0) + \varepsilon\ \mathbf{slp}(s) = \varepsilon\ \mathbf{slp}(s)$$

and, since $-\varepsilon \in \Delta$,

$$\varepsilon^2 = (-\varepsilon)^2 = s(-\varepsilon) = s(0) - \varepsilon\ \mathbf{slp}(s) = -\varepsilon\ \mathbf{slp}(s).$$

Hence $\varepsilon^2 = -\varepsilon^2$, so that $\varepsilon^2 = 0$.

Notice that we do not claim that Δ comprises *all* nilsquare quantities. This, although usually assumed in presentations of **SIA**, is not needed.

Here is a further fact of interest. Without assuming the symmetry of Δ around 0, the following conditions are equivalent:

(i) $\forall x\ (x \in \Delta \Rightarrow x^2 = 0)$.
(ii) $\mathbf{slp}(s) = 0$.

(i) $\Rightarrow$ **(ii)**. Assuming **(i)**, we have, for all $\varepsilon \in \Delta$, $0 = \varepsilon^2 = s(\varepsilon) = s(0) + \varepsilon\ \mathbf{slp}(s) = \varepsilon.\ \mathbf{slp}(s)$. The Cancellation Principle yields $\mathbf{slp}(s) = 0$.

(ii) $\Rightarrow$ **(i)** Assuming **(ii)**, we have, for all $\varepsilon \in \Delta$, $\varepsilon^2 = s(\varepsilon) = s(0) + \varepsilon\ \mathbf{slp}(s) = 0 + 0 = 0$. Hence **(i)**.

The Differential Calculus in SIA. Let us call a *real function* any real-valued function defined on a closed interval into **R**. The derivative of an arbitrary real function can be formulated as follows. Given a closed interval **I** in **R** and a function f: $\mathbf{I}\rightarrow\mathbf{R}$, for each $x \in \mathbf{I}$ define the function f_x: $\Delta\rightarrow\mathbf{R}$ by $f_x(\varepsilon) = f(x + \varepsilon)$. (Intervals are defined by providing **R** with a suitable order relation. We assume that closed intervals are stable under the addition of infinitesimals: see below.)

The *derivative* f': $\mathbf{I}\rightarrow\mathbf{R}$ of f is defined by $f'(x) = \mathbf{slp}\ (f_x)$. It follows easily that

$(*)$ $\qquad f(x + \varepsilon) = f(x) + \varepsilon\, f'(x)$.

This is the *Fundamental Equation of the Differential Calculus* in **SIA**. The quantity $f'(x)$ is the slope at x of the curve determined by f and the infinitesimal

$$\varepsilon\ f'(x) = f(x + \varepsilon) - f(x)$$

is the infinitesimal change or *increment* in the value of f on passing from x to $x + \varepsilon$.

The fundamental equation (*) can be thought of as a rigorous formulation of the engineers' "practical" version of the Taylor series of a function. For (*) is what we get if in the Taylor series

$$f(x + \varepsilon) = f(x) + \varepsilon\ f'(x) + 1/2!\varepsilon^2 f''(x) + \ldots$$

we follow the practice of engineers to treat as "vanishingly small" the square and higher powers of the "small" quantity ε.

Derivatives of elementary functions are easily calculated in **SIA** using the Cancellation Principle. For example, here is the calculation of the derivative of the function x^n:

$$\varepsilon(x^n)' = (x+\varepsilon)^n - x^n = \varepsilon n\, x^{n-1} + \text{terms in } \varepsilon^{2-} \text{ and higher powers} = \varepsilon n\, x^{n-1}.$$

Hence, by the Cancellation Principle,

$$(x^n)' = n\, x^{n-1}.$$

And here is the calculation of the derivative of the function $1/x$ (for $x > 0$):

$$\begin{aligned}\varepsilon(1/x)' &= 1/x + \varepsilon - 1/x = -\varepsilon/x(x+\varepsilon) = -\varepsilon(x-\varepsilon)/x(x+\varepsilon)(x-\varepsilon)\\ &= -\varepsilon x + \varepsilon^2/x\left(x^2-\varepsilon^2\right)\\ &= -\varepsilon x/x^3\\ &= -\varepsilon/x^2.\end{aligned}$$

Cancelling ε on both sides of the equation gives

$$(1/x)' = -1/x^2.$$

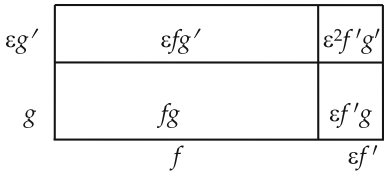
Let us derive in **SIA** a basic law of the differential calculus, the *product rule*:

$$(fg)' = f'g + fg'.$$

To do this we compute

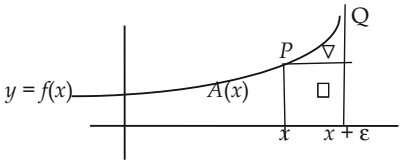
$$\begin{aligned}(fg)(x+\varepsilon) &= (fg)(x) + (fg)'(x) \quad = f(x)g(x) + (fg)'(x),\\ (fg)(x+\varepsilon) &= f(x+\varepsilon)g(x+\varepsilon) = [f(x) + f'(x)].[g(x) + g'(x)]\\ &= f(x)g(x) + \varepsilon(\, f'g + fg') + \varepsilon^2 f'g'\\ &= f(x)g(x) + \varepsilon(\, f'g + fg'),\end{aligned}$$

since $\varepsilon^2 = 0$. Therefore $\varepsilon(fg)' = \varepsilon(f'g + fg')$, and the result follows by the Cancellation Principle. This calculation is depicted in the diagram below.



Infinitesimals in **SIA** have two fundamental aspects, *algebraic* and *geometric*. Algebraically, they are real numbers whose squares vanish; geometrically, they are straight segments of curves. Both of these aspects are used in applications.

The interplay of these aspects is well illustrated by the derivation in **SIA** of the *Fundamental Theorem of the Calculus*. To this end, let **I** be a closed interval $\{x: a \leq x \leq b\}$ in **R** – or **R** itself – and let f: **I**→**R**; also let $A(x)$ be the area under the curve $y = f(x)$ as indicated in the figure:



Then

$$\varepsilon A'(x) = A(x+\varepsilon) - A(x) = \square + \nabla = \varepsilon f(x) + \nabla.$$

Now, by the Infinitesimal Straightness Principle, the arc PQ is a straight line; accordingly ∇ is a triangle of area ½ $\varepsilon.\varepsilon f'(x) = 0$. It follows that $\varepsilon A'(x) = \varepsilon f(x)$, and the Cancellation Principle gives

$$A'(x) = f(x).$$

Thus, if we regard $A(x)$ as the integral of $f(x)$, the above equation asserts that differentiation is the inverse of integration – the Fundamental Theorem of the Calculus.

*Smoothness of Functions in **SIA***. From the Fundamental Equation the *Principle of Continuity* can be deduced, namely, that in **SIA** *all real functions are continuous*, in the sense of *sending neighboring points to neighboring points*. (Here two points x, y on **R** are said to be *neighbors* if $x - y$ is in $\mathbf{\Delta}$, that is, if x and y differ by an infinitesimal.) To see this, given a real function f and neighboring points x, y, note that $y = x + \varepsilon$ with ε in $\mathbf{\Delta}$, so that

$$f(y) - f(x) = f(x+\varepsilon) - f(x) = \varepsilon f'(x).$$

Since $\varepsilon f'(x)$ is infinitesimal, the result follows.

From the fact that in **SIA** every real function has a derivative, which is itself differentiable, it follows that any real function is arbitrarily many times differentiable. This fact justifies the use of the term "smooth."

In **SIA** there is a sense in which *everything is generated by the domain of infinitesimals*. For consider the set $\mathbf{\Delta}^{\mathbf{\Delta}}$ of all maps $\mathbf{\Delta}$→$\mathbf{\Delta}$. It follows from the Linearity Principle that **R** can be identified as the subset of $\mathbf{\Delta}^{\mathbf{\Delta}}$ consisting of all maps vanishing at 0. In this sense **R** is "generated" by $\mathbf{\Delta}$.

Explicitly, $\mathbf{\Delta}^{\mathbf{\Delta}}$ is a monoid under composition which may be regarded as acting on $\mathbf{\Delta}$ by composition: for $f \in \Delta^{\Delta}$, $f \cdot \varepsilon = f(\varepsilon)$. The subset V consisting of all maps

vanishing at 0 is a submonoid naturally identified as the set of *ratios of infinitesimals*. The identification of **R** and V made possible by the principle of infinitesimal linearity thus leads to the characterization of **R** itself as the set of ratios of infinitesimals. This was essentially the view of Euler, who regarded infinitesimals as formal zeros and real numbers as representing the possible values of 0/0. For this reason Lawvere has suggested that **R** in **SIA** should be called the space of *Euler reals*.

Once one has **R**, Euclidean spaces of all dimensions may be obtained as powers of **R**, and arbitrary Riemannian manifolds may be obtained by patching together subspaces of these.

*The "Internal Logic" of **SIA** is Intuitionistic*. We observe that the postulates of **SIA** are *incompatible with the Law of Excluded Middle of classical logic* (LEM) – the assertion that, for any proposition p, either p holds or not p holds. This incompatibility can be demonstrated in two ways, one informal and the other rigorous. First the informal argument. Consider the function f defined for real numbers x by $f(x) = 1$ if $x = 0$ and $f(x) = 0$ whenever $x \neq 0$. If LEM held, each real number would then be either equal or unequal to 0, so that the function f would be defined on the whole of **R**. But, considered as a function with domain **R**, f is clearly discontinuous. Since, as we know, in **SIA** every function on **R** is continuous, f cannot have domain **R** there. (The domain of f is in fact $(\mathbf{R} - \{0\}) \cup \{0\}$, which, because of the failure of the law of excluded middle in **SIA**, is provably unequal to **R**.)

So LEM fails in **SIA**. To put it succinctly, *universal continuity implies the failure of the Law of Excluded Middle*.

Here now is the rigorous argument. We derive the failure of LEM from the Cancellation Principle. To begin with, if $x \neq 0$, then $x^2 \neq 0$, so that, if $x^2 = 0$, then necessarily not $x \neq 0$. This means that

$$\text{for all infinitesimal } \varepsilon, \text{ not } \varepsilon \neq 0. \qquad (*)$$

Now suppose that LEM were to hold. Then we would have, for any ε, either $\varepsilon = 0$ or $\varepsilon \neq 0$. But (*) allows us to eliminate the second alternative, and we infer that, for all ε, $\varepsilon = 0$. This may be written

$$\text{for all } \varepsilon, \varepsilon.1 = \varepsilon.0,$$

from which we derive by Cancellation Principle the falsehood $1 = 0$. So again LEM must fail.

The "internal" logic of **SIA** is accordingly not full classical logic. It is, instead, *intuitionistic* logic, that is, the logic derived from the constructive interpretation of mathematical assertions. In practice when working in **SIA** one does not notice this "change of logic" because, like much of elementary mathematics, the topics discussed there are naturally treated by constructive means such as direct computation.

It is worth noting that the refutability of the Law of Excluded Middle in **SIA** leads to the refutability of an important principle of set theory, the *Axiom of Choice*. This is the assertion

(AC) for any family **A** of inhabited sets, there is a *choice function on* **A**, that is, a function f: $\mathbf{A} \to \cup\mathbf{A}$ for which $f(X) \in X$ whenever $X \in \mathbf{A}$. (A set t is said to be *inhabited* if it can be constructively shown to have a member. In intuitionistic logic this is a stronger condition than the assertion that the set be nonempty.)

Now the Law of Excluded Middle can be derived merely from the very special case of the Axiom of Choice which asserts merely that any *doubleton* $\{U, V\}$ has a choice function. For let p be any proposition, define

$$U = \{x \in 2 : x = 0 \vee p\} V = \{x \in 2 : x = 1 \vee p\}$$

and let f be a choice function on $\{U, V\}$. Writing $a = f(U)$, $b = f(V)$, we have $a \in U$, $b \in V$, i.e.,

$$[a = 0 \vee p] \wedge [b = 1 \vee p].$$

It follows that

$$[a = 0 \vee b = 1] \vee p,$$

whence

$$(*) \qquad\qquad a \neq b \vee p.$$

Now clearly

$$p \to U = V = 2 \to a = b,$$

whence

$$a \neq b \to \neg p,$$

but this and (*) together imply $\neg p \vee p$.

Thus LEM is derivable from this special case of the Axiom of Choice. Since LEM is refutable in **SIA**, so equally, then, is **AC**.

The refutability of the Axiom of Choice in **SIA**, and hence its incompatibility with the Principle of Continuity which prevails there, is not surprising in view of the Axiom's well-known "paradoxical" consequences. One of these is the famous *Banach-Tarski paradox* which asserts that any solid sphere can be decomposed into finitely many pieces which can themselves be reassembled to form two solid spheres each of the same size as the original, or into one solid sphere of any preassigned size. Paradoxical decompositions such as these become possible only when continuous geometric objects are, recalling Dedekind's words, "dissolved to atoms ... [through a] frightful, dizzying discontinuity" into discrete sets of points which the Axiom of Choice then allows to be rearranged in an arbitrary (discontinuous) manner. Such procedures violate the Principle of Continuity.

*The Link Between the Infinitesimal and the Real in **SIA**; the Constancy Principle.* Suppose that we are investigating the behavior of some variable quantity represented by a function F. The approach taken in **SIA**, as (implicitly) in the differential calculus, is to begin the investigation by confining it initially to the infinitesimal world. Life in the infinitesimal world is beautifully simple: curves are just straight lines, and the squares of incremental changes vanish. This makes the determination of infinitesimal increments equally simple, enabling the increment $\varepsilon F'(x)$ in $F(x)$ to be presented in the form $\varepsilon k(x)$, where $k(x)$ is some explicit function whose form has been obtained by "infinitesimal" analysis. Thus we obtain an "infinitesimal" equation of the form $\varepsilon F'(x) = \varepsilon k(x)$. Applying the Cancellation Principle in turn yields the "differential" equation

$$(*) \qquad F'(x) = k(x)$$

which holds in the world "in the large."

The Cancellation Principle thus provides a formal, astonishingly simple link between the infinitesimal world and the real world, the world "in the large." The idea of a linkage between these two worlds was the animating principle behind applications of the calculus throughout the seventeenth and eighteenth centuries.

In practice, of course, the equation (*), while of fundamental importance, is only the first step in determining the explicit form of the function F. For this, it is necessary to "integrate" k, that is, to provide k with an *antiderivative,* an explicit function G such that $G' = k$. It will then follow that $F' = G'$, from which we will be able to conclude that $F = G$. (Strictly speaking, F and G may differ by a constant function but we shall ignore this here.)

To carry out this procedure in **SIA**, we need to introduce an additional postulate. Let **I** be a closed interval. We define a *stationary point* of a function f: **I**→**R** to be a point $a \in \mathbf{I}$ in whose vicinity "infinitesimal changes in the value of the argument fail to change the value of f'', that is, for which $f(a + \varepsilon) = f(a)$ for all ε. This means that $f(a) + \varepsilon f'(a) = f(a)$, so that $\varepsilon f'(a) = 0$ for all ε, from which it follows by the Cancellation Principle that $f'(a) = 0$. Thus a stationary point of a function is precisely a point at which the derivative of the function vanishes.

In classical analysis, if the derivative of a function is identically zero, the function is constant. This fact is the source of the following postulate concerning stationary points adopted in **SIA**:

Constancy Principle. If every point in a closed interval **I** is a stationary point of f: **I**→**R** (that is, if f' is identically 0), then f is constant.

It follows from the Constancy Principle that two functions with identical derivatives differ by at most a constant.

Put succinctly, the Constancy Principle asserts that "universal infinitesimal (or "local") constancy implies global constancy," or "infinitesimal behaviour determines global behaviour." The Constancy Principle brings into sharp focus the difference in **SIA** between points and infinitesimals. For if in the Constancy Principle one replaces

"infinitesimal constancy" by "constancy at a point" the resulting "Principle" is false because *any function whatsoever* is constant at every point. But since in **SIA** all functions on **R** are smooth, the Constancy Principle embodies the idea that for such functions local constancy is sufficient for global constancy, that a nonconstant smooth function must be somewhere nonconstant over arbitrarily small intervals.

The Constancy Principle provides another bridge between the infinitesimal world and the world "in the large." Hermann Weyl could not see such a direct linkage between the two worlds, and inferred that this absence doomed the idea of infinitesimal, leading to its inevitable replacement by the limit concept. In his *Philosophy of Mathematics and Natural Science* he says:

> [In its struggle with the infinitely small] *the limiting process was victorious. For the* limit *is an indispensable concept, whose importance is not affected by the acceptance or rejection of the infinitely small. But once the limit concept has been grasped, it is seen to render the infinitely small superfluous. Infinitesimal analysis proposes to draw conclusions by integration from the behavior in the infinitely small, which is governed by elementary laws, to the behavior in the large; for instance, from the universal law of attraction for two material "volume elements" to the magnitude of attraction between two arbitrarily shaped bodies with homogeneous or non-homogeneous mass distribution. If the infinitely small is not interpreted 'potentially' here, in the sense of the limiting process, then the one has nothing to do with the other, the process in infinitesimal and finite dimensions become independent of each other, the tie which binds them together is cut.*

In **SIA**, the Constancy Principle reconnects the infinitesimal and the extended. Behavior "in the large" is completely determined by behavior "in the infinitely small."

*Cohesiveness of the Continuum in **SIA**.* In classical analysis the continuum and its closed intervals are connected in the sense that they cannot be split into two nonempty subsets neither of which contains a limit point of the other. In **SIA**, the Constancy Principle ensures that these have the vastly stronger property of *cohesiveness* (or indecomposability) which, we recall, is the property of not being splittable into two disjoint nonempty parts *in any way whatsoever.* This is clearly equivalent to the condition that any map defined on **R** or one of its closed intervals to $\{0, 1\}$ takes a constant value.

To show that this condition holds, let A be **R** or any closed interval, and suppose $f: A \to \{0, 1\}$. We claim that f is constant. For we have, for any $x \in A$,

$$(\, f(x) = 0 \text{ or } f(x) = 1) \& (\, f(x + \varepsilon) = 0 \text{ or } f(x + \varepsilon) = 1).$$

This gives four possibilities:

(i) $f(x) = 0$ & $f(x + \varepsilon) = 0$
(ii) $f(x) = 0$ & $f(x + \varepsilon) = 1$
(iii) $f(x) = 1$ & $f(x + \varepsilon) = 0$
(iv) $f(x) = 1$ & $f(x + \varepsilon) = 1$

Possibilities (ii) and (iii) may be ruled out because f is continuous. This leaves (i) and (iv), in either of which $f(x) = f(x + \varepsilon)$. So f is locally, and hence globally, constant, that is, constantly 1 or 0.

From the cohesiveness of closed intervals it can be inferred (Bell 2001) that in **SIA** *all intervals in* **R** *are cohesive.*

Smooth Infinitesimal Analysis as an Axiomatic Theory: Consequences for the Continuum.

SIA can be *axiomatized* as a theory formulated within (higher-order) intuitionistic logic. Here are the basic axioms of the theory (Moerdijk and Reyes 1991).

Axioms for the continuum, or smooth real line R. These include the usual axioms for a commutative ring with unit expressed in terms of two operations + and · (we usually write xy for $x \cdot y$) and two distinguished elements $0 \neq 1$. In addition we stipulate that **R** is an *intuitionistic field*, i.e., satisfies the following axiom:

$$x \neq 0 \text{ implies} \exists y(xy = 1).$$

Axioms for the strict order relation < on R. These are:

O1. $a < b$ and $b < c$ implies $a < c$.
O2. $\neg(a < a)$
O3. $a < b$ implies $a + c < b + c$ for any c.
O4. $a < b$ and $0 < c$ implies $ac < bc$
O5. either $0 < a$ or $a < 1$.
O6. $a \neq b$ implies $a < b$ or $b < a$.
O7. $0 < x$ implies $\exists y\ (x = y^2)$.

Arithmetical Axioms. These govern the set **N** of *Archimedean* (or smooth) *natural numbers,* and read as follows:

1. **N** is a cofinal or Archimedean subset of **R**, i.e., $\mathbf{N} \subseteq \mathbf{R}$ and $\forall\ x \in \mathbf{R}\ \exists n \in \mathbf{N}$ $x < n$.
2. *Peano axioms*:

$$0 \in \mathbf{N}$$

$$\forall x \in \mathbf{R}(x \in \mathbf{N} \rightarrow x + 1 \in \mathbf{N})$$

$$\forall x \in \mathbf{R}(x \in \mathbf{N} \rightarrow x + 1 \neq 0)$$

3. *Restricted Induction scheme.* For every formula $\varphi(x)$ Involving Just $=$, $\wedge$, $\vee$, $\top$ ("True"), $\perp$ ("False"), $\exists$

$$\varphi(0) \wedge \forall x \in \mathbf{N}[\varphi(x) \rightarrow \varphi(x + 1)] \rightarrow \forall x \in \boldsymbol{N}\varphi(x).$$

Using restricted induction it follows that

- **N** has *decidable equality,* i.e., $\forall x \in \mathbf{N} \forall y \in \mathbf{N}(x = y \vee x \neq y)$.
- **N** is *linearly ordered,* i.e., $\forall x \in \mathbf{N} \forall y \in \mathbf{N}(x < y \vee x = y \vee y < x)$.
- **N** satisfies *decidable induction:* for any formula $\varphi(x)$,

$$\forall x \in \mathbf{N}(\varphi(x) \vee \neg\varphi(x)) \rightarrow [[\varphi(0) \wedge \forall x \in \mathbf{N}(\varphi(x) \rightarrow \varphi(x+1)] \rightarrow \forall x \varphi(x)].$$

The relation $\leq$ on **R** is defined by $a \leq b \Leftrightarrow \neg(b < a)$. The open interval (a, b) and closed interval $[a, b]$ are defined as usual, viz., $(a, b) = \{x: a < x < b\}$ and $[a, b] = \{x: a < x < b\}$; similarly for half-open, half-closed, and unbounded intervals. It can be shown from the axioms introduced so far that closed intervals are stable under the addition of infinitesimals.

We have written $\mathbf{\Delta}$ for the subset $\{x: x^2 = 0\}$ of **R** consisting of (nilsquare) *infinitesimals.* As before, we use the letter ε as a variable ranging over $\mathbf{\Delta}$.

The two final axioms are:

Infinitesimal Linearity Axiom. *For any map g:* $\mathbf{\Delta} \rightarrow \mathbf{R}$ *there exist unique* $a, b \in \mathbf{R}$ *such that, for all* ε, *we have*

$$g(\varepsilon) = a + b\varepsilon.$$

Constancy Axiom. *If* $A \subseteq \mathbf{R}$ *is any closed interval on* **R**, *or* **R** *itself, and* $f: A \rightarrow \mathbf{R}$ *satisfies* $f(a + \varepsilon) = f(a)$ *for all* $a \in A$ *and* $\varepsilon \in A$, *then f is constant.*

It follows easily from the Infinitesimal Linearity Axiom that $\mathbf{\Delta}$ is *nondegenerate,* i.e., $\mathbf{\Delta} \neq \{0\}$ (It should be noted that, while Δ does not reduce to $\{0\}$, nevertheless 0 is the only explicitly nameable element of Δ. For it is easily seen to be inconsistent to assert that Δ actually contains an element $\neq 0$.). For if $\mathbf{\Delta} = \{0\}$, then the identity map $i: \mathbf{\Delta} \rightarrow \mathbf{\Delta}$ can be represented as $i(\varepsilon) = b\varepsilon$ for any b, in violation of the uniqueness condition on b. It should be noted that, while $\mathbf{\Delta}$ does not reduce to $\{0\}$, nevertheless 0 is the only *explicitly nameable* element of $\mathbf{\Delta}$. For it is easily seen to be inconsistent to assert that $\mathbf{\Delta}$ actually contains an element $\neq 0$.

From the nondegeneracy of $\mathbf{\Delta}$ we can also (again) refute the Law of Excluded Middle in **SIA**, more particularly, we can prove

$$(*) \qquad \neg\forall\varepsilon(\varepsilon = 0 \vee \varepsilon \neq 0).$$

For we have, for $\varepsilon \in \mathbf{\Delta}$, $\varepsilon^2 = 0$, whence $\neg(\varepsilon \neq 0)$, and (*) would give $\varepsilon = 0$. So $\mathbf{\Delta}$ would be degenerate, contrary to what we have already shown. It follows from (*) that, using x and y as variables ranging over **R**,

$$\neg\forall x \forall y(x = y \vee x \neq y).$$

In a word, the identity relation is *undecidable* on **R**.

Call a binary relation S on **R** *balanced* if it satisfies

$$\forall x \forall y(\neg\neg xRy \rightarrow xRy).$$

Then the nondegeneracy of **Δ** implies that, in **SIA** the equality relation is *unbalanced.* For suppose that $=$ were balanced. Then, for any ε, it would be the case that $\neg\varepsilon \neq 0 \rightarrow \varepsilon = 0$. But we have already shown above that $\neg(\varepsilon \neq 0)$, so it would follow that $\varepsilon = 0$. This being the case for any ε, **Δ** would be degenerate.

On the other hand, in **SIA** the order relation $<$ is balanced. For suppose $\neg\neg a < b$. Then certainly $a \neq b$, since $a = b \rightarrow \neg a < b$ by irreflexivity. Therefore, by axiom O6, $a < b$ or $b < a$. The second disjunct together with $\neg\neg a < b$ and transitivity gives $\neg\neg a < a$, which contradicts $\neg a < a$. Accordingly, we are left with $a < b$. Hence $<$ is balanced.

Even allowing for the presence of intuitionistic logic, we note that the *algebraic structure* on **R** in **SIA** differs little from that of the usual classical system of real numbers. In **SIA**, **R** is equipped with the usual addition and multiplication operations under which it is a field. In particular, **R** satisfies the condition that each $x \neq 0$ has a multiplicative inverse. Notice, however, that since in **SIA** no infinitesimal (apart from 0 itself) is provably $\neq 0$, they are not required to have multiplicative inverses (a requirement which would lead to inconsistency). From a strictly algebraic standpoint, **R** in **SIA** differs from its classical counterpart only in being required to satisfy the Cancellation Principle.

The situation is otherwise, however, as regards the order structure of **R** in **SIA**. Since infinitesimals do not have multiplicative inverses, and **R** is an intuitionistic field, it must be the case that $\forall\varepsilon\neg(\varepsilon \neq 0)$, whence

$$\forall\varepsilon(\neg\varepsilon < 0 \wedge \neg\, \varepsilon > 0),$$

or equivalently

$$\forall\varepsilon(\varepsilon \leq 0 \wedge \varepsilon \geq 0).$$

It follows easily from this and the nondegeneracy of **Δ** that

$$\neg\forall x\forall y(x < y \vee y < x \vee x = y).$$

In other words the order relation $<$ on **R** in **SIA** *fails to satisfy the trichotomy law;* it is a *partial,* rather than a *total* ordering.

The axioms of **SIA** entail that **R** differs in certain key respects from its counterpart in *constructive analysis* **CA** (Sect. 4). For example, in **CA** the equality relation is balanced, while we have shown above that in **SIA** it is unbalanced. Also in **CA** the ordering relation $<$ satisfies

$$(*) \qquad \neg(x < y \vee y < x) \rightarrow x = y;$$

and this is incompatible with the axioms of **SIA**. For (*) implies

$$(**) \qquad \forall x\neg(x < 0 \vee 0 < x) \rightarrow x = 0.$$

But in **SIA** it is easy to derive

$$\forall x \in \mathbf{\Delta} \neg (x < 0 \vee 0 < x),$$

and this, together with (**), would give $\mathbf{\Delta} = \{0\}$, contradicting the nondegeneracy of $\mathbf{\Delta}$.

In **CA** the object $\mathbf{\Delta}$ is degenerate while the nondegeneracy of $\mathbf{\Delta}$ in **SIA** is one of its characteristic features.

In **SIA** cohesive subsets of $\mathbf{R}$ correspond, *grosso modo,* to connected subsets of $\mathbb{R}$ in classical analysis, that is, to intervals. This is borne out by the fact that any puncturing of $\mathbf{R}$ is *decomposable,* for it follows immediately from Axiom O6 that

$$\mathbf{R} - \{a\} = \{x : x > a\} \cup \{x : x < a\}.$$

The set $\mathbf{Q}$ of (smooth) *rational numbers* is defined as usual to be the set of all fractions of the form m/n with $m, n \in \mathbf{N}$, $n \neq 0$. The fact that $\mathbf{N}$ is cofinal in $\mathbf{R}$ ensures that $\mathbf{Q}$ is dense in $\mathbf{R}$.

The set $\mathbf{R} - \mathbf{Q}$ of *irrational* numbers is decomposable as

$$\mathbf{R} - \mathbf{Q} = [\{x : x > 0\} - \mathbf{Q}] \cup [\{x : x < 0\} - \mathbf{Q}\}.$$

This is in sharp contrast with the situation in *intuitionistic analysis* that is, **CA** augmented by Kripke's scheme, Brouwer's Continuity Principle, and bar induction. For we have observed (Sect. 4) that in intuitionistic analysis not only is any puncturing of $\mathbf{R}$ cohesive, but that this is even the case for the irrational numbers. This would seem to indicate that in some sense the continuum in **SIA** is considerably less "syrupy" than its counterpart in **SIA**.

*Comparing the Smooth and Dedekind Real Lines in **SIA**.* The usual set $\mathbb{R}$ of real numbers can be constructed in **SIA** as Dedekind cuts. A *Dedekind real* is a pair $(U, V) \in \mathcal{P}\,\mathbf{Q} \times \mathcal{P}\mathbf{Q}$ (here $\mathcal{P}\mathbf{Q}$ is the power set of $\mathbf{Q}$, the set of rational numbers) satisfying the conditions:

$$\exists x \exists y (x \in U \wedge y \in V)$$

$$U \cap V = \varnothing$$

$$\forall x (x \in U \leftrightarrow \exists y \in U . x < y)$$

$$\forall x (x \in V \leftrightarrow \exists y \in V . y < x)$$

$$\forall x \forall y (x < y \leftrightarrow x \in U \vee y \in x).$$

The set $\mathbb{R}$ of Dedekind reals as thus defined can be turned into an ordered ring (Johnstone 1977). This ring is always *constructively complete,* that is, satisfies the condition: Let A be an inhabited subset of $\mathbb{R}$ that is bounded above. Then sup A exists if and only if for all $x, y \in \mathbb{R}$ with $x < y$, either y is an upper bound for A or there exists $a \in A$ with $x < a$. (A real number b is called a *supremum*, or *least upper bound*, of A if it is an upper bound for A and if for each $\varepsilon > 0$ there exists $x \in A$ with $x > b - \varepsilon$.)

Although $\mathbb{R}$ is constructively complete, it is not conditionally complete in the classical sense. (This is because of the failure in **SIA** of the logical law $\neg p \vee \neg\neg p$. The failure of the latter in **SIA** follows immediately from the cohesiveness of **R** by considering the predicate $x \neq 0$.)

But $\mathbb{R}$ shares some features of the constructive reals not possessed by **R**, e.g.,

$$\neg\neg x = y \rightarrow x = y$$

$$x \leq y \wedge y \leq x \rightarrow x = y$$

$$x^n = 0 \rightarrow x = 0.$$

There is a natural order preserving homomorphism φ: $\mathbf{R} \rightarrow \mathbb{R}$ given by

$$\varphi(r) = (\{q \in \mathbf{Q} : q < r\}, \{q \in \mathbf{Q} : q > r\})$$

This is injective on **Q**, and embeds **Q** as the rational numbers in $\mathbb{R}$. Moreover, the kernel of φ coincides with the ideal **I** of strict infinitesimals in **R**, so φ induces an embedding of the quotient ring **R/I** into $\mathbb{R}$. **R/I** is **R** *shorn of its nilpotent infinitesimals:* it is both an intuitionistic field and an integral domain, that is, satisfies

$$\forall x(x \neq 0 \rightarrow x \text{ is invertible}) \qquad \forall x \forall y[xy = 0 \rightarrow x = 0 \vee y = 0].$$

It can be shown that φ is *surjective* – so that $\mathbf{R/I} \cong \mathbb{R}$ – precisely when **R** is constructively complete in the sense above. In that event $\mathbb{R}$ is both an intuitionistic field and an integral domain.

In **SIA** the usual *open interval topology* can be defined on $\mathbb{R}$. It can be shown (Stout 1976) that with this topology $\mathbb{R}$ is always *connected* in the sense that it cannot be partitioned into two disjoint inhabited *open* subsets. In **SIA** $\mathbb{R}$ actually inherits a stronger cohesiveness property from **R**. To see this, call a subset X of a set A *detachable* if there is a subset Y of A such that $X \cap Y = \emptyset$, $X \cup Y = A$. Now we can show that, if X is a detachable subset of the set of Dedekind reals, then $\varphi[\mathbf{R}] \subseteq X$ or $X \cap \varphi[\mathbf{R}] = \emptyset$. For suppose $X \subseteq \mathbb{R}$ detachable and define f: $\mathbb{R} \rightarrow \{0, 1\}$ by $f(x) = 1$ if $x \in X$, $f(x) = 0$ if $x \notin X$. Then $f \circ \varphi : \mathbf{R} \rightarrow \{0, 1\}$ must be constant since **R** is cohesive. If $f \circ \varphi$ is constantly 1, then $\varphi[\mathbf{R}] \subseteq X$; if constantly 0, then $A \cap \varphi[\mathbf{R}] = \emptyset$. It follows easily that if φ is surjective then $\mathbb{R}$ is itself cohesive.

*Nonstandard Analysis in **SIA**.* In certain formulations of **SIA** the system of *natural numbers* possesses some intriguing features which make it possible to introduce another type of infinitesimal – the so-called *invertible* infinitesimals – resembling those of Nonstandard Analysis.

We recall that the set **N** of smooth natural numbers is required to satisfy not the full principle of mathematical induction for arbitrary properties but only the weaker restricted induction scheme. This raises the possibility that **N** may not coincide with the set $\mathbb{N}$ of *standard natural numbers,* which is defined to be the *smallest* subset of **R** containing 0 and closed under the operation of adding 1. Now, models of **SIA** have

been constructed (Moerdijk and Reyes 1991) in which $\mathbb{N}$ is a *proper subset* of $\mathbf{N}$; accordingly the members of $\mathbf{N} - \mathbb{N}$ may be considered *nonstandard integers.* Multiplicative inverses of nonstandard integers are infinitesimals, but, being themselves invertible, they are of a different type from the (necessarily noninvertible) nilpotent infinitesimals which are basic to **SIA**.

Proceeding formally, we define the set $\mathbb{N}$ of *standard* natural numbers to be the intersection of all *inductive* subsets of $\mathbf{N}$, i.e., $\mathbb{N} = \{n \in \mathbf{N} : \forall X \subseteq \mathbf{N}[0 \in X \wedge \forall m(m \in X \rightarrow m + 1 \in X) \rightarrow n \in X]$.

$\mathbb{N}$ evidently satisfies *full induction*:

$$\forall X \subseteq \mathbf{N}[0 \in X \wedge \forall m(m \in X \rightarrow m + 1 \in X) \rightarrow X = \mathbf{N}].$$

The space of *arithmetical infinitesimals* is the set

$$\mathbf{AIN} = \{x \in \mathbf{R} : \forall n \in \mathbb{N}(-1/(n+1) < x < 1/(n+1)\}.$$

This may be considered the largest infinitesimal neighborhood of zero in **SIA**: it contains the space $\mathbf{\Delta}$ of nilsquare infinitesimals as well as the space of *invertible* or *Robinsonian infinitesimals*

$$\mathbf{IN} = \{x \in \mathbf{AIN} : x \textit{ is invertible}\}.$$

As inverses of "infinitely large" reals (i.e., reals r satisfying $\forall n \in \mathbb{N}.\ n < r \vee \forall n \in \mathbb{N}.\ r < -n$) invertible infinitesimals are the counterparts in **SIA** of the infinitesimals of Nonstandard Analysis.

To assert the existence of invertible infinitesimals is to assert that **I**N be inhabited: this is equivalent to asserting that the set $\mathbf{N} - \mathbb{N}$ of *nonstandard integers* be inhabited, or equivalently, that the following holds:

$$\exists n \in \mathbf{N} \forall m \in \mathbb{N}\ m < n.$$

When this condition is satisfied, as it is in certain models of **SIA**, we shall say that nonstandard integers, or invertible infinitesimals, are *actually present.* It is consistent to assert the actual presence of invertible infinitesimals, i.e., that **I**N be inhabited,.

One may also postulate the condition

$$\forall n \in \mathbf{N}[\forall x \in \mathbf{N} - \mathbb{N}(x < n) \rightarrow n \in \mathbb{N}],$$

i.e., "a natural number which is smaller than all nonstandard natural numbers must be standard." This is in fact equivalent to the condition that $\mathbb{N}$ be a stable subset of $\mathbf{N}$, i.e., $\mathbf{N} - (\mathbf{N} - \mathbb{N}) = \mathbb{N}$. Assuming that nonstandard integers are actually present, this latter may be understood as asserting that as many as possible of these are actually present.

In the presence of invertible infinitesimals, $\mathbb{R}$ is a nonstandard model of the reals lacking nilsquare infinitesimals. The passage via the function φ from $\mathbf{R}$ to $\mathbb{R}$

eliminates the nilsquare infinitesimals but preserves the invertible infinitesimals. When φ is onto, $\mathbb{R}$ is a cohesive *nonstandard model of the reals.*

*Consistency and Models of **SIA***. Smooth Infinitesimal Analysis is a fascinating theory with many attractive, even exotic features, but one may well ask whether is it *consistent.* Now the consistency of an axiomatic theory is usually established by producing a *model* for it, that is, providing an interpretation of the basic constituents of the theory under which all the axioms of the theory can be shown to be true. The consistency of **SIA** can be established by these means. *Models* of **SIA** are *categories* of a certain kind known as *toposes* (or *topoi*). These originated through the work of Alexander Grothendieck in algebraic geometry in the 1950s and 1960s as a generalization of the idea of a category of sheaves on a topological space (hence the term *topos*). In the late 1960s F William Lawvere and Myles Tierney formulated a simple but powerful set of first-order axioms characterizing the idea underlying Grothendieck's concept of topos. A category satisfying the Lawvere-Tierney axioms was initially known as an *elementary topos*, and later just a *topos.* Although the origin of the topos concept lay in topology and algebraic geometry, elementary toposes have deep connections with logic and set theory. In particular, a topos can be thought of as a "generalized model of set theory." The remarkable thing is that the body of laws satisfied by the logical operations in a topos – its "internal logic" – does not, in general, correspond to classical logic, but rather to the *intuitionistic* or *constructive* logic of Brouwer and Heyting in which the law of excluded middle is not affirmed.

Within a topos, mathematical concepts can be formulated, arguments carried out and constructions performed much as one does in "ordinary" set theory, only observing the rules of intuitionistic logic.

In fact, any topos may be regarded as a model of *intuitionistic set theory* **IST**. (By intuitionistic set theory **IST** is meant the theory in intuitionistic first-order logic whose axioms are the "usual" axioms of Zermelo set theory (without the Axiom of Choice), namely: Extensionality, Pairing, Union, Power set, Infinity, and Separation. For an exposition of **IST**). This means that, within a topos, mathematical constructions can be carried out as if within **IST**. In particular, the rational numbers can be defined as usual and the real numbers constructed by employing either Dedekind's procedure of making cuts in the rationals or Cantor's procedure employing equivalence classes of Cauchy sequences of rationals. While classically these two constructions lead to isomorphic results, this is *not true* in **IST**: indeed, a number of toposes have been constructed in which the ordered rings of Dedekind and Cantor reals fail to be isomorphic (Johnstone 2002).

Accordingly, in a topos-theoretic or constructive universe there is more than one candidate for the role of the mathematical continuum. The classical view that the linear continuum is a uniquely determined entity gives way to a pluralistic conception under which the continuum has a number of embodiments with essentially different properties. Even when one decides to choose the Dedekind reals as one's continuum, its properties may fall far short of those possessed by its classical counterpart. For example, in intuitionistic logic it cannot be proved that the

Dedekind reals satisfy the least upper bound property. It can be shown (Johnstone 2002) that, in a topos, the Dedekind reals possess this property exactly when the logical law $\neg(\alpha \wedge \beta) \rightarrow (\neg\alpha \vee \neg\beta)$ holds there. This is an arresting instance of the connection between logic and the properties of the mathematical continuum made visible by the shift from classical to constructive logic.

The practice of topos theory quickly spawned an associated philosophy – jocularly known as "toposophy" – whose chief tenet is the idea that, like a model of set theory, any topos may be taken as an autonomous universe of discourse or "world" in which mathematical concepts can be interpreted and constructions performed (Bell 1986, 1988). These "worlds" opened up by topos theory can have startlingly different properties from the world of classical mathematics. For example, recall that in 1924 Brouwer proved from his intuitionistic principles that every real-valued function on a closed interval of the intuitionistic continuum of real numbers is uniformly continuous; in short, that the intuitionistic continuum has the *Brouwer Property.* This is of course inconsistent with the classical account of the continuum. But a number of toposes have been constructed in which the continuum of Dedekind reals has the Brouwer property (Mac Lane and Moerdijk 1992). In these all closed intervals in the Dedekind continuum are cohesive.

Toposes which are models of **SIA** have also been constructed. The basic ideas here go back to Lawvere. In the 1960s he conceived the ideas of developing the concept of smoothness in category-theoretic terms and of employing nilpotent infinitesimals in the calculus and differential geometry. The framework he formulated is known as *synthetic differential geometry* (**SDG**). The key principle of **SIA**, what we have termed the Infinitesimal Linearity Principle, was given its explicit form by Lawvere and Anders Kock, and, as remarked above, is often called the *Kock-Lawvere axiom*. The explicit construction of topos models of **SIA** in which differential geometry can be fully developed – so-called *smooth toposes* (Moerdijk and Reyes 1991), was first achieved by Dubuc in 1979. This, no easy task, established fully the consistency of **SIA**.

*Contrasting Nonstandard Analysis with **SIA.*** **SIA** shares with Nonstandard Analysis (**NSA**) the feature that continuity is represented by the idea of "preservation of infinitesimal closeness." Nevertheless, there are a number of differences between the two approaches:

- In models of **SIA**, only smooth maps between objects are present. In models of **NSA**, all set-theoretically definable maps (including, in particular, discontinuous ones) appear.
- The logic of **SIA** is intuitionistic, while the logic of SIA is classical. (It should be pointed out, however, that constructive versions of **NSA** have been developed. See Palmgren 1998).
- In **SIA**, all curves are infinitesimally straight. Nothing resembling this is present in **NSA**.
- The nilpotency of the infinitesimals of **SIA** reduces the differential calculus to simple algebra. In **NSA** the use of infinitesimals is a disguised form of the classical limit method.

- The hyperreal line in **NSA** is obtained by augmenting the classical real line with infinitesimals (and infinite numbers), while the smooth real line **R** comes already equipped with infinitesimals.
- In any model of **NSA**, the hyperreal line $\mathbb{R}^\star$ has exactly the same set-theoretically expressible properties as does the classical real line: in particular $\mathbb{R}^\star$ is an archimedean field in the sense of that model. This means that the infinitesimals (and infinite numbers) of **NSA** are not intrinsically so in the sense of the model in which they "live," but only relative to the "standard" model with which the construction began. That is, speaking figuratively, an inhabitant of a model of **NSA** would be unable to detect the presence of infinitesimals or infinite numbers in $\mathbb{R}^\star$. This contrasts with **SIA** in two respects. First, in models of **SIA** containing invertible infinitesimals, the real line is nonarchimedean with respect to the set of standard natural numbers, which is itself an object of the model. In other words, the presence of (invertible) infinitesimals and infinite numbers would be perfectly detectable by an inhabitant of the model. And secondly, the characteristic property of nilpotency possessed by the microquantities of a model of **SIA** is an intrinsic property, perfectly identifiable within the model. In **NSA** the hyperreals have precisely the same algebraic properties as do the classical real numbers, but the smooth reals in **SIA** do not.

The differences between **NSA** and **SIA** arise because the former is essentially a theory of infinitesimal numbers designed to provide a succinct formulation of the limit concept, while the latter is, by contrast, a theory of infinitesimal geometric objects, designed to provide an intrinsic formulation of the concept of differentiability.

***SIA** and Physics*. In the past, physicists showed no hesitation in employing infinitesimal methods, In this connection we recall the words of Hermann Weyl:

> *The principle of gaining knowledge of the external world from the behaviour of its infinitesimal parts is the mainspring of the theory of knowledge in infinitesimal physics as in Riemann's geometry and, indeed, the mainspring of all the eminent work of Riemann* (Weyl 1922, p. 92).

The use of infinitesimals relied on the implicit assumption that the (physical) world is smooth, or at least that the maps encountered there are differentiable as many times as needed. For this reason **SIA** provides an ideal framework for the rigorous derivation, using infinitesimals, of results in classical physics (Bell 1998). We present two of these here.

First, we derive the *equation of continuity for fluids,* first derived in 1757 by Euler. The derivation in **SIA** will follow Euler's very closely, but the use of nilsquare infinitesimals and the Cancellation Principle will render the argument entirely rigorous.

Before we begin we require a few observations on *partial derivatives* in **SIA**. Given a function f: $\mathbf{R}^n \to \mathbf{R}$ of n variables $x_1, \ldots, x_n$, the partial derivative $\frac{\partial f}{\partial x_i}$ is defined as usual to be the derivative of the function $f(a_1, \ldots, x_i, \ldots, a_n)$ obtained by

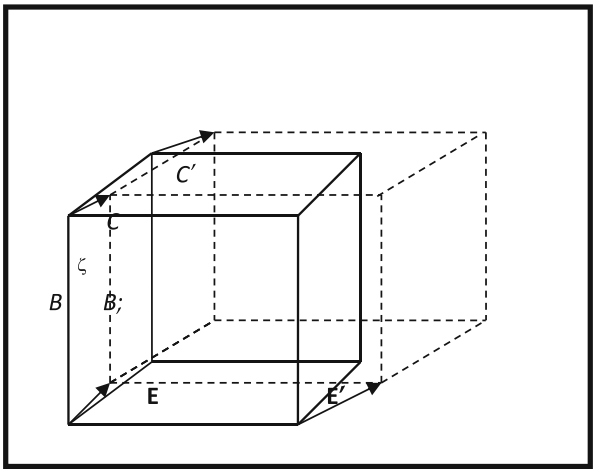
fixing the values of all the variables apart from x_i. In that case, for an arbitrary infinitesimal ε, we have

$$f(x_1, \ldots, x_i + \varepsilon, \ldots, x_n) = f(x_1, \ldots, x_n) + \varepsilon \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n). \tag{2}$$

Using the fact that $\varepsilon^2 = 0$, it is then easily shown that

$$f(x_1 + a_1\varepsilon, \ldots, x_n + a_n\varepsilon) = f(x_1, \ldots, x_n) + \varepsilon \sum_{i=1}^{n} a_i \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n). \tag{3}$$

These equations are pivotal in deriving the equation of continuity. Here we are given a inviscid fluid of varying density flowing smoothly in space. At any point $O = (x, y, z)$ in the fluid and at any time t, the fluid's density ρ and the components u, v, w of the fluid's velocity are given as functions of x, y, z, t. Following Euler, we consider the elementary volume element **E** – an infinitesimal parallelepiped – with origin O and edges OA, OB, BC of infinitesimal lengths ε, η, ζ and so of mass εηζρ:



Fluid flow during the infinitesimal time τ transforms the volume element **E** into the infinitesimal parallelepiped **E**′ (see figure above) with vertices O', A', B', C'. We first calculate the length of the side $O'A'$. Now, using (1), the rate at which A is moving away from O in the x-direction is

$$u(x + \varepsilon, y, z, t) - u(x, y, z, t) = \varepsilon \frac{\partial u}{\partial x}.$$

The change in length of OA during the infinitesimal time τ is thus $\varepsilon\tau \frac{\partial u}{\partial x}$, so that the length of $O'A'$ is

$$\varepsilon + \varepsilon\tau\frac{\partial u}{\partial x} = \varepsilon\left(1 + \tau\frac{\partial u}{\partial x}\right).$$

Similarly, the lengths of $O'B'$ and $O'C'$ are, respectively,

$$\eta\left(1 + \tau\frac{\partial v}{\partial y}\right), \zeta\left(1 + \tau\frac{\partial w}{\partial z}\right).$$

The volume of $\mathbf{E}'$ is the product of these three quantities, which, using the fact that $\tau^2 = 0$, comes out as

$$\varepsilon\eta\zeta\left[1 + \tau\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)\right]. \tag{4}$$

Since the coordinates of O' are $(x + u\tau, y + v\tau, z + w\tau)$, the fluid density ρ' at time $t + x$ is, using (3),

$$\rho + \tau\left(\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z}\right). \tag{5}$$

The mass of $\mathbf{E}'$ is then the product of (4) and (5), which, again using the fact that $\tau^2 = 0$, comes out as

$$\varepsilon\eta\zeta\rho + \varepsilon\eta\zeta\tau\left(\frac{\partial \rho}{\partial t} + \rho\frac{\partial u}{\partial x} + \rho\frac{\partial v}{\partial y} + \rho\frac{\partial w}{\partial z} + u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z}\right). \tag{6}$$

Now by the principle of conservation of mass, the masses of the fluid in $\mathbf{E}$ and $\mathbf{E}'$ are the same, so equating the mass $\varepsilon\eta\zeta\rho$ of $\mathbf{E}$ to the mass of $\mathbf{E}'$ given by (6) yields

$$\varepsilon\eta\zeta\tau\left(\frac{\partial \rho}{\partial t} + \rho\frac{\partial u}{\partial x} + \rho\frac{\partial v}{\partial y} + \rho\frac{\partial w}{\partial z} + u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z}\right) = 0.$$

The Cancellation Principle gives

$$\frac{\partial \rho}{\partial t} + \rho\frac{\partial u}{\partial x} + \rho\frac{\partial v}{\partial y} + \rho\frac{\partial w}{\partial z} + u\frac{\partial \rho}{\partial x} + v\frac{\partial \rho}{\partial y} + w\frac{\partial \rho}{\partial z} = 0,$$

i.e.,

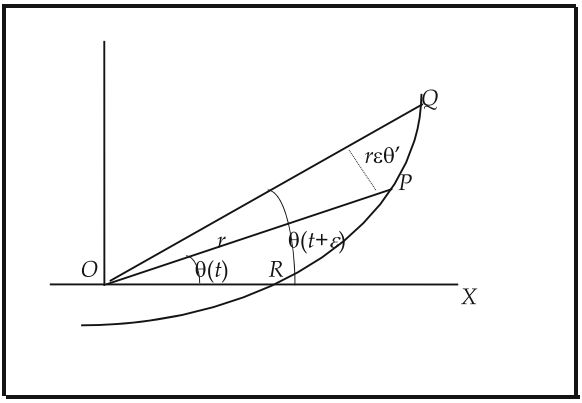
$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) + \frac{\partial}{\partial y}(\rho v)\frac{\partial}{\partial z}(\rho w) = 0.$$

This is Euler's equation of continuity.

Next, we derive the *Kepler-Newton areal law of motion under a central force.* We suppose that a particle executes plane motion under the influence of a force directed

toward some fixed point O. If P is a point on the particle's trajectory with coordinates $x, y,$ we write r for the length of the line PO and θ for the angle that it makes with the x-axis OX. Let A be the area of the sector ORP, where $\boldsymbol{R}$ is the point of intersection of the trajectory with OX. We regard x, y, r, θ as functions of a time variable t: thus

$$x = x(t), y = y(t), r = r(t), \theta = \theta(t), A = A(t).$$



Now let Q be a point on the trajectory at which the time variable has value $t + \varepsilon$, with ε in $\boldsymbol{\Delta}$ (figure above). Then by Infinitesimal Straightness the sector OPQ is a triangle of base $r(t + \varepsilon) = r + \varepsilon r'$ and height

$$r \sin [\theta(t + \varepsilon) - \theta] = r \sin \varepsilon\theta' = r\varepsilon\theta'.$$

Here we have used the fact in **SIA** $\sin \varepsilon = \varepsilon$ for infinitesimal ε (Bell 1998).

The area of OPQ is accordingly

$$2 \text{ base} \times \text{height} = 2(r + \varepsilon r')r\varepsilon\theta' = 2\left(r^2\varepsilon\theta' + \varepsilon^2 r r'\theta'\right) = 2r^2\varepsilon\theta'.$$

Therefore

$$\varepsilon A'(t) = A(t + \varepsilon) - A(t) = \textit{area } OPQ = 2\varepsilon r^2\theta',$$

so that, cancelling ε,

$$A'(t) = 2r^2\theta'. \qquad (*)$$

Now let $H = H(t)$ be the acceleration toward O induced by the force. Resolving the acceleration along and normal to OX, we have

$$x'' = H \cos\theta \qquad y'' = H \sin\theta.$$

Also $x = r \cos\theta$, $y = r \sin\theta$. Hence

$$yx'' = Hy\cos\theta = Hr\sin\theta\cos\theta \qquad xy'' = Hx\sin\theta = Hr\sin\theta\cos\theta,$$

from which we infer that

$$(xy' - yx')' = xy'' - yx'' = 0.$$

Hence

$$xy' - yx' = k, \qquad (**)$$

where k is a constant.

Finally, from $x = r\cos\theta$, $y = r\sin\theta$, it follows in the usual way that

$$xy' - yx' = r^2\theta',$$

and hence, by (**) and (*), that

$$A'(t) = 2k.$$

Assuming $A(0) = 0$, we conclude that

$$A(t) = 2kt.$$

Thus the radius vector joining the body to the point of origin sweeps out equal areas in equal times (Kepler's law).

There is an intriguing use of infinitesimals in Einstein's celebrated 1905 paper *On the Electrodynamics of Moving Bodies* (Einstein et al. 1952) in which the special theory of relativity is first formulated. In deriving the Lorentz transformations from the principle of the constancy of the velocity of light, Einstein obtains the following equation for the time coordinate $\tau(x', y, z, t)$ of a moving frame:

(i)
$$\frac{1}{2}\left[\tau(0,0,0,t) + \tau\left(0,0,0,t + \frac{x'}{c-v} + \frac{x'}{c+v}\right)\right] = \tau\left(x',0,0,t + \frac{x'}{c-v}\right).$$

He continues:

Hence, if x' be chosen infinitesimally small,

(ii)
$$\frac{1}{2}\left(\frac{1}{c-v} + \frac{1}{c+v}\right)\frac{\partial\tau}{\partial t} = \frac{\partial\tau}{\partial x'} + \frac{1}{c+v}\frac{\partial\tau}{\partial t},$$

or

$$\frac{\partial\tau}{\partial x'} + \frac{v}{c^2 - v^2}\frac{\partial\tau}{\partial t} = 0.$$

Now the derivation of equation (ii) from equation (i) can be simply and rigorously carried out in **SIA** by *choosing x' to be a nilsquare infinitesimal* ε. For then (i) becomes

$$\frac{1}{2}\left[\tau(0,0,0,t)+\tau\left(0,0,0,t+\varepsilon\left(\frac{1}{c-v}+\frac{1}{c+v}\right)\right)\right]=\tau\left(\varepsilon,0,0,t+\frac{\varepsilon}{c-v}\right).$$

From this we get, using Eq. (1) above,

$$\tau(0,0,0,t)+\frac{1}{2}\varepsilon\left(\frac{1}{c-v}+\frac{1}{c+v}\right)\frac{\partial\tau}{\partial t}=\tau(0,0,0,t)+\varepsilon\left(\frac{\partial\tau}{\partial x'}+\frac{1}{c-v}\frac{\partial\tau}{\partial t}\right).$$

So

$$\frac{1}{2}\varepsilon\left(\frac{1}{c-v}+\frac{1}{c+v}\right)\frac{\partial\tau}{\partial t}=\varepsilon\left(\frac{\partial\tau}{\partial x'}+\frac{1}{c-v}\frac{\partial\tau}{\partial t}\right),$$

and (ii) follows by the Cancellation Principle.

Spacetime metrics have some intriguing properties in **SIA**. In a spacetime the metric can be written in the form

$$(*)\qquad ds^2=\sum g_{\mu\nu}dx_\mu dx_\nu \qquad \mu,\nu=1,2,3,4.$$

In the classical setting (*) is in fact an abbreviation for an equation involving derivatives and the "differentials" ds and dx_μ are not really quantities at all. What form does this equation take in **SIA**? Notice that the "differentials" cannot be taken as infinitesimals in the sense of **SIA** since all the squared terms would vanish. But the equation does have a very natural form in **SIA**. Here is an informal way of obtaining it.

Think of the dx_μ as being multiples $k_\mu e$ of some small quantity e. Then (*) becomes

$$ds^2=e^2\sum g_{\mu\nu}k_\mu k_\nu,$$

so that

$$ds=e\sqrt{\sum g_{\mu\nu}k_\mu k_\nu}.$$

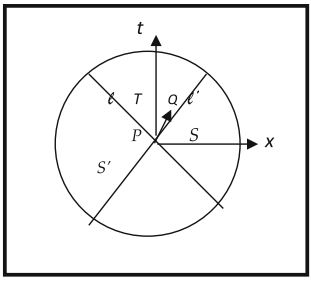
Now replace e by a nilsquare infinitesimal ε. Then we obtain the *metric relation:* in **SIA**

$$ds=\varepsilon\sqrt{\sum g_{\mu\nu}k_\mu k_\nu}.$$

This tells us that the "infinitesimal distance" ds between a point P with coordinates (x_1, x_2, x_3, x_4) and an infinitesimally near point Q with coordinate (x_1 + $k_1\varepsilon$,

$x_2 + k_2\varepsilon, x_3 + k_3\varepsilon, x_4 + k_4\varepsilon)$ is $\varepsilon\sqrt{\Sigma g_{\mu\nu}k_\mu k_\nu}$. Here a curious situation arises. For when the "infinitesimal interval" ds between P and Q is timelike (or lightlike), the quantity $\Sigma g_{\mu\nu}k_\mu k_\nu$ is nonnegative, so that its square root is a real number. In this case ds may be written as εd, where d is a real number. On the other hand, if ds is spacelike, then $\Sigma g_{\mu\nu}k_\mu k_\nu$ is negative, so that its square root is imaginary. In this case, then, ds assumes the form $i\varepsilon d$, where d is a real number (and, of course $i = \sqrt{-1}$). On comparing these we see that, if we take ε as the "infinitesimal unit" for measuring infinitesimal *timelike* distances, then iε serves as the "imaginary infinitesimal unit" for measuring infinitesimal *spacelike* distances.

For purposes of illustration (see figure below), let us restrict the spacetime to two dimensions (x, t), and assume that the metric takes the simple form $ds^2 = dt^2 - dx^2$. The infinitesimal light cone at a point P divides the infinitesimal neighborhood at P into a timelike region T and a spacelike region S bounded by the null lines l and l, respectively (see figure below). If we take P as origin of coordinates, a typical point Q in this neighborhood will have coordinates ($a\varepsilon$, $b\varepsilon$) with a and b real numbers: if $|b| > |a|$, Q lies in T; if $a = b$, P lies on l or l′; if $|a| < |b|$, P lies in S. If we write $d = \sqrt{|\, a^2 - b^2 \,|}$, then in the first case, the infinitesimal distance between P and Q is εd, in the second, it is 0, and in the third it is $i\varepsilon d$.



At the beginning of the twentieth century Poincaré and Minkowski introduced "ict" to replace the "t" coordinate so as to make the metric of relativistic spacetime positive definite. This was purely a matter of formal convenience and was later rejected by (general) relativists. In conventional physics one never works with nilpotent quantities so it is always possible to replace formal imaginaries by their (negative) squares. But spacetime theory in **SIA** *forces* one to use imaginary units, since, infinitesimally, one cannot "square oneself out of trouble." This being the case, it would seem that, infinitesimally the dictum *Farewell to ict* (Misner et al. 1973) needs to be replaced by

$$Vale\text{“}ict\text{”}, ave\text{“}i\varepsilon\text{”}!$$

We conclude this section with a speculation. The infinitesimal domain Δ can be seen to be "tiny" in the order-theoretic sense. For, using ε, η, as variables ranging over Δ, it is not hard to show that

$$\forall\varepsilon\forall\eta\neg(\varepsilon < \eta \vee \eta < \varepsilon)^*.$$

whence

$$\forall\varepsilon\forall\eta(\varepsilon \leq \eta \wedge \eta \leq \varepsilon).$$

In particular, the members of $\mathbf{\Delta}$ are all simultaneously ≤ 0 and ≥ 0 but cannot (because of the nondegeneracy of $\mathbf{\Delta}$) be shown to coincide with zero.

In his book *Just Six Numbers* (Rees 2001) the astrophysicist Martin Rees comments on the microstructure of space and time, and the possibility of developing a theory of quantum gravity. In particular he says:

> *Some theorists are more willing to speculate than others. But even the boldest acknowledge the "Planck scales" as an ultimate barrier. We cannot measure distances smaller than the Planck length [about 10^{19} times smaller than a proton]. We cannot distinguish two events (or even decide which came first) when the time interval between them is less than the Planck time (about 10^{-43} seconds).*

On this account, Planck scales seem very similar in certain respects to $\mathbf{\Delta}$. In particular, the sentence (*) above seems to be an exact embodiment of the idea that we cannot decide of two "events" in $\mathbf{\Delta}$ which came first; in fact it makes the stronger assertion that actually neither comes "first."

Could $\mathbf{\Delta}$ serve as a suitable model for "Planck scales"? While $\mathbf{\Delta}$ is unquestionably small enough to play the role, it inhabits a domain in which everything is smooth and continuous, while Planck scales live in the quantum world which, if not outright discrete, is far from being universally continuous. So if Planck scales could indeed be modeled by microneighborhoods in **SIA**, then one might begin to suspect that the quantum microworld, the Planck regime – smaller, in Rees's words, "than atoms by just as much as atoms are smaller than stars" – is not, like the world of atoms, discrete, but instead continuous like the world of stars. This would be a major victory for the Continuous in its long struggle with the Discrete.

References

Bell JL (1986) From absolute to local mathematics. Synthese 69:409–426

Bell JL (1988) Toposes and local set theories: an introduction. Oxford logic guides 14. Clarendon Press, Oxford. Reprinted by Dove, 2008

Bell JL (1998) A primer of infinitesimal analysis. Cambridge University Press. Second edition 2008

Bell JL (2001) The continuum in smooth infinitesimal analysis. Schuster, Berger and Osswald 2001:19–24

Bell JL, Machover M (1977) A course in mathematical logic. North-Holland

Benacerraf P, Putnam H (1977) Philosophy of mathematics: selected readings, 2nd edn. Cambridge University Press

Bishop E, Bridges D (1985) Constructive analysis. Springer, Berlin

Brentano F (1974) Psychology from an empirical standpoint. Humanities Press

Brentano F (1988) Philosophical investigations on space, time and the continuum. Tr. Smith, B. Croom Helm

Bridges D (1999) Constructive mathematics: a foundation for computable analysis. Theor Comput Sci 219:95–109
Bridges D, Richman F (1987) Varieties of constructive mathematics. Cambridge University Press
Cantor G (1872) Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math Ann 5:123–132
Cantor G (1878) Ein Beitrag zur Mannigfaltigkeitslehre. J Reine Angew Math 84:242–258
Dauben J (1979) Georg Cantor: his mathematics and philosophy of the infinite. Harvard University Press
Dummett M (1977) Elements of intuitionism. Clarendon Press, Oxford
Ehrlich P (ed) (1994) Real Numbers, generalizations of the reals, and theories of continua. Kluwer
Ehrlich P (1994a) All numbers great and small. In Ehrlich, ed (1994), pp 239–258
Einstein A et al (1952) The Principle of Relativity: a Collection of Original Memoirs on the Special and General Theory of Relativity. Tr. Perrett and Jeffrey, Dover. Reprint of 1923 Methuen edition
Ewald W (1999) From Kant to Hilbert, A source book in the foundations of mathematics. Vols I and II. Oxford University Press
Fine HB (1917) Ratio, proportion and measurement in the elements of Euclid. Ann Math Second Series 19(1):70–76
Fisher G (1981) The infinite and infinitesimal quantities of du Bois-Reymond and their reception. Arch Hist Exact Sci 24(2):101–163
Folina J (1992) Poincaré and the philosophy of mathematics. St. Martin's Press
Fraenkel A, Bar-Hillel Y, Levy A (1973) Foundations of set theory, 2nd edn. North-Holland, Amsterdam
Hallett M (1984) Cantarian set theory and limitation of size. Clarendon Press, Oxford
Hobson EW (1957) The theory of functions of a real variable. Dover
Johnstone PT (1977) Topos theory. Academic Press
Johnstone PT (2002) Sketches of an elephant: a topos theory compendium, Vols I and II. Oxford Logic Guides Vols 43 and 44. Clarendon Press
Keisler H (1994) The hyperreal line. In Ehrlich (1994), pp 207–238
Kline M (1972) Mathematical thought from ancient to modern times. 3 volumes. Oxford University Press
Kneebone GT (1963) Mathematical logic and the foundations of mathematics: an introductory survey. Van Nostrand, London
Laugwitz AH, Schmieden C (1958) Eine Erweiterung der Infinitesimalrechnung. Math Z 69:1–39
Mac Lane S, Moerdijk I (1992) Sheaves in geometry and logic: a first introduction to topos theory. Springer
Mancosu P (1998) From Brouwer to Hilbert: the debate on the foundations of mathematics in the 1920s. Clarendon Press
McShane EJ, Botts TA (1959) Real analysis. Van Nostrand
Misner C, Thorne K, Wheeler J (1973) Gravitation. Freeman
Moerdijk I (1995) A model for intuitionistic non-standard arithmetic. Ann Pure Appl Logic 73:297–325
Moerdijk I, Reyes GE (1991) Models for smooth infinitesimal analysis. Springer
Palmgren E (1998) Developments in constructive nonstandard analysis. Bull Symb Log 4:233–272
Palmgren E (2001) Unifying constructive and nonstandard analysis. Schuster Berger Osswald 2001:167–183
Peirce CS (1976) The new elements of mathematics, Vol III. Edited by Carolyn Eisele. Mouton Publishers and Humanities Press
Peirce CS (1992) Reasoning and the logic of things. Edited by Kenneth Laine Ketner. Harvard University Press
Poincaré H (1946) Foundations of science. Tr. Halstead. Science Press
Pyle A (1997) Atomism and its critics. Thoemmes Press
Rees M (2001) Just six numbers: the deep forces that shape the universe. Basic Books
Robinson A (1996) Non-standard analysis. Princeton University Press

Russell B (1964) The principles of mathematics. Allen and Unwin
Stout L (1976) Topological properties of the real numbers object in a topos. Cahiers Topologie Géom. Différentielle 17(3):295–326
Weyl H (1921) On the new foundational crisis in mathematics. (English translation of 'Über der neue Grundslagenkrise der Mathematik,' Mathematische Zeitschrift 10, 1921, 37–79.) In Mancosu (1998) pp 86–122
Weyl H (1949) Philosophy of Mathematics and Natural Science. Princeton University Press. (An expanded Engish version of Philosophie der Mathematik und Naturwissenschaft, Leibniz Verlag, 1927)
Weyl H (1950) Space-Time-Matter. Tr. Brose, H. Dover. (English translation of Raum, Zeit, Materie, Springer, 1918)

On Bolzano and Greek Concepts of Continuity

Piotr Błaszczyk and Marlena Fila

Contents

1 Introduction 1564
2 Continuity in the Greek Philosophy 1566
3 Continuous vs Discrete in the Greek Mathematics 1568
3.1 Magnitudes 1568
3.2 Numbers 1569
3.3 Mechanical Curves 1572
4 Toward Modern Mathematics 1573
5 Curves Consisting of Points 1574
5.1 Descartes' Curves Described by Rotating Lines 1574
6 Continuity and Intermediate Value Theorem 1577
6.1 Intermediate Value Theorem 1578
7 The Structure of *Rein Analytischer Beweis* 1579
7.1 Two-Functions Pattern 1580
7.2 Lagrange on IVT_p 1581
7.3 From Movement to Continuous Function 1582
8 Continuous Function 1582
8.1 *Preface's* Version of a Continuous Function 1582
8.2 Continuity and Sign Preserving Property 1582
8.3 From the Continuity of Polynomial to $\varepsilon\delta$ Continuity 1583
9 Continuity of Totally Ordered Set 1584
9.1 Proof of IVT_f 1584
9.2 Least Upper Bound Principle 1585
9.3 Magnitudes as Real Numbers 1586
10 Function vs Line 1589

The first author is supported by the National Science Centre, Poland grant 018/31/B/HS1/03896. The second author is supported by the National Science Centre, Poland grant 2019/33/N/HS1/02045.

P. Błaszczyk · M. Fila (✉)
Institute of Mathematics, Pedagogical University of Cracow, Cracow, Poland
e-mail: piotr.blaszczyk@up.krakow.pl; marlena.fila@up.krakow.pl

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_18

11 Topological Continuum 1589
11.1 Bolzanian Continuum 1589
11.2 Cantor on a Continuum in the Metric Space $\mathbb{R}^n$ 1590
12 Summary 1591
13 Cross-Reference 1592
References 1592

Abstract

In ancient Greek, continuity referred to lines, figures, and solids, or time, place, and motion. All these objects were considered continuous in the same sense: they were divisible into infinitely divisible parts. Bolzano 1817 *Rein analytischer Beweis* (Bolzano, Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein ent-gegengesetzes Resultat gewäihren, wenigstens eine reelle Wurzel der Gleichung liege. Gottlieb Hasse, Prague. In: [53], 253–277, 1817a) gave rise to two brand new ideas: $\varepsilon\delta$ concept of continuous function and continuity of total order in terms of the least upper bound principle. The first was related to motion, the second to a straight line. The chapter aims to reveal the mathematical motives that lead Bolzano to these ideas. In the first part, we discuss Aristotle's definition and examples of continuity, Euclid's concept of magnitude and number, and Pappus' account of mechanical curves. Then we show how the idea of the punctiform curve emerged in Descartes' (Descartes, La Géométrie. Jan Maire, Leiden, 1637). The second part includes a detailed analysis of (Bolzano, Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein ent-gegengesetzes Resultat gewäihren, wenigstens eine reelle Wurzel der Gleichung liege. Gottlieb Hasse, Prague. In: [53], 253–277, 1817a).

Keywords

Continuity · Continuous function · Least upper bound · Magnitudes · Real numbers · Topological continuum

1 Introduction

In modern mathematics, continuity refers to a function or a total order. Continuity of function, whether voiced in terms of $\varepsilon\delta$ or limits of sequences, is the core concept of calculus. Advanced courses prefer a more general, topological definition: a function is continuous when the inverse image of an open set is open.

The characteristic of continuous order depends on whether one considers a totally ordered set or a totally ordered algebraic structure. For the first case, $(X, <)$ has to be a separable space (in the order topology) and satisfy Dedekind cut principle (C1, as formalized below).[1] For the second, we focus on an ordered field, then the continuity means the so-called completeness of real numbers. Regarding real numbers, two

[1] This result is due to Cantor (Cantor 1895), though, originally stated in alternative terms; see (Błaszczyk and Fila 2020a), Sect. 7.

approaches developed at the turn of the nineteenth and twentieth century still play a role in mathematical practice, namely, the axiomatic account initiated by (Hilbert 1900) and the construction of real numbers, as exemplified in (Cantor 1872; Dedekind 1872; Heine 1872; Méray 1869; Weber 1895). Due to analysis of Bolzano (Bolzano 1817a), we prefer the first option. Thus,

Definition 1
The field of real numbers is an ordered field $(\mathbb{F}, +, \cdot, 0, 1, <)$, *in which every Dedekind cut* (L, U) *in* $(\mathbb{F}, <)$ *satisfies the following condition*.[2]

(C1) $(\exists x \in \mathbb{F})(\forall y \in L)(\forall z \in U)(y \leqslant x \leqslant z)$.

This definition refers to the fact that there is only one (up to isomorphism) complete ordered field (see (Birkhoff and Mac Lane 1977), 105–106).

Here are other equivalent versions of C1 we will refer to in sections dedicated to Bolzano (Bolzano 1817a).

(C2) If $A \subset \mathbb{F}$ is a nonempty set which is bounded above, then there exists $a \in \mathbb{F}$ such that $a = \sup A$.
(C3) The field is Archimedean and every Cauchy (fundamental) sequence $(a_n) \subset \mathbb{F}$ has a limit in $\mathbb{F}$.

The Archimedean axiom can also take alternative forms. We will need the following two:

(A1) $(\forall x \in \mathbb{F})(\exists n \in \mathbb{N})(n > x)$.
(A2) $(\forall x, y \in \mathbb{F})(\exists q \in \mathbb{Q})(x < y \Rightarrow x < q < y)$.

Another equivalent form of C1 which plays a role when (Bolzano 1817a) is under scrutiny is the Intermediate Value Theorem (see Sect. 6 below).

By combining the continuity of function and the concept of an ordered field, one gets a topological field, that is, an algebraic field $(\mathbb{F}, +, \cdot, 0, 1, \tau)$, where τ is a topology on $\mathbb{F}$, in which addition, multiplication, and the operation of the inverse element x^{-1} are all continuous in the topology τ. Pontriagin (Pontriagin 1932) provides an insightful characterization of real and complex numbers in terms of topological fields: any locally compact and connected topological field is isomorphic to the field of real or complex numbers.[3] Pontriagin's theorem, thus, provides grounds for the uniqueness of real numbers in the class of topological fields.[4]

[2] It is the standard definition, let us just recall that a pair of nonempty sets A,B is a Dedekind cut in $(\mathbb{F}, <)$ iff (1) $A \bigcup B = \mathbb{F}$, (2) $(\forall x \in A)(\forall y \in B)(x < y)$. See Cohen and Ehrlich (1963).

[3] See Pontriagin (1932), p. 164. Pontriagin theorem also provides a topological characterization of the field of quaternions, which is a non-commutative.

[4] The concept of topological dimension enables determining the difference between fields of real and complex numbers.

Since axioms characterize both ordered and topological fields, there are two unique characteristics of real numbers: in terms of an ordered and topological field.

The idea of continuity is also related to a topological continuum, that is, a compact and connected set. Yet another thread in this history of continuity concerns curves. The concepts of continuous function and topological continuum cover different aspects of that issue (see (Błaszczyk and Fila 2020b), §3).

This variety of ideas mirrors the variety of techniques applied in modern studies of the ancient old phenomenon of continuity. In Greek philosophy, however, it was simple. Examination of continuous objects meant the geometrical studies of lines, figures, and solids, or physical of time, place, and motion. Moreover, all these objects were continuous in the same sense: they were divisible into infinitely divisible parts, while divisibility, interpreted in modern terms, meant addition.

From ancient to modern times, mathematics has undergone a tremendous transformation. Our analysis of (Bolzano 1817a) discusses only two issues. First, the replacement of the ancient technique of proportion by processing formulas according to the rules of an ordered field. Second, the replacement of the mechanical curves of ancient mathematics by punctiform functions. In chapter *Descartes' transformation of the Greek notion of proportionality* (in this *Handbook*) we explore in part the first topic. Throughout (Bolzano 1817a), Bolzano applies laws of an ordered field regularly, some of them explicitly, others implicitly. Considering his explicit use of the LUB principle, variables x, y, z in (Bolzano 1817a) range over the domain of real numbers. In this way, real numbers frame Bolzano's study of continuous function and continuous line.

The plan of the chapter is as follows. In Sects. 2 and 3, we discuss the idea of continuity in Aristotle's philosophy and ancient mathematics. Section 4 is a methodological interlude. In Sect. 5, we show how the Greek concept of a curve was turned into a punctiform line of modern mathematics. Sections 6, 7, 8, 9, and 10 include a detailed analysis of (Bolzano 1817a). Section 11 presents Bolzano and Cantor's concept of the topological continuum in a metric space.

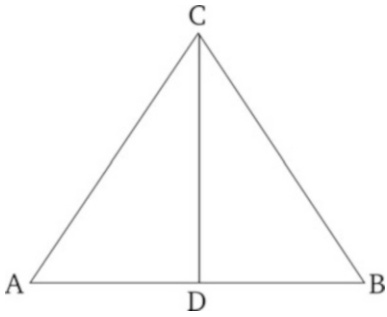
2 Continuity in the Greek Philosophy

Boyer (1949) overviews competing trends in the ancient philosophy of continuity. In this section, we present a sample of Aristotle's statements concerning continuity and discreteness. His definitions and examples parallel the mathematical treatment of magnitudes and numbers. We show they are a philosophical characterization of entities studied by Euclid, Archimedes, and Pappus, to mention only those who we refer to in sections that follow.

Let us start with the famous definition: "I call continuous (συνεχές) that which is divisible into infinitely (ἀεί) divisible divisibles" (*Physics* VI.2, 232b24–25).[5]

[5] English translations of Aristotle by Benjamin Wilck.

Fig. 1 The *Elements*, I.10



A model example of a continuous object is a segment line. In proposition I.10 of the *Elements*, Euclid shows how to cut a given line in half. Figure 1 represents the accompanying diagram.

Line segment *AB* is divided into other line segments *AD*, *DB*. Given the same construction applies to *AD*, and, say D_1, cuts *AD* in half, *AD* will be divided into AD_1, and D_1D. This procedure can be iterated ad infinitum, and on each stage, it will output new line segments. That is how Euclid's proof supports Aristotle's definition: a line segment is divisible into infinitely divisible line segments.

In the following quotation, Aristotle exemplifies continuous and discrete entities: "Of quantities (ποσόν) some are discrete (διωρισμένον), others continuous. [. . .] Discrete are, for instance, number and language, continuous are lines, surfaces, bodies, and, besides these, time and place" (*Categories* 6, 4b20–25). Here, Aristotle characterizes geometrical objects and numbers as objects of substantially different nature. Elsewhere he alludes, they need a different way of studying.[6] Indeed, in Euclid's *Elements*, geometrical objects and numbers are studied by different techniques in separate books. In the next section, we point out substantial mathematical differences between these domains.

In *Physics*, books V to VII, Aristotle also takes motion to be continuous. At *Physics* V.4, 228a20 he says plainly "every motion is continuous." The same is dedicated to Zeno paradox *Physics*, VI.2. It starts as follows: "Since every magnitude (μέγεθος) is divisible into magnitudes (for we have shown that it is impossible for anything continuous (συνεχές) to be composed of indivisibles, and every magnitude is continuous), it is necessary for that which is faster to traverse a greater magnitude in an equal time, and equal magnitude in less time, and a greater magnitude in less time, just as some people have defined 'faster'" (*Physics* VI.2, 232a23–27).

Continuity of motion will turn to be relevant in our discussion of mechanical curves and continuity of function.

[6] At *Posterior Analytics* I.7, 75a38–39, Aristotle states that one cannot cross from one kind (e.g., numbers) to another (e.g., magnitudes). From *Posterior Analytics* I.7, 75b9 follows that this applies to mathematical proofs: one cannot derive theorems on magnitudes from principles in the domain of numbers. Interestingly, Bolzano, raised in the scholastic tradition, refers to that methodological rule in (Bolzano 1817a).

3 Continuous vs Discrete in the Greek Mathematics

3.1 Magnitudes

In ancient Greek mathematics, the general term μέγεθος; (magnitude, quantity) covers line segments, triangles, convex polygons, circles, angles, arcs, and solids.[7] Euclid's book V develops rigorous treatment of magnitudes, ratios, and proportions.

We formalize magnitudes of the same kind (line segments being of one kind, triangles being of another, etc.) as an additive semi-group with a total order, $(M, +, <)$, characterized by the following five axioms:

E1 $(\forall x, y)(\exists n \in \mathbb{N})(nx > y)$,
E2 $(\forall x, y)(\exists z)(x < y \Rightarrow x + z = y)$,
E3 $(\forall x, y, z)(x < y \Rightarrow x + z < y + z)$,
E4 $(\forall x)(\forall n \in \mathbb{N})(\exists y)(x = ny)$,
E5 $(\forall x, y, z)(\exists v)(x : y :: z : v)$.
The term nx is defined by $nx = \underbrace{x + \ldots + x}_{n-times}$.

The transitivity and the trichotomy laws characterize the relation *greater than*, $<$. Equality of line segments, angles, and arcs means congruence. Equality of figures should be interpreted in Euclid's theory of area, as explained in (Błaszczyk 2018); equality of solids involves a similar theory developed in book XII of the *Elements.*

E1 interprets definition V.4, the so-called Archimedean axiom: "Magnitudes [x, y] are said to have a ratio with respect to one another which, being multiplied [nx] (*πολλαπλασιαζόμενα*), are capable of exceeding one another [$nx > y$]."[8] The term *nx,* as it stands here, interprets the phrase *multiple of the magnitude.*

Axioms E2 and E3 apply all through book V, for example, in V.8 and V.25. E2 also occurs in Aristotle's writings. They are, then, easily identified.

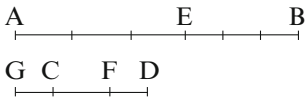
From the logical point of view, E4 is inessential since it follows from the remaining four axioms, that is, E1–E3, and E5. Yet, it bears a deep philosophical sense; we consider it a formalization of Aristotle's definition of continuity.

E4 implicitly applies in proposition V.5. Therein Euclid proves the following relation (see Fig. 2): "[L]et the magnitude AB be the same multiple of the magnitude CD that the (part) taken away AE (is) of the (part) taken away CF (respectively). I say that the remainder EB will also be the same multiple of the remainder FD as that which the whole AB (is) of the whole CD (respectively)." In symbols: $AB = kCD$, $AE = kCF \Rightarrow EB = kFD$.

[7] Euclid's definition I.8 provoked a philosophical debate whether angles are magnitudes; see (Wilck 2020), p. 381. From the perspective of book V, angles are terms in proportions on a par with line segments or circles. However, one can raise strictly mathematical objections whether angles satisfy axiom E3.

[8] English translations of the Elements after (Euclid 2007).

Fig. 2 The *Elements*, V.5



In the proof, he assumes such a point G that the relation obtains $EB = kGC$. In other words, for any EB and any k, there exists GC such that $EB = kGC$.

Proposition VI.9, where Euclid shows how to divide a line segment into any number of parts, demonstrates the infinite divisibility of line segments. In Book V, line segments in diagrams represent magnitudes of any kind, yet, E4 stands for divisibility of magnitudes rather than line segments alone.

E5 represents the so-called fourth proportional axiom. In book V, it applies in proposition V.18. Next, it is explored in the method of exhaustion (see, e.g., XII.1).

For a symbolic representation of proportion, we adopt the term x: y:: z: v. It was introduced already in the seventeenth century. The definition of proportion, V.5, is verbal. We interpret it by the following formula:

$$x : \ y :: \ \mathrm{z} : \ v \Leftrightarrow_{df} (\forall m, n \in \mathbb{N})[(nx >_1 my \Rightarrow nz >_2 mv) \wedge$$

$$\wedge (nx = my \rightarrow nz = mv) \wedge (\mathrm{nx} <_1 \mathrm{my} \Rightarrow \mathrm{nz} <_2 \mathrm{mv})].$$

The assumption regarding magnitudes x, y, on the one hand, and z, v, on the other, being of the same kind is formalized by $x, y \in \mathfrak{M}_1 = (M_1, +, <_1)$, and z, $v \in \mathfrak{M}_2 = (M_2, +, <_2)$.

Definition V.7 introduces a relation *greater than* between pairs of magnitudes. We interpret it as follows:

$$x : \ y \succ z : \ v \Leftrightarrow_{df} (\exists m, n \in \mathbb{N})[(nx >_1 my) \wedge (nz \leqslant_2 mv)].$$

Thus, the structure of magnitudes is an ordered semigroup equipped with the 4-ary relation of proportion $(M, +, ::, <)$.[9]

3.2 Numbers

Euclid's *Elements* books VII to IX provide foundations for the ancient Greek number theory.[10] Number (ἀριθμός) means the modern integer greater than 1. Zero does not occur in the *Elements,* while 1 is the so-called monad (μονάς). The first three definitions of book VII clarify the relation between monads and numbers. They read:

[9] For detailed analysis leading to axioms for book V, see (Beckmann 1967; Błaszczyk and Mrówka 2013), or (Mueller 2006). See also Descartes' transformation of Greek notion of proportionality in this Handbook.

[10] More advanced arithmetic was developed by Diophantus, see (Christianidis and Oaks 2021).

"A unit (μονάς) is (that) according to which each existing (thing) is said (to be) one. And a number (is) a multitude (πλῆθος) composed of units. A number is part of a number, the lesser of the greater [$n < m$], when it measures the greater."

We formalize definition VII.2 as follows:

$$n = \underbrace{1 + \ldots + 1}_{n-times}. \tag{1}$$

The term *n measures m* means that for some $k \in \mathbb{N}$ the equality holds

$$m = kn = \underbrace{n + \ldots + n}_{k-times}.$$

Multiplication of numbers as introduced by definition VII.15, we formalize by

$$n \cdot m = \underbrace{m + \ldots + m.}_{n-times}$$

Hence, *n measures m* means *n* divides *m* in the ring of integers.

Euclid's numbers, then, form a structure (A, +, ·, <), where addition and relation *greater than* are primitive notions characterized by the above axioms E2, E3. *Greater than* is a transitive relation that satisfies the trichotomy law. Both addition and multiplication are associative and commutative operations.

The infinite divisibility of magnitudes, demonstrated as the axiom E4, does not hold in the domain of numbers. On an ontological level, a number is a discrete object for it consists of units, and units have no parts.

Next to these arithmetic operations, in definition VII.20, Euclid introduces a proportion of numbers. It reads:

> Numbers are proportional when the first is the same multiple, or the same part, or the same parts, of the second that the third (is) of the fourth.

We formalize it by

$$m : \ n :: \ k : \ l \Leftrightarrow \exists \mathrm{p}(pn = m, pl = k) \vee \exists p(pm = n, pk = 1) \vee$$

$$\vee (\exists u_1, u_2)(\exists \mathrm{p}, q)(m = pu_1, n = qu_1, k = pu_2, l = qu_2).$$

The proportion of numbers, contrary to magnitudes, does not refer to the relation *greater than.* There is also no counterpart of definition V.7 in Euclid's arithmetic. Moreover, axiom E5 does not hold in the domain of numbers. Indeed, the fourth proportional usually is linked with continuous rather than discrete structures. Weber (Weber 1895), for example, develops a theory of real numbers based on a tricky account of proportions. In his approach, the fourth proportional follows from the

Dedekind cut principle. Then, through the fourth proportional, Weber defines addition and multiplication of ratios, and as a result, real numbers.[11]

Finally, Euclidean numbers are well-ordered, that is, in every non-empty subset of A, there is the least element. This principle, explicitly or implicitly, applies in books VII–IX.[12] Let us start with the train of propositions VII.30–32. They anticipate the fundamental theorem of arithmetic. VII.31 reads:

> Every composite number is measured by some prime number.

The *reductio ad absurdum* proof runs as follows. Let A be a composite number. By definition VII.13, it is measured by some number, say B. If B is prime, the work is done. If it is composite, another number, say C, measures it. Since C measures B, it also measures A. Reiterations of this procedure produce an endless chain of composite numbers $A > B > C > \ldots$ such that a less divides each one greater. At this stage, Euclid concludes: "the number A will be measured by an infinity of numbers (ἄπειροι ἀριθμοί), each of which is less than the preceding. The very thing is impossible for numbers."

Thus, Euclid explicitly assumes that there are no infinite decreasing sequences of numbers. This condition is equivalent to the well-ordering of numbers.[13] The well-ordering of numbers plus E2 and E3 entail the axiom E1.

In proposition VII.2, Euclid determines the greatest common measure of numbers AB and CD (not prime to one another). The crucial step in his procedure includes the division algorithm (in modern terms), or an algorithm of continuous subtraction (ἀνθυφαίρεσις), as Euclid refers to it. It runs as follows: "if CD does not measure AB then some number will remain from AB and CD, the lesser being continually subtracted, in turn, from the greater, which will measure the (number) preceding it. For a unit will not be left."[14] In symbols

$$AB = kCD + EA, CD > EA > 1.$$

Given the continuous subtraction applies to CD and EA, another equality obtains

$$CD = mEA + FC, EA > F\,C.$$

Finally, Euclid assumes "let CF measure AE," which is

[11] For Weber's concept of magnitude see (Błaszczyk and Fila 2020a).

[12] In proposition V.8, Euclid assumes there is the least element in the set {k: kx > y}, given x, y are magnitudes. Euclid's argument is as follows. He considers multiples of D, L = 2D,M = 3D,N = 4D, and declares "the quadruple multiple N of D — the first greater than K." It means $M \leqslant K \leqslant N$. In this context, 'any multiples' are 2D, 3D, 4D because Euclid does not use the big quantifier. In a modern interpretation, that argument would be like this N = (k + 1)D, M = kD.

[13] This equivalence is easily demonstrable, yet it requires the use of the axiom of choice (see (Kuratowski and Mostowski 1976), 225).

[14] E1 and the well-ordering of numbers justify the division algorithm.

$$EA = nF\ C.$$

Then he shows that CF measures both AB and CD, and it is also their greatest common measure.

The proof, thus, includes what we currently know as the so-called Euclidean algorithm (actually three while loop iterations of that algorithm). Euclid assumes that the algorithm stops regardless of input numbers. Indeed, it is guaranteed by the continuous decreasing of remainders, plus the assumption that every decreasing series of numbers is finite.[15]

Interestingly, in book X of the *Elements*, the algorithm of continuous subtraction is applied to line segments. When it stops, two lines are considered commensurable, otherwise incommensurable (see proposition X.2). According to Knorr (Knorr 1975), that algorithm was known long before Euclid. It looks like Euclid applied it also to numbers. With the rule that every decreasing series of numbers is finite, it turned into an effective procedure. One could speculate further that definition VII.2 could play a heuristic role in that process. Although it hardly bears any mathematical sense, it could have contributed to establishing the well-ordering of numbers.

To sum up this section, Euclidean numbers form a structure $(A, +, \cdot, <)$ which satisfies axioms E1–E3. Moreover, the total order of numbers is well-ordering.[16]

3.3 Mechanical Curves

Book IV of Pappus' *Collection* introduces quadratrix, spiral, and conchoid. A characteristic of each curve repeats the same pattern of genesis followed by symptomata. The genesis part determines a movement of a point, say C, being a composition of a turning straight line and a rectilinear motion of another straight line or point. As a result, it is the moving point C that describes a curve. The symptomata part defines the decisive feature of the curve employed in any further analysis. While genesis refers to the non-mathematized concept of movement, symptomata provides purely mathematical characteristics. Pappus applies symptomata straightforwardly. However, he seeks to relate phenomena of different kinds: non-mathematized movement and mathematized genesis of a curve.

Here is how Pappus sets forth conchoida:

[15] For the modern mathematical account of the Euclidean algorithm see (Birkhoff and Mac Lane 1977), 18–22.

[16] (Malmendier 1975) provides a concise axiomatic interpretation of Euclid's book VII. (Mueller 2006) interprets Euclid's books VII to IX, though it is rather an exercise in lengthy discussions. Both (Malmendier 1975) and (Mueller 2006) are focused on a formalization of proportions of numbers. Neither (Malmendier 1975) nor (Mueller 2006) discusses the role of well-ordering in Euclid's arithmetic.

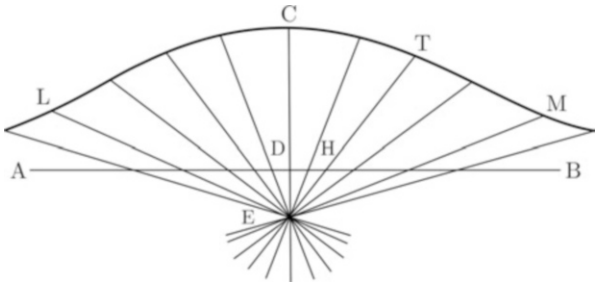


Fig. 3 Pappus, *Collection*, Book IV, p. 126

Genesis: "the straight line CDEZ travels along the straight line ADB, dragged via the point E in such a way that [. . .] it is obvious that the point C will describe a line such as LCM is" (Pappus of Alexandria 2010, 126).

Symptomata: "[. . .] whenever some straight line from the point E toward the line meets it, the [straight line] cut off between the straight line AB and the line LCM is equal to the straight line CD" (Pappus of Alexandria 2010), 126) (Fig. 3).

Compared with current courses, a definition of curve plays the role of symptomata, while genesis is usually only a heuristic supplement. Moreover, the modern curve is a set of points that meet some conditions. For instance, the modern definition of conchoid reads: "Let S be any curve and let A be a fixed point. If a straight line is drawn through A to meet S at Q, and if P and P' are points on this line such that $P'Q = PQ = k$ (constant), the locus of P and P' is called *a conchoid of S with respect to A*" (Lockwood 1961), 127).

When S is a straight line, and A does not lie on S, the resulting curve is the so-called conchoid of Nichomedes. On the Cartesian plane $\mathbb{R}^2$, when the straight line rotates around the point $O = (0, 0)$, and S is the vertical line $x = a$, the conchoid is the following set of points $\{(x, y) \in \mathbb{R}^2 : (x^2 + y^2)(x - a)^2 = k^2x^2\}$. In that context, the early modern term *locus* takes the meaning of set. In the set-theoretical framework, the cardinality of the set $\mathbb{R}^2$ represents a multitude of points on a curve.[17]

4 Toward Modern Mathematics

Ancient philosophy characterization of magnitudes and numbers concern individual objects: a line segment and a number. Their mathematical counterparts seem very simple. It is the axiom E4 and formula (1), respectively. However, the above characteristic of magnitudes and numbers consists of axioms E1–E5 and a list of properties including E1–E3 plus the well-ordering. It is because we adopt the modern perspective and shift a focus from individual objects to the algebra (structure) of magnitudes

[17] Here is a video presenting the ancient conchoida http://www.eudoxos.pl/kartezjusz/konstrukcje-nieklasyczne/.

(M, +, ::, <) or numbers (A, +, ·, <). There is no doubt that in Greek mathematics, a comparison *A greater than B* did not pair with the concept of a totality of As, and Bs ordered by a relation *greater than*. On the other hand, it is also obvious that thinking in terms of operations and their properties precedes the set-theoretic paradigm.

The modern concept of continuity refers to structures, whether an ordered field, a topological field, a totally ordered set, or a topological space. As regards continuity of a function, it assumes that a function consists of points. This idea predates Bolzano's definition of a continuous function by 180 years.

5 Curves Consisting of Points

Brouwer's choice sequences aiming to represent real numbers sparked the early twentieth-century philosophical debate over the nature of continuum. Intuitionists proudly declared that contrary to classical mathematics, their continuum did not fall apart into individual points. Bell (Bell 2005) essentially reiterates the same criticism regarding the entire twentieth-century mathematics.

The term continuum in that debate stands for the domain of functions studied by the calculus techniques and implicitly is related to the set-theoretical account of mainstream mathematics. Indeed, 1872s papers of Heine, Cantor, and Dedekind introducing real numbers implied mathematical objects were point sets. The twentieth-century achievements have completed the project of founding mathematics on the set theory. Nevertheless, the second wave of the foundational studies, that is, the axiomatic account of real numbers developed by[18] (Hilbert 1900; Artin and Schreier 1926; Tarski 1931) does not rely on the idea of a punctiform number-line. Moreover, in the context of elementary geometry, straight lines consisting of points occur in axiomatic systems developed by (Birkhooff 1932; Borsuk and Szmielew 1960). Yet (Hilbert 1903) does not go beyond the ancient Greek idea of points lying on a line, while (Tarski 1959) develops geometry without the concept of a straight line.

Bolzano's ideas of continuous total order and continuous function apply to point set objects. In the context of elementary geometry, the idea of straight lines consisting of points was explicitly introduced by Bolzano already in (Bolzano 1804). His later works applied it in his studies of the topological continuum, time, and space (see (Bolzano 1851), §38). Yet, it is not Bolzano who introduced this new perspective. It all started with Descartes and his formula representing a curve.

5.1 Descartes' Curves Described by Rotating Lines

In ancient Greek mathematics, a line is not a composition of points, yet we find specific relations between lines and points. (1) A straight line is usually a line segment with its

[18] Hilbert 1900; Artin and Schreier 1926; Tarski 1931.

ends. In a proposition and on an accompanying diagram, it is represented by its endpoints such as AB. Conics and other curves, when bounded, are also taken with their endpoints. Points lying on a half-line with one end or an infinite line with no ends represent these objects. Circles, as a rule, are identified by three random points. (2) On lines, there are random points employed in proofs. Capital letters such as *A* and *B* represent them. (3) Each intersection of lines determines a point represented by a capital letter. When, for example, in proposition I.10, a finite straight line *AB* is cut in half, a new point occurs, *D*. It is the intersection of *AB* and the bisector of the angle $\angle ACB$. New straight lines, *AD* and *DB* share point *D* in common (see Fig. 1 above).

These three relations significantly affected Descartes' concept of the curve. In book II of (Descartes 1637), he studies curves described by intersecting lines: a turning straight line and a curve that moves along some other straight line. That procedure carries out three times in (Descartes 1637).[19] The resulting curves are represented in Figs. 4 and 5.

In the first case, the turning line *GCL* intersects with the straight line *CKN*. In the second, it intersects with the circle. Here is how Descartes portrays these moving lines and their intersections.[20]

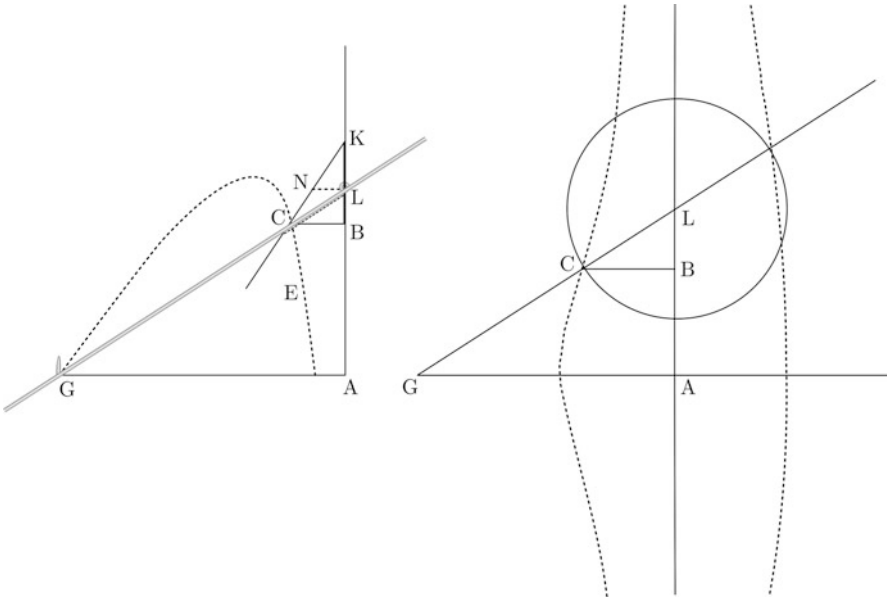


Fig. 4 *La Géométrie*, p. 321 (left). Descartes' conchoid (right)

[19] One can also add to that group curves drawn by mezolabium of Descartes.

[20] Here is a video presenting Descartes' hyperbola http://www.eudoxos.pl/hiperbola-kartezjusza/. Here is another one presenting his conchoida http://www.eudoxos.pl/konchoida/.

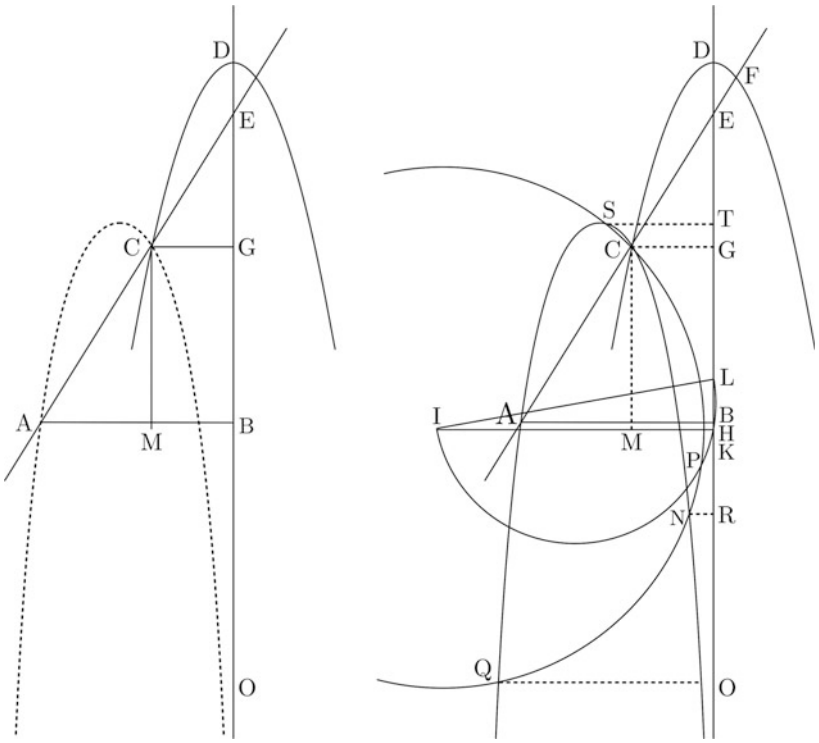


Fig. 5 Descrates' cubic parabola (left). *La Géométrie*, p. 404 (right)

> Suppose the curve EC to be described by the intersection (*decrite par l'intersection*) of the ruler GL and the rectilinear plane figure CNKL [. . .] the curve EC belongs to the first class, it being, in fact, a hyperbola (Smith and Lathan 2007, pp. 51, 55).
>
> [I]f CNK be a circle having its center at L, we shall describe the first conchoid of the ancients (Smith and Lathan 2007, p. 55).

In Fig. 5, the diagram on the left, the turning line ACE intersects with the parabola CD. The resulting cubic parabola (marked with a dashed line) applies in the construction of roots of a sixth-degree polynomial. The genesis of the curve plays a role in that problem (see Fig. 5, the diagram on the right).

In (Descartes 1637), the formula of a curve replaced Pappus' symptomata. Moreover, the genesis of the curve is not relevant in deriving the formula. In the case of hyperbola, it relies on a relationship between coordinates x, y of a specific intersection point, $C = C(x, y)$. That relation is encoded in a polynomial equation $f(x, y)$ – it is the formula of the curve. Yet, $f(x, y)$ is derived from relations between similar triangles, as represented in Fig. 6 below. Here is how Descartes describes the derivation process:

"Then I say that as NL is to LK, or as c is to b, so CB, or y, is to BK, which is therefore equal to $\frac{b}{c}y$. Then BL is equal to $\frac{b}{c}y - b$, and AL is equal to $x + \frac{b}{c}y - b$.

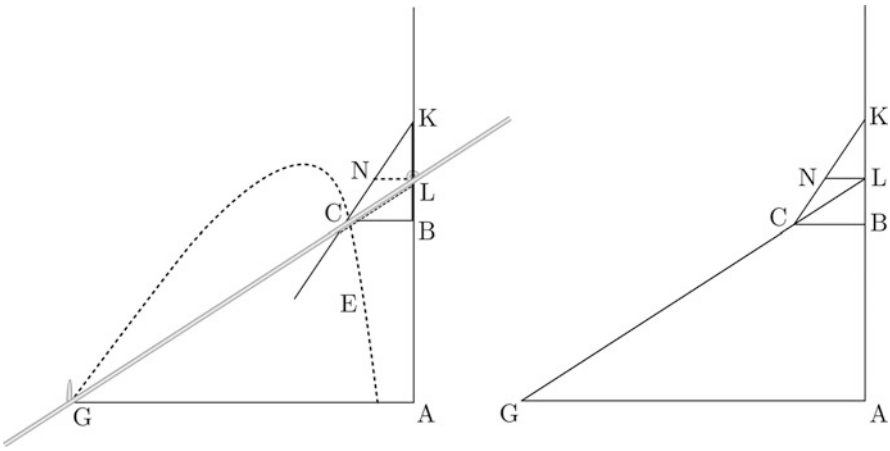


Fig. 6 *La Géométrie*, p. 320 (left). Similar triangles used in deriving the formula of the hyperbola (right)

Moreover, as CB to LB, that is, as y is to $^{-}\frac{b}{c}y - b$, so AG or a is to LA or $x + \frac{b}{c}y - b$ Multiplying the second by the third, we get $\frac{ab}{c}y - ab$ equal to $xy + \frac{b}{c}yy - by$, which is obtained by multiplying the first by the last. Therefore, the required equation is $yy = cy - \frac{cx}{b}y + ay - ac$" (Smith and Lathan 2007, 51).

In sum, the hyperbola consists of intersection points $C = C(x, y)$, and no other points. A required *locus* is characterized by the formula $f(x, y)$. However, the continuity of that curve comes from the continuous motion of the point C tracing the curve. Although the formula $f(x, y)$ represents the curve, there is no idea of a domain of f. Curve is still continuous in the ancient style. The modern concept of continuous function requires both formula and domain.

6 Continuity and Intermediate Value Theorem

(Bolzano 1817a) plays a crucial role in the bifurcation of the ancient concept of continuity. In this paper, Bolzano provides modern characteristics of a continuous function and continuity of a totally ordered set. These two concepts are involved in his proof of the Intermediate Value Theorem, IVT, in short. Interestingly, Bolzano seeks to prove what we would call a specific case of this theorem, namely, the IVT for a polynomial, IVT for a continuous function is a lemma in his study. Nowadays, his lemma is the main theorem. Arguably, it is so, for Bolzano's characterization of a continuous function was a brand new idea at that time, while it is the basic concept of current calculus courses. To be clear, it took more than five decades to rephrase Bolzano's interpretation of continuous function into the modern $\varepsilon\delta$ continuity. Even throughout his paper, the account of continuous function evolves significantly: it starts with the *Preface's* definition adopting the two quantifiers form $\forall\exists$, then, through the proof of the so-called sign preserving principle where it finds a three

quantifiers version ∀∃∀, finally it reaches yet another recast in proving the continuity of a polynomial. Moreover, in (Bolzano 2000), Bolzano's sense of discontinuity is by no means a negation of any of these three versions of continuity. Therefore, instead of focusing on the explicit definition, we prefer to study its actual applications.

While the paper discusses the concept of a continuous function carefully, the import of continuity of total order disappears in a maze of inferences between the least upper bound principle, the so-called Cauchy completeness, and implicit use of the Archimedean axiom. In fact, from the perspective of the general plan of (Bolzano 1817a), the continuous order plays but a side role. Yet, it is hard to overestimate its mathematical importance. We address this topic in Sect. 9.

To build a perspective, we start this part with a remainder of the modern proof of IVT.[21]

6.1 Intermediate Value Theorem

Here is the standard wording of IVT: If a continuous function f changes its sign in an interval of real numbers [a, b], then there exists an argument c in (a, b) such that $f(c) = 0$. In symbols,

$$f : [a, b] \mapsto \mathbb{R}, f(a) \cdot f(b) < 0 \Rightarrow \exists c \in (a, b)[f(c) = 0]. \tag{2}$$

The classic proof emulates the proof developed by Bolzano. It goes like that. Assume that $f(a) < 0 < f(b)$. Put $S = \{x \in [a, b] : f(x) < 0\}$. The set S is not empty, since $a \in S$; it is also bounded above by b. Therefore we can apply the least upper bound principle, LUB. Let

$$c = \sup S. \tag{3}$$

There are three possibilities: $f(c) < 0, f(c) = 0, f(c) > 0$. If $f(c) < 0$, then due to the assumed continuity of f, the sign of $f(c)$ is preserved in some neighborhood of c, that is:

$$(\exists \delta)(\forall x)[x \in (c - \delta, c + \delta) \Rightarrow f(x) < 0]. \tag{4}$$

From (4) it follows that there is x such that $c < x$ and $f(x) < 0$. This contradicts (3). The case $f(c) > 0$ is treated similarly. Hence, $f(c) = 0$ (see Fig. 7).

This proof, thus, relies on two different concepts of continuity: the definition (3) appeals to the continuity (completeness) of an order of real numbers, formula (4)

[21] Błaszczyk (2016) shows that IVTf is equivalent to (C1). This result enables a new interpretation of (Bolzano 1817a) in terms of real-closed fields. For the concept of real-closed field and its relation to IVTp see (Jacobson 1969), ch. VI. In what follows, we present, say, a conservative reading of (Bolzano 1817a).

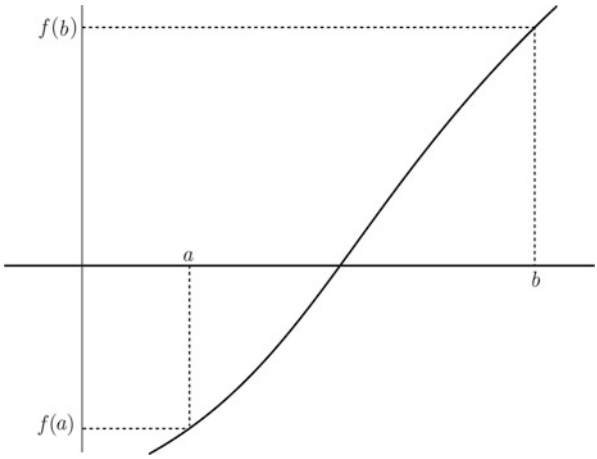


Fig. 7 Diagram for IVT

follows from the assumed continuity of *f*. In fact, (Bolzano 1817a) provides the first-ever formulation of the LUB principle and the first-ever εδ-kind characteristic of continuous function.

7 The Structure of *Rein Analytischer Beweis*

Figure 8 represents the logical structure of (Bolzano 1817a). In this paper, Bolzano aims to show the Intermediate Value Theorem for polynomials, IVT_P in short. The proof applies, on the one hand, the theorem that a polynomial is a continuous function, *cont of poly*, on the other, the Intermediate Value Theorem for a continuous function, IVT_f. Both of these results apply the definition of a continuous function, *df of cont*; the first directly, the other via the so-called sign preserving property, SPP, as explained by formula (4) above. The proof of IVT_f also employs LUB. Bolzano proves it by referring to Cauchy Completeness, CC. He also seeks to prove CC, which seems incorrect from the modern axiomatic perspective. Throughout the paper, he implicitly refers to two different forms of the Archimedean axiom, namely: (1) $(\forall r \in \mathbb{R})(\exists n \in \mathbb{N})(n > r)$, (2)$(\forall D, d \in \mathbb{R}_+)(\exists n \in \mathbb{N})\left(\frac{D}{2^n} < d\right)$.[22]

Hypothetically, it was possible to prove IVT_P by applying the concept of continuity straight to a polynomial with no reference to the general concept of function.[23]

[22] In (Bolzano 1817b), Bolzano calls axioms of Archimedes the following conditions: "I. Every curved line is longer than the straight line lying between the same end-points. II. Of two curved lines which are both concave on one side the enclosing one is longer than the enclosed one. III. If a curved and a plane surface have the same boundaries the former is greater than the latter. IV. Of two curved surfaces which are both concave on one side the enclosing one is greater than the enclosed one" (Russ 2004), 281. These axioms mimic assumptions of Archimedes' treaty On the Sphere and Cylinder, see (Heath 2002), pp. 3–4.

[23] That is how Birkhoff and Mac Lane prove the Fundamental Theorem of Algebra 150 years after Bolzano; see (Birkhoff and Mac Lane 1977), ch. 4, or (Błaszczyk 2016), p. 19.

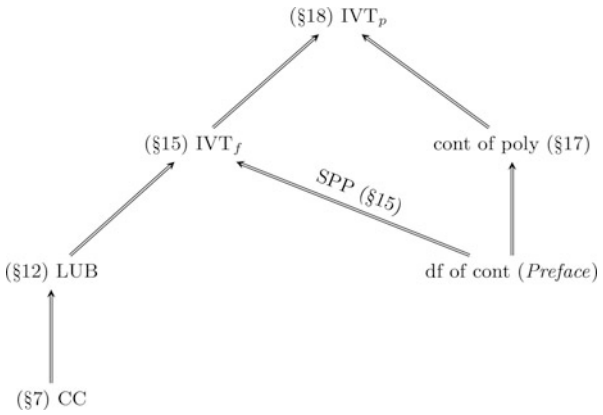


Fig. 8 Plan of Bolzano's *Rein analytischer Beweis*

Whether applied to a function or a polynomial, the analytic form of continuity introduced by Bolzano was then a brand new idea.[24]

7.1 Two-Functions Pattern

Contrary to the modern version, Bolzano's IVT$_f$ involves two functions as it reads: "If two functions of x, fx and φx, vary according to the law of continuity either for all values of x or for all those lying between α and β, and furthermore if $f\alpha < \varphi\alpha$ and $f\beta > \varphi\beta$, then there is always a certain value of x between α and β for which $fx = \varphi x$" (Russ 2004), 274).

Thus, for continuous functions f, φ on the interval $[\alpha, \beta]$ the following implication holds

$$f(\alpha) < \varphi(\alpha), f(\beta) > \varphi(\beta) \Rightarrow \exists x \in (\alpha, \beta)[f(x) = \varphi(x)]. \tag{5}$$

Putting $f := f - \varphi$, we get IVT in version (2). In Sect. 9.3 below, we will address the issue of whether f, φ are real maps.

As for the main theorem of (Bolzano 1817a), it reads: "If a function of the form $x^n + ax^{n-1} + bx^{n-2} + \ldots + px + q$, in which n denotes a positive integer, takes a *positive* value for $x = \alpha$, but a *negative* value for $x = \beta$, then the equation $x^n + ax^{n-1} + bx^{n-2} + \ldots + px + q = 0$ has at least one *real root* lying between α and β" (Russ 2004), 276).

Its proof is as follows. Let $W(x) = x^n + ax^{n-1} + bx^{n-2} + \ldots + px + q$, let $0 < \alpha < \beta$, and $W(\alpha) > 0 > W(\beta)$. Then, some terms of $W(\alpha)$ are positive, some are negative.

[24] There is an ongoing debate over a relation between Bolzano and Cauchy definitions of a continuous function; see (Bair et al. 2020) for the latest position in this dispute. In Sect. 8.3, we show how Otto Stolz obtained the modern $\varepsilon\delta$ definition of a continuous function by interpreting Bolzano's proof of the continuity of polynomial.

Moreover, "it is clear that exactly the same terms of the function which are positive or negative for $x = \alpha$ also keep this sign for $x = \beta$ and for all values of x lying between α and β" (Russ 2004, 277).

Indeed, since x^k are positive for every integer k, due to the condition $W(\alpha) > 0 > W(\beta)$, some coefficients among $a,b,\ldots,p,q$ are positive, some are negative. For instance, suppose $a,b > 0$, $p,q < 0$, then $W(x) = x^n + ax^{n-1} + bx^{n-2}$ — (| p| x+| q|). The sum of *positive* terms forms a polynomial $\varphi(x)$, the sum of *negative* terms forms another polynomial $f(x)$. To give a more specific example, when $W(x) = x^2 - x^3$, $\alpha = ½$, $\beta = 2$, then $\varphi(x) = x^2$, $f(x) = x^3$.

Since polynomials are continuous functions, and $f(\alpha) < \varphi(\alpha)$, $f(\beta) > \varphi(\beta)$, by applying (5) Bolzano gets the result $f(x) = \varphi(x)$, for some x. The other two cases, that is, $\alpha < 0 < \beta$, and $\alpha < \beta < 0$ are reduced to the first one.

This seemingly odd strategy, both in the proof of IVT_f and IVT_p, results from the fact that Bolzano adopts Lagrange's scheme of proof introduced in his 1808s (Lagrange 1808).

7.2 Lagrange on IVT_p

In *Preface* to (Bolzano 1817a), Bolzano provides detailed analysis of previous proofs of IVT_P developed by Euler, Gauss, and others. Some of them apply mathematical assumptions questioned by Bolzano others rely on diagrams or non-mathematical concepts such as movement.

Regarding proofs referring to movement, Bolzano alludes to Lagrange's (Lagrange 1808). In this treaty, Lagrange interprets IVT_p in terms of a root of a polynomial $P - Q$. Yet, in the crucial step of his proof, he refers to the following intuition of race: when a quicker runner outruns a slower one, they must meet at some point. It is a point x, Lagrange concluded, in which $(P - Q)(x) = 0$. Here is the essential part of that argument.

> Let us present the generally proposed equation as $P - Q = 0$, where P is the sum of all terms that has a plus sign, and Q is the sum of all those that have a minus sign. Let's first assume that two numbers p and q are positive and that q is greater than p. Assuming further $x = p$, we have $P - Q < 0$, assuming $x = q$, we have $P - Q > 0$. In the first case, P will clearly be $< Q$, and in the second, P will be $> Q$. Thus, from the forms of P and Q, which contain only positive quantities and integral and positive powers, it is obvious that these quantities necessarily increase as x increases, and increasing x through all invisible degrees, from p to q, also increase by invisible degrees, but in this way that P will increase more than Q, because of the smaller it was, the larger it becomes. Thus, there must necessarily be a term between the two values p and q, where P will be equal to Q, **just like two moving bodies supposed to move on the same line in the same direction, starting from two different points at the same time, reach two other points at the same time, but in such a way that the one who, at first, was behind the other then is ahead, must necessarily meet on their way**. The value x, which makes P equal to Q, will, therefore, be one of the roots of the equation, and will be between the values of p and q. Similarly, if we take $x = p$, we get $P - Q > 0$, and when we take $x = q$, we get $P - Q < 0$; in the first case we will get $Q < P$, and in the second $Q > P$; and by increasing x from p up to q, the amount of Q will increase more than the amount of P, and equals it at a point between p and q. (Lagrange 1808, 98–99; emphasis added, English translation by P.B.)

7.3 From Movement to Continuous Function

Bolzano summarizes Lagrange's two-polynomial trick involved in the proof and then concludes: "This is further illustrated by the example of the *motion* of two bodies, of which one is initially *behind* the other and later *ahead* of the other. It necessarily follows that at one time it must have *passed* the other" (Russ 2004, 255). Then, he adds: "No one will want to deny that the concept of *time*, as well as that of *motion*, is just as alien to general mathematics as the concept of *space*" (Russ 2004, 255). In this context, Bolzano brings in his definition of continuous function. It is designed to supersede Lagrange's argument based on the parallelism between polynomials and moving bodies.

8 Continuous Function

8.1 *Preface's* Version of a Continuous Function

The definition of continuity (*Stetigkeit*) of function at point x given in *Preface* reads: *"function fx varies according to the law of continuity for all values of x inside or outside certain limits means only that, if x is any such value* ***the difference*** $f(x+\omega) - fx$ ***can be made smaller than any given quantity****, provided ω can be taken as small as we please* [...] or $f(x+\omega) = fx + \Omega$" (Russ 2004, 256; emphasizes added). Regarding the emphasized phrase and its supposed clarification, this definition is ambiguous. We formalize it in two ways as follows:

$$(\forall\Omega)(\exists\omega)(\ f(x+\omega) - fx < \Omega), \tag{6}$$

$$(\forall\Omega)(\exists\omega)(\ f(x+\omega) - fx = \Omega). \tag{7}$$

In this way, regardless of which formula is adopted, continuity of function replaces continuity of motion. Due to this substitution, Bolzano managed to introduce a kind of algebra of movements: assuming functions f, g represent movements, terms $f+g$, or $f-g$ have clear mathematical meaning, even though the concept of function is not yet defined.

8.2 Continuity and Sign Preserving Property

In (Bolzano 1817a) section § 15, the definition of a continuous function is employed, firstly, to infer SPP, secondly, to show the continuity of a polynomial. In the first case, Bolzano seeks to derive the following statement: "Now because $f\alpha < \varphi\,\alpha$, if u denotes a positive quantity which can become as small as we please, then also $f(\alpha+\omega) < \varphi(\alpha+\omega)$" (Russ 2004, 274).

Here the definition (6/7) is applied at the point α to two functions f and φ to get the same ω for each function. That is why Bolzano modifies the original formula.[25] To this end, he specifies the phrase "ω can be taken as small as we please." The amendment reads: "whenever we take $\omega < i$, then it must be possible to make $f(\alpha + \omega) - f\alpha$ and $\varphi(\alpha + \omega) - \varphi\alpha$ as small as we please if ω is taken small enough" (Russ 2004), 274). In symbols

$$(\forall\Omega)(\exists i)(\forall\omega)(0 < \omega < i \Rightarrow f(\alpha + \omega) - f(\alpha) = \Omega). \tag{8}$$

In what follows, Bolzano develops version (7) of his definition. Nonetheless, from (8), through some intricate arguments, he finally gets

$$(\exists i)(\forall\omega)(0 < \omega < i \Rightarrow f(\alpha + \omega) < \varphi(\alpha + \omega)) \tag{9}$$

Formula (9) represents Bolzano's version of SPP.

8.3 From the Continuity of Polynomial to *εδ* Continuity

In (Bolzano 1817a), the continuity of a polynomial function is stated as follows: "Every function of the form $a + bx^m + cx^n + \ldots + px^r$, in which $m,n,\ldots,r$ denote positive integer exponents, is a quantity which varies *according to the law of continuity* for all values of x" (Russ 2004, 275).

Put $W(x) = a + bx^m + cx^n + \ldots + px^r$. By binomial theorem, the difference $W(x + \omega) - W(x)$ takes the following form:[26]

$$W(x + \omega) - W(x) = \omega S = \omega\left(mbx^{m-1} + \ldots + p\omega^{r-1}\right).$$

where S represents a polynomial of two variables $S = S(x,\omega)$. As for S, Bolzano observes: "the whole expression is always *finite* for every value of x and ω" (Russ 2004, 276). Since the difference $W(x + \omega) - W(x)$ is determined at some point x, the sum S depends only on ω, that is $S = S(\omega)$. Now, the inequality obtains

$$S(\omega_1) \leq \mid mbx^{m-1} \mid + \ldots + \mid p\omega_1^{r-1} \mid . \tag{10}$$

The formula (10) represents the following sentence: "So if we denote by S the quantity arrived at by putting a definite value, say ω_1, for ω in all the individual terms of the expression, and summing as if they all had the same sign, then the actual value of this expression for this particular ω_1 is certainly not $> S$" (Russ 2004, 276).

[25] In this paragraph, Bolzano also copes with the claim that f − φ is continuous at α, given f and φ are continuous at α.

[26] There is a misprint in the formula representing W(x + ω) −W(x) in (Bolzano 1817a), p. 57. Instead of $p\omega^{r-1}$, there is ω^{r-1}, similarly, instead of ω^{m-1}, there should be $b\omega^{m-1}$. And instead of ω^{n-1} there should be $c\omega^{n-1}$.

Finally, Bolzano concludes with this claim: "Hence if it is required that the change in the function $a + bx^m + cx^n + \ldots + px^r$ should be $< D$, then we take an ω that is at the same time $< \omega_1$ and also $< \frac{D}{S}$. Then $\omega.S$ and, all the more, the product of ω with a quantity which is $< S$, must be $< D$" (Russ 2004, 276).

We formalize this argument by the following formula:

$$(\forall D)(\exists \omega_1)(\forall \omega)[0 < \omega < \omega_1 \Rightarrow |W(x+\omega) - W(x)| < D]. \tag{11}$$

It includes the now standard trick of the theory of limits to take ω fulfilling two conditions: it is smaller than ω_1 and $\frac{D}{S}$ at the same time.

From Bolzano's perspective, formula (11) follows, in some way, from (6/7). Logically, within the framework of an ordered field, it does not. However, it provides both a specific interpretation of continuity of a function and arithmetic of continuous functions. The latter consists of proofs of this kind: $f \pm g$, fg, and f/g are continuous at some point a, given f and g are continuous at a.

Interestingly, it is just in the context of the arithmetic of continuous functions that Stolz (Stolz 1885) refers to (Bolzano 1817a) and presents the following definition: "function $f(x)$ defined for all x in the interval $(a - d, a + d)$ is called continuous for the finite value $x = a$ when continuous limit processes is obtained: [if] $\lim x = a \pm 0$, [then] $\lim f(x) = f(a)$ [...] that is, for any $\varepsilon > 0$ there must be assigned a number $\delta > 0$ such that whenever $|x - a| < \delta$ then $|f(x) - f(a) < \varepsilon$" (Stolz 1885, 179, English translation by P.B.). Obviously, the second clause is formalized by

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)(|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon).$$

Stolz declares that he adopts Bolzano's definition (6/7); see (Stolz 1885), pp. 179, 340. In fact, he takes a formalized version of Bolzano's proof of continuity of a polynomial, that is, formula (11).

9 Continuity of Totally Ordered Set

In the set theory, the theory of ordered fields, or real analysis, continuity usually gets the name of completeness. Specified, it turns into Dedekind cut principle (Dedekind completeness), the least upper bound principle, or Cauchy completeness. Since these three versions are not equivalent (see Sect. 1 above) we prefer to use specific terms rather than the general term continuity.

9.1 Proof of IVT$_f$

In Bolzano (1817a) §15, Bolzano proves the theorem summarized by our formula (5). Figure 9 represents its general scheme.

Bolzano considers cases depending on how α, β are located relative to zero. We will trace the case $0 < \alpha < \beta$. Let $f(\alpha) < \varphi(\alpha)$. From SPP it follows that there is i such that

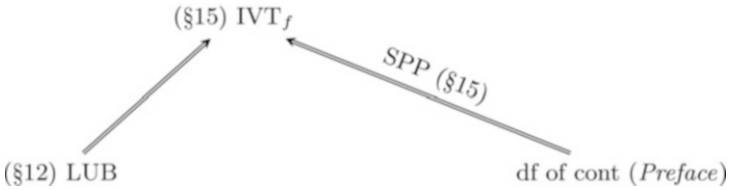


Fig. 9 Sections of *Rein analytischer Beweis* involved in the proof of IVT$_f$

$$(\forall\omega)[0 < \omega < i \Rightarrow f(\alpha + \omega) < \phi(\alpha + \omega)].$$

Then, Bolzano concludes: "Therefore it can be asserted that for *all* values of *u smaller* than a certain value the two functions $f(\alpha + \omega)$ and $\varphi(\alpha + \omega)$ stand in the relationship of smaller quantity to greater quantity. Let us denote this property of the variable quantity ω by M, then we can say that all ω that are smaller than a certain one possess the property M" (Russ 2004, 274); in symbols:

$$M(\omega) = \{0 < \omega < i : f(\alpha + \omega) < \varphi(\alpha + \omega)\}.$$

By applying LUB to the set $M(\omega)$, he gets the number U, which is

$$U = \sup M(\omega).$$

As a consequence of the theorem of Sect. 12 there must therefore be a certain quantity U which is the greatest of those of which it can be asserted that all ω which are $< U$ have the property M (Russ 2004, 274).

Now, Bolzano compares values $f(\alpha + U)$ and $\varphi(\alpha + U)$. Cases $f(\alpha + U) < \varphi(\alpha + U)$ and $f(\alpha + U) > \varphi(\alpha + U)$ are reduced to contradiction. For instance, Bolzano writes: "it cannot be that $f(\alpha + U) < \varphi(\alpha + U)$, for this would also give $f(\alpha + U + \omega) < \varphi(\alpha + U + \omega)$, if ω were taken small enough, and consequently $\alpha + U$ would not be the *greatest* value of which it can be asserted that all x below it have the property M" (Russ 2004, 274–275).

In conclusion, he finds: "Nothing else therefore remains but that $f(\alpha + U) = \varphi(\alpha + U)$, and so it is proved that there is a value of x lying between α and β, namely $\alpha + U$, for which $fx = \varphi(x)$" (Russ 2004, 275).

Thus, Bolzano proof falls under the same pattern as the proof presented in Sect. 6.1 above. The only difference is the status of LUB. While in modern mathematics it is the axiom characterizing real numbers, Bolzano seeks to prove it. At this stage, his reasoning does not meet modern standards for rigorous proof.

9.2 Least Upper Bound Principle

Bolzano proves LUB in (Bolzano 1817a) section § 12. The theorem reads: "If a property M does not apply to *all* values of a variable quantity x but does apply to *all*

values *smaller* than a certain u, then there is always a quantity U which is the greatest of those of which it can be asserted that all smaller x possess the property M" (Russ 2004, 269).

This thesis has the form of the following implication:

If

$$(\exists x)(\neg M(x)) \wedge (\exists u)(\forall x)(x < u \Rightarrow M(x)),$$

then

$$(\exists U)[(\forall x)(x < U \Rightarrow M(x)] \wedge (\forall u)(\forall x)[x < u \Rightarrow M(x) \Rightarrow u \leq U).$$

The set M is defined by criteria the least upper bound U does not meet. Thus, this version is not as general as the modern one worded in terms of any subset M of a totally ordered set – Bolzano considers sets defined by property. Still, it is much more precise than Dedekind's 1872 version. In section § 7 of his (Dedekind 1872), Dedekind seeks to show that the "essence of continuity" formulated in terms of cuts implies the following statement: "If a magnitude x grows continually but not beyond all limits it approaches a limiting value" (Dedekind 1901, 12). Dedekind's proof gives room to interpret this claim as the least upper bound principle, yet, it is not self-evident.

Bolzano's modern characterization of LUB contrasts with his scholastic methodology. He seeks to prove LUB by applying the so-called Cauchy completeness (CC). Then, however, he also attempts to prove CC. Current knowledge is that assuming axioms of an ordered field the equivalence holds $LUB \Leftrightarrow CC + Archimedean\ axiom$. Thus, viewed from the modern axiomatic perspective, Bolzano's argument is circular.

Yet, it took more than 50 years to point out unprovable mathematical truths. (Dedekind 1872) provides the first spectacular example. Regarding the "essence of continuity," Dedekind, first, states his cuts principle: "If all points of the straight line fall into two classes such that every point of the first class lies to the left of every point of the second class, then there exists one and only one point which produces this division of all points into two classes, this severing of the straight line into two portions" (Dedekind 1901, 5). Then, he adds: "I am utterly unable to adduce any proof of its correctness, nor has anyone the power. The assumption of this property of the line is nothing else than an axiom by which we attribute to the line its continuity, by which we find continuity in the line" (Dedekind 1901, p. 5).

9.3 Magnitudes as Real Numbers

Throughout (Bolzano 1817a), the adjective *reell* (real in (Russ 2004)) refers to a root of a polynomial, which means it is not an imaginary root. As for the domains of functions, the variable x is called *Größe (magnitude* or *quantity* in (Russ 2004)). In the second half of the nineteenth century, a series of mathematical papers occurred

dedicated to the concept of *magnitude*. They aimed to interpret the ancient Greek concept of μέγεθος. The common feature of these studies was to present magnitudes as a totally ordered, Archimedean group. During the first decades of the twentieth century, real numbers replaced the concept of magnitude.[27] Bolzano's magnitudes form a more complex structure than an ordered group: they are added and compared in terms of greater-lesser, they also form products. We will argue that they can be interpreted as real numbers.

In mathematics, real numbers are considered: (1) via a construction, (2) via an axiomatic approach. Historically, constructions have been studied first. Méray, Heine, Cantor, Dedekind, then Peano, and Weber complete the list of authors of the nineteenth-century constructions.[28] They rested on techniques of Cauchy sequences (or *Fundamentalreihen*, as Germans called them), Dedekind cuts, or ancient proportions smartly adjusted to a new role by Weber. Then, in 1900, Hilbert introduced an axiomatic approach. Yet, still, new constructions applying novel techniques have been discovered throughout the twentieth century and beyond.

Here is how Bolzano formulates what we currently call Cauchy completeness: "If a series of quantities

$$F^1x, F^2x, F^3x, \ldots, F^nx, \ldots, F^{n+r}x, \ldots$$

has the property that the difference between its nth term F^nx and every later one $F^{n+r}x$, however far this latter term may be from the former, remains smaller than any given quantity if n has been taken large enough, then there is always a certain *constant quantity*, and indeed only *one*, which the terms of this series approach and to which they can come as near as we please if the series is continued far enough" (Russ 2004, 266–267).

Now, while the continuity axiom is explicitly in the form of LUB as formulated in (Bolzano 1817a) section § 12, axioms for an ordered field are applied implicitly in Bolzano's proofs. Here is how one can reveal them.

In (Bolzano 1817a), the transitivity and the trichotomy laws characterize the total order of magnitudes. Let us start with the trichotomy.

The logical scheme of the final step in the proof of IVT_f as presented in Sect. 9.1 above is as follows:

$$\neg(f(\alpha+U)<\varphi(\alpha+U))\Lambda\neg(f(\alpha+U)>\varphi(\alpha+U))\Rightarrow f(\alpha+U)=\varphi(\alpha+U)$$

This implication is equivalent to the following alternative:

$$f(\alpha+U)<\varphi(\alpha+U)\vee f(\alpha+U)>\varphi(\alpha+U)\wedge f(\alpha+U)=\varphi(\alpha+U).$$

Implicitly it is also assumed that only one of these cases is satisfied.

[27] See (Błaszczyk and Fila 2020a), § 5–6.

[28] See (Cantor 1872; Dedekind 1872; Heine 1872; Méray 1869; Peano 1891; Weber 1895), respectively. See also Sect. 1 above.

Now, let us consider a sub-argument in the proof of SPP as discussed in Sect. 8.2 above: "assume that α and β are both *positive* and that (because it does not matter) β is the *greater* of the two, so $\beta = \alpha + \mathrm{i}$, where i denotes a positive quantity. [...] and $\alpha + \omega$ lies between α and β whenever we take $\omega < i$" (Russ 2004, 274).

Here, first, Bolzano applies the standard ordered field theory argument to the effect

$$\beta > \alpha \Rightarrow (\exists i > 0)(\beta = \alpha + i).$$

Note, the axiom E2 in Sect. 3.1 above also formalizes this rule. Thus, it was well-known to the ancient Greeks.

Then, the following axiom of an ordered field, the so-called compatibility of the order with addition,

$$(\forall x, y, z)(x < y \Rightarrow x + z < y + z)$$

is adopted as a self-evident truth. It occurs in the following form:

$$\omega < i \Rightarrow \alpha + \omega < \alpha + i.$$

Let us go back to the argument which proceeds formula (11). It reads: "Hence if it is required that the change in the function $\alpha + bx^m + cx^n + \ldots + px^r$ should be $<$ D, then we simply take an ω that is at the same time $< \omega_1$ and also $< \frac{D}{S}$. Then $\omega.S$ and, all the more, the product of ω with a quantity which is $< S$, must be $< D$" (Russ 2004, 276). Here Bolzano's argument is this

$$\omega < \frac{D}{S} \Rightarrow \omega S < S\frac{D}{S}$$

Meaning, he implicitly applies another axiom of an ordered field, the so-called compatibility of the order with multiplication

$$(\forall x, y, z)(x < y, 0 < z \Rightarrow xz < yz)$$

Moreover, considering the condition "simply take an ω that is at the same time $< \omega_1$ and also $< \frac{D}{S}$," the phrase "at the same time" should be interpreted as follows. Suppose that $\omega_1 < \frac{D}{S}$. Then

$$\omega < \omega_1 \Rightarrow \omega < \frac{D}{S}.$$

Here, the transitivity of the greater-lesser relation between magnitudes is assumed.

Similarly, we can analyze other parts of (Bolzano 1817a) and show that sums and products of magnitudes satisfy axioms of a commutative field.

10 Function vs Line

Bolzano considers two versions of IVT: for a geometric line and for a function. The first is stated like that: "every continuous [continuirlich] line of simple curvature of which the ordinates are first positive and then negative (or conversely), must necessarily intersect the abscissae-line somewhere at a point lying between those ordinates" (Russ 2004, 254). The version for two functions is as follows: "If two functions of x, fx and φx, vary *according to the law of continuity [nach dem Gesetze der Stetigkeit*] either for *all* values of x or for all those lying between α and β, and furthermore if $f\alpha < \varphi\alpha$ and $f\beta > \varphi\beta$, then there is always a certain value of x between α and β for which $fx = \varphi$x" (Russ 2004, 273). Bolzano also provides a version for one function, namely: "between any two Values, which give Results of Opposite Sign, there lies at least one real Root" (Russ 2004, 251).

The paper includes neither a definition of a geometric line nor a definition of a function. Moreover, it does not explain what continuity of line means.

Now, Bolzano declares that the truth [Wahrheit] of the geometric version of IVT "lies in nothing other than that general truth, as a result of which every continuous function" (Russ 2004, 255) has the IVT property. This claim presupposes that for every continuous line there is a representation by a continuous function. Yet, he never addressed this question. Anyway, that was the basis for the reduction of IVT for a geometric line to IVT_f.

In modern mathematics, the concept of function prevails, and the geometrical aspect of IVT is reduced to the graph of a function.

11 Topological Continuum

In this final section, we discuss Bolzano's attempts to characterize a topological continuum. Both Bolzano and Cantor refer to the ancient Greek model of a continuum, yet, they operate in a different mathematical framework, as their continuum consists of points. Nevertheless, they both compare their definitions of a continuum with Aristotle's idea of infinite divisibility.

11.1 Bolzanian Continuum

In section §38 of (Bolzano 1851), Bolzano discusses the paradoxes of "applied mathematics" concerning time and space. The paradox of extension is that when parts of a whole have no extension so the whole itself can not be extended: "a contradiction in the assumption that extension arises from parts which have no extension but are absolutely simple (points in time, or space, atoms, i.e., simple substances in the universe in the realm of reality" (Russ 2004, 642). Yet, Bolzano's solution is simple: "Every whole has and must have, some properties which the parts lack. [. . .] every two instants in time are separated by an infinite multitude of instants

in between, that likewise between every two points in space there is an infinite multitude of points lying in between them" (Russ 2004, 642).

Thus, first of all, a continuum is an "infinite multitude of points." Nevertheless, these points need some special arrangement, namely: "a finite multitude can never provide an extension and an infinite multitude only does so if the condition [...] that every point has certain neighbours at every sufficiently small distance" (Russ 2004, 644).

In this monograph (Kraus 2014) dedicated to Bolzano's concept of continuum, Kraus interprets the phrase "every sufficiently small distance" by two quantifiers. His reading, based on a thorough analysis of other fragments concerning continuum scattered across Bolzano's writings, offers the following definition: "Subset A of a metric space is *(Bolzanian) continuum* iff for every point p in A there is a real number $\varepsilon > 0$ such that for every $\eta \in (0, \varepsilon]$ there is at least one point q in A such that its distance from p equals η:

$$(\forall p \in A)(\exists \in > 0)(\forall \eta \in (0, \varepsilon])(\exists q \in A)[(d(p, q) = \eta]'' \ ([44], 31)."$$

Although Bolzano's idea of continuum consisting of "simple parts" (i.e., points) is foreign to the ancient Greek mathematics, he still refers to the Aristotelian characteristics: "it should be admitted that not every division of a given spatial thing into parts, namely no division into such parts whose number is only finite, not even every one such as goes on indefinitely (e.g. by continued bisection) [...] reaches the simple parts. [...] ultimately every continuum can arise from nothing else but from points and only points" (Russ 2004, 644).

11.2 Cantor on a Continuum in the Metric Space $\mathbb{R}^n$

Section § 9 of Cantor (Cantor 1883) summarizes some nineteenth-century theories of real numbers, namely: Weierstrass' as explained in (Kossak 1872), Dedekind's as developed in (Dedekind 1872), and Cantor's theory presented in (Cantor 1872). Section § 10 is dedicated to the concept of the continuum. Tracing back its history, Cantor discusses ancient Greek and Medieval philosophers. He believes that "the underlying idea has taken on different meanings." Therefore, his definition is comparable, for instance, with Aristotle's *ex partibus sine fine divisibilibus*. This section initiated a new branch of point-set topology – the theory of continua, a study of compact and connected spaces. Yet, Cantor's original definition is a bit different.

Thus, according to Cantor, a subset T of the space $\mathbb{R}^n$ with Euclidean metric is *connected* iff for any of its points t_0 and t^0 and any positive real number e there are finitely many points $t_1, t_2, \ldots, t_n$ of T such that the distances $d(t_0, t_1)$, $d(t_1, t_2)$, ..., $d(t_n, t^0)$ are all less than ε. T is perfect iff $T = T'$, where T' is defined as follows:

$$x \in T' \Leftrightarrow \lim_{n \to \infty} x_n = x$$

for some $(x_n) \subset \mathbb{R}^n \setminus \{x\}$; convergence of a sequence is defined in the metric space $\mathbb{R}^n$. The latter condition is equivalent to the following assertion: Every convergent series $(t_n) \subset$ T has a limit in T, and for every $t \in T$, there exists a sequence $(t_n) \subset T$ such that $\lim_{n\to\infty} t_n = t$.

Then comes the well-known definition: "I, therefore, define a point-continuum inside G_n [$\mathbb{R}^n$ in our notation] as *a perfect-connected*" (Ewald 1996, 906).

Indeed, one can show that in a metric space, Cantor's continuum agrees with the modern one, that is, it is a compact and connected set.

Based on this definition, Cantor rebukes the definition of the continuum as given in Bolzano (Bolzano 1851), §38. He claims that non-connected sets exemplify Bolzano's definition. In fact, for instance, set $A = [0,1) \cup (1, 2]$ is a *Bolzanian continuum* in the sense of Kraus while it is not a *Cantorian continuum.*[29]

12 Summary

In Greek mathematics, straight lines, curves, figures, solids, and movements involved in the genesis of mechanical curves were continuous, meaning divisible into infinitely divisible parts. The evolution of the concept of a curve is the most intriguing in the history of mathematical continuity. At first, Descartes turned the ancient scheme genesis-symptomata, involved in studies of mechanical curves, into genesis-formula, where genesis played a role in the derivation of a formula. In *La Geometire*, formula, inextricably linked to the curve, was an analytic description of the curve previously given by some other means. Then, with time, the formula evolved into the concept of a function and its domain. It enabled a separation of function and curve. This function-curve duality finds a spectacular manifestation in the very arrangement of the foremost analysis textbook in history – Euler's 1748 *Introductio in analysin infinitorum.* Its first volume explores functions, the second curves.

In Bolzano's *Rein analytischer Beweis*, a curve is an aspect – the graph, to use a modern term – of a function. His discussion on the geometric and analytic version of IVT echoes this function-curve duality. Due to methodological reasons, Bolzano assumes that the geometric version "lies in nothing other than that general truth [Wahrheit], as a result of which every continuous function" (Russ 2004, 255) has the IVT property. This solution paved his way to a purely analytic proof of IVT.

Like predecessors, Bolzano seeks to prove IVT for polynomials. To this end, he adopts a two-functions scheme originated from Lagrange (Lagrange 1808). Since the crucial step in Lagrange's proof involves the non-analytic concept of motion, Bolzano replaces it with a new concept of continuous function. This idea refers to a general function and evolves throughout the paper. The version applied in the proof of continuity of polynomial inspired Stolz to formulate the modern $\varepsilon\delta$ definition.

[29] Błaszczyk and Fila (2020b) discusses Cantor's definition of topological continuum in a broader context of his struggle with the concept of continuity.

When Bolzano managed to show the continuity of a polynomial, he could apply the lemma of IVT for continuous function. Proving this lemma, he devised another fundamental law of mathematics. His proof, in a simplified, one- rather than two-functions form, is like the modern one, given $f(a)f(b) < 0$, he determines a point c such that $c \in (a, b)$ and $f(c) = 0$. This point is defined as the least upper bound of the set $\{x \in [a, b] : f(x) < 0\}$. Bolzano does not realize the fundamental role of the LUB principle and seeks to prove it. Furthermore, since it applies to definable sets, it is not as general as the modern version. Despite imperfections, it gave rise to a new idea of a continuous (complete) total order.

In this way, in *Rein analytischer Beweis* (Bolzano 1817a) the Greek idea of a continuity of a curve bifurcates into the continuity of a function and continuity of its domain.

13 Cross-Reference

▶ Descartes' Transformation of Greek Notion of Proportionality

References

Artin E, Schreier O (1926) Algebraische Konstruction reeller Körper. Abhandlungen aus dem Mathema-tischen Seminar der Hamburgischen Universität 5:85–99

Bair J, Błaszczyk P, Fuentes-Guillen E, Heining P, Kanovei V, Katz M (2020) Continuity between Cauchy and Bolzano: issues of antecedents and priority. Br J Hist Math 35(3):207–224

Beckmann F (1967) Neue Geschichte zum 5. Buch Euklids. Arch Hist Exact Sci IV:1–144

Bekker I (1931) Aristotelis Opera. Georg Reimer, Berlin

Bell JL (2005) The continuous and the infinitesimal. Polimetrico, Milano

Birkhoff G, Mac Lane S (1977) A survey of modern algebra. Macmillan, New York

Birkhooff G (1932) A set of postulates for plane geometry (based on the scale and protractor). Ann Math 33:329–345

Błaszczyk P (2016) A purely algebraic proof of the fundamental theorem of algebra. Annales Universitatis Paedagogicae Cracoviensis Studia ad Didacticam Mathematicae Pertinentia 8:6–22

Błaszczyk P (2018) From Euclid's Elements to the methodology of mathematics. Two ways of viewing mathematical theory. Annales Universitatis Paedagogicae Cracoviensis Studia ad Didacticam Mathematicae Pertinentia 10:5–15

Błaszczyk P, Fila M (2020a) Cantor on infinitesimals. Historical and modern perspective. Bull Sect Logic 49(2):149–179

Błaszczyk P, Fila M (2020b) Modes of continuity in diagram for intermediate value theorem. In: Pietari-nen A, Chapman P, Bosveld de Smet L, Giardino V, Corter J, Linker S (eds) Diagram-matic representation and inference – 11th international conference, diagrams 2020, proceedings, Lecture Notes in Artificial Intelligence (LNAI) 12169. Springer, pp 34–49

Błaszczyk P, Mrówka K (2013) Euklides, Elementy. Księgi V-VI. Tlumaczenie i komentarz. Copernicus Center Press, Krakow

Błaszczyk P, Fila M, Mrówka K (2017) Wprowadzenie do rozprawy Bernarda Bolzano Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege. Annales Universitatis Paedagogicae Cracoviensis Studia ad Didacticam Mathematicae Pertinentia 9:167–176

Bolzano B (1804) Betrachtungen über einige Gegenstände der Elementargeometrie. Karl Barth, Prague. In: [53], 25–82
Bolzano B (1817a) Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein ent-gegengesetzes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege. Gottlieb Hasse, Prague. In: [53], 253–277
Bolzano B (1817b) Die drey Probleme der Rectification, der Complanation und der Cubirung, ohne Betrach-tung des unendlich Kleinen, ohne die Annahmen des Archimedes, und ohne irgend eine nicht streng erweisliche Voraussetzung gelöst: zugleich als Probe einer gänzlichen Umstaltung der Raumwissenschaft, allen Mathematikern zur Prüfung vorgelegt. Leipzig . In: [53], 278–344
Bolzano B (1851) Paradoxien des Unendlichen. Reclam, Leipzig
Bolzano B (2000) Functionenlehre. Hrsg. Bob van Rootselaar: Theory of Functions with Improvements and Additions to the Theory of Functions Bernard Bolzano. Bernard Bolzano Gesamtausgabe Bd. 2A10/1, Stuttgart. In: [53], 429–572
Borsuk K, Szmielew W (1960) Foundations of geometry. North Holland Publishing Company, Amsterdam
Boyer C (1949) The history of the calculus and its conceptual development. Dover, New York
Cantor G (1872) Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math Ann 5:123–132. In: [26], 92–102
Cantor G (1883) Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Teubner, Leipzig. In: [26], 165–208
Cantor G (1895) Beiträge zur Begrundung der transfiniten Mengenlehre. Math Ann 46:481512. In: [26], 282–311
Cantor G (1932) Gesammelte Abhandlungen mathematischen und philosophischen Inhalts. Springer, Berlin
Christianidis J, Oaks J (2021) The Arithmetica of Diophantus. A complete translation and commentary. Routledge, Abingdon/Oxfordshire
Cohen L, Ehrlich G (1963) The structure of the real number system. Von Nostrand, Englewood Cliffs
Dedekind R (1872) Stetigkeit und irrationale Zahlen. Fredrich Vieweg & Sohn, Brunschweig
Dedekind R (1901) Continuity and irrational numbers (Trans. Beman WW). Open Court, Chicago
Descartes R (1637) La Géométrie. Jan Maire, Leiden
Euclid (2007) Euclid's Elements of Geometry. Edited, and provided with a modern English translation, by Richard Fitzpatrick; http://farside.ph.utexas.edu/Books/Euclid/Elements.pdf
Euler L (1748) Introductio in analysin infinitorum. Bousquet, Lausanne
Ewald W (1996) From Kant to Hilbert, vol II. Clarendon Press, Oxford
Heath TL (2002) The works of Archimedes. Dover, New York
Heiberg J (1883–1888) Euclidis Elementa. Teubner, Lipsiae
Heine E (1872) Die Elemente der Functionenlehre. J Reine Angew Math 74:172–188
Hilbert D (1900) Über den Zahlbegriff. Jahresbericht der Deutschen Mathematiker-Vereinigung 8: 180–184
Hilbert D (1903) Grundlagen der Geometrie. Teubner, Leipzig
Jacobson N (1969) Lectures in abstract algebra, vol III. Von Nostrad, Toronto
Knorr W (1975) The evolution of the Euclidean elements. D. Reidel Publishing Company, Dordrecht
Kossak E (1872) Die Elemente der Arithmetik. Nicolapsche Verlagbuchhandlung, Berlin
Kraus LB (2014) Der Begriff des Kontinuums bei Bernard Bolzano. Academia Verlag, Sankt Augustin
Kuratowski K, Mostowski A (1976) Set theory. Polish Scientific Publishers and North Holland Publishing Company, Warsaw/Amsterdam
Lagrange JL (1808) Traité de la résolution des équations numériques de tous les degrés. Courcies, Paris
Lockwood EH (1961) A book of curves. Cambridge Univeristy Press, Cambridge

Malmendier N (1975) Eine Axiomatik zum 7. Buch der Elemente von Euklid. Mathematisch-Physikalische Semesterberichte 22:240–254
Méray C (1869) Remarques sur la nature des quantités déjinies par la condition de servir de limites à des variable donnèes. Revue des Sociétés savantes, Sciences mathém phys et naturelles 2:281–289
Mueller I (2006) Philosophy of mathematics and deductive structure in Euclid's Elements. Dover, New York. (reprint of the first edition: MIT Press, Cambridge, MA, 1981)
Pappus of Alexandria (2010) Book 4 of the collection. Edited with translation and commentary by Heike Sefrin-Weis. Springer, New York
Peano G (1891) Sul concetto di numero. Rivista di matematica 1:256–267
Pontriagin L (1932) Über stetige algebraische Körper. The Annals of Mathematics 33(1):163–174
Russ S (2004) The mathematical works of Bernard Bolzano. Oxford University Press, Oxford
Smith ED, Lathan M (2007) Geometry of Rene Descartes. CosimoClassics, New York
Stolz O (1885) Vorlesungen uber allgemeine Arithmetik. Teubner, Leipzig
Tarski A (1931) Sur les ensembles definissables de nombres réels. Fundam Math 17:210–239
Tarski A (1959) What is elementary geometry? In: Henkin L, Suppes P, Tarski A (eds) The axiomatic method. With special reference to geometry and physics. Proceedings of an international symposium held at the University of California, Berkeley, December 26, 1957 January 4, 1958, Studies in logic and the foundations of mathematics, North-Holland/Amsterdam, pp 16–29
Weber H (1895) Lehrbuch der Algebra. Vieweg, Braunschweig
Wilck B (2020) Euclid's kinds and (their) attributes. Hist Philos Logic Anal 23:362–397

Bolzano's Theory of *meßbare Zahlen*: Insights and Uncertainties Regarding the Number Continuum

Elías Fuentes Guillén

Contents

1 Introduction 1596
2 Bolzano's Uncertainties Regarding the Notions of Quantity and Number 1598
3 The Conceptual Framework of Bolzano's Measurable Numbers 1607
4 Bolzano's *meßbare Zahlen* 1614
5 Final Remarks 1625
References 1629

Abstract

During the first half of the 1830s, and as part of his project for a *Größenlehre*, Bernard Bolzano worked on a manuscript entitled *Reine Zahlenlehre* in which he introduced the notion of what he called "*meßbare Zahlen*." The various additions and corrections to its three extant versions are evidence of an unfinished work, the definitive edition of which was not published until 1976. The present chapter casts light upon the links between, on the one hand, his theory of "measurable numbers" and its conceptual framework, and, on the other hand, his insights and uncertainties with regard to the notions of number and quantity prior to the writing of that work. While Bolzano's proposal has usually been considered as an attempt at a theory of what nowadays is called the real-number continuum, this chapter shows that a more faithful reading must consider it as a pioneering and transitional theory of the number continuum which provided relevant insights into this latter but which remained, nonetheless, still bound to a not-yet-modern conception of mathematics and numbers.

E. Fuentes Guillén (✉)
Institute of Philosophy of the Czech Academy of Sciences, Prague, Czech Republic
Department of Mathematics, Faculty of Sciences, UNAM, Mexico City, Mexico

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_96

Keywords

Bernard Bolzano · Measurable numbers · Number continuum · Real numbers · Nineteenth-century mathematics

1 Introduction

Throughout the eighteenth century mathematics was usually defined as the "science of measuring everything that can be measured" or the "science of the quantities, that is, all those things that can be augmented or reduced."[1] However, by the end of the nineteenth century such an understanding of mathematics, which entailed the division of quantities into discrete (i.e., numbers) and continuous, was no longer advocated by most mathematicians. So, while certain appellations used by some of the authors who developed the first theories of real numbers, such as *Zahlengrössen* (numerical quantities; cf. Cantor 1872; Weierstraß 1868/1986), still evoked that earlier understanding, they nevertheless involved a different, relatively abstract conception of quantities and, ultimately, mathematics (cf. Ferreirós 2007, p. 125).

The disagreement between Richard Dedekind and Rudolf Lipschitz regarding a note included by the former in the introduction to his *Sur la théorie des nombres entiers algébriques* (1876–77) reveals the prevalence of the idea that Euclid's theory of proportions, and in particular definition 5 from Book V of his *Elements*, was enough "to guarantee the continuity of the domain of 'incommensurable magnitudes'" (Benis-Sinaceur 2015, p. 25; cf. Lipschitz 1876/1986, p. 72). Whereas Lipschitz advocated that idea, Dedekind explicitly rejected it and pointed out that, from his point of view, "arithmetic is to be kept free of any [mixture] of foreign elements,"[2] something on which he insisted in his own 1888 work (cf. Dedekind 1888, p. X).

This chapter draws attention to some of the complexities involved in the transition from a traditional notion of numbers to the definitions of real numbers published from c. 1870 onward, including the ones by Dedekind and Cantor. This is done by focusing specifically on the work of Bernard Bolzano, whose theory of "measurable numbers" (*meßbare Zahlen*) is considered to be an early "attempt at a theory of real numbers" (Russ 2004, p. 348; cf. Šebestík 1992, p. 354; Laugwitz 1982, p. 667; Rusnock 2000, p. 177). The work by Bolzano which is considered by some scholars to contain such an attempt is entitled *Pure Theory of Numbers* (*Reine Zahlenlehre*; henceforth *RZ*). Although this work was written in the first half of the 1830s and its manuscript was rediscovered at the beginning of the twentieth century, its definitive

[1](Wolff 1716, p. 863): "Ist eine Wissenschafft alles auszumessen, was sich ausmessen läst. Insgemein beschreibet man sie *per scientiam quantitatum*, durch eine Wissenschafft der Größen, das heisset, aller derjenigen Dinge, die sich vergrössern oder verkleinern lassen."

[2](Dedekind 1876, p. 284): "l'Arithmétique doit être maintenue exempte de tout mélange d'éléments étrangers."

edition was published only in 1976 as part of the *Bernard Bolzano Gesamtausgabe* (*Bernard Bolzano: The Complete Edition*; henceforth *BBGA*),[3] publication of which, by the publisher Frommann-Holzboog, has been ongoing from 1969 up to the present day.[4]

As it will be argued, Bolzano's *RZ*, which was developed within the framework of his project for a theory of quantities (*Größenlehre*), and in particular his theory of measurable numbers, was still tied to a not-yet-modern conception of numbers, quantities and, ultimately, mathematics. Thus, Sect. 2 discusses his insights and uncertainties with regard to the notions of quantity and number mainly as expressed in the mathematical works that preceded *RZ*, while also paying attention to certain documents (both posthumously published and as yet unpublished) preceding, contemporary with, and succeeding *RZ*, such as Bolzano's mathematical notebooks (*Miscellanea Mathematica*), which were published for the first time in the *BBGA*. Then, Sect. 3 goes on to focus on the introductory volume to Bolzano's theory of quantities, which precedes *RZ*, as well as on the first six sections of the latter work in order to better understand his project of a pure theory of numbers, while also paying attention to his mathematical notebooks written during the 1820s. Finally, Sect. 4 addresses his insights and uncertainties with regard to the number continuum by bringing to the fore the particularities of his "measurable numbers" and the other core notions of the seventh and last section of *RZ*.

The need to study Bolzano's manuscripts has recently been highlighted by both Steve Russ and Kateřina Trlifajová, with specific reference to *RZ* (cf. Russ and Trlifajová 2016, p. 44), and Jan Šebestík, with reference to Bolzano's project of a *Größenlehre*. As noted by Šebestík in the "Afterword" to *Bernard Bolzano: His Life and Work*, which he co-wrote with Paul Rusnock, in some cases such a study could "modify our understanding of Bolzano's thought" (Rusnock and Šebestík 2019, p. 597). This chapter will show that this is the case with Bolzano's notion of numbers and that the present research not only contributes to its elucidation but also accounts for the complexities of the emergence of our number continuum insofar as, to

[3]For a detailed account of the publication of *RZ* and its reception throughout the second half of the twentieth century, cf. (Spalt 1991, pp. 16–24; Russ 2004, pp. XX & 348). According to two letters written by Bolzano in the early 1840s and quoted by Eduard Winter, Bolzano resumed work on his project on mathematics between mid-1840 and mid-1841 (cf. Winter 1933, pp. 215–216).

[4]Founded by Jan Berg, Friedrich Kambartel, Jaromír Louzil, Bob van Rootselaar and Eduard Winter, and currently edited by Edgar Morscher, the *BBGA* comprises some introductory volumes (*Einleitungsbände*) and, up to the present day, about 100 published volumes divided into four series: (I) *Schriften*, the "works published during Bolzano's lifetime"; (II) *Nachlaß*, his "posthumous works" (subpart A) and his "scientific diaries" (subpart B); (III) *Briefwechsel*, his correspondence with various correspondents; and (IV) *Dokumente*, which includes various documents ranging from portraits and biographies to documents on Bolzano's trial (https://www.frommann-holzboog.de/editionen/20?lang=de). All references to volumes of the *BBGA*, therefore, contain: (a) the initials *BBGA*; (b) the letter "E" (for the introductory volumes) or the number of the series (1–4) and, in the case of the *Reihe* II, the subpart letter (i.e., 2A or 2B); and (c) the volume number and, in some cases, the supplement number.

paraphrase Russ, Bolzano's theory epitomizes those complexities vividly (cf. Russ 2004, p. 348). In this sense, this chapter follows Epple's recommendation to thoroughly investigate the "eventful atmosphere" that led to "the end of the paradigm of the science of quantity"[5] by addressing the subtleties of the work of an author writing somewhat earlier than those whom Epple discusses in *Das Ende der Grössenlehre. Eine Einführung in die Geschichte der Grundlagen der Analysis, 1860–1930*.

2 Bolzano's Uncertainties Regarding the Notions of Quantity and Number

Bolzano entered the Faculty of Philosophy at the Carl-Ferdinandischen Universität (Prague) as a student in 1796 and, as it is known from his lecture notes for the first academic year, was taught there the notions that were customary throughout the eighteenth century: mathematics was the "science of quantities," that is, "that which is capable of increase or decrease," while numbers were defined as "a multitude of things of one kind."[6] Such traditional notions can be found in almost all German-language mathematics textbooks of the time, including Abraham Gotthelf Kästner's *Anfangsgründe*, which had been the official textbook in Prague since 1784 (cf. *Handbuch* 1786, p. 407).[7]

As Bolzano explained in his 1804 *Considerations on Some Objects of Elementary Geometry* (henceforth *BG*), such a conception of "quantity" entailed that a thing could only be quantified if "regarded as consisting of a *number* (plurality) of things that are equal to the *unit* (or the measure)."[8] Thus, just as natural numbers would be a plurality of things equal to 1, a straight line would be a plurality of things equal to a certain straight line segment taken as unit of measurement. However, as he noted in his written examination for the chair of Elementary Mathematics at Prague University, which took place shortly after the publication of that work, certain things that at the time were regarded as "quantities" in fact posed a problem both conceptually and operationally, as in the case of the so-called infinitely small quantities. For Bolzano,

[5](Epple 1996, pp. 3 & 1): "Ende des Paradigmas der Größenlehre"; "bewegten Atmosphäre."

[6](LA PNP, C II 14/2, [2r] & [4v]): "Die Größe ist das, was einer Vermehrung, oder einer Verminderung fähig ist"; "Zahl ist eine Menge von Dingen einerley Art." LA PNP is the abbreviation for the Literární archiv Památníku národního písemnictví (Literary Archive of the Museum of National [Czech] Literature), in Prague, where Bolzano's notes are preserved.

[7]The above-cited definitions can be found almost verbatim in (Kästner 1786, pp. 1 & 21). I refer to the fourth edition of Kästner's work since it is the one which is known, with certainty, to have been available at the time at Prague's Public and University Library, which eventually became the National Library of the Czech Republic (cf. *Catalogus Mathematicorum* IX A 19 1781ff., p. 130 [r]).

[8](Bolzano 1804, pp. 3–4; cf. Russ 2004, p. 37): "Größe heißt ein Ding, insofern es angesehen wird als bestehend aus einer *Anzahl* (Vielheit) von Dingen, die der *Einheit* (oder dem Maße) gleich sind." The title of Bolzano's work is *Betrachtungen über einige Gegenstände der Elementargeometrie*.

the notion of these quantities was "incomprehensible" (*unverständlich*) inasmuch as they were either to be considered as "nothing" (*Nichts*), in which case they would not be a quantity but rather the absence of a quantity, or to be considered as something, in which case the usual procedure, according to which a piece of a curved line equals a piece of a straight line if their difference is an infinitesimal, would be strictly speaking unacceptable (*ČG Publicum* 1796–1805 98/755, p. [13v]). As he went on to explain further, such a procedure as this latter "does not fundamentally remove the difficulty; rather it only removes it for the eye, though not for the understanding."[9]

Such reluctance to accept infinitesimals can also be found in Bolzano's second published mathematical work, namely his 1810 *Contributions to a Better-Grounded Presentation of Mathematics* (henceforth *BD*).[10] There, addressing the so-called infinitely small and great quantities, he wrote that it was yet to be decided whether the "*infinite*" (whose notion comprised both), "or the *differential*," was "nothing else but a *symbolic expression*," as in the case of "$\sqrt{-1}$ and suchlike [expressions]."[11] At the time, "symbolic expression" was an appellation commonly used to refer to so-called imaginary quantities, as is attested to by the corresponding entry in the supplement to the famous mathematical dictionary initially developed by Georg Simon Klügel (cf. Grunert 1836, p. 387), as well as by the term's usage in Augustin-Louis Cauchy's *Cours d'Analyse* (cf. Cauchy 1821, p. 173ff.). Nevertheless, in the case of these two references, as in the case of so many others, there seems to have been no deep reflection, as there was in Bolzano, on why mathematical infinites and imaginaries could not legitimately be regarded as quantities, let alone as numbers, but at most as "symbolic expressions."[12]

Interestingly, among the extant notes for a section entitled "On the Properties of Numbers," which was to be part of the second installment of *BD* and would have been written around 1810 (cf. Berg 1977, pp. 7–8), Bolzano stated that "*if a number*

[9](*ČG Publicum* 1796–1805 98/755, p. [13v]): "hebt die Schwierigkeit nicht im Grunde; sondern entrückt sie nur dem Auge, nicht aber dem Verstande." For a detailed account of Bolzano's examination, as well as its transcription and English translation, cf. (Fuentes Guillén and Crippa 2021).

[10]The title of Bolzano's work is *Beyträge zu einer begründeteren Darstellung der Mathematik*. This was the first installment (*Lieferung*) of that work, the second—and unfinished—installment of which was reconstructed by Jan Berg, based on two versions, and not published until 1977 in *BBGA* 2A5.

[11](Bolzano 1810, p. 30): "das *Unendliche*, oder das *Differenzial*, nichts anders als ein *symbolischer Ausdruck* sey, gerade wie $\sqrt{-1}$, dgl."

[12]A noteworthy precedent here may possibly be found in Lambert, who in a letter to Kant, dated October 13, 1770, wrote: "The sign $\sqrt{-1}$ represents an unthinkable non-thing. And yet it can be used very well in finding theorems" (Zweig 1999, p. 118). It is worth mentioning that it had been only a short time before, in 1767/1768, that Lambert had composed and published his famous *Mémoire sur quelques propriétés remarquables des quantités transcendantes, circulaires et logarithmiques*, in which he seems to avoid the use of imaginaries (cf. Lambert 1761/1768, p. 319; Barnett 2004, pp. 24–25). I am grateful to Eduardo Dorrego López for bringing this to my attention.

is conceived as augmented by 1, *a new number arises*," so that, as he added in a marginal note, a new number would arise whenever two numbers were "*conceived together*."[13] As noted by Johan Blok, "such a statement would seem akin to the modern notion of a successor function" (Blok 2016, p. 256), inasmuch as the addition of the unit to each number would originate the subsequent one. Nonetheless, considering the aforementioned traditional notions of both number and quantity, the extension of the former beyond the sequence of the naturals still required, from Bolzano's point of view, some justification—something that, in the case of the irrationals, posed problems.

In a couple of notes included in his mathematical notebooks and dating from late 1814, Bolzano addressed the question of how to determine irrational quantities, irrationality being for him, as in the corresponding cases of the infinite and of fractions, a "predicat[e] [that] refer[s] to quantities" and not to numbers.[14] Thus, on October 26, he pointed out the possibility of considering an irrational as "a quantity for which the rule to determine it by numbers is indeed given but requires an operation which can never be brought to an end," which will allow one to approximate to said quantity "as closely as one desires."[15] Moreover, he said, an irrational would be a quantity that lies between two given quantities which involve two such rules, so that, for example, in the case of $\sqrt{2}$, one could establish the two following "rules (manners)" for determining it:

(a) to "[l]ook at any power of 10 as a denominator, and seek out the greatest possible number which, considered as numerator, gives a fraction the square of which would be smaller than 2";
(b) to "look at any power of 10 as a denominator, and seek out the smallest possible number which, considered as numerator, gives a fraction the square of which would be greater than 2."[16]

This would mean that $\sqrt{2}$ could be determined as the "limit" (*Grenze*) to which, following these rules, the quantities 1.4 (or $\frac{14}{10}$, i.e., the greatest fraction the square of which—according to the first rule—would be smaller than 2) and 1.5 (or $\frac{15}{10}$, i.e., the smallest fraction the square of which—according to the second rule—would be

[13](*BBGA* 2A5, p. 61 & fn. r): "*Wenn zu einer Zahl noch 1 hinzugedacht wird, entsteht eine neue Zahl*"; "*Wenn* 2 *Zahlen zusammengedacht werden, entsteht eine neue Zahl*." The title of the note is *Von den Eigenschaften der Zahlen*.

[14](*BBGA* 2B6/2, p. 195): "alle diese Prädicate beziehen sich meiner Ansicht nach auf *Größen*."

[15](*BBGA* 2B6/2, p. 195): "ist eine Größe, für welche die Regel sie durch Zahlen zu bestimmen, zwar gegeben ist, aber eine nie zu beendigende Operation erfordert"; "als man nur immer will."

[16](*BBGA* 2B6/2, p. 196): "*Regeln* (*Arten*)"; "Man sehe irgend eine Potenz der 10 als Nenner an, und suche dazu die möglichst größte Zahl, welche als Zähler betrachtet einen Bruch gibt, dessen Quadrat *kleiner* als 2 wäre"; "man sehe irgend eine Potenz der 10 als Nenner an, und suche dazu die möglichst kleinste Zahl, welche als Zähler betrachtet einen Bruch gibt, dessen Quadrat *größer* als 2 wäre."

greater than 2), 1.41 and 1.42, and so on, could approximate as closely as one desired (cf. *BBGA* 2B6/2, pp. 195–197).[17]

But then, in a note dated three days later (i.e., October 29, 1814) and entitled "*Irrational*, *transcendental quantities*," Bolzano added that, according to his previous note, "the *irrational* quantities are to *numbers* (arithmetic) what the *transcendent* ones are to functions, that is to say, to *finite* functions, i.e. to *algebraic* functions (algebra)," inasmuch as the former "cannot be expressed by a finite multitude of *numbers*," while the latter "cannot be expressed by a finite multitude of powers and products."[18] This was a quite interesting remark and meant that, strictly speaking, irrationals were not to be regarded as numbers but rather as quantities that could only be determined numerically according to the aforementioned rules. Indeed, he even suggested that his previous explanation could be shortened "by using only one instead of two *bounding* [or *limit*] *quantities*," so that an irrational could be regarded as a "*constant* quantity to which the variable [quantity] *X*, which arises from a rule that contains a number of repetitions of one and the same operation which can be increased to any extent one wishes, can approximate as closely as one desires."[19]

As we will discuss in Sect. 4, Bolzano's later procedure for determining what he would call "measurable numbers" resembles the one first outlined in his notes of 1814. However, in the first place (as noted by Bob van Rootselaar in the introduction to the volume of *BBGA* containing the first of these notes), Bolzano's early procedure "provides a means of deciding whether a given quantity is measurable [...] [,] not a means of constructing a measurable quantity" since, "[f]or Bolzano, the quantities are already there [and] one needs only to determine them by numbers."[20] In the second place, moreover, such a procedure still required some further clarifications in order to operate with numerically determined irrationals, in particular with regard to the difference between these latter and their limits, that is to say, concerning

[17]This procedure could be rooted in the successful use of continuous fractions, by means of which approximations of irrationals, both by excess and by default, could be provided. For a detailed account of continuous fractions at the time, as well as a historical sketch of their use, cf. (Klügel 1808, pp. 43–87; cf. also Dorrego López and Fuentes Guillén, forthcoming).

[18](*BBGA* 2B7/1, p. 35): "*Irrationale, transcendente Größen*"; "sind die *irrationalen* Größen das für die *Zahlen* (Arithmetik), was die *transcendenten* für die Functionen, die *endlichen*, d.h. *algebraischen* Functionen, (Algebra) sind"; "durch keine endliche Menge von *Zahlen* ausgedrückt werden kann; so kann [...] durch keine endliche Menge von Potenzen und Producten ausgedrückt werden."

[19](*BBGA* 2B7/1, p. 35): "In Ansehung beyder aber werfe ich die *Frage* auf: ob man ihre Erklärung nicht dadurch bedeutend verkürzen könnte, daß man statt ihren Werth durch zwey *Grenzgrößen* [...], zu bestimmen vielmehr nur eine einzige gebrauchte"; "*beständige* Größe, der die veränderliche *X*, welche durch eine Regel entsteht, die eine beliebig »zu vermehrende Anzahl von Wiederholungen einer und derselben Operation enthält, so nahe kommen »kann, als man nur immer will." It should be noted that, immediately afterward, Bolzano wrote that "however, sometimes one has to use another quantity *Y* as a help."

[20](van Rootselaar 1995, pp. 27–28): "sie ergibt ein Mittel zu entscheiden, ob eine vorgegebene Größe meßbar sei. Es ist kein Mittel, eine meßbare Größe zu konstruieren. Für Bolzano sind die Größen schon da, man muß sie nur durch Zahlen bestimmen."

the decreasing difference between an irrational quantity and its corresponding rational limits. As Bolzano stated in another note included in his mathematical notebooks and written at the turn of 1815, he had, up to this point, still not been able to "get things clear in his mind" either as regards the concept of the infinite or as regards the concepts of 0 and $\sqrt{-1}$: all of them, he wrote, "denote a kind of nothing," even though "they are essentially different."[21] This last note was to be followed by the discussion of an example showing the nature of the difference between these three concepts; in the end, however, Bolzano did not write it.

Of course, this does not mean that Bolzano did not resort to, for example, the irrationals and zero. For example, in his 1816 work on the binomial theorem (henceforth *BL*),[22] he argued that the validity of the binomial series expansion for $(1 + x)^n$ could also be proved for an irrational value $n = i$, on the basis that such a value could be approximated to by a fraction $\frac{n}{m}$, with both n and m integers, so that "$(1+x)^{\frac{n}{m}}$ comes as close as one might desire to the value $(1 + x)^i$."[23] Thus, considering "$x < \pm 1$" (or, as noted by Paul Rusnock, "$|x| < 1$" (Rusnock 2000, 68)), Bolzano obtained the series

$$(1+x)^{\frac{n}{m}} = 1 + \frac{n}{m}x + \frac{n}{m}\frac{\frac{n}{m}-1}{2}x^2 + \ldots + \frac{n}{m}\frac{\frac{n}{m}-1}{2}\ldots\frac{\frac{n}{m}-r+1}{r}x^r + \Omega^2,$$

where Ω^2 can become smaller than any given quantity (or, from a modern perspective, $\Omega^2 \to 0$), from which, considering that the difference between i and $\frac{n}{m}$ can therefore be made as small as desired, he substituted the latter by the former and obtained the series[24]

$$(1+x)^i = 1 + ix + i\frac{i-1}{2}x^2 + \ldots + i\frac{i-1}{2}\ldots\frac{i-r+1}{r}x^r + \Omega^2.$$

Leaving aside the fact that, as Rusnock has pointed out, there is a subtle mistake in Bolzano's proof—since Bolzano "uses the expansion for $n = -1$ as the basic convergence lemma [,] [b]ut this series diverges for $x = \pm 1$" (Rusnock 2000, p. 69)—there are two things that are particularly interesting about it nonetheless. Firstly, although the meaning of Ω^2 might seem clear, it remains to be explained just what this sign signified for Bolzano. Secondly, at the very end of that work, he stated that the concept of "the *irrationality* of a quantity," along with that of "*imaginary*

[21](*BBGA* 2B7/1, p. 68): "kann ich noch immer nicht ins Reine kommen"; "bezeichnen eine Art von *Nichts*"; "sie doch *wesentlich* verschieden sind."

[22]The title of Bolzano's work is *Der binomische Lehrsatz, und als Folgerung aus ihm der polynomische, und die Reihen, die zur Berechnung der Logarithmen und Exponentialgrößen dienen, genauer als bisher erwiesen*.

[23](Bolzano 1816, p. 76): "$(1+x)^{\frac{n}{m}}$ dem Werthe $(1+x)^i$ so nahe, als man will."

[24]It should be noted, firstly, that in the original the superscript in Ω^2 is placed at the top and that, as Steve Russ wrote, "[a]s often the case with Bolzano, these superscripts only distinguish, they are not powers" (Russ 2004, p. 349). Secondly, in the printed version the substitution of $1 + \frac{n}{m}x$ by $1 + ix$ appears as "1" (cf. Bolzano 1816, p. 77).

expressions" and some others, had yet to be clearly developed and that, in fact, this lack of clarity about these concepts themselves prevented him from addressing some cases in which they were used, for example, those involving the so-called imaginary quantities or "when a negative quantity is to be raised to a power with an irrational exponent."[25]

On the one hand, throughout his works of 1816–17 Bolzano referred to the irrationals as quantities and not as numbers.[26] Such an unequivocal use, then, in these works—which include his *Purely Analytic Proof* (henceforth *RaB*)[27]—of the appellation "quantity" to refer to the irrationals would not be merely a conceptual or even a nominal issue but would have practical, operational implications. On the other hand, the notion of quantities ω and Ω, which he introduced in *BL*, was a key element not only in that work, but also in his works of 1817 and in particular in *RaB*, a work in which he took very significant and ground-breaking steps toward demonstrating highly relevant theorems, such as the one known as the intermediate value theorem (i.e., if a continuous function of x is positive for a value of x and negative for another value, it must be 0 for an intermediate value).

As Bolzano explained in the preface to *BL*, he was opposed to the notion of infinitesimals or quantities that "are supposed *to be* smaller than any [...] *alleged*, i.e. conceivable quantity," instead of which he proposed the use of the notion of variable quantities "*which can become smaller than any given* [*one*]."[28] This reluctance with regard to infinitesimals was nothing strange among mathematicians at the time (cf. Lagrange 1797; du Bourguet 1810). However, as discussed before in connection with his 1804 written exam for a chair at Prague University, Bolzano's criticism had to do not only with what he considered the lack of rigor in the procedures associated with infinitesimals and the "attempted calculation of the infinite," but also with the concept itself, which, he said, was "self-contradictory."[29] Thus, as the first version of his alternative notion to infinitely small quantities explicitly states (a notion that would correspond to the quantities that in §14 of *BL* he designated by the symbol ω, Ω being the "algebraic *sum* or *difference*" of a "(finite) multitude" of ω's),[30] the idea underlying it was that "one can assume an

[25](Bolzano 1816, pp. 143–144; cf. Russ 2004, p. 247): "wenn eine *negative* Größe auf eine Potenz mit *irrationalem* Exponenten erhoben werden sollte"; "*imaginäre*[n] *Ausdrücke*"; "*Irrationalität* einer Größe."

[26]Indeed, in his early mathematical works (1804–17) Bolzano only once used the expression "irrational[e] Zahl," namely in the formulation of one of the last problems in his 1816 work, in the solution of which, however, he again referred to irrational quantities (Bolzano 1816, p. 137).

[27]The full title of Bolzano's work is *Rein analytischer Beweis des Lehrsatzes, daß zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege*.

[28](Bolzano 1816, p. V): "schon kleiner *seyn* soll[en], als eine jede [...] *angebliche* d. h. gedenkbare Größe"; "*die kleiner als jede gegebene werden können*."

[29](Bolzano 1816, pp. IV & XI): "versuchte Berechnung des Unendlichen"; "selbst widersprechenden."

[30](Bolzano 1816, p. 15): "algebraische *Summe* oder *Differenz*"; "(endliche) Menge."

even smaller ([or] greater) [quantity] for each already assumed one."[31] In other words, Bolzano stressed the fact that one could, at most, use quantities that "can be" smaller than a "given" one, but not quantities that are *de facto* smaller than any "conceivable" one, while resisting the idea of identifying such quantities (or any quantity) with 0 (cf. Fuentes Guillén and Martínez Adame 2020).

It should be emphasized that at the time, and in spite of the innovative features of the mathematics that he was developing, Bolzano still advocated a traditional notion of number. As he wrote in "On the Concept of Quantity and the Different Kinds Thereof,"[32] dated 1816, "only that which mathematicians usually call a *whole number* deserves to be called a *true* number," while the rest were "not actually numbers, but quantities, or not even quantities."[33] Moreover, in that paper Bolzano pointed out that in previous works, *BL* among them, he had relied on the "common use of language" to refer to numbers and quantities, despite having already "recognized its incorrectness," since he "really feared" that by deviating from it his works would be "completely incomprehensible."[34] As he went on to explain, for him:

> *Any thing X is to be called a quantity (**Quantum**) insofar as it possesses a property, in consideration of which another thing M, called measure, behaves with respect to X in the same way as two divisible things of the same kind behave with respect to each other, one of which is the unit [and] the other an arithmetically determinable whole N, i.e. a thing for which a rule can be given, according to which, for any arbitrary multitude of equal and aliquot parts into which the unit is decomposed, one can determine how many of those parts are contained in N.*[35]

From Bolzano's point of view, such a relation could also be expressed by pointing out that "*X is measured* [. . .] *by N*," the latter being called "the *number* that expresses the quantity of *X*," albeit "very improperly" since "[o]nly in the rarest cases is *N* an actual [or real] number, although it is always something *determinable* by *numbers* (several,

[31](*BBGA* 2B7/1, p. 79): "man zu jedem schon angenommenen ein noch kleineres (größeres) annehmen kann."

[32]The title is "*Uiber den Begriff der Größe und die verschiedenen Arten derselben*."

[33](*BBGA* 2A5, p. 191): "nur dasjenige, was die Mathematiker gewöhnlich eine *ganze Zahl* nennen, den Nahmen einer *wahren* Zahl verdiene"; "sind eigentlich nicht *Zahlen*, sondern *Größen*, oder nicht einmahl *diese*."

[34](*BBGA* 2A5, p. 191): "gewöhnlichen Sprachgebrauch"; "die Unrichtigkeit desselben [. . .] erkannt"; "auch wirklich befürchten mußte, hiedurch ganz unverständlich zu werden."

[35](*BBGA* 2A5, p. 199): "*Größe (**Quantum**) heißt jede Ding X, so fern es irgend eine Eigenschaft besitzt, in Betracht deren ein anderes M, das Maß genannt, sich zu X so verhält, wie sich zwey theilbare Dinge derselben Art verhalten, deren eines die Einheit, das andere ein arithmetisch bestimmbares Ganze N ist, d.h. ein Ding, von dem eine Regel angeblich ist, nach welcher man bey jeder beliebigen Menge gleicher und aliquoter Theile, in die man die Einheit zerlegt, bestimmen kann, wie viele dergleichen in N enthalten sind*." This passage makes it clear that, at least in the case of Bolzano, the appropriate translation for "Größe" is "quantity" and not "magnitude" (from the Latin *magnitūdō*).

often infinitely many)."[36] As will be discussed in the following sections, this is basically the idea of measurement that underlies Bolzano's later project of a theory of quantities and, within it, a theory of numbers. This means that for him, both then and later, "the concept of *quantity* contain[ed] that of *number* as a characteristic."[37]

Interestingly, in his 1816 paper Bolzano maintained that, according to such a definition of quantity and as he had already stated in *BL*, the concept of infinitely small quantities was self-contradictory, while the infinitely great and the imaginaries were "*improperly* so-called" ("*uneigentlich* so genannt") due to certain similarities with quantities (cf. *BBGA* 2A5, pp. 201–203). However, two years later, in a note included in his mathematical notebooks and dated September 9, 1818, he showed himself inclined to accept the notion of infinitely small quantities (cf. *BBGA* 2B9/2, p. 126).

Many years later, in *Paradoxes of the Infinite* (henceforth *PU*),[38] the final version of which dates from the late 1840s and was published posthumously, Bolzano went back to the notion of the measurement of a quantity by another one with equal status standing in a certain ratio to it (cf. Bolzano 1851/1920, pp. 22 & 59) and insisted on the fact that, in a sense, "all numbers [are] also quantities [and] there are even more quantities than numbers," "because also the fractions $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \ldots..$, as well as the so-called *irrational* expressions $\sqrt{2}, \sqrt[3]{2}, \ldots..., \pi, e, \ldots..$ designate quantities."[39] Furthermore, in that work he addressed the discussion that he had left hanging in his 1815 note on the similarities and differences between 0, $\sqrt{-1}$ (both "objectless" quantity ideas or expressions)[40] and the infinitely small quantities, which, he insisted, should not be regarded as equivalent (Bolzano 1851/1920, p. 56).

In particular, Bolzano devoted several paragraphs in *PU* to determining, in a precise manner, the concept of zero[41] and, in connection with this, to discussing the

[36](*BBGA* 2A5, p. 199): "*X werde* (bey dem Maße *M*) *durch N gemessen*"; "die *Zahl*, welche die Größe von *X* ausdrückt"; "sehr uneigentlich"; "Nur in den seltensten Fällen ist *N* eine wirkliche Zahl; wohl aber ist es jederzeit etwas durch *Zahlen* (mehrere, oft selbst unendlich viele) *Bestimmbares*."

[37](*BBGA* 2A5, p. 208): "der Begriff der *Größe* jenen der *Zahl* als ein Merkmahl."

[38]The full title of Bolzano's work is *Paradoxien des Unendlichen herausgegeben aus dem schriftlichen Nachlasse des Verfassers von Dr. Fr. Příhonský*.

[39](Bolzano 1851/1920, p. 21): "alle Zahlen zugleich auch Größen, sondern es gibt noch weit mehr Größen als Zahlen, weil auch die Brüche $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \ldots..$, ingleichen die sogenannten *irrationalen* Ausdrücke $\sqrt{2}, \sqrt[3]{2}, \ldots..\pi, e, \ldots..$ Größen bezeichnen."

[40]Concerning the translation of "gegenstandlos" as "objectless" (*gegenstand-los*), it is worth noting, as Russ does, that Bolzano uses it "in contrast to *gegenständlich* meaning, of an idea, that it does have objects associated with it" (Russ 2004, p. XXIX).

[41]Here Bolzano acknowledges Martin Ohm for "having been the first who drew the attention of the mathematical public to the difficulties in the concept of zero" (Bolzano 1851/1920, p. 55). Interestingly, in the work to which Bolzano refers, Ohm stated that "so würde die 0, als benannte Zahl gedacht, Nichts vorstellen," having established before that a "benannte Zahl" would be "diejenige Zahl, bey welcher wir uns eine völlig bestimmte und genannte Einheit denken," so that "[e]ine Größe *A* wird gemessen, wenn man sie als eine benannte Zahl *aE* ausdrückt, oder doch die Elemente findet, aus denen diese benannte Zahl *aE* leicht gebildet werden kann" (Ohm 1828, pp. 48–49).

idea of infinitely small quantities. Thus, on the one hand, he explained that while 0 could only be considered a quantity in a *translatitious* or figurative sense (*uneigentlichem Sinne*), it could be determined as the "symbol" (*Zeichen*) for which the equations $A - A = 0$ and $A \pm 0 = A$ hold in general (i.e., regardless the quantity expression A corresponds to an actual quantity or is "objectless"), so that $A \times 0 = 0$ (the concept of multiplication) and $B \times \left(\frac{A}{B}\right) = A$ (the concept of division) could be formed, as long as $B \neq 0$ (Bolzano 1851/1920, pp. 50–51 & 56–57).[42] In other words, Bolzano drew attention to the fact that, although strictly speaking 0 only accounts for the absence of a quantity, one might be entitled to consider it as a quantity since, leaving aside division by 0, the four basic operations can be carried out using it.

On the other hand, Bolzano insisted that the notion of infinitely small quantities led to contradictions (Bolzano 1851/1920, pp. 59–60). However, he pointed out that problems could be overcome if, in the expression $\frac{\Delta y}{\Delta x}$, both Δy and Δx were regarded as $=0$, while the corresponding notation for the derivative $\frac{dy}{dx}$ was not considered "as a *quotient* of *dy* by *dx*," that is, not as a ratio between zeros, "but only as a *symbol* of the derivative of *y* by *x*."[43] From Bolzano's point of view, this justified the practice of calculating using such signs inasmuch as these latter were considered "not as signs of actual quantities, but rather as equal to zero," and he added that, "on entirely similar principles," the same could be said about "calculation[s] using so-called *imaginary quantities*,"[44] the values of which he had already identified—along with the value of the "infinitely great"—as non-measurable (cf. Bolzano 1930, p. 16) in a work composed prior to *PU*, namely his *Theory of Functions* (*Functionenlehre*; henceforth *F*).

The following sections will discuss Bolzano's project of a *Pure Theory of Numbers* (*RZ*) in which he presented the notion of "measurable numbers" and which, coupled with *F*, is to be understood within the framework of his more ambitious project of a general theory of quantities. The development of such a project was carried out precisely between the publication of his early mathematical works (i.e., the period 1804–17) and the publication of *PU* (i.e., 1851), in which latter work he noted, as regards his considerations on the legitimacy of making calculations with certain symbols, that, ultimately, such considerations had led to "important truths of the general theory of quantities."[45]

[42]Bolzano noted that 0 should only be used as divisor in identity equations, such as $\frac{A}{0} = \frac{A}{0}$ (Bolzano 1851/1920, pp. 57–58).

[43](Bolzano 1851/1920, p. 67; Russ 2004, p. 639): "als einen *Quotienten* von *dx* in *dy*, sondern nur eben für ein *Symbol* der abgeleiteten von *y* nach *x*."

[44](Bolzano 1851/1920, pp. 69 & 71): "nicht als Zeichen wirklicher Größen, sondern sie vielmehr als der Null gleichgeltend betrachten"; "ganz ähnlichen Grundsätzen"; "sogenannten *imaginären Größen*."

[45](Bolzano 1851/1920, pp. 64 & 69): "*meistens ganz richtige Ergebnisse*"; "wichtige Wahrheiten der allgemeinen Größenlehre."

3 The Conceptual Framework of Bolzano's Measurable Numbers

According to Jan Berg, who is the editor of both the *Introduction to the Theory of Quantities and First Concepts of the General Theory of Quantities* (henceforth *EG*)[46] and *RZ* in *BBGA*, it is known from Bolzano's correspondence that he worked on his project of a theory of quantities from around 1830 onward and that by 1835 he had almost finished writing *EG* and *RZ* (Berg 1975, pp. 9–10; cf. Winter 1933, pp. 213–216; Russ 2004, p. 347). Thus, in order to better understand Bolzano's proposal of "measurable numbers," which he introduced in the seventh and last section of *RZ*, it seems necessary first to shed light on the conceptual groundwork of such a proposal—which Bolzano developed in *EG* and the previous sections of *RZ*—as well as to address some of the reflections on quantities and numbers included in Bolzano's mathematical notebooks from the early 1820s through to the early 1830s.

In an extant note from his mathematical notebooks entitled "Concept and Division of Mathematics," dated May 1830, Bolzano explicitly refers to the fact that, whereas in *BD* he had favored an alternative definition of mathematics,[47] by the time of this note, 20 years later, mathematics had become for him "once again a theory of quantities alone."[48] As noted by Russ in his English translation of Bolzano's mathematical works, the very title "for his mature comprehensive work on mathematics" (i.e., *Größenlehre*) makes it clear that Bolzano had embraced, by this time, once again the traditional definition of mathematics as the "science of quantities" (Russ 2004, p. 347).

Accordingly, Bolzano began *EG* by pointing out that "the word *quantity*" could be understood in both a narrow sense (*enger Sinn*), to refer only to so-called continuous quantities, and a broad sense (*weiter Sinn*), which would include also so-called discrete quantities, and that it was in this latter, broader sense that mathematics could be defined as "a science of quantities," a science which would therefore comprise the theory of numbers, "one of the most important mathematical disciplines."[49] For Bolzano, however, this did not mean that all mathematical disciplines dealt only with quantities, as would be shown by the case of the theory of combinations, which dealt with the combinations of which any given plurality of

[46]The title is *Einleitung in die Größenlehre und erste Begriffe der allgemeinen Größenlehre*.

[47]In 1810 Bolzano defined mathematics as "*eine Wissenschaft* [. . .], *die von den allgemeinen Gesetzen (Formen) handelt, nach welchen sich die Dinge in ihrem Daseyn richten müssen*" (Bolzano 1810, p. 11). For Paola Cantù, referring to Bolzano's *Größenlehre*, "[i]t is exactly for the sake of generality that Bolzano abandons the previous definition of mathematics and modifies the meaning of 'Größe' from quantity to quantity in general, as it was common in the tradition of the *mathesis universalis*" (Cantù 2014, p. 313).

[48](*BBGA* 2B12/1, p. 132): "Mathematik ist mir jetzt wieder nur Größenlehre, und zerfällt zunächst in die reine und angewandte." The title of Bolzano's note is "*Begriff und Eintheilung der Mathematik*." For an earlier note of Bolzano on this subject, cf. (*BBGA* 2B2/1, p. 23).

[49](*BBGA* 2A7, p. 25): "eine Wissenschaft der Größen"; "eine der wichtigsten mathematischen Disciplinen."

objects were susceptible (*BBGA* 2A7, pp. 27–28); this was a consideration, it must be said, that he had already introduced in *BD* (cf. Bolzano 1810, p. 4).

In addition to this, and with respect to the particular disciplines of the general theory of quantities, Bolzano noted that the theory of the infinitely small or great quantities, which for him represented a "special type" (*besondere Gattung*) of quantities, should be regarded as a peculiar but useful part of pure mathematics, of which the theory of numbers (another special type of quantities) would also be a part (cf. *BBGA* 2A8, pp. 33–34). Moreover, he explained that the latter could be used to accurately determine "all other quantities," provided only that—and this was certainly a proviso he intended to abide by in his own work—*numbers* were not considered just in the strict sense of the term, that is, provided that "special forms" (*Arten*) such as rationals and irrationals were also taken into account and thus "the properties of numbers [were] mixed with the properties common to *all* quantities, as well as with those that [were] only common to certain other forms."[50]

But, had Bolzano by this point already sorted out his uncertainties regarding the notions of number and quantity? In a broad sense, he wrote in *EG*, quantity would be every object "belonging to a kind of things such that, for each two, only one of the following two relations to each other can ever hold: they must either be *equal* to one another or one of them must contain a *part equal* to the other."[51] Thus, on the one hand, an irrational i could be regarded either as a quantity equal to a part of a rational quantity from which a certain quantity would have to be subtracted in order to obtain i, or as a quantity to one of whose parts a rational quantity could be equated and to which a certain quantity would have to be added in order to obtain i; whereas, on the other hand, the corresponding irrational number would be the numerical determination of i, which should comply with certain properties common to all quantities and to numbers.

The mathematical notebooks of Bolzano written throughout the previous decade seem very precisely to account for the process which eventually led him to such a stance in his *Größenlehre*. For example, in a note from 1821 he wrote that, while infinite quantities could be determined by saying that "a quantity is ∞ great if every number is a part of it and ∞ small if it is [. . .] $<$ than every fraction," imaginary quantities (*Größen*) such as "$\sqrt{-1}$, $\sqrt{-2}$, etc." should be regarded as "attributes that are determined by the numbers 1, 2. . ., but [. . .] cannot be assigned to actual objects" (in other words, while such expressions somehow involve numbers, there is no real object to which they correspond and so they are "objectless"), both the latter and the

[50](*BBGA* 2A7, p. 34): "aller übrigen Größen"; "wo (wie wir in diesem Buche selbst zu thun gesonnen sind) die Eigenschaften der Zahlen untermischt mit den Eigenschaften, die *allen* Größen gemeinschaftlich, ingleichen auch mit solchen, die nur gewissen anderen Arten (z.B. den irrationalen) Größen zukommen, abgehandelt werden."

[51](*BBGA* 2A7, pp. 25–26): "als gehörig zu einer solchen Art von Dingen betrachten, deren je zwei immer nur eines von folgenden zwey Verhältnissen gegen einander behaupten können: sie müssen entweder einander *gleich* seyn, oder das Eine derselben muß einen dem andern *gleichen Theil* enthalten." Bolzano was to go back to the principle of trichotomy in his theory of measurable numbers (*BBGA* 2A8, pp. 136–137 & 139; cf. Rychlík 1962, p. 5 & 56–59; Berg 1990, p. 151).

former being "*unmeasurable quantities*."[52] Then, in a note written in 1823, he defined quantity as "a collection of parts, among which two parts that are not equal to each other have instead the relation in which one of them contains a part that is equal to the other,"[53] but he noted that such a definition would rule out certain ideas (*Vorstellungen*), so that, in order to include "$0, \infty, \sqrt{-1}, \ldots$," a more general "concept of *quantity* or rather of a quantity *expression* (a quantity *idea*)" would be required.[54] Because of that, in a later note written between 1823 and 1824, Bolzano asked himself if it would not "be easier to say: A *quantity* (in abstracto) is such an *attribute* of a thing which is determinable by the thought of a *sum*, i.e. by the thought of a collection of parts in which no order among the same is perceived,"[55] from which he concluded that the so-called imaginary quantities, zero and the infinite (both great and small) would not be quantities but "mere *quantity ideas*" ("bloße *Größenvorstellungen*") (*BBGA* 2B11/2, p. 91).

Finally, in a note on "[t]he theorem *that factors with modified order give one and the same product*" (written between 1828 and 1829), Bolzano argued that such a theorem was valid not only for whole numbers, but also for fractions and "irrational quantities" such as "$i = \frac{m}{n} + \omega$, where m, n are whole numbers and ω can become as small as one desires," while adding that imaginary quantities "are not actual quantities at all, but only signs that have been obtained according to the rules that apply to all real quantities."[56] And then, in a note on "*imaginary quantities*" (written between 1829 and 1830), he insisted that these latter were not quantities but "*quantity concepts*," inasmuch as they represented a quantity "of which there is none."[57]

[52](*BBGA* 2B11/2, pp. 22–23): "eine Größe ist ∞groß, wenn jede Zahl ein Theil von ihr; und ∞klein, wenn sie ein Theil von jedem Bruch ist oder wie man sagt $<$ als jeder Bruch"; "$\sqrt{-1}, \sqrt{-2}$, etc sind Beschaffenheiten, die durch Zahlen 1, 2... bestimmt werden; obgleich Beschaffenheiten, die keinen wirklichen Gegenständen zukommen können"; "*unermeßlichen Größen*." This note is entitled "*Begriff der Größe*," dated September 1.

[53](*BBGA* 2B11/2, p. 86): "Die *Größe* ist ein Inbegriff von Theilen, zwischen deren je zweyen, die einander nicht gleich sind, das Verhältniß Statt hat, daß der Eine derselben einen dem anderen gleichen Theil enthält." This note is entitled "*Begriff einer Größe*" and is dated March 10, 1823.

[54](*BBGA* 2B11/2, p. 86): "einer *Größe* oder vielmehr eines Größen*ausdrucks* (einer Größen*vorstellung*)." I opted to translate "Vorstellung" by the term "idea," and not "representation," in order to emphasize its mental character and, at the same time, to distinguish it from "Darstellung," which can be associated with a "representation" by signs (cf. *BBGA* 2A8, p. 22). For a detailed account of "Vorstellung" in the work of both Kant and Bolzano, cf. (Rusnock 2000, p. 102ff.).

[55](*BBGA* 2B11/2, p. 91): "Sollte man denn nicht einfacher sagen können: Eine *Größe* (in abstracto) sey eine solche *Beschaffenheit* einer Sache, welche bestimmbar ist durch den Gedanken einer *Summe*, d.h. durch den Gedanken eines Inbegriffs von Theilen, bei welchen auf keine Ordnung unter denselben gesehen wird?" This note is entitled "*Begriff der Größe*" and it is not dated.

[56](*BBGA* 2B12/1, pp. 84–85): "Der Lehrsatz, *daß Faktoren mit veränderter Ordnung einerlei Produkt geben*"; "Irrationalgrößen"; "$i = \frac{m}{n} + \omega$ wo m, n ganze Zahlen, und ω so klein werden kann, als man nur immer will"; "sind nähmlich gar keine wirklichen Größen, sondern nur Zeichen, die nach denjenigen Regeln, welche für alle reellen Größen gelten erhalten worden sind."

[57](*BBGA* 2B12/1, p. 103): "*Imaginäre Größen*"; "*Größenbegriffe*"; "deren es keine gibt."

All this helps to understand the position taken by Bolzano on the notion of number in *RZ* and, in particular, in the first section of this latter work, entitled "On the Concept of Numbers, Their most General Properties and the Manner of Their Designation."[58] There, he began by stating that numbers are sums formed from a unit (i.e., 2, 3, 4, 5, 6, . . .), which meant that, strictly speaking, the "number sequence" (*Zahlenreihe*) was the "*natural sequence of numbers*" or the "*sequence of natural numbers*,"[59] which would be infinite, although in a broad sense one could also consider "objectless ideas of numbers" and, ultimately, numbers could be identified with "whole or actual" (*wirkliche*) ones (*BBGA* 2A8, p. 17). On the one hand, this explains Bolzano's subsequent discussion of several definitions of *number*, ranging from the definition common at the time (i.e., that of Kästner, although he only mentions Stahl) and the traditional definition in which "things of the same kind" was replaced by "units" (Euclid, Kraushaar), through to those definitions in which the meaning of "number" was extended also to fractions (Schröder, Stein) and even to the irrationals, the infinitely great and the infinitely small, but excluded zero and imaginaries (Stevin, Newton) (*BBGA* 2A8, pp. 18–20).

On the other hand, this led Bolzano to declare that by the "appellation" (*Benennung*) of a "number idea" he wished his readers to understand "any idea of the form '[a] number, which has the attributes b', b'', b'''', regardless of whether these attributes can in fact be found combined [*vereinigt*] in a number or not; i.e. regardless of whether this idea has a real [or actual] object or not."[60] Moreover, he made it clear that he would also call these ideas "number concepts for a change" (*zur Abwechslung*), the representation (*Darstellung*) of which was carried out by means of "signs" which he called "number expressions" (*BBGA* 2A8, p. 22). He then went on to write:

> Finally, to abbreviate, in cases where there is no misunderstanding to be worried about, we will also allow ourselves to simply call the number ideas and number expressions *numbers* themselves, and to do so indeed even at times when there are no numbers corresponding to these ideas or signs in reality.[61]

[58]The title is "Von dem Begriffe, den allgemeinsten Beschaffenheiten und der Bezeichnungsart der Zahlen."

[59](*BBGA* 2A8, p. 15): "die *natürliche Reihe der Zahlen*, oder mit Einigen auch *die Reihe der natürlichen Zahlen*."

[60](*BBGA* 2A8, pp. 18–21): "eine jede Vorstellung von der Form: »*Eine Zahl, welche die Beschaffenheiten* b', b'', b'''. . .*hat*,« verstehen wollen, gleichviel ob diese Beschaffenheiten bei einer Zahl in der That vereinigt angetroffen werden können oder nicht; d.h. gleichviel ob diese Vorstellung einen wirklichen Gegenstand hat oder nicht."

[61](*BBGA* 2A8, p. 22): "Zur Abkürzung endlich werden wir uns in Fällen, wo eben kein Mißverstand zu besorgen stehet, erlauben, die Zahlenvorstellungen und Zahlenausdrücke auch schlechtweg *Zahlen* selbst zu nennen, u. dieß zwar zuweilen selbst dann, wenn es dergleichen dieser Vorstellung oder Zeichen entsprechende Zahlen in Wahrheit gar nicht gibt."

After such explanations as these, Bolzano goes on in the second section of *RZ* to address the addition and subtraction of—natural—numbers (which, for him, did not comprise 0 or the unit), and shows that, given two numbers S and M:

(a) with $S > M$, the difference $S - M$ will be another number;
(b) with $S = M$, their difference (i.e., 0) will be an "objectless idea," as he had already pointed out in *EG* (cf. *BBGA* 2A7, p. 164);
(c) with $S < M$, and given the "arbitrary proposition" (*willkürliche Satz*) or "convention" (cf. Russ 2004, p. 106) according to which "in all numerical expressions we want to regard the unit to which they refer as a unit capable of *opposition*," the difference $S - M$ will nevertheless represent an "actual number" (*wirkliche Zahl*), although from an "opposite i.e. negative" (*entgegengesetzt d.h. negativ*) unit (*BBGA* 2A8, pp. 52–53).

In the third section of *RZ* he explains the product (and proves some properties, including distributivity and commutativity) and "quotients" of numbers, introducing the "designation" or "symbol" (*Zeichnung*) $\frac{p}{m}$ (*BBGA* 2A8, p. 67). Moreover, in the fourth section (on *Rationale Zahlen*), at the beginning of which Berg (editor of *RZ* in *BBGA*) notes that there is a marginal note in the manuscript in which Bolzano states that in these sections things could be "shortened by considering *quantities* instead of just numbers,"[62] Bolzano turns to look at number ideas (*Zahlenvorstellungen*) in which only a finite "multitude" (*Menge*)[63] of the four basic operations are to be performed and explains "simple fractions" (*einfache Brüche*), addressing their calculation rules (avoiding division by 0) and stating that an "elementary expression" will be any "number expression" (*Zahlenausdruck*) in which only a finite or infinite multitude of these operations occur (cf. *BBGA* 2A8, pp. 73–84). He does not discuss the latter case in this section (namely, expressions which involve an infinite multitude of the four basic operations), although he gives the following examples:

$$\text{“}1 - 1 + 1 - 1 + 1 - ..\textbf{in inf}:$$

or

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + ..\textbf{in inf}:\text{”}\ (BBGA\ 2A8, p.73).$$

[62] (*BBGA* 2A8, p. 73 fn. u): "In allen diesen Abschnitten läßt sich Manches verkürzen, indem statt bloßer Zahlen, *Größen* überhaupt betrachtet werden."

[63] Here I agree with both Peter Simons and Russ about the translation of "Menge" as "multitude," and not as "set," which emphasizes its primitive meaning as a "bunch" of something and avoids its association with the set theory developed decades later (cf. Simons 1997, pp. 95–96; Russ 2004, pp. XXVIII–XXIX; see also Ferreirós 2007, pp. XX–XXI & 72–73).

The fifth section of *RZ* focuses on the "Relations of Height Between Rational Numbers,"[64] meaning that form of comparison between different rational numbers whereby one will be "*greater* or *higher*" and the other "*smaller* or *lower*," so that for two unequal rational numbers one is always higher than the other and "there always lies [another] one *between* them," indeed there always lies an infinite "multitude" of such numbers.[65] For Bolzano this section was necessary in order to "justify" the extension of the meaning of that "relation of *being greater*" ("Verhältniß des *Größerseyns*"), constitutive of the "concept of number" "taken as a basis in this theory of numbers," so that it could be applied to rationals, which he regarded as "mere number ideas" with "no [associated] object, because in fact there is no number that is $=\frac{3}{5}$ or $=\frac{2}{5}$ or $=\frac{1}{5}$."[66] Otherwise, he points out, one would be led to consider relations of height involving "objectless" (*gegenstandlos*) number ideas, as in the case of "$1+\sqrt{-1}>1$, which no mathematician will admit" insofar as the combination of numbers and imaginaries would not admit such a relation, or expressions such as "$1+\frac{1}{1+1+1+\text{in inf.}}$," which, not being a "finite elementary number expression," would entail an extension of the concept of number.[67]

Such an extension is performed in the sixth section of *RZ*, entitled "Rational Numbers Which Can Increase or Decrease to Infinity,"[68] where Bolzano explains the "concepts" of these numbers which, he writes, should not "[be confused] with the concepts of an infinitely great and an infinitely small number": whereas the latter

[64]The title is "Verhältnisse der Höhe zwischen den rationalen Zahlen." In his translation of the seventh section of *RZ*, Russ translates "Höhe" as "order" (Russ 2004, p. 395). Undoubtedly, the notion of "height" involves a certain relation of order among the multitude of rational numbers. However, in order to provide a faithful account of Bolzano's proposal and to avoid any possible—and misleading—set-theoretical interpretation, I opted not to translate "Höhe" as "order." Similarly, and with regard to a paragraph of Bolzano's *Theory of Science* (*Wissenschaftslehre*; henceforth *WL*) on the relation of subordination (*Unterordnung*), Jan Šebestík has translated "Höhe" as "*hauteur*" (Šebestík 1992, p. 177; cf. Bolzano 1837, p. 451). Later, Dedekind and Cantor were to use what they called the "height" (*Höhe*) of algebraic numbers in order to enumerate these latter (cf. Cantor 1874, pp. 259–260; Ferreirós 2007, pp. 178–180).

[65](*BBGA* 2A8, pp. 86, 90 & 92): "*größer* oder *höher*"; "*kleiner* oder *niedriger*"; "so gibt es jederzeit eine, die *zwischen* ihnen liegt"; "Die Menge der Rationalzahlen, die zwischen je zwey von einander verschiedenen Rationalzahlen *A* und *Z* eingeschaltet werden können, geht ins Unendliche." Nowadays this can be expressed by saying that the set of rational numbers are densely ordered in the reals. Such an approach, however, is foreign to Bolzano's proposal.

[66](*BBGA* 2A8, p. 87): "ich in dieser Zahlenlehre einmahl zu Grunde gelegt habe"; "bloße Zahlenvorstellungen"; "die keinen Gegenstand haben, weil es in der That keine Zahl gibt, die $=\frac{3}{5}$ oder $=\frac{2}{5}$ oder $=\frac{1}{5}$ wäre." Indeed, Bolzano once again acknowledges Ohm "for having first recognized this need for a particular conceptual determination [*Begriffsbestimmung*] of the signs $>$ and $<$ when applied to general numbers [*allgemeine Zahlen*]" (*BBGA* 2A8, p. 87). It should be noted that, in the paragraph that Bolzano refers to, Ohm discusses the relations of "being greater than" and "being smaller than" between "real numbers" (*reele Zahlen*), which for him, as he explained before, consisted of the integers, fractions and zero (Ohm 1828, pp. 123–124).

[67](*BBGA* 2A8, pp. 87 & 88): "was doch kein Mathematiker zugeben wird"; "endliche[r] elementare[r] Zahlenausdru[ck]."

[68]The title is "Rationale Zahlen, die ins Unendliche zu- oder ab-nehmen können."

involve "thinking" of a "multitude that *consists* of infinitely many units," the former involve thinking of variable numbers that "*can* take on infinitely many values, including those that are greater [or smaller] than any other given number."[69] Consequently, from Bolzano's point of view, whereas the "multitude of clear ideas that can be cultivated over time in the mind of a thinking being" would be a number idea that *can* increase to infinity while "being and always remaining finite,"[70] "the multitude of points lying on a given line," thought "as a number," would "form the (admittedly self-contradictory) concept of an infinitely great number."[71] Or, as he later explained in *WL*, no "attempted enumeration" (*versuchten Abzählung*) of the members of an infinite multitude of points on a bounded line could ever be achieved (Bolzano 1837, p. 413).

The idea of rational numbers that *can* increase or decrease to infinity is a very interesting one not only because the concept of such numbers, along with Bolzano's criticism of self-contradictory concepts of number, reveals a link with Bolzano's earlier works (cf. Sect. 2), but also because of what he goes on to explain about the former in this section of *RZ*. For him, the idea associated with a "sign" (*Zeichen*) designating a *number* that increases or decreases to infinity was a "multi-form number idea" (*mehrförmige Zahlenvorstellung*), given that it involves "infinitely many values," none of which would be the greatest or smallest (*BBGA* 2A8, pp. 95–96). This meant that, as he had already explained in the first section of that work, such a number idea and its concept would comprise infinite numbers and thus a particular sign would be required to denote said number idea, while, for example, a seven-form number concept would be that of the number between 3 and 11, or 21 and 29, comprising in each case seven different numbers (4, 5, 6, 7, 8, 9, 10 and 22, 23, 24, 25, 26, 27, 28, respectively; cf. *BBGA* 2A8, p. 24).

As a consequence, in the case of the aforementioned number idea involving infinitely many small values, Bolzano introduced an arbitrary proposition or convention (*willkürliche Satz*) according to which the signs "ω, Ω, ω^1, ω^2, etc. [denote]

[69](*BBGA* 2A8, p. 95): "verwechsel[t] [werden] mit den Begriffen einer *unendlich großen* und *unendlich kleinen* Zahl"; "eine Menge denke, die aus unendlich vielen Einheiten *bestehet*"; "die der Werthe unendlich viele und darunter auch solche annehmen *kann*, die größer als eine jede andere gegebene Zahl sind" (emphasis added in "bestehet" and "kann").

[70]It is interesting to note the subtle contrast between Bolzano's stance on this issue and the proof of Dedekind's theorem 66 in his *Was sind und was sollen die Zahlen?*, where the latter stated that "the totality S of all things that can be the object of my thought is infinite" (Dedekind 1888, p. 17): whereas for Dedekind this proves that there *exist* "infinite systems" (*unendliche Systeme*), as in the case of "my world of thoughts," for Bolzano the number idea of one's own clear ideas *can* increase to infinity but remains finite. Nevertheless, as Dedekind himself pointed out in a footnote with regard to (Bolzano 1851/1920, pp. 13–14), Bolzano does consider "infinite multitudes" (cf. Bolzano 1817a, pp. 4–6).

[71](*BBGA* 2A8, p. 95): "[d]ie Menge der klaren Vorstellungen, welche sich in dem Gemüthe eines Denkenden Wesens mit der Zeit ausbilden können, ist und bleibt immer endlich"; "die Menge der Punkte, die in einer gegebenen Linie liegen; und denken wir uns sonach diese Menge als eine Zahl; so bilden wir uns den (freylich sich selbst widersprechenden) Begriff einer unendlich großen Zahl."

numbers that can decrease indefinitely."[72] The theorem that follows this proposition, which deals with "the algebraic sum of a finite and non-variable multitude of rational numbers $\omega^1, \omega^2, \ldots \omega^n$," makes it clear that, at least at that point, those signs stand for rational numbers that decrease and thus vary to infinity, which in turn allows Bolzano to hold that each non-variable rational number is to be enclosed by a pair of variable numbers that "can come as close to each other as one desires."[73] In other words, as the proof (*Beweis*) of the latter theorem shows, the non-variable rational number A would be the limit to which the pair of rational numbers

$$X = A - \frac{1}{2N} \text{ and } Y = A + \frac{1}{2N}$$

would approximate, with N denoting a number that can become as great as one desires (*BBGA* 2A8, p. 99), a procedure which resembles that which he had already proposed in his mathematical notebooks as a procedure for determining an irrational quantity. As it will be discussed in the following section, it is with reference to the proposition about enclosed non-variable numbers that, in the last section of *RZ*, Bolzano goes back to the possibility pointed out at the beginning of this book's fourth section regarding numerical expressions in which an infinite multitude of the four basic operations occur, which he calls "infinite quantity concepts."

4 Bolzano's *meßbare Zahlen*

As pointed out by Berg, whereas the final title that Bolzano assigned to the seventh section of *RZ* was "*Infinite Quantity Concepts* (Quantity Expressions),"[74] he had initially used the expression "Infinite Number Concepts" (*Unendliche Zahlenbegriffe*; *BBGA* 2A8, p. 100 fn. m). In this section, Bolzano sets out to consider expressions that represent numerical concepts involving an infinite multitude of basic operations, that is, infinite quantity expressions, such as the sum of the infinite "multitude of all actual [or real] numbers," the expression of which would be "1 + 2 + 3…in inf.," or even such as the sum of the fractions (*Brüche*) "$\frac{1}{2} - \frac{1}{4} + \frac{1}{8} - \frac{1}{16} + ..$**in inf**:," which arises by the division of the unit by 2 and the multiplication of the denominators of the subsequent fractions by -2.[75]

[72](*BBGA* 2A8, p. 97): "*Willk. S.* Zahlen, die ins Unendliche abnehmen können, wollen wir zur Abkürzung zuweilen durch Zeichen wie ω, Ω, ω^1, ω^2, u.s.w. andeuten."

[73](*BBGA* 2A8, pp. 97 & 98): "Die algebraische Summe aus einer endlichen und unveränderlichen Menge von Rationalzahlen $\omega^1, \omega^2, \ldots \omega^n$"; "Zu jeder unveränderlichen Rationalzahl A gibt es ein Paar veränderliche Zahlen, die sie als ihre Grenzen zwischen sich einschließen (d.h. deren die Eine höher, die andern niedriger ist) während sie selbst einander so nahe rücken können, als man nur immer will."

[74]The title is "*Unendliche Größenbegriffe*. (Größenausdrücke)."

[75](*BBGA* 2A8, p. 100): "die Menge aller wirklichen Zahlen." A similar statement can be found in (Bolzano 1851/1920, p. 20).

However, Bolzano notes in §3 that an infinite number concept does not require that the idea of every one of the operations that it involves be comprised in it as "singular constituent parts," in which case it would be "unthinkable for a finite mind such as ours."[76] This remark is connected with his previous and subsequent discussion of concepts and determinable objects. For example, in *BD* (1810) he established that composite concepts were decomposable (*zerlegbaren*) and definable, unlike simple concepts (e.g., a point) or "indeterminate or infinite" concepts (*unbestimmte oder unendliche*; concepts of the form "[e]*verything which is not A* [...] *is also not M*"; Bolzano 1810, pp. 43, 56–57 & 84–85), so that, in the case of a concept composed of an "infinite number of parts" (*unendlich vielen Theilen*), this would not be definable (*BBGA* 2A8, p. 101). Later, in *The Three Problems* (henceforth *DP*; 1817),[77] he introduced the notion of an infinite multitude of quantities and explained that such a concept was not contradictory and indeed that such a multitude could be "determinable" (*bestimmbar*) given a rule (or law) or a finite multitude of rules that determines all its parts, as in the case of the functions "fx, $\overline{f}x$, $\overline{\overline{f}}x$, ...," which "vary according to the law of continuity," or a straight line, "determined by the mere specification of its two endpoints and by a single law."[78] As he explained further, in such cases "only the *multitude* of its parts is indeterminable, but not the multitude itself, if there is a rule that determines every single one of those parts."[79] And, finally, when addressing, in *PU*, the "paradox" whereby, according to the concept established for them, the—natural—numbers would constitute an infinite multitude of all the finite multitudes that said numbers are, he pointed out that this multitude was determined by "the *rule of formation*" according to which "each of its terms has a *succeeding one*."[80]

More importantly, it is in §5 that Bolzano posits for the first time the property that, as he explains in §6, will allow him to consider what he calls "measurable expressions," namely:

> *Theo*[rem]. Among the infinite number concepts there are also some that are of such a nature that, for any arbitrary actual number q that we want to regard as the denominator of a

[76](*BBGA* 2A8, p. 101): "eigenthümlichen Bestandtheile"; "und eben deßhalb für einen endlichen Verstand, wie es der unsrige ist, undenkbar."

[77]The full title of Bolzano's work is *Die drey Probleme der Rectification, der Complanation und der Cubirung, ohne Betrachtung des unendlich Kleinen, ohne die Annahmen des Archimedes, und ohne irgend eine nicht streng erweisliche Voraussetzung gelöst; zugleich als Probe einer gänzlichen Umstaltung der Raumwissenschaft, allen Mathematikern zur Prüfung vorgelegt*.

[78](Bolzano 1817a, p. 4): "nach dem Gesetze der Stetigkeit ändern." For Bolzano's use of superscripts, cf. footnote 24.

[79](Bolzano 1817a, p. 6): "nur die *Menge* seine Theile unbestimmbar, nicht aber er selbst, wenn anders irgend ein Gesetz vorhanden ist, das jeden einzelnen aus diesen Theilen bestimmt"; "durch die bloße Angabe ihrer zwey Endpunkte und durch ein einziges Gesetz [...] bestimmt."

[80](Bolzano 1851/1920, p. 21; cf. Russ 2004, p. 611): "Denn nach dem, in der Erklärung jener Reihe (§8) angegebenen *Bildungsgesetze* derselben hat jedes ihrer Glieder wieder ein *folgendes*."

> fraction, a numerator p can be found which is again a positive or negative actual number, or sometimes even a zero, [and] with the outcome that we obtain the two equations $S = \frac{p}{q} + P$ and $S = \frac{p+1}{q} - P^1$, in which the sign S denotes the infinite number expression, but the P and P^1denote a pair of strictly positive number expressions or, [in the case of] the former, sometimes even a mere zero.[81]

As Bolzano explains further in the following paragraph, this "means of determining infinite number concepts" was in fact very useful, since it implied "infinitely many operations" that would otherwise be impossible for us.[82] Moreover, he says, such a means could be used to "*determine or measure by approximation*," that is, "*up to a* $\frac{1}{q}$," "every not only infinite but also finite number expression S," which he calls "measurable or estimable" (*meßbar* oder *ermeßlich*).[83] Therefore, whereas in the case of rational numbers the "*measuring*" (*messenden*) fraction $\frac{p}{q}$ would be "*complete* or *perfect*" (*voll* oder *vollkommen*), so that the "complement" (*Ergänzung*) P would consequently be $=0$, in other cases (leaving aside the "unmeasurable or inestimable" expressions, that is, those for which the aforementioned equations do not occur) one would require not only the measuring fraction plus a complement to measure S but also the "*next greater* fraction" (*nächst*[e] *größer*[e] Bruch) $\frac{p+1}{q}$ (*BBGA* 2A8, p. 105). In other words, those equations would constitute the rules that would allow one to determine numerically a certain quantity, either completely or approximately.

In order to better understand Bolzano's proposal, let us consider the concrete case of $\sqrt{2}$, an example of a non-rational measurable expression (or "number," as he states in §7) and an example, moreover, which, as discussed in Sect. 2, he had addressed years earlier in his mathematical notebooks. He is no longer here, as he was back in 1814, providing some rules for finding out if a quantity is measurable; instead, he is requiring that the aforementioned equations be satisfied in order to determine or measure the number expression S, without, however, identifying the latter with anything specific, as for example Cantor (with a fundamental sequence of

[81](*BBGA* 2A8, p. 102; cf. Russ 2004, pp. 358–359): "*Lehrs*. Unter den unendlichen Zahlenbegriffen gibt es auch einige, die von einer solchen Beschaffenheit sind, daß sich zu jeder beliebigen wirklichen Zahl q, die wir als Nenner eines Bruches betrachten wollen, ein Zähler p, der abermahls eine positive oder negative wirkliche Zahl, oder zuweilen auch eine Null ist, mit dem Erfolge auffinden läßt, daß wir die beyden Gleichungen erhalten $S = \frac{p}{q} + P$ und $S = \frac{p+1}{q} + P^1$, in welchen das Zeichen S den unendlichen Zahlenausdruck, die P und P^1 aber ein Paar durchaus positive Zahlenausdrücke oder das erstere zuweilen auch eine bloße Null bedeutet."

[82](*BBGA* 2A8, pp. 103–104): "dieß Mittel zur Bestimmung unendlicher Zahlenbegriffe"; "unendlich vielen Verrichtungen."

[83](*BBGA* 2A8, p. 104): "*näherungsweise bestimmen oder messen*"; "*bis auf ein* $\frac{1}{q}$"; "jeden, nicht nur unendlichen, sondern auch endlichen Zahlenausdrucke S." Here I follow Russ' translation of "ermeßlich" as "estimable"—in measure—and therefore of "unermeßlich" as "inestimable" (Russ 2004, p. 361). According to the corresponding entry in the last volume of the *Deutsches Wörterbuch* by Jacob and Wilhelm Grimm, "*ermeszlich*" meant "*mensurabilis*" (Grimm and Grimm 1862, p. 915).

rational numbers)[84] and Dedekind (with cuts)[85] were to do a few decades later. Thus, given $S = \sqrt{2}$, one would be able to measure it up to $\frac{1}{100}$ by means of, on the one hand, the measuring fraction $\frac{141}{100}$ plus the complement P and, on the other hand, $\frac{142}{100}$ (the next greater fraction) from which P^1 should be subtracted, with $P + P^1 < \frac{1}{100}$. Clearly, one would be able to carry out such a measurement as far as desired, so that, following the "visualization" proposed by Russ and Trlifajová, one would always be able to approximate more closely to the rational division marks enclosing its value on a line of length $>S$ (cf. Russ and Trlifajová 2016, p. 43).

Throughout the following paragraphs Bolzano proves several theorems, including the following: that "[a]ny *rational number* is a [complete] measurable number" and vice versa (§7); that "[i]f A is measurable, then $-A$ is also measurable" (§10); and that "there are [...] not two different [numerators] p^1 and p^2" of a pair of measuring fractions with the same q as denominator and which correspond to the same number expression S so that there are two different pairs of what could be called measuring equations (although this is not Bolzano's terminology) corresponding to S (§11).[86] Then, however, in the theorem of §21 and the definition of §22, and with reference to an addendum (*Zusatz*) to the previous theorem in which he had pointed out that the numbers

$$A = 1 - \frac{1}{1 + 1 + 1 + \ldots \textbf{ in inf.}} \text{ and } B = 1$$

have different measuring fractions (§20; cf. *BBGA* 2A8, pp. 111–112), Bolzano notes that "[a]mong the infinite number concepts there are also some" that "represent" numbers that can be called "*infinitely small*," such as

$$S = \frac{1}{1 + 1 + 1 + \ldots \textbf{ in inf.}},$$

and which, "according to the concepts [introduced] so far" (Bolzano once more refers to (*BBGA* 2A7, p. 164)), should not be considered as $= 0$.[87] In other words, such number concepts were those which, "in the attempted process of measuring"

[84] A fundamental sequence $a_1, a_2, \ldots a_n, \ldots$ of rational numbers has the property that the difference $|a_{n+m} - a_n|$ becomes as small as desired as n increases, b (a numerical quantity or *Zahlengrösse*) being the limit of the sequence and B (the totality of such numerical quantities) being the domain of real numbers (Cantor 1872, pp. 123–124).

[85] A Dedekind's cut decomposes all rational numbers into two classes, such that each rational number in the class to the left is smaller than each one in the class to the right, the cut being either a rational or an irrational number (Dedekind 1872, pp. 19–21).

[86] (*BBGA* 2A8, pp. 105–107; cf. Russ 2004, pp. 361 & 363): "Jede *Rationalzahl* ist eine meßbare Zahl, und zwar läßt sich ein volles Maß für sie angeben"; "Ist A meßbar, so ist auch $-A$ meßbar"; "gibt es [...] nicht zwey von einander verschiedene Zahlen p^1 und p^2, welche als Zähler des messenden Bruches."

[87] (*BBGA* 2A8, pp. 112–113): "Unter den unendlichen Zahlenbegriffen gibt es auch einige"; "vorstell[en]"; "*unendlich kleine*"; "nach den bisherigen Begriffen."

and with the numerator of the measuring fraction = 0 (so that the measuring fraction is $\frac{0}{q} + P^1 = P^1$ or, following Bolzano's initial notation, $= P$), "the two equations $S = P^1 = \frac{1}{q} - P^2$ [or $-P^1 = -\frac{1}{q} + P^2$, for negative S] hold, where q is taken to be as great as one desires."[88]

Bolzano goes on to say, first of all, that he will call a "*finite* number" (*endliche* Zahl, which should not be confused with the appellation "finite expression") any non-infinitely-small measurable number (end of §22), and, secondly, that infinitely small numbers are not "actual numbers" (*wirkliche Zahlen*; §23). Then, in the theorem of §26 and the definition of §27, he notes that, conversely, "[a]mong the infinite number concepts there are also some" for which, "in the attempted process of measuring," "there is a p for every q, such that one of the two equations $S = \frac{p}{q} + P^1 = \frac{p+1}{q} - P^2$ is satisfied, but not both of them at once,"[89] infinite number concepts the representation of which he calls "infinitely great" positive or negative numbers, depending on whether they satisfy the first or second equation, respectively (*BBGA* 2A8, p. 114). Examples of these *numbers* would be those designated by the expression $S = 1 + 2 + 3 + 4 + \ldots$ **in inf**. and, for negative S, $-1 - 2 - 3 - 4 - \ldots$ **in inf**. (cf. *BBGA* 2A8, p. 114). However, unlike the infinitely small numbers, the infinitely great *numbers* would be "unmeasurable" (*unermeßlich*) inasmuch as they would not comply with the requirement of the two measuring equations set out in §§5–6 (cf. *BBGA* 2A8, p. 115).

Interestingly, Bolzano points out that there would, nevertheless, still also be unmeasurable numbers that are not infinitely great, such as the *number* represented by the expression $1 - 1 + 1 - 1 + \ldots$ **in inf**. and 0 (§§29–30; cf. *BBGA* 2A8, p. 115). Thus, and at least up to this point in *RZ*, his theory of numbers (in the broad sense of this term) would not entirely correspond to our modern continuum of real numbers. This inasmuch as it comprises:

(a) the unit, that is, 1;
(b) the numbers in the strictest sense of the term, that is, 2, 3, 4, 5, 6, . . .;
(c) the negative numbers;
(d) the rational numbers;
(e) the non-rational measurable numbers, which so far seem to hint at our irrational numbers but also include the infinitely small numbers;

[88](*BBGA* 2A8, p. 113): "bey dem versuchten Geschäfte des Messens"; "die beyden Gleichungen $S = P^1 = \frac{1}{q} - P^2$ bestehen, q werde so groß als man nur immer will genommen." Bolzano introduces the denomination of "infinitely small *negative* number" (unendlich kleine *negative* Zahl) and its corresponding measuring equations in this very same paragraph.

[89](*BBGA* 2A8, pp. 112–114): "Unter den unendlichen Zahlenbegriffen gibt es auch einige"; "bey dem versuchten Geschäfte des Messens [. . .] für welche die beyden Gleichungen $S = P^1 = \frac{1}{q} - P^2$ bestehen, q werde so groß als man nur immer will genommen"; "für jedes q wohl ein p gibt, das Einer der beyden Gleichungen $S = \frac{p}{q} + P^1 = \frac{p+1}{q} - P^2$, keines aber, das beyden zugleich entspräche."

(f) the unmeasurable numbers, such as the infinitely great numbers, 0 and what nowadays would be look on as "divergent series."

Furthermore—although, as it must once again be emphasized, this holds true of Bolzano's proposal in *RZ* only up to this point, namely §30 of the seventh section—his "measurable numbers" would not strictly correspond either to our modern continuum of real numbers, since they would include the infinitely small numbers but not 1 and 0, or to our nonstandard reals, since Bolzano's measurable numbers would indeed not include 1, 0 or the infinitely great numbers.

In the following paragraphs Bolzano focuses on operations involving infinitely great and small numbers, without saying anything further about $1 - 1 + 1 - 1 + \ldots$ **in inf**. and suchlike expressions. To mention just a few results, in §32 he proves that the sum of an infinitely great and a rational number is a number of the first kind; in §§34–35 he proves that the sum of two, or of a "finite multitude" (*endlich*[e] *Menge*) of either infinitely small numbers or infinitely great numbers of the same sign, is infinitely small or infinitely great, respectively; in §36 he states that the difference between a pair of infinitely great numbers can be zero (if they are equal), infinitely small (if an infinitely small number is added to one or the other of them), finite (if a finite number is added to one or the other of them), or even infinitely great (if an infinitely great number is added to one or the other of them); in §§37–38 he proves that the product of an infinitely small or great number by a finite one (thus excluding 0) or by another infinitely small or great number is infinitely small or great, respectively; in §39 he proves that the product of a finite multitude of what he calls finite numbers cannot be either infinitely small or infinitely great, while in §40 he proves that the quotient of a finite number divided by another finite number cannot be either infinitely small or infinitely great; and in §44 he states that the "signs" (*Zeichen*) P^1 and P^2 that appear in the measuring equations are also measurable numbers (cf. *BBGA* 2A8, pp. 115–122).

But then in §53, after arguing that the sum, the difference (although he neglects this case) or the product of two—or of a finite multitude of—measurable numbers must necessarily be another measurable number (§§45–46 & 49–50), Bolzano states that, "according to [the] concepts [introduced] so far," there are number expressions that cannot be compared with one another, that is, between which there obtains no relation of equality or of height.[90] As he goes on to say, such would be the case, for example, of

$$A \text{ and } A + \frac{1}{1 + 1 + \ldots \textbf{ in inf.}}$$

since, according to a definition set out in the fifth section of *RZ*, the difference $A - B$, with $A > B$ and with both of these being rational numbers, "can be reduced to a

[90](*BBGA* 2A8, p. 128): "nach unsern bisherigen Begriffen."

fraction of the form $+\frac{m}{n}$, where m and n denote actual numbers,"[91] which is not the case here since

$$A + \frac{1}{1+1+\ldots \textbf{ in inf.}} - A = \frac{1}{1+1+\ldots \textbf{ in inf.}},$$

with n (i.e., $1+1+\ldots$ **in inf.**) not being an actual number. As a consequence, in §54 he poses the "problem" (*Aufg*[abe]) of investigating whether the previously discussed concepts could be extended in order to allow us to establish such a relation, and he presents the following solution: given any two finite or infinite number expressions A and B, if both are measurable and if "for any arbitrary q, the same positive or negative p is always found for both expressions A and B, which satisfies the equations $A = \frac{p}{q} + P^1 = \frac{p+1}{q} - P^2$ and $B = \frac{p}{q} + P^3 = \frac{p+1}{q} - P^4$,"[92] with P^1 and P^3 zero or positive expressions and P^2 and P^4 positive expressions, then $A = B$; but if the difference between those two number expressions is positive and not infinitely small, then the one being subtracted will be smaller than the other (cf. *BBGA* 2A8, pp. 133–134).

The problem with such a solution, as Bolzano notes in an additional sheet which was not included in the 1962 edition of *RZ* (cf. Rychlík 1962) and which indicates that §54—and by extension the definition in §55, as well as perhaps some other paragraphs—should be rewritten (*BBGA* 2A8, p. 129 fn. n), arises when one considers that "numbers that differ only by an infinitely small [number?] can behave differently in the process of measuring."[93] This would be the case, for example, of the sum of

$$.33333\ldots \text{ and } .66666\ldots,$$

which, according to Bolzano's initial formulation of the relations of equality and height, would be

$$= .99999\ldots \text{ and so } < 1,$$

which in turn would result in

$$\frac{1}{3} + \frac{2}{3} < 1.$$

[91] (*BBGA* 2A8, p. 128; cf. Russ 2004, p. 390): "auf einen Bruch von der Form $+\frac{m}{n}$ bringen läßt, wo m und n wirkliche Zahlen bedeuten."

[92] (*BBGA* 2A8, p. 131): "zu jedem beliebigen q immer dasselbe positive oder negative p für beyde Ausdrücke A und B vorfindet, welches die Gleichungen $A = \frac{p}{q} + P^1 = \frac{p+1}{q} - P^2$ und $B = \frac{p}{q} + P^3 = \frac{p+1}{q} - P^4$ erfüllet."

[93] (*BBGA* 2A8, p. 130): "Zahlen, die sich nur um ein unendlich Kleines unterscheiden, können sich bei dem Geschäfte des Messens verschiedentlich verhalten."

Or, following Bolzano, this would also be the case for

$$1 \text{ and } 1 - \frac{1}{1+1+\ldots \textbf{ in inf.}},$$

which, "for every measuring fraction q, give the numerators $p = q$ and $p = q - 1$ [respectively]"[94] and yet should be regarded as equal. Because of this, he adds, if the difference between A and B, "considered absolutely," "behaves like zero," "one finds that the numerator of the measuring fraction [is] $=0$, and so [...] $A = B$."[95]

First of all, it should be stressed that Bolzano's corrections and amendments, such as the one just presented, are evidence of a work the final version of which he did not manage to finish. Secondly, the amendment introduced in §54 not only allows Bolzano to revisit the case posited in §20, namely

$$A \neq B, \text{with } A = 1 - \frac{1}{1+1+1+\ldots \textbf{ in inf.}} \text{ and } B = 1,$$

and state that

$$A + J = A,$$

with J an infinitely small number or $=0$ (*BBGA* 2A8, p. 136). Such an amendment, as he explains further in the note of §58, involves reconsidering the notion of infinitely small numbers insofar as, "in the process of measuring," they "behave" (he uses the noun *Verhalten*) as "equal to zero," even though they are not, and thus, he says, in order to distinguish them from zero in the strict sense of the term, they should be called "*relative* or *respective* zero."[96]

As pointed out by Detlef Laugwitz in his second paper on *RZ*, this last formulation of the equality relation between measurables is crucial in order to overcome some significant difficulties in Bolzano's theory (cf. Laugwitz 1982, p. 669), although, as noted by Berg and—following him—Rusnock, it may at the same time pose some difficulties, in particular due to Bolzano's reference to the difference $A - B$ "considered absolutely" (cf. *BBGA* 2A8, pp. 134–135 fn. 64; Rusnock 2000, p. 185). Moreover, as pointed out by Rusnock, "[t]he condition on P^1 reveals an important weakness in Bolzano's approach" (Rusnock 2000, p. 182), not only from the modern perspective of how our real numbers must behave, but also from the perspective of what Bolzano's own theory of numbers dictates. As noted by Russ and Trlifajová, whereas according to §§45–46 the sum of a finite multitude of measurable numbers will also be a measurable number (so, from a modern point of view, there is closure under addition), problems emerge when one considers an

[94](*BBGA* 2A8, p. 130): "bei jedem messenden Bruche q, den Zähler $p = q$, und $p = q - 1$ geben."
[95](*BBGA* 2A8, p. 130): "absolut betrachtet"; "wie Null sich verhält"; "man zu Jedem auch noch so großen Nenner q den Zähler des messenden Bruches $= 0$ findet, und somit [...] daß $A = B$ sey."
[96](*BBGA* 2A8, p. 136): "bey dem Geschäfte des Messens"; "Null gleichgelten"; "*relative* oder *beziehungsweise* Null."

expression that Bolzano quotes in §2 as an example of an infinite quantity expression, namely "$\frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\frac{1}{16}+\ldots$ **in inf.**" (*BBGA* 2A8, p. 100; Russ and Trlifajová 2016, p. 49). Following Russ and Trlifajová, we might interpret such an expression "as the sequence of partial sums [so that] we obtain the sequence $\frac{1}{2}, \frac{1}{4}, \frac{3}{8}, \frac{5}{16}, \ldots$ that converges to $\frac{1}{3}$" (Russ and Trlifajová 2016, p. 49), but for which the two corresponding measuring equations cannot be given. Nonetheless, if this type of expression is unmeasurable, the sum of "$A = 2-2+\frac{1}{2}-\frac{1}{2}+\frac{1}{8}-\frac{1}{8}+\ldots$ in inf." and "$B = -1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots$ in inf.," "$A+B = 1-\frac{3}{2}+\frac{3}{4}-\frac{3}{8}+\frac{3}{16}-\ldots$ in inf.," would likewise be unmeasurable and, as a consequence, measurable numbers would not be closed under addition (Russ and Trlifajová 2016, p. 49).[97]

Whereas from a modern perspective it may be desirable to *repair* Bolzano's proposal in such a way that his "measurable numbers" turn out to be equivalent to our real numbers, we consider that it would be a more faithful reading of *RZ* to understand it as a transitional work that was not leading inexorably to the modern post-Weierstraßian or nonstandard frameworks.[98] It is undeniable that Bolzano himself was struggling to overcome some problems in his theory of numbers, as is shown by a handwritten note that he added at the very end of *RZ*:

> *For the theory of measurable numbers*
> Should it not be possible to simplify the theory of measurable numbers if one establishes the definition of them so that A is called measurable if one has 2 equations of the form $A = \frac{p}{q} + P = \frac{p+n}{q} - P$, where for the same n, q can be increased to infinity [or indefinitely]?[99]

In his first work on *RZ*, which was based on the 1962 edition, Laugwitz suggests an amendment in line with those which Bolzano introduced but which were only included in the second edition of that work (as in the case of the note just quoted), whereby "[a]n expression A is called measurable if, for every natural (or 'actual') number q, there is an integer $p = p(q)$, so that $\frac{p}{q} < A < \frac{p+2}{q}$."[100] Berg interprets the last of Bolzano's amendments as meaning that a measurable number expression corresponds "to a sequence S $(=\langle S_n\rangle)$ of rational numbers in such a way that, for every natural number q, there is an integer $p(q, S)$ so that, for all n, holds $S_n =$

[97] Previously, Rusnock posed the example of "$1-\frac{3}{2}+\frac{3}{4}-\frac{3}{8}+\frac{3}{16}-\ldots+\frac{3\cdot(-1)^n}{2^n}$ which gives rise to the sequence of values 1, $-\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \frac{1}{16}, \ldots,$" a sequence that converges to zero and that would not be measurable even though "it can be obtained as the sum or difference of two expressions each of which is measurable" (Rusnock 2000, p. 182).

[98] For similar stance with regard to both the modern post-Weierstraßian and nonstandard interpretations of Augustin-Louis Cauchy's analytical proposal, cf. (Lützen 2003, p. 164).

[99] (*BBGA* 2A8, p. 168): "*Zur Lehre von den meßbaren Zahlen*
Sollte die Lehre von den meßbaren Zahlen nicht vielleicht vereinfacht werden können, wenn man die Erklärung derselben so errichtet, daß A meßbar heißt, wenn man 2 Gleichungen von der Form, $A = \frac{p}{q} + P = \frac{p+n}{q} - P$ hat, wo bey einerlei n, q ins Unendliche zunehmen kann?"

[100] (Laugwitz 1965, p. 407): "Ein Ausdruck A heißt meßbar, wenn es zu jeder natürlichen (oder „wirklichen") Zahl q eine ganze Zahl $p = p(q)$ gibt, so daß $\frac{p}{q} < A \frac{p+2}{q}$."

$\frac{p(q,S)}{q} + P_n(q,S) = \frac{p(q,S)+m}{q} - P_n(q,S)$, where m is a natural number and $\langle P_n(q, S)\rangle$ is a purely positive sequence."[101] Similarly, and following (Rychlík 1962) and (van Rootselaar 1964), who also suggested such an interpretation of Bolzano's measurable numbers, Rusnock proposes that A be considered as measurable "iff for all q there is exactly one integer p and number concepts $P_1 \gtrsim 0$ and $P_2 > 0$ such that $A = \frac{p}{q} + P_1 = \frac{p+1}{q} - P_2$" (Rusnock 2000, p. 186). And, finally, Russ and Trlifajová suggest that Bolzano's definition of measurable numbers can be repaired either by considering the numerator p in both the measuring fraction and the next greater fraction of the measuring equations as a rational number, or by interpreting that Bolzano "generally considered number expressions with oscillating values and approaching a rational number as being equal to that rational number" (Russ and Trlifajová 2016, p. 50).[102]

Undoubtedly, each of those proposals has its advantages; but they all involve certain conjectures regarding issues that are not entirely clear or explicit in Bolzano's theory. In particular, whereas the consideration of measurable numbers as sequences of rational numbers would be "in harmony with Bolzano's remarks on infinite series from the paper on the binomial theorem" (Rusnock 2000, p. 181), where—from a modern point of view—Bolzano treated infinite series in a finite manner by focusing on their partial sums or finite segments of them (cf. Rusnock 2000, p. 63; Russ 2004, p. 144), it should be stressed that the latter does not explicitly endorse such an interpretation in *RZ*. So, while this chapter is not intended to demean the challenging and praiseworthy work of reinterpreting Bolzano's theory, it rather concurs with the suggestion of both Spalt and Šebestík that this theory should be understood, first and foremost, "on its own terms,"[103] that is, "in its specificity, including its stylistic peculiarities."[104]

Thus, on the one hand, in the case of convergent (in modern terminology) infinite number expressions with alternating values, Bolzano—as Rusnock has already suggested—seems to have "never fully mastered the use of inequalities in

[101](*BBGA* 2A8, p. 168 fn. 77): "entspricht nun eine Folge S ($=\langle S_n\rangle$) von rationalen Zahlen derart, daß es für jede natürliche Zahl q eine ganze Zahl $p(q, S)$ gibt, so daß für alle n gilt: $S_n = \frac{p(q,S)}{q} + P_n(q,S) = \frac{p(q,S)+m}{q} - P_n(q,S)$, wobei m eine natürliche Zahl und $\langle P_n(q, S)\rangle$ eine rein positive Folge ist."

[102]Russ and Trlifajová base the second interpretation on an example provided by Bolzano in §45, "which could be interpreted as non-monotonic sequences," namely "$A + B = \frac{p^1+p^2}{q} + \frac{1}{q} - \Omega^1 + P^{13} = \frac{p^1+p^2+2}{q} - \frac{1}{q} + \Omega^2 - P^{14}$," from which "$A + B = \frac{p^1+p^2+1}{q}$" (there would be a mistake here, either in the manuscript or in the 1976 edition, since the numerator is written "$p^1 + p^1 + 1$"; *BBGA* 2A8, p. 123; Russ and Trlifajová 2016, pp. 50–51; Russ 2004, p. 384).

[103](Spalt 1991, p. 68): "Sie [Bolzano's Zahlenlehre] muß nur in ihrem eigenen Recht verstanden werden."

[104](Šebestík 1992, p. 358): "l'historien doit de prime abord restituer une théorie du passé dans sa spécificité, y compris dans ses particularités stylistiques; c'est ensuite seulement qu'il peut tenter un rapprochement avec la pratique mathématique de notre temps."

conjunction with absolute values, and never seems to have understood the behaviour of alternating series" (Rusnock 2000, p. 183). On the other hand, as regards the infinitely great and other unmeasurable *numbers* (i.e., those that cannot be placed in a relation of equality or height), such as the unmeasurable *number* represented by the divergent (in modern terminology) number expression $1 - 1 + 1 - 1 + \ldots$ **in inf.** (§§29–30), Bolzano says very little more about these throughout the whole of the rest of *RZ*. First, in §62 he notes that "we are not entitled, *even according to the broader concepts now established*," to maintain that all infinite number expressions can be placed in a relation of equality or height, as shown by "the difference $B - A = 1 - 2 + 3 - 4 + 5 - \ldots$ **in inf.**," with A being "an arbitrary expression."[105] Then, in §69 he points out that it can be said that every positive or negative infinitely great number will be greater or smaller than any "actual number" (*wirkliche Zahl*), respectively (*BBGA* 2A8, p. 138), so that

(a) it can be considered that every positive or negative measurable number lies between 0 and a positive or negative infinitely great number, respectively (theorem of §78; *BBGA* 2A8, p. 140);
(b) given two numbers L and R, and a "property B [that] is to belong to all measurable numbers from L to R," if both are infinitely great and one is positive and the other negative, then B will belong to all measurable numbers; but if one of them is $=0$ and the other a positive or negative infinitely great number, B will belong to all positive or negative measurable numbers, respectively (theorem of §89).[106]

Throughout the remaining paragraphs of the seventh section of *RZ*, proceeding in a way similar to the way in which he proceeds in the previous sections with regard to natural numbers, integers and rationals (cf. Sect. 3; Šebestík 1992, pp. 355–356), Bolzano discusses the extension of several properties to measurable numbers (cf. Russ and Trlifajová 2016, p. 45). For example, he states that

$$\text{"[i]f } A > B \text{ and } B > C, \text{ then also } A > C\text{"} \quad (\S 63)$$

and that[107]

[105](*BBGA* 2A8, p. 137; cf. Russ 2004, p. 396): "wir *selbst nach den jetzt aufgestellten weiteren Begriffen* nicht berechtiget sind"; "der Unterschied $B - A = 1 - 2 + 3 - 4 + 5 - \ldots$ **in inf.**"; "beliebiger Ausdruck."

[106](*BBGA* 2A8, p. 142; Russ 2004, p. 403): "die Beschaffenheit B allen meßbaren Zahlen von L bis zu R ein- oder ausschließlich zukommen soll." Bolzano's formulation of this theorem contains a fairly subtle mistake since, strictly speaking, he says at the beginning of the paragraph that if L and R are infinitely great numbers, with one of them positive and the other negative, and the property B belongs "to all measurable numbers from L to R, inclusively or exclusively," then it belongs to all measurable numbers. However, it is clear that L and R cannot be counted among the measurable numbers and thus L and R cannot be included among "all measurable numbers."

[107](*BBGA* 2A8, pp. 137 & 139): "Wenn $A > B$ und $B > C$, so ist auch $A > C$"; "entweder $A = B$ oder $A > B$ oder $A < B$."

"either $A = B$ or $A > B$ or $A < B$" (theorem of §73),

a proposition which, as in other cases, is false with his first formulation of the equality relation (cf. Russ 2004, p. 399 fn. q; *BBGA* 2A8, pp. 134–135 fn. 64). In addition to this, he takes up once again some important theorems which he had already addressed in *RaB*, such as the theorem on what nowadays is called a least upper bound property (§109) and that about what—depending on whether or not Bolzano's infinitesimals are considered to be equivalent to 0—can be considered equivalent to so-called Cauchy's convergence criterion (§107). Nevertheless, inasmuch as the elucidation of the extent to which Bolzano succeeds to ground these properties in a rigorous manner (from his own point of view) in *RZ* is something that requires and deserves a thorough study,[108] whereas the objective of this section has been to address the insights and uncertainties with regard to the number continuum underlying his notion of "measurable numbers," readers are referred to (Spalt 1991), (Šebestík 1992) and (Rusnock 2000) for a careful assessment of Bolzano's discussion of such theorems.

5 Final Remarks

In his introduction to the 1976 edition of *RZ*, Berg pointed out a dilemma faced by the scholar of Bolzano's theory of numbers: either one updates Bolzano's proposal to bring it into line with present-day knowledge, thereby running the risk of misinterpreting some of its basic characteristics, or, if one attempts to be faithful to Bolzano's proposal, some of such characteristics remain unclear (cf. Berg 1976, p. 8). This chapter has attempted to cast light upon Bolzano's insights and uncertainties regarding the number continuum by paying attention not only conceptual framework of *RZ*, but also to the introductory volume to his general theory of quantities (i.e., *EG*), his mathematical notebooks and some other of his works, both those published during his lifetime and those published only posthumously. Moreover, this chapter has tried to provide a more faithful assessment of *RZ* by showing how Bolzano's stance on numbers evolved from his student days through to his maturity, when he wrote this latter work.

It should be pointed out, however, that there are some volumes of Bolzano's *Größenlehre* and *Miscellanea Mathematica* (comprising, indeed, those notebooks which cover the period from 1830 onward) which have not yet been published and which might provide further evidence regarding certain aspects of *RZ*. Coupled with this, it is worth noting that, due to its limited scope, this chapter has left out of account certain interesting and not uncontroversial issues that are either addressed in *RZ* or linked to it. For example, this chapter did not delve into Bolzano's

[108] In her presentation at the symposium "Bolzano's Mathematics and the General Methodology of the Sciences," which was organized by Steve Russ and held as part of the 16th CLMPST, in Prague, Anna Bellomo examined §107 and provided appealing evidence in favor of a non-set-theoretical interpretation of *RZ* by focusing on Bolzano's reasoning "in terms of parts and wholes."

characterization of numbers as abstract (the property or nature to which the numbers are due) and concrete (the "actual multitude of things"), or "nominated" (*benannte* or *genannte*; e.g., "the concept three points") and "unnominated" (*unbenannte* or *ungenannte*; in themselves, e.g., "the concept three," so there is no unnominated number idea that has only one object, because "one" is not a number), these latter being those which are "primarily" (*vornehmlich*) studied in the "pure theory of numbers" (cf. *BBGA* 2A8, pp. 19–26). For a careful study of these topics, readers are referred to (Benis-Sinaceur 1975) and (Šebestík 1992, pp. 345–350).

Likewise, this chapter has not engaged in any deeper discussion of Bolzano's distinction between multitudes and collections (not sets), for which readers are referred to (Simons 1997), nor of Bolzano's vindication of the so-called part-whole principle (cf. Trlifajová 2018). Regarding this last issue: (a) in *PU*, Bolzano defended the idea that an infinite multitude is greater than those infinite multitudes contained as parts of it (cf. Bolzano 1851/1920, pp. 26–27 & 30–31); (b) in *WL*, he stated that in the "infinite series of concepts [of whole numbers] [. . .] n, n^2, n^4, n^8, . . .," "every succeeding [concept] is always subordinated to [i.e., smaller than] the preceding one";[109] and (c) he is said to have retracted his previous stance on the part-whole principle in a letter dated a few months before his death (cf. Mancosu 2009, p. 625). Paolo Mancosu has challenged the interpretation according to which "Bolzano saved his mathematical soul in extremis and joined the rank of the blessed Cantorians by repudiating his previous sins," and he has pointed out that "recent mathematical work gives us theories that can (cum grano salis) be seen as formally capturing (parts of) the intuitive concept of infinity found in [. . .] all those who believed that the size of the natural numbers is larger than the size of the even numbers" (Mancosu 2009, pp. 626–627). Instead, and based not only on the passages alluded to but also on a paper by Bolzano dated 1840–41, Rusnock and Šebestík have recently argued that "[m]aintaining the whole–part principle would thus appear to force Bolzano" to pay a price that he presumably would not have been willing to pay (Rusnock and Šebestík 2019, pp. 537–538).

There is a well-known passage—"well-known" at least among Bolzano's scholars—toward the end of the preface to *RaB* where, with reference to his theorem on what today is designated as a "least upper bound," Bolzano claims that, according to a "correct concept of *quantity*," "the thought [or idea] of an i which is the greatest of those [quantities] of which it may be said that all those standing below it possess the property M, is the thought [or idea] of a real, i.e. of an *actual* quantity."[110] From the point of view expressed in this chapter, the fact that in this passage Bolzano refers to i as a "quantity" and not as a "number," whereas in §109 of *RZ* he refers to it as a "measurable number," not only constitutes evidence of both a terminological and a

[109](Bolzano 1837, p. 474): "unendlichen Reihe von Begriffen"; "ist also jeder folgende immer den vorhergehenden untergeordnet."

[110](Bolzano 1817a, p. 23): "für Jeden, der einen richtigen Begriff von *Größe* hat, [. . .] daß der Gedanke eines i, welches das größte derjenigen ist, von denen gesagt werden mag, daß alle unter ihm stehende die Eigenschaft M besitzen, der Gedanke einer reellen d.h. *wirklichen* Größe sei."

conceptual change, but ultimately points to the substratum on which his later notion of "measurable numbers" was to rest, namely that of quantities. Therefore it may be concluded that, on the one hand, from his early works onward Bolzano showed "a profound insight into the nature of [quantities]" (Russ 2004, p. 150)—indeed, one would even want to be entitled to say "into the nature of *numbers*"—which eventually led him to the development of his notion of "measurable numbers." On the other hand, however, even in the case of the latter notion, which reflects a transition to a non-traditional notion of numbers and represents "a surprising achievement for its time" (Russ 2004, p. 348), one still cannot help but note significant traits of a not-yet-modern understanding of the number continuum, of quantities and of mathematics itself.

In this sense, Šebestík's remark that the relevance of Bolzano's theory of numbers can be better appreciated if it is compared both with later and with earlier proposals in this subject seems to be extremely appropriate (cf. Šebestík 1992, p. 385). Following Šebestík, it is worth considering, for example, "Kästner's decomposition of irrational numbers,"[111] first published in 1758. For the latter, irrational—algebraic—numbers, "which originate from roots, can be approximated to just as closely as one desires," inasmuch as they are composed of both a rational part and of another part that can be considered "an innumerable multitude of infinitely small parts."[112] Such an approach, which was not novel,[113] certainly had its shortcomings, starting with the fact that Kästner only referred to the algebraic irrationals and took for granted the operations on them (Šebestík 1992, pp. 353–355). Nonetheless, as pointed out by Šebestík, Kästner's presentation can be regarded as a step forward, inasmuch as he would seem to have intended to develop "[a]ll the concepts of arithmetic" on the basis of the notion of whole numbers.[114]

The ending of "the paradigm of the science of quantity" (to echo Epple's expression mentioned in the introduction), which was still strongly advocated in the early nineteenth century and which, as has been shown to some extent in this chapter, entailed profound changes, was actually a protracted process which went on well beyond Bolzano's own lifetime, extending for the whole of that century. As a consequence, remnants of that previous understanding of mathematics as *Größenlehre* were to continue to prevail, in one way or another, in several key proposals that are usually linked to the development of the modern notion of real

[111](Šebestík 1992, p. 353): "Lorsque nous parlerons de la représentation bolzanienne des «grandeurs mesurables», on devra se rappeler la décomposition des nombres irrationnels selon Kästner."

[112](Kästner 1786, pp. 126–127; cf. Šebestík 1992, pp. 352–353): "Irrationalzahlen, die von Wurzeln herrühren, kann man sich, so weit man will, nähern"; "eine unzählige Menge unendlich kleiner Theile."

[113]In the mid-1810s, Bolzano explicitly rejected such an approach, which was still common at the turn of the nineteenth century (cf. van Rootselaar 1995, p. 28; Šebestík 1992, p. 355), because otherwise "one would arrive at the *consideration of the infinite*" ("käme man da in die *Betrachtung des Unendlichen*"; *BBGA* 2B6/2, p. 195).

[114](Kästner 1786, p. [4]): "Alle Begriffe der Arithmetik gründen sich meines Erachtens auf die von ganzen Zahlen." Cf. (Šebestík 1992, pp. 52 & 54).

numbers. A couple of examples shall be mentioned. Firstly and as noted by Ferreirós, whereas Weierstraß continued to differentiate between, on the one hand, numbers in the strict sense of the term and, on the other, quantities—leading to "[h]is theory of rational and irrational numbers [which dates from the 1860s] [being] formulated as a theory of 'numerical [quantities]'"—, there nevertheless lies "under [such a] traditional denomination [. . .] an abstract conception of [quantities]" (Ferreirós 2007, pp. 64–65 & 125). Secondly, as pointed out by Epple, whereas in his *Theorie der complexen Zahlensysteme* Hermann Hankel intended to develop a formal concept of real numbers following "the British tradition of Peacock and de Morgan," he nevertheless claimed that "[t]he irrational require[d] for its systematic treatment the concept of quantity"[115] and that irrationals, therefore, should be developed "within the theory of quantities."[116]

Nevertheless, as the works of Dedekind, Cantor, and Heine show, by the 1870s the notion of number had changed significantly. Hence Dedekind's insistence in 1872 that the "gradual extension of the concept of number" should be carried out without using "foreign ideas," such as that of "measurable quantities."[117] As he was later to write, in the preface to the first edition of *Was sind und was sollen die Zahlen?*, even though he agreed that the "ancient conviction" that an "irrational number [could be] understood as a ratio of measurable quantities" was ultimately "the source" of any attempt at a theory of irrational numbers, for Dedekind there was an essential difference between his own theory based on cuts and certain earlier proposals.[118] In particular, Dedekind disagreed with Jules Tannery's claim that the latter's theory (which the former recognized as equivalent to his) represented "the development of [the] idea originating from J. Bertrand [. . .] of defining an irrational number by indicating all rational numbers which are smaller and all those which are greater than the number to be defined."[119] As he went on to explain, Joseph Bertrand's proposal, according to which any "incommensurable number" "is necessarily placed between two [commensurable numbers], $\frac{x}{n}$ and $\frac{x+1}{n}$, and [the integer]

[115](Epple 1996, pp. 6–7; cf. Hankel 1867, p. 47): "die britische Tradition Peacocks und de Morgans"; "Das Irrationale verlangt zu seiner systematischen Fassung den Grössenbegriff."

[116]The third and fourth sections of Hankel's work are entitled "Die reellen Zahlen in ihrem formalen Begriffe" and "Die reellen Zahlen in der Grössenlehre," respectively (Hankel 1867, pp. 35 & 48), while the full title of that work is *Theorie der complexen Zahlensysteme insbesondere der gemeinen imaginären Zahlen und der Hamilton'schen Quaternionen nebst ihrer geometrischen Darstellung*.

[117](Dedekind 1888, p. X): "die shrittweise Erweiterung der Zahlbegriffes"; "fremdartiger Vorstellungen"; "der meßbaren Größen."

[118](Dedekind 1888, p. XIV): "diese uralte Ueberzeugung"; "die irrationale Zahl als Verhältniß meßbarer Größen auffaßt"; "die Quelle."

[119](Dedekind 1888, pp. XIII–XIV): "die Entwickelung eines von Herrn J. Bertrand herrührenden Gedankens nennt, welcher in dessen *Traité d'arithmétique* enthalten sei und darin bestehe, eine irrationale Zahl zu definiren durch Angabe aller rationalen Zahlen, die kleiner, und aller derjenigen, die größer sind als die zu definirende Zahl."

n can be taken great enough, so that their difference $\frac{1}{n}$ is as small as desired,"[120] was still a proposal which "ha[d] recourse to the existence of measurable quantities" and "present[ed] some essential gaps."[121]

Bolzano's proposal of "measurable numbers," which was unknown to all these mathematicians, certainly constitutes a step forward, made already within the first half of the nineteenth century, in the elucidation of the number continuum. However, several crucial steps remained still to be taken, beginning with the abandonment of the substratum of quantities and by extension of the core notion of variable quantities (cf. Fuentes Guillén and Martínez Adame 2020). The position defended in this chapter, therefore, is that a proper assessment of Bolzano's theory of numbers should always take into account its original setting and, as fruitful as such an "updating" approach may sometimes appear to be, should always avoid dislocating this theory by reinterpreting it within a modern, post-Weierstraßian or nonstandard framework.

Acknowledgments The author would like to thank Steve Russ, José Ferreirós and Carlos Álvarez for their very helpful remarks and suggestions, as well as Michèle Friend and Bharath Sriraman for all their support for the publication of this chapter. He would also like to thank Davide Crippa for his help in obtaining some of the documents cited in this chapter, as well as Alejandrina Viesca Ramírez for her careful revision of it and Alexander Reynolds for his excellent proofreading.

Funding This chapter was made possible thanks to the support by the Postdoctoral Scholarship Program of the Dirección General de Asuntos del Personal Académico (DGAPA) and the Faculty of Sciences at Universidad Nacional Autónoma de México, as well as thanks to the support for the grant project "Mathematics in the Czech Lands: From the Jesuit Teaching to Bernard Bolzano" (GA ČR 19-03125Y) and the support by the Centre for Science, Technology and Society Studies of the Filosofický ústav Akademie věd České republiky.

References

Barnett JH (2004) Enter, Stage Center: The Early Drama of The Hyperbolic Functions. Math Mag 77(1):15–30. https://www.tandfonline.com/doi/abs/10.1080/0025570X.2004.11953223

Benis-Sinaceur H (1975) Bolzano est-il le précurseur de Frege? Archiv für Geschichte der Philosophie 57(3):286–303

Benis-Sinaceur H (2015) Is Dedekind a Logicist? Why Does Such a Question Arise? In: Benis-Sinaceur H, Panza M, Sandu G (eds) Functions and generality of logic. Reflections on Dedekind's and Frege's Logicisms. Springer, Cham, pp 1–57

Berg J (1975) Einleitung des Herausgebers. In: Berg J (ed) Bernard Bolzano Gesamtausgabe, Band 2A7. Frommann-Holzboog, Stuttgart-Bad Cannstatt, pp 9–16

Berg J (1976) Einleitung des Herausgebers. In: Berg J (ed) Bernard Bolzano Gesamtausgabe, Band 2A8. Frommann-Holzboog, Stuttgart-Bad Cannstatt, pp 7–9

[120](Bertrand 1849, pp. 204–205): "nombre incommensurable"; "est nécessairement compris entre deux d'entre eux, $\frac{x}{n}$ et $\frac{x+1}{n}$, et l'on peut prendre *n* assez grand, pour que leur différence $\frac{1}{n}$ soit aussi petite qu'on le voudra."

[121](Dedekind 1888, pp. XIV–XV): "sie sogleich ihre Zuflucht zu der Existenz einer meßbaren Größe nimmt"; "einige so wesentliche Lücken darzubieten."

Berg J (1977) Einleitung des Herausgebers. In: Berg J (ed) Bernard Bolzano Gesamtausgabe, Band 2A5. Frommann-Holzboog, Stuttgart-Bad Cannstatt, pp 7–11

Berg J (1990) Zur logischen und mathematischen Ontologie. Geneseologie und Resultatismus in der Analysis der Grundlagen der Bolzanoschen Zahlenlehre. In: Spalt DD (ed) Rechnen mit dem Unendlichen. Beiträge zur Entwicklung eines kontroversen Gegenstandes. Springer, Basel, pp 123–155

Bertrand J (1849) Traité d'arithmétique. L. Hachette et c^{ie}, Paris

Blok J (2016) Bolzano's Early Quest for A Priori Synthetic Principles. Mereological Aspects of the Analytic-Synthetic Distinction in Kant and the Early Bolzano. PhD thesis. Rijksuniversiteit Groningen

Bolzano B (1804; BG) Betrachtungen über einige Gegenstände der Elementargeometrie. In Commission bey Karl Barth, Prag

Bolzano B (1810; BD) Beyträge zu einer begründeteren Darstellung der Mathematik. Erste Lieferung. Caspar Widtmann, Prag

Bolzano B (1816; BL) Der binomische Lehrsatz, und als Folgerung aus ihm der polynomische, und die Reihen, die zur Berechnung der Logarithmen und Exponentialgrößen dienen, genauer als bisher erwiesen. C.W. Enders, Prag

Bolzano B (1817a; DP) Die drey Probleme der Rectification, der Complanation und der Cubirung, ohne Betrachtung des unendlich Kleinen, ohne die Annahmen des Archimedes, und ohne irgend eine nicht streng erweisliche Voraussetzung gelöst; zugleich als Probe einer gänzlichen Umstaltung der Raumwissenschaft, allen Mathematikern zur Prüfung vorgelegt. Paul Gotthelf Kummer, Leipzig

Bolzano B (1817b; RaB) Rein analytischer Beweis des Lehrsatzes, daß zwischen je zwey Werthen, die ein entgegensetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege. Gottliebe Haase, Prag

Bolzano B (1837; WL) Wissenschaftslehre. Versuch einer ausführlichen und größtentheils neuen Darstellung der Logik mit steter Rücksicht auf deren bisherige Bearbeiter. Erster Band. Seidelschen Buchhandlung, Sulzbach

Bolzano B (1851/1920; PU) Paradoxien des Unendlichen herausgegeben aus dem schriftlichen Nachlasse des Verfassers von Dr. Fr. Příhonský, A. Höfler (ed). Felix Meiner, Leipzig

Bolzano B (1930; F) Functionenlehre. In: Rychlík K (ed) Spisy Bernarda Bolzana/Bernard Bolzano's Schriften, Svazek 1. Královská česká společnost nauk, Praha

Bolzano B (1975; EG BBGA 2A7) Einleitung in die Größenlehre und erste Begriffe der allgemeinen Größenlehre. In: Berg J (ed) Bernard Bolzano Gesamtausgabe, Band 2A7. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (1976; RZ BBGA 2A8) Größenlehre II: Reine Zahlenlehre. In Berg J (ed) Bernard Bolzano Gesamtausgabe, Band 2A8. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (1977; BBGA 2A5) Mathematische und philosophische Schriften 1810–1816. In: Berg J (ed) Bernard Bolzano Gesamtausgabe, Band 2A5. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (1995; BBGA 2B6/2) Miscellanea Mathematica 10. In: van Rootselaar B, van der Lugt A (eds) Bernard Bolzano Gesamtausgabe, Band 2B6/2. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (1996; BBGA 2B7/1) Miscellanea Mathematica 11. In: van Rootselaar B, van der Lugt A (eds) Bernard Bolzano Gesamtausgabe, Band 2B7/1. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (2001; BBGA 2B9/2) Miscellanea Mathematica 16. In: van Rootselaar B, van der Lugt A (eds) Bernard Bolzano Gesamtausgabe, Band 2B9/2. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (2005; BBGA 2B11/1) Miscellanea Mathematica 19. In: van Rootselaar B, van der Lugt A (eds) Bernard Bolzano Gesamtausgabe, Band 2B11/1. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (2006; BBGA 2B11/2) Miscellanea Mathematica 20. In: van Rootselaar B, van der Lugt A (eds) Bernard Bolzano Gesamtausgabe, Band 2B11/2. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Bolzano B (2007; BBGA 2B12/1) Miscellanea Mathematica 21. In: van Rootselaar B, Berg J (eds) Bernard Bolzano Gesamtausgabe, Band 2B12/1. Frommann-Holzboog, Stuttgart-Bad Cannstatt

Cantor G (1872) Ueber die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen. Math Ann 5:123–132

Cantor G (1874) Ueber eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. J Reine Angew Math 77:258–262

Cantù P (2014) Bolzano versus Kant: Mathematics as a Scientia Universalis. In: Reboul A (ed) Mind, Values, and Metaphysics. Philosophical essays in Honor of Kevin Mulligan, vol 1. Springer, Cham, pp 295–316

Catalogus Mathematicorum IX A 19 (1781ff) Catalogus Mathematicorum Anno Era Christiano 1781 elucubratus. Katalogy Národní knihovna České republiky

Cauchy A-L (1821) Cours d'Analyse de l'École Royale Polytechnique. 1er Partie, Analyse Algébrique. Debure frères, Paris

Dedekind R (1872) Stetigkeit und irrationale Zahlen. Friedrich Vieweg und Sohn, Braunschweig

Dedekind R (1876) Sur la théorie des nombres entiers algébriques. Bull Sci Math Astron 11:278–288

Dedekind R (1888) Was sind und was sollen die Zahlen? Friedrich Vieweg und Sohn, Braunschweig

Dorrego López E, Fuentes Guillén E (forthcoming) Irrationality and transcendence in the 18th and 19th centuries. A contextualised study of J. H. Lambert's work. Prologue by José Ferreirós. Springer

Du Bourguet JBE (1810) Traités élémentaires de calcul différentiel et de calcul intégral, indépendans de toutes notions de quantités infinitésimales et de limites, Tome premier. Courcier, Paris

Epple M (1996) Das Ende der Grössenlehre. Eine Einführung in die Geschichte der Grundlagen der Analysis, 1860-1930. Preprint-Reihe des Fachbereichs Mathematik 8. Johannes Gutenberg-Universität Mainz, Mainz

Ferreirós J (2007) Labyrinth of Thought. A History of Set Theory and Its Role in Modern Mathematics, 2nd edn. Birkhäuser, Germany

Fuentes Guillén E, Crippa D (2021) The 1804 examination for the chair of Elementary Mathematics at the University of Prague. Hist Math 57:24-54.e18.

Fuentes Guillén E, Martínez Adame C (2020) The notion of variable quantities ω in Bolzano's early work. Hist Math 50:25–49

Grimm J, Grimm W (1862) Deutsches Wörterbuch, Dritter Band, E – Forsche. S. Hirzel, Leipzig

Grunert JA (ed) (1836) Supplemente zu Georg Simon Klügel's Wörterbuche der reinen Mathematik, Zweite Abtheilung, E bis Z. E. B. Schwickert, Leipzig

Handbuch (1786) Handbuch aller unter der Regierung des Kaisers Joseph II für die K.K. Erbländer ergangener [sic] Verordnungen und Gesetze in einer Systematischen Verbindung enthält die Verordnungen und Gesetze von Jahre 1784, Sechster Band. Georg Moesle, Wien

Hankel H (1867) Theorie der complexen Zahlensysteme insbesondere der gemeinen imaginären Zahlen und der Hamilton'schen Quaternionen nebst ihrer geometrischen Darstellung. Leopold Voss, Leipzig

Kästner AG (1786) Anfangsgründe der Arithmetik, Geometrie, ebenen und sphärischen Trigonometrie und Perspectiv, 1sten Theils erste Abtheil, Vierte vermehrte Auflage. Wittwe Vandenhoeck, Göttingen

Klügel GS (1808) Mathematisches Wörterbuch oder Erklärung der Begriff, Lehrsätze, Aufgaben und Methoden der Mathematik, Erste Abtheilung, Dritter Theil von K bis P. Schwickertschen Verlage, Leipzig

Lagrange J-L (1797) Théorie des fonctions analytiques, contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits ou d'évanouissans, de limites

ou de fluxions, et réduits a l'analyse algébrique des quantités finies. Imprimerie de la République, Paris
Lambert JH (1761/1768) Mémoire sur quelques propriétés remarquables des quantités transcendantes, circulaires et logarithmiques. Mémoires de l'Académie Royale des Sciences et Belles-Lettres, Tom. XVII. Haude et Spener, Berlin, pp 265–322
Laugwitz D (1965) Bemerkungen zu Bolzanos Größenlehre. Arch Hist Exact Sci 2(5):398–409
Laugwitz D (1982) Bolzano's infinitesimal numbers. Czechoslov Math J 32(4):667–670
Lipschitz R (1876/1986) Briefwechsel mit Cantor, Dedekind, Helmholtz, Kronecker, Weierstrass und anderen, Winfried Scharlau (ed). Springer, Braunschweig/Wiesbaden
Lützen J (2003) The foundation of analysis in the 19th century. In: Jahnke HN (ed) A History of Analysis. American Mathematical Society/London Mathematical Society, pp 155–195
Mancosu P (2009) Measuring the size of infinite collections of natural numbers: was Cantor's theory of infinite number inevitable? Rev Symb Logic 2(4):612–646
Ohm M (1828) Versuch eines vollkommen consequenten Systems der Mathematik, Erster Theil. T. H. Riemann, Berlin
Rusnock P (2000) Bolzano's Philosophy and the Emergence of Modern Mathematics. Rodopi, Amsterdam
Rusnock P, Šebestík J (2019) Bernard Bolzano. His Life and Work. Oxford University Press, Oxford
Russ S (2004) The Mathematical Works of Bernard Bolzano. Oxford University Press, Oxford
Russ S, Trlifajová K (2016) Bolzano's measurable numbers: are they real? In: Zack M, Landry E (eds) Research in History and Philosophy of Mathematics, The CSHPM 2015 annual meeting in Washington D.C. Cham, Birkhäuser, pp 39–56
Rychlík K (1962) Theorie der reellen Zahlen im Bolzanos handschriftlichen Nachlasse. Verlag der Tschechoslowakischen Akademie der Wissenschaften, Prag
Šebestík J (1992) Logique et mathématique chez Bernard Bolzano. Vrin, Paris
Simons P (1997) Bolzano on collections. In: Künne W, Siebel M, Textor M (eds) Bolzano and Analytic Philosophy. Rodopi, Amsterdam, pp 87–108
Spalt DD (1991) Bolzanos Lehre von den meßbaren Zahlen 1830–1989. Arch Hist Exact Sci 42: 15–70
Trlifajová K (2018) Bolzano's Infinite Quantities. Found Sci 23:681–704
van Rootselaar B (1964) Bolzano's Theory of Real Numbers. Arch Hist Exact Sci 2(2):168–180
van Rootselaar B (1995) Einleitung. In: van Rootselaar B, van der Lugt A (eds) Bernard Bolzano Gesamtausgabe, Band 2B6/2. Frommann-Holzboog, Stuttgart-Bad Cannstatt, pp 11–32
Weierstraß K (1868/1986) Einführung in die Theorie der analytischen Functionen nach einer Vorlesungsmitschrift von Wilhem Killing aus dem Jahr 1868, Remmert R (ed). Universität Münster, Münster
Winter E (1933) Bernard Bolzano und sein Kreis. Jakob Hegner, Leipzig
Wolff C (1716) Mathematisches Lexicon. Johann Friedrich Gleditschens seel. Sohn, Leipzig
Zweig A (1999) Immanuel Kant. Correspondence. Cambridge University Press, Cambridge

Abel's Approach to Elliptic Integrals

The Addition Theorems

John K. Dagsvik

Contents

1 Introduction 1634
2 Elliptic, Hyperelliptic, and Abelian Integrals 1636
3 Early Approaches to Addition Theorems 1639
3.1 Addition Theorems for Inverse Trigonometric Functions 1640
3.2 Development of the Theory of Elliptic Integrals in the Eighteenth and early Nineteenth Century 1641
3.3 Results Obtained by Fagnano 1642
3.4 Results Obtained by Euler 1643
3.5 Lagrange's Approach 1646
3.6 Results Achieved by Legendre 1649
4 Abel's Addition Theorem: An Elementary Exposition 1651
4.1 The Addition Theorem for Elliptic and Hyperelliptic Integrals 1651
4.2 A More General Version of Abel's Addition Theorem 1658
4.3 Abel's Addition Theorem in Its Most General Form 1661
4.4 Elliptic Integrals and Multiplication 1662
4.5 Examples 1664
5 A Short Review of Abel's Work on Elliptic Integrals and the Addition Theorem 1669
6 Conclusion 1674
References 1674

Abstract

The theory of elliptic integrals and functions was a major research topic during the nineteenth century. Great mathematicians such as Euler, Lagrange, and Legendre made important contributions in this field. The Norwegian mathematician Niels Henrik Abel revolutionized this theory as he introduced novel ideas and approaches. Specifically, he proved what is called Abel's addition theorem which is a sweeping extension of previous addition theorems for elliptic integrals obtained in the eighteenth century.

J. K. Dagsvik (✉)
Research Department, Statistics Norway, Oslo, Norway
e-mail: john.dagsvik@ssb.no

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_99

This chapter provides an elementary review of addition theorems for elliptic integrals before 1830 with special focus on Abel's addition theorem. An important aim of the chapter is to bring out the intuition behind the various methodological approaches. The chapter also contains a short review of Abel's life with special reference to his contribution to the theory of elliptic integrals and beyond.

Keywords

Elliptic integrals · Hyperelliptic integrals · Abelian integrals · Elliptic functions · Addition theorems · Theory of transformations · Multiplication of elliptic integrals

1 Introduction

The life of the Norwegian mathematician Niels Henrik Abel (1802–1829) was short and dramatic. Nevertheless, he was able to make pathbreaking contributions in several fields within mathematics. A major part of Abel's research was on the theory of elliptic functions and integrals. He revolutionized this theory as he introduced new ideas and approaches that later served as the starting point for seminal works by mathematicians such as Karl Theodor Wilhelm Weierstrass (1815–1897) and Georg Bernhard Riemann (1826–1866), among many others. One of Abel's most important achievement is his addition theorem.

The mathematician Carl Gustav Jacob Jacobi (1804–1851) described Abel's theorem as *"die grosste mathematische Entdeckung unserer Zeit, obgleich erst eine kunftige grosses Arbeit ihre ganze Bedeutung aufweise könne" ("the greatest mathematical discovery of our time, even though only a great work in the future will reveal its full significance").* The most general and complete exposition of the addition theorem is given in Abel's famous Paris memoir. Due to a number of unfortunate circumstances, it was first published 12 years after his death (Abel 1841).

An integral is called elliptic if the integrand is a rational function of u and y where y^2 is a polynomial in u of degree three or four, $p(u)$, (say). Specifically, when $p(u) = (1 - u^2)(1 - k^2u^2)$ where $k^2 < 1$ is a constant, one obtains what is called an elliptic integral of the first kind. The purpose of this chapter is to give a brief elementary review of the development of the theory of elliptic integrals before 1830 with focus on Abel's addition theorem. Special emphasis is placed on bringing out the key idea and intuition of Abel's approach by means of elementary calculus. Accordingly, this exposition should be accessible to a general audience interested in the development of mathematics but who only have just an elementary knowledge about calculus.

A number of survey articles and books on Abel's work cover the addition theorem. Some of them are simple without the ambition to cover essential details of the mathematical developments and proofs whereas others require considerable

background knowledge of mathematics. The works of Houzel (1986, 2004) and Sørensen (2004) provide deeper analyses and deal with Abel's entire work within several areas. The paper written by Huntington Barnum (1910) discusses aspects of the theory of elliptic integrals and Abel's addition theorem, but it is not that much easier to read than Abel's original papers. The same might be said about the work by Houzel (2004). The paper written by Mittag-Leffler (1923) provides a nice introduction to the theory of elliptic integrals and early works on addition theorems. He also compares aspects of Abel's work with the approach Weierstrass developed to study elliptic functions and integrals. However, he does not go through the details of the proof of Abel's addition theorem. Hopefully, this chapter will fill a gap.

This chapter starts by showing how elliptic integrals emerged historically from problems in geometry and mechanics. During the seventeenth century, it became clear that this type of integrals could not be expressed on closed form by elementary functions. Subsequently, it was discovered that this type of integrals possesses specific invariance properties and the theory of such integrals became a central research area by the end of the eighteenth century. Before Abel, it was Giulio Carlo Fagnano (1682–1766), Leonhard Euler (1707–1783), Joseph Louis Lagrange (1736–1813), and Adrien-Marie Legendre (1752–1833) who were the main contributors to this theory.

However, when Abel entered the scene, he introduced a totally new and powerful approach that enabled him to generalize the theory far beyond the previous contributions. Consider, for example, a set of m elliptic integrals of the first kind with the same parameter k and a constant lower integration limit and upper limits x_j, $j = 1, 2, \ldots, m$. As will be discussed in detail in Sect. 4, Abel' s approach starts by making a change of variable where the variable of integration u (say) is transformed to z by $u = r_j(z)$ where $r_j(z)$ is chosen as the root of a suitably selected polynomial $\psi(u)$ (say), where one or several of the coefficients of $\psi(u)$ is allowed to depend on z. The upper integration limit is changed from x_j to z_1 (say) where z_1 does not depend on j. After this change of variable, Abel shows that the integrand of the j-th elliptic integral can be expressed as $r'_j(z)/\sqrt{p\left(r_j(z)\right)} = q\left(r_j(z)\right)/\psi'\left(r_j(z)\right)$, where $q(u)$ is a polynomial with degree less than the degree of $\psi(u)$ minus one. According to the theorem of partial fraction decomposition, the sum over j of the fractions $\{q(r_j(z))/\psi'(r_j(z))\}$ is equal to zero and therefore the sum of the corresponding elliptic integrals becomes equal to a constant. In other words, an addition theorem for this special case has been established.

The chapter is organized as follows: Section 2 contains a description of what is meant by elliptic integrals and we discuss examples of how elliptic integrals arise in selected applications in mechanics and geometry. Section 3 contains a review of the theory of elliptic integrals before Abel's contribution. In Sect. 4, several versions of Abel's addition theorem are presented, including examples. Section 5 contains a short review of Abel's life with special reference to his contribution to the theory of elliptic integrals and extensions.

2 Elliptic, Hyperelliptic, and Abelian Integrals

The class of elliptic integrals can be described as follows: let $p(x)$ be a polynomial of degree three or four and $R(x, y)$ be a rational function of x and y, i.e., $R(x, y)$ is a quotient with numerator and denominator that are polynomials in x and y. An integral of the form

$$\int R\left(x, \sqrt{p(x)}\right)dx$$

it is called an *elliptic integral*. The historical reason for this is that when calculating the arch length of an ellipse, one gets an integral of this type. After Gottfried Wilhelm Leibniz (1646–1716) and Isaac Newton (1642–1727) had developed the field of infinitesimal calculus in the 1660s, and Johannes Kepler (1571–1630) had established the laws of celestial mechanics, intensive research into pure as well as applied mathematics, such as mechanics and spherical geometry (navigation), followed. Because Kepler established that planets move around the sun according to elliptic orbits, the calculation of the length of these orbits corresponds to the calculation of the arch length of an ellipse. According to Hoffman (1949), Leibniz contacted Newton and the mathematician James Gregory (1638–1675) in 1675, asking whether it would be possible for him to compute the arch length of an ellipse. Leibniz was told that this problem could only be solved approximately, not in an exact way. At the time, Leibniz thought that he himself would be able to solve this problem, but he eventually realized that he could not. Much later, it was Joseph Liouville (1809–1882) who proved that elliptic integrals cannot be expressed by elementary functions.

When the degree of the polynomial $p(x)$ is greater than 4, the above integral is called hyperelliptic. Abel considered hyperelliptic integrals as well as much more general cases called Abelian integrals. An Abelian integral has the form

$$\int R(x, y(x))dx$$

where $y(x)$ is an algebraic function. Remember that an algebraic function $y(x)$ is the root of a polynomial in y, $P(x, y)$, (say) where the coefficients are polynomials in x with integer coefficients. In other words, the algebraic function $y(x)$ is determined by $P(x, y(x)) = 0$.

Example 2.1. Arch Length of an Ellipse
Consider the problem of computing the arch length of an ellipse. Recall that the arch length, $s(u,v)$, from $x = u$ to $x = v$ of a differentiable function $y = f(x)$ can be computed by means of the formula

$$s(u, v) = \int_u^v \sqrt{1 + f'(x)^2}dx. \tag{1}$$

Recall also that the ellipse with parameters a and b can be expressed by

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1. \tag{2}$$

From (2), it follows by applying (1) that

$$s(u,v) = \frac{1}{a}\int_u^v \frac{a^4 + (b^2 - a^2)x^2}{\sqrt{(a^2 - x^2)(a^4 + (b^2 - a^2)x^2)}}dx. \tag{3}$$

As mentioned above, Liouville's work has made clear that the integral in (3) cannot be "solved" in the sense that it can be expressed by elementary functions. It may be instructive to see how the integral in (3) can be transformed to an elliptic integral of the second kind (see Sect. 3.5). Assume first that $b > a$. Let us make a change of variable, $x \to z, x = a\sqrt{1 - z^2}$. Then $dx = -az(1 - z^2)^{-1/2}dz$. The corresponding transformed integral limits become $z_1 = \sqrt{1 - (u/a)^2}$ and $z_2 = \sqrt{1 - (v/a)^2}$. As a result, it follows from (3) that

$$s(u,v) = ab\int_{z_1}^{z_2} \frac{(1 - k^2z^2)dz}{\sqrt{(1 - z^2)(1 - k^2z^2)}}$$

where $k^2 = (b^2 - a^2)/b^2$. The last integral is an elliptic integral of the second kind with $m = -k^2$. In the case where $b < a$, the proof is similar.

Example 2.2. Oscillation of the Pendulum

There is another well-known problem in mechanics that also leads to an elliptic integral, namely: the analysis of the oscillation of the simple pendulum. Let g denote the acceleration due to gravity, l the length of the pendulum cord, and $\theta(t)$ the angle between the y-axis and the pendulum cord at time t. The challenge is to establish how $\theta(t)$ depends on t. By applying the standard laws of mechanics, one can show, see for example, (Stephenson 1960), that the following differential equation holds:

$$\theta''(t) = -\frac{g}{l}\sin\theta(t). \tag{4}$$

By multiplying (4) by $2\theta'(t)$, it follows that

$$2\theta''(t)\theta'(t) = -\frac{2g}{l}\sin(\theta(t))\theta'(t). \tag{5}$$

The left-hand side of (5) is the derivative of $\theta'(t)^2$and the right-hand side is the derivative of

$$2\cos(\theta(t))g/l.$$

Hence

$$\theta'(t)^2 = \kappa\cos\theta(t) + c \tag{6}$$

where c is a constant and $\kappa = 2g/l$. Let α denote the greatest angle between the y-axis and the pendulum cord: that is, the angle for which $\theta'(t) = 0$. From this, we obtain $c = -\kappa\cos\alpha$. By solving the differential equation in (6) and by making the change of variable $u = \cos\theta(t)$, it follows that the inverse solution can be written as

$$t(\theta) = \int_{\cos\theta}^{1} \frac{du}{\sqrt{(1-u^2)(u-\cos\alpha)\kappa}} \tag{7}$$

where $t(\theta)$ is the time at which the pendulum cord makes an angle θ with the vertical axis. Note that this integral is an elliptic integral. The integral in (7) can be transformed into an elliptic integral of the first kind (see Sect. 3.5) by making the change of variable $u \to z$, $u = 1 + (\cos\alpha - 1)z^2$, which gives

$$\int_{\cos\theta}^{1} \frac{du}{\sqrt{(1-u^2)(u-\cos\alpha)\kappa}} = \sqrt{\frac{2}{x}} \int_{h(\theta)}^{1} \frac{dz}{\sqrt{(1-z^2)(1-0.5(1-\cos\alpha)z^2)}}$$

where

$$h(\theta) = \sqrt{\frac{1-\cos\theta}{1-\cos\alpha}} = \frac{\sin(0.5\theta)}{\sin(0.5\alpha)}.$$

This last integral is an elliptic integral of the first kind with $k^2 = 0.5(1-\cos\alpha)$.

Example 2.3. The Arch Length of the Lemniscate

A third example where an elliptic integral occurs is when computing the arch length of the lemniscate. The graph of the lemniscate has the same form as the infinity sign, instead of w, and can be expressed in Cartesian coordinates by the following formula

$$\left(x^2 + y^2\right)^2 = x^2 - y^2$$

or in polar coordinates as

$$r(\theta)^2 = \cos(2\theta).$$

Here, only the case where $\theta \in [-\pi/4, \pi/4]$ is considered where θ is the angle between the x-axis and the radius $r(\theta)$ from the origin to the graph. Since $x = r(\theta)\cos\theta$ and $y = r(\theta)\sin\theta$, this implies that $dx = (r'(\theta)\cos\theta - r(\theta)\sin\theta)d\theta$ and $dy = (r'(\theta)\sin\theta + r(\theta)\cos\theta)d\theta$, so that

$$\sqrt{1+\left(\frac{dy}{dx}\right)^2}dx = \sqrt{dx^2+dy^2} = \sqrt{r'(\theta)^2 + r(\theta)^2}d\theta = \frac{d\theta}{\sqrt{\cos(2\theta)}} = \frac{d\theta}{r(\theta)}.$$

Thus, the arch length $s(\theta_0)$ which corresponds to the angle from 0 to θ_0 is given by

$$s(\theta_0) = \int_0^{\theta_0} \frac{d\theta}{r(\theta)}.$$

By the change of variable $\theta \to r, r = \sqrt{\cos 2\theta}$, one gets

$$dr = -\frac{\sin(2\theta)}{\sqrt{\cos(2\theta)}}d\theta = -\frac{\sqrt{1-\cos^2(2\theta)}}{\sqrt{\cos 2\theta}}d\theta = -\frac{\sqrt{1-r^4}}{r}d\theta$$

which yields

$$s(\theta_0) = \int_0^{\theta_0} \frac{d\theta}{r(\theta)} = -\int_1^{r_0} \frac{dr}{\sqrt{1-r^4}} = \int_{r_0}^{1} \frac{dr}{\sqrt{1-r^4}} \tag{8}$$

where $r_0 = \sqrt{\cos 2\theta_0}$. By making the change of variable $r \to w, r = \sqrt{1-w^2}$, it follows that

$$s(\theta_0) = \frac{1}{\sqrt{2}} \int_0^{\sin 2\theta_0} \frac{dw}{\sqrt{(1-w^2)(1-0.5w^2)}}.$$

This last integral is an elliptic integral of the first kind with $k^2 = 0.5$.

3 Early Approaches to Addition Theorems

By addition theorems, we mean theorems on how the sum or product of two or several functions (integrals) of given arguments can be expressed as a function of the same type with argument that is an explicit function of the given arguments. Familiar examples are the logarithm, exponential, and trigonometric functions.

3.1 Addition Theorems for Inverse Trigonometric Functions

The inverse trigonometric functions are special cases of elliptic integrals, cf. Sects. 1 and 2. One example of an addition theorem is given by

$$\arcsin x + \arcsin z = \arcsin\left(x\sqrt{1-z^2} + z\sqrt{1-x^2}\right), \tag{9}$$

for $xz \leq 0$ or $x^2 + z^2 \leq 1$,

$$\arcsin x + \arcsin z = \pi - \arcsin\left(x\sqrt{1-z^2} + z\sqrt{1-x^2}\right), \tag{10}$$

for $x > 0$, $z > 0$, $x^2 + z^2 > 1$, and

$$\arcsin x + \arcsin z = -\pi - \arcsin\left(x\sqrt{1-z^2} + z\sqrt{1-x^2}\right), \tag{11}$$

for $x < 0$, $z < 0$, $x^2 + z^2 > 1$, where

$$\arcsin x = \int_0^x \frac{du}{\sqrt{1-u^2}}.$$

Thus, arcsin$x$ is an elliptic integral of the first kind with $k = 0$. The addition theorem in (9), (10), and (11) can be proved by applying the well-known formula

$$\sin(u+v) = \sin u \cos v + \cos u \sin v$$

where $u = \arcsin x$ and $v = \arcsin z$. A similar addition theorem holds for the integral

$$arctgx = \int_0^x \frac{du}{1+u^2}.$$

By the change of variable, $u \to z$, $u = z(1-z^2)^{-1/2}$, the integral above becomes

$$\int_0^x \frac{du}{1+u^2} = \int_0^{x\left(1+x^2\right)^{-1/2}} \frac{dz}{\sqrt{1-z^2}}$$

which demonstrates that arctg$x$ can be expressed by an elliptic integral of the first kind with $k = 0$.

It can easily be demonstrated that the following addition theorem, namely

$$\text{arctg}x + \text{arctg}z = \text{arctg}\left(\frac{x+z}{1-xz}\right)$$

holds for $xz < 1$, and an analogous formula holds for $xz > 1$.

3.2 Development of the Theory of Elliptic Integrals in the Eighteenth and early Nineteenth Century

Since it seemed impossible to reduce elliptic integrals to elementary functions, mathematicians started to investigate what might be called the intrinsic properties of these transcendental functions. Note that in the eighteenth and early nineteenth century, elliptic integrals were often called elliptic functions, in contrast to the terminology established after Abel and Jacobi where the inverse of elliptic integrals are called elliptic functions. In the eighteenth century, the properties of the trigonometric and inverse trigonometric functions were well known by Euler and others, including the addition relations reviewed in Sect. 3.1. It is not known if Euler was led to his approach in the study elliptic integrals by considering the properties of the inverse trigonometric functions reviewed in Sect. 3.1, but it may very well have been so. To elaborate on this point, consider the relation in (9) and let

$$y = x\sqrt{1-z^2} + z\sqrt{1-x^2} \tag{12}$$

where z now is a given constant. By differentiation, it follows from (9) that

$$\frac{dy}{\sqrt{1-y^2}} = \frac{dx}{\sqrt{1-x^2}}. \tag{13}$$

The solution of (13) is given by $y = \sin(\arcsin x + C)$ which is equivalent to

$$y = \sin(\arcsin x + \arcsin z)$$

where $z = \sin C$. But it is known from (9) that y is also given by (12), which implies that x and y are related by the second-degree polynomial equation

$$x^2 + y^2 - 2\alpha xy - \beta = 0 \tag{14}$$

where α and β are constants, $\beta = z^2$ and $\alpha^2 + \beta = 1$. In other words, the solution (y) of (13) can be expressed as an algebraic function of x.

As mentioned above, it was Fagnano and Euler who were the first to study elliptic integrals systematically, see Ayoub (1984). When Euler started his investigations of elliptic integrals, inspired by the results obtained by Fagnano, his point of departure was the differential equation

$$\frac{dy}{\sqrt{1-y^4}} = \frac{dx}{\sqrt{1-x^4}}. \tag{15}$$

Euler's conjecture was that the solution of (15) would satisfy the polynomial equation

$$x^2 + y^2 + \beta x^2 y^2 - 2\alpha xy - \beta = 0 \tag{16}$$

where α, and β, are constants. In fact, he proved that this is so with $\beta > 0$ and $\alpha^2 + \beta^2 = 1$. Since the differential equations in (13) and (15) are quite similar, Euler's conjecture might very well have been motivated by the fact that the solution of (13) is determined by (14). Apart from the term $\gamma x^2 y^2$, the latter equation is identical to (16). Subsequently, Euler studied more general differential equations of the same type where the terms under the radicals of (15) were general polynomials of degree four, to be discussed below.

Among the approaches to addition theorems of elliptic functions in the nineteenth century, it is the approach of Lagrange (1766–9) that appears to be the most interesting one. His method is constructive whereas the approaches of Fagnano (1718), Euler (1761), and Legendre (1825–1828) were based on conjectures of how the addition theorem might look like.

In this context, it is worth mentioning that John Landen (1719–1790) obtained what is known as Landen's theorem (Landen 1775). Landen's theorem expresses the length of the arc of a hyperbola in terms of the lengths of the arcs of two ellipses. For details, see Cayley (1961) who goes through Landen's result. It seems, however, that Landen did not establish addition theorems explicitly and did not fully realize the value of his discovery (Mittag-Leffler 1923).

Both Euler and Lagrange tried in vain to extend their approaches to more general situations where the terms under the radicals of (15) were polynomials of a higher degree than the fourth. Much later, Richelot (1842) demonstrated how Lagrange's approach can be used to prove more general addition theorems.

After Euler and Lagrange, Legendre continued research on the theory of elliptic integrals. Although he produced a different proof of Euler's addition theorem, he did not come up with fundamental new ideas beyond those of Euler and Lagrange. One of his major results was that the demonstration that all elliptic integrals can be reduced to the three fixed canonical forms given in Sect. 3.5 below.

3.3 Results Obtained by Fagnano

Fagnano wrote several papers that dealt with integrals that had geometric interpretations but could not be expressed by elementary functions. In 1718, he published the following remarkable relation for the arch length of the lemniscate (Fagnano 1718; compare Example 2.3),

$$2s(x) = s\left(\frac{2x\sqrt{1-x^4}}{1+x^4}\right) \tag{17}$$

for $x \in [-1, 1]$,where $s(u)$ is as defined in (8). The problem of establishing (17) is equivalent to showing that

$$y = \frac{2x\sqrt{1-x^4}}{1+x^4} \tag{18}$$

is a solution of the differential equation

$$\frac{dy}{\sqrt{1-y^4}} = \frac{2dx}{\sqrt{1-x^4}}.$$

Apparently, it was the following problem formulated by John Bernoulli (1667–1748) that motivated Fagnano in his investigations, namely: what is the curve with the property that the time taken for a particle to traverse the curve is proportional to the distance from a given point? The solution of this problem yields the integral given in (8). Fagnano's result in (17) can in fact be viewed as a special case within the theory of transformations of elliptic integrals. This topic was to become central in the competition between Abel and Jacobi. Although Fagnano demonstrated that (18) was a solution to (14), he was not able to find the general solution to this differential equation. For further discussion on Fagnano's work on elliptic integrals, see Ayoub (1984).

3.4 Results Obtained by Euler

Euler first saw Fagnano's work on 23 December 1751, which explains why Jacobi referred to this date as the birth of elliptic functions. According to Ayoub (1984), it was then that Euler was asked to examine the collected papers of Fagnano, who was being put forward for membership of the Berlin Academy. Euler had already encountered elliptic integrals through his work on elasticity and had published several papers on the topic before 1751. He had also written papers on the rectification of the ellipse, which leads to an elliptic integral, see Example 2.3. Ayoub (1984) notes that Euler had apparently not written anything about addition theorems for elliptic integrals before reading the work of Fagnano. Subsequently, Euler (1761) obtained the general solution to (15) and he established the following addition theorem

$$s(x) + s(y) = s\left(\frac{x\sqrt{1-y^4} + y\sqrt{1-x^4}}{1+x^2y^2}\right) \tag{19}$$

which generalizes (17), since (19) implies (17) when $x = y$. In the same paper, Euler demonstrated that the approach applied to prove (19) can also be applied to prove addition theorems for more general integrals $F(x)$ of the form

$$F(x) = \int_0^x \frac{du}{\sqrt{p(u)}}$$

where $p(x)$ is a polynomial of degree three or four.

Next, consider Euler's proof of his addition theorem in the case with $p(x) = 1 + ax^2 + bx^4$ where a and b are real constants such that $p(x) > 0$, $x \in [0, c]$, for some constant c. It will be clear that the corresponding addition theorem can be obtained by solving the differential equation

$$\frac{dy}{\sqrt{1 + ay^2 + by^4}} = \frac{\varepsilon dx}{\sqrt{1 + ax^2 + bx^4}} \tag{20}$$

where y is an unknown function $f(x)$ (say), and ε is equal to 1 or -1. Here, it is sufficient to consider the case with $\varepsilon = 1$. However, it will be convenient to consider the case with $\varepsilon = -1$ below when discussing Legendre's approach. With $\varepsilon = 1$, the equation in (20) becomes equivalent to $F(y) = F(x) + C$ where C is a constant and y is a function of x and C. Thus, the solution can be expressed as $y = F^{-1}(F(x) + C)$. Without loss of generality, we can write $C = F(z)$, where z is a constant. Hence, (20) becomes equivalent to

$$F(y) = F(x) + F(z) \tag{21}$$

where now y is a function of x and z. Euler (1761) proved that the solution of (20) also satisfies the equation

$$a\left(x^2 + y^2\right) = 2\beta xy + \gamma x^2 y^2 + \delta \tag{22}$$

where α, β, γ, and δ are constants. We shall next prove this. By creating the total differential of (22), one obtains

$$\alpha(xdx + ydy) = \beta(xdy + ydx) + \gamma\left(xy^2 dx + x^2 ydy\right)$$

which is equivalent to

$$\left(\alpha x - \beta y - \gamma xy^2\right)dx + \left(\alpha y - \beta x - \gamma x^2 y\right)dy = 0. \tag{23}$$

When solving (22) for x and y, respectively, two possible solutions are

$$x = \frac{\beta y - \sqrt{\alpha\delta + \left(\beta^2 - \alpha^2 - \gamma\delta\right)y^2 + \alpha\gamma y^4}}{\alpha - \gamma y^2} \tag{24}$$

and

$$y = \frac{\beta x + \sqrt{\alpha\delta + \left(\beta^2 - \alpha^2 - \gamma\delta\right)x^2 + \alpha\gamma x^4}}{\alpha - \gamma x^2} \tag{25}$$

There are in fact four possible solutions of (22) of x (y) in terms of y (x). However, it is essential for the subsequent results to pick from among these solutions the largest one for x and the smallest one for y.

From the equations above, by multiplying by the respective denominators in (24) and (25), it follows that

$$\alpha x - \beta y - \gamma x y^2 = -\sqrt{\alpha\delta + (\beta^2 - \alpha^2 - \gamma\delta)y^2 + \alpha\gamma y^4} \tag{26}$$

and

$$\alpha y - \beta x - \gamma x^2 y = \sqrt{\alpha\delta + (\beta^2 - \alpha^2 - \gamma\delta)x^2 + \alpha\gamma x^4}. \tag{27}$$

By inserting (26) and (27) into (23), one obtains

$$\frac{dy}{\sqrt{\alpha\delta + (\beta^2 - \alpha^2 - \gamma\delta)y^2 + \alpha\gamma y^4}} = \frac{dx}{\sqrt{\alpha\delta + (\beta^2 - \alpha^2 - \gamma\delta)x^2 + \alpha\gamma x^4}} \tag{28}$$

which proves that the relation in (23) in fact represents a solution of (20) with $\alpha\delta = 1$, $a = \beta^2 - \alpha^2 - \gamma\delta$, $b = \alpha\gamma$ and $\alpha = 1/z$. Hence, $\delta = z$, $\gamma = bz$, and $\beta = \sqrt{a + z^{-2} + bz^2}$ where z is an arbitrary variable that for the moment is kept fixed. By using the above notation, it follows that (25) can be expressed as

$$y = \frac{x\sqrt{1 + az^2 + bz^4} + z\sqrt{1 + ax^2 + bx^4}}{1 - bz^2x^2}. \tag{29}$$

Above, z is treated as a constant, but since the relations above hold for any value of z, one can treat z as a free variable. From (29) and (21), it is therefore the case that

$$F\left(\frac{x\sqrt{1 + az^2 + bz^4} + z\sqrt{1 + ax^2 + bx^4}}{1 - bz^2x^2}\right) = F(x) + F(z). \tag{30}$$

Note that the equation in (19) is a special case of (30). By letting $a = -(1 + k^2)$ and $b = k^2$, the addition equation in (30) can be reduced to

$$H\left(\frac{x\Delta(z) + z\Delta(x)}{1 - k^2z^2x^2}\right) = H(x) + H(z) \tag{31}$$

where $H(x)$ denotes the (definite) elliptic integral of the first kind and $\Delta(x)^2 = (1 - x^2)(1 - k^2x^2)$. Thus, an addition theorem for elliptic integrals has been established. Note that when $z = x$ in (31), an extension of Fagnano's invariance relation in (17) is obtained.

Moreover, Euler (1761) demonstrated that the same technique as above works for solving differential equations of the type

$$\frac{mdx}{\sqrt{p(x)}} = \frac{ndy}{\sqrt{q(y)}} \tag{32}$$

where m and n are integers, and $p(x)$ and $q(x)$ are general fourth-order polynomials in x that are not necessarily equal. The solutions of (32) constitutes part of the theory of transformations, see Cayley (1961). For example, if in (32), $m = n = 1$, and $p(x) = (1 - \alpha^2x^2)(1 - \beta^2x^2)$, then the differential $dx/\sqrt{p(x)}$ is carried into a differential of the same form $dy/\sqrt{q(y)}$ by means of the transformation

$$y = \frac{x}{1 + \alpha\beta x^2}$$

where $q(y) = \left(1 - \alpha_1^2y^2\right)\left(1 - \beta_1^2y^2\right), \alpha_1 = \alpha + \beta$ and $\beta_1 = 2\sqrt{\alpha\beta}$. Similarly, Landen (1775) and Mittag-Leffler (1923) discovered a similar result, namely that when $m = n = 1, \alpha_1 = \alpha + \sqrt{\alpha^2 - \beta^2}$ and $\beta_1 = \alpha - \sqrt{\alpha^2 - \beta^2}$, then

$$y = x\sqrt{\frac{1 - \alpha^2x^2}{1 - \beta^2x^2}}.$$

Recall that part of Euler's method is the trick of letting the coefficients of the original differential equation in (20) be suitable functions of the free variable z. At the outset, this variable is understood to be fixed, but it becomes clear later that it may vary freely since the corresponding differential equation is supposed to hold for *any* value of z. Although this observation is trivial, it is nevertheless very important because it makes it possible to express y in the symmetric relation given in (29), which is a key element in the statement of the corresponding addition theorem. Below it will be shown that in Abel's approach also a similar trick of allowing one or several coefficients to depend on a free variable is essential for establishing Abel's addition theorem.

3.5 Lagrange's Approach

Lagrange introduced a method for solving a specific type of differential equations that can be applied constructively to obtain the addition theorem for elliptic integrals (Cayley 1961). Lagrange's approach for solving (20) with $\varepsilon = 1$ starts with the introduction of a new variable t, defined by

$$t = t(x) = \int_0^x \frac{du}{\sqrt{p(u)}}. \tag{33}$$

Equations (20) and (33) imply that

$$dt = \frac{dx}{\sqrt{p(x)}} = \frac{dy}{\sqrt{p(y)}} \tag{34}$$

and

$$\frac{dx}{dt} = \sqrt{p(x)} \text{ and } \frac{dy}{dt} = \sqrt{p(y)}. \tag{35}$$

By differentiating the first equation in (35) with respect to t, it follows that

$$\frac{d^2x}{dt^2} = \frac{d(dx/dt)}{dx} \cdot \frac{dx}{dt} = \frac{d\left(\sqrt{p(x)}\right)}{dx} \cdot \frac{dx}{dt}$$

$$= \frac{2ax + 4bx^3}{2\sqrt{p(x)}} \cdot \sqrt{p(x)} = ax + 2bx^3.$$

Similarly, it follows from (35) that

$$\frac{d^2y}{dt^2} = ay + 2by^3.$$

At this point, it becomes convenient to introduce the notation $r = x + y$ and $q = x - y$. By means of this notation, one gets

$$\frac{d^2q}{dt^2} = \frac{d^2x}{dt^2} - \frac{d^2y}{dt^2} = aq + \frac{b}{2}\left(q^3 + 3r^2q\right) \tag{36}$$

and

$$\frac{dr}{dt} \cdot \frac{dq}{dt} = \left(\frac{dx}{dt} + \frac{dy}{dt}\right) \cdot \left(\frac{dx}{dt} - \frac{dy}{dt}\right) = \left(\frac{dx}{dt}\right)^2 - \left(\frac{dy}{dt}\right)^2$$

$$= a\left(x^2 - y^2\right) + b\left(x^4 - y^4\right) = arq + \frac{b}{2}rq\left(r^2 + q^2\right). \tag{37}$$

From (36) and (37), it follows that

$$r\frac{d^2q}{dt^2} - \frac{dr}{dt} \cdot \frac{dq}{dt} = arq + \frac{brq^3}{2} + \frac{3r^3q}{2} - arq - \frac{brq^3}{2} - \frac{br^3q}{2} = br^3q. \tag{38}$$

By dividing both sides of (38) by $r^3$and multiplying both sides by 2 dq/dt lead to

$$\frac{2}{r^2}\cdot\frac{d^2q}{dt^2}\cdot\frac{dq}{dt}-\frac{2}{r^3}\left(\frac{dq}{dt}\right)^2\cdot\frac{dr}{dt}=2br\cdot\frac{dq}{dt}. \tag{39}$$

Equation (39) can readily be integrated, which gives

$$\frac{1}{r^2}\cdot\left(\frac{dq}{dt}\right)^2=bq^2+K \tag{40}$$

where K is an arbitrary constant. When inserting for r, q, and dq/dt, as functions of x and y, into (40) and applying the equation in (35), the resulting equation becomes

$$\left(\frac{\sqrt{p(x)}-\sqrt{p(y)}}{x+y}\right)^2=b(x-y)^2+K. \tag{41}$$

Equation (41) expresses the relationship between x and y in implicit form. It is, however, possible to simplify this equation considerably. After multiplication on both sides of (41) by $(y+x)^2$and rearranging, it is obtained that

$$-2\sqrt{p(x)p(y)}=(y+x)^2\left(K+(x-y)^2b\right)-2-a\left(x^2+y^2\right)-b\left(x^4+y^4\right). \tag{42}$$

By raising both sides of (42) to the power of 2 yields

$$\begin{aligned}0&=\left((y+x)^2\left(K+(x-y)^2b\right)-2-a\left(x^2+y^2\right)-b\left(x^4+y^4\right)\right)^2\\&\quad-4\left(1+ax^2+bx^4\right)\left(1+ay^2+by^4\right)\\&=\left(a^2-4b\right)\left(y^2-x^2\right)^2+K^2(y+x)^4-2aK(y+x)^2\left(x^2+y^2\right)\\&\quad-4K(y+x)^2-4bK(y+x)^2x^2y^2.\end{aligned} \tag{43}$$

Note that the right-hand side of (43) has $(y+x)^2$ as a common factor. By dividing through and rearranging, (43) can be reduced to

$$\alpha\left(x^2+y^2\right)=2\beta xy+\gamma x^2y^2+\delta \tag{44}$$

where $\delta=4K$, $\beta=a^2-4b-K^2$, $\gamma=4bK$, and $\alpha=K^2-4b+a^2-2aK$.

Equation (44) represents the solution y as a function of x, of the differential equation in (20). Note that the equation (44) is equivalent to the equation (22). This means that there are no other solutions to Euler's differential equation (20) than the one Euler postulated and subsequently verified. The rest of the analysis proceeds in the same way as Euler's deductions.

3.6 Results Achieved by Legendre

As mentioned above, it was Legendre who most systematically and extensively conducted research on elliptic integrals before Abel and Jacobi. In his monumental three-volume work, *Traité des fonctions elliptiques et des intégrales éulériens,* which was published in the period 1825–1828, Legendre presented what was the state of the theory of elliptical integrals (which he called elliptic functions) at that time, and he also discussed a number of applications in geometry and mechanics. Furthermore, he has given a comprehensive contribution to numerical analysis that leads to practical approximation formulas that could be used to calculate numerical values of various functions, such as the Gamma function and elliptical integrals. In volume 2 of Legendre's work, there are approximately 141 pages with tables calculated with a precision of between 10 and 15 decimals. It thus is an impressive amount of work that has been put into making these tables. Some of the results presented by Legendre in 1825–1828 were previously published by Legendre (1793) and Legendre (1811–1817). In 1825, Legendre proved that elliptic integrals can always be expressed by linear combinations of elementary functions and the following three types of elliptic integrals, namely:

$$\int \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \int \frac{(1-k^2x^2)dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \text{and} \int \frac{dx}{(1-nx^2)\sqrt{(1-x^2)(1-k^2x^2)}}$$

where k and n are constants and $|k| \leq 1$. In Legendre's classification, such integrals are called elliptic integrals of the first, second, and third kind, respectively.

In Chapter 6 of Legendre (1825), a proof is given of the addition theorem established by Euler and Lagrange for elliptic integrals. He shows that Euler's differential equation in (20) with $a = 1$, $b = -(1 + k^2)$, $c = k^2$ and $\varepsilon = -1$ can, by making the substitution $x = \sin \varphi$ and $y = \sin \psi$, the differential equation becomes

$$\frac{d\psi}{\sqrt{1-k^2 \sin^2\psi}} + \frac{d\varphi}{\sqrt{1-k^2 \sin^2\varphi}} = 0. \tag{45}$$

Here, it is understood that ψ is a function of φ. The problem now is to solve (45), that is, to establish the relationship between φ and ψ. By integrating (45), it follows that

$$G(\psi) + G(\varphi) = C \tag{46}$$

where

$$G(\varphi) = \int_0^{\phi} \frac{dv}{\sqrt{1-k^2 \sin^2 v}}$$

and C is a constant. As with Euler, Legendre's proof is based on conjecture about what a solution of (45) might look like. Specifically, Legendre's conjecture was that the relationship between y and q can be expressed by

$$\cos\varphi\cos\psi - \cos\mu = (\sin\varphi\sin\psi)\sqrt{1-k^2\sin^2\mu} \tag{47}$$

where μ is a constant. The equation in (47) is equivalent to each of the equations

$$\cos\psi - \cos\mu\cos\varphi = (\sin\mu\sin\varphi)\sqrt{1-k^2\sin^2\psi} \tag{48}$$

and

$$\cos\varphi - \cos\mu\cos\psi = (\sin\mu\sin\psi)\sqrt{1-k^2\sin^2\varphi}. \tag{49}$$

The simplest way of verifying that (47), (48), and (49) are equivalent is to raise both sides of the equations to the power of 2 and then make suitable rearrangements. By dividing (46) by $\sin\varphi\sin\psi$ and subsequently differentiating with respect to ψ, treating φ as a function of ψ, one gets

$$\frac{1}{\sin\varphi}(\cos\psi - \cos\mu\cos\varphi) + \frac{d\psi/d\varphi}{\sin\psi}(\cos\varphi - \cos\mu\cos\psi) = 0.$$

Moreover, by using (47) and (48), it follows that

$$\frac{d\varphi}{\sin\varphi}\cdot(\sin\mu\sin\varphi)\sqrt{1-k^2\sin^2\psi} + \frac{d\psi}{\sin\psi}\cdot(\sin\mu\sin\psi)\sqrt{1-k^2\sin^2\varphi} = 0$$

which reduces to the differential equation in (45). By letting $\varphi = 0$, it follows from (47) that $\psi(0) = \mu$, implying that $C = G(\mu)$. If one instead considers μ as a function of φ and ψ, determined by (48) and (49), it follows that one can establish addition theorems by solving the equations in (48) and (49) for μ. Specifically, by dividing the equation in (47) by the equation in (48) $\sin\mu$ cancels on the right-hand side and one can solve for $\cos\mu$, which yields

$$\cos\mu = \frac{(\sin\psi\cos\psi)\Lambda(\varphi) - (\sin\varphi\cos\varphi)\Lambda(\psi)}{(\sin\psi\cos\varphi)\Lambda(\varphi) - (\sin\varphi\cos\psi)\Lambda(\psi)} \tag{50}$$

where

$$\Lambda(\varphi) = \sqrt{1-k^2\sin^2\varphi^2}.$$

By multiplying the denominator and numerator of (50) by $(\sin\psi\cos\varphi)\Lambda(\varphi) - (\sin\varphi\cos\psi)\Lambda(\psi)$ and rearranging, we get that

$$\cos\mu = \frac{\cos\varphi\cos\psi - (\sin\varphi\sin\psi)\Lambda(\varphi)\Lambda(\psi)}{1 - k^2\sin^2\varphi\sin^2\psi} \tag{51}$$

Similarly, it is possible to show that

$$\sin\mu = \frac{(\sin\varphi\cos\psi)\Lambda(\psi) + (\sin\psi\cos\varphi)\Lambda(\varphi)}{1 - k^2\sin^2\varphi\sin^2\psi} \tag{52}$$

and

$$tg\mu = \frac{(tg\varphi)\Lambda(\psi) + (tg\psi)\Lambda(\varphi)}{1 - (tg\varphi tg\psi)\Lambda(\varphi)\Lambda(\psi)}. \tag{53}$$

Next, make the change of variable back to the original variables: that is, $\varphi \to x$ and $\psi \to y$, where $\varphi = \arcsin x$ and $\psi = \arcsin y$. Accordingly, $G(\varphi) = H(x)$, $G(\psi) = H(y)$, and $G(\mu) = H(\sin\mu)$, where $H(x)$ is the elliptic integral of the first kind where the integration goes from zero to x. Thus, it is possible to write (46) as $H(x) + H(y) = H(\sin\mu)$. Furthermore, (52) can be expressed as

$$\sin\mu = \frac{x\Delta(y) + y\Delta(x)}{1 - k^2x^2y^2}$$

which is equivalent to the result in (31).

Legendre has also given an alternative proof in which he shows that the relations in (46) to (49) follow from spherical geometry, and more specifically, the relations between the sides of a triangle on the surface of a ball with respective lengths φ, ψ, and μ. In addition to being of independent interest, this proof also provides a good illustration of how problems within spherical geometry lead to elliptic integrals.

4 Abel's Addition Theorem: An Elementary Exposition

When Abel began working on the theory of elliptic integrals, his approach was radically different from previous contributions in this field. His results are not limited to elliptic integrals but also extend to hyperelliptic integrals and the much more general case of Abelian integrals. The next section discusses the addition theorem in the elliptic and hyperelliptic case.

4.1 The Addition Theorem for Elliptic and Hyperelliptic Integrals

Consider an integral of the form given by

$$F(x) = \int_c^x \frac{du}{\sqrt{p(u)}}$$

where $p(u)$ is a polynomial of degree higher than 2 and c is a constant. When the degree of $p(u)$ is equal to 3 or 4, F reduces to an elliptic integral. As discussed above, Euler, Lagrange, and Legendre proved the addition theorem

$$F(x_1) + F(x_2) = F(x_3)$$

where x_3 is a particular algebraic function of (x_1, x_2) where (x_1, x_2) are two arbitrary distinct (real) variables subject to the condition that $p(x_j) > 0, j = 1, 2$. In the general case, $p(x)$ may be negative, have complex roots, and have complex coefficients. It is, however, beyond the scope of this chapter to analyze the complex case. In this section, integrals of the form

$$F(x, a) = \int_c^x \frac{\mu(u)du}{(u-a)\sqrt{p(u)}}$$

will also be treated, where a is a constant smaller than c and $\mu(u)$ is a polynomial such that the degree of $\mu(u)^2$ is less than the degree of $p(u)$.

As will be clear shortly, the method of partial fraction decomposition of rational functions plays a key role in Abel's approach. It was Gottfried Leibniz (1646–1716) and Johann Bernoulli (1667–1748) who, independently, in 1702 discovered the method known as partial fraction decomposition of rational functions, which is stated in Lemma 1 below. Abel was of course familiar with this result.

Lemma 1
(Partial Fraction Decomposition)

Let $A(x)$ and $B(x)$ be polynomials where the degree of $B(x)$ is less than the degree of $A(x)$.

Furthermore, assume that all the roots $x_1, x_2, \ldots, x_K$ of $B(x)$ are distinct and different from the roots of $A(x)$. Then

$$\sum_{j=1}^{K} \frac{A(x_j)}{(x - x_j)B'(x_j)} = \frac{A(x)}{B(x)}. \tag{i}$$

If $B(0) \neq 0$ and $A(x) = xg(x)$, then

$$\sum_{j=1}^{K} \frac{g(x_j)}{B'(x_j)} = 0. \tag{ii}$$

A proof of Lemma 1(i) is given in several mathematical textbooks, see, for example, Miller et al. (1990). The proof of Lemma 1(ii) follows by letting $x = 0$ in (i).

The next theorem is a special case of Abel's general addition theorem.

Theorem 1
Let $f(x)$, $\varphi(x)$, and $p(x)$ be polynomials where the degree of $p(x)$ is greater than 2, let $x_j, j = 1, 2, \ldots, K$ be the roots of the polynomial $f(x)^2 - \varphi(x)^2 p(x)$ and let $\varepsilon_j = \pm 1$ be determined by

$$f(x_j) = \varepsilon_j \varphi(x_j) \sqrt{p(x_j)}.$$

If all the roots $\{x_j\}$ are real, distinct, and different from zero, and $p(x) > 0$ for $x \in [c, \max_j x_j]$, then

$$\sum_{j=1}^{K} \varepsilon_j F(x_j) = C \tag{54}$$

where C is a constant.

As demonstrated below, some of the roots $\{x_j\}$ can be chosen freely, but not all. Note that the roots $\{x_j\}$ must either satisfy $f(x_j) = \varphi(x_j)\sqrt{p(x_j)}$ or $f(x_j) = -\varphi(x_j)\sqrt{p(x_j)}$. The role of ε_j is thus to identify which of these two relations applies. Note also that the degree of the polynomial $p(x)$ can be greater than 4 which means that Theorem 1 holds in the hyperelliptic case. This shows that Abel's result is far more general than the previous contributions by Euler, Lagrange, and Legendre.

A very brief summary of Abel's approach goes as follows. A key step is the introduction of an auxiliary variable z and the change of variable, $u = r_j(z)$, where $r_j(z)$ is a suitable function of z. This change of variable implies that it is possible to express the integrand of the elliptic integral $F(x)$, namely $r'_j(z)/\sqrt{p(r_j(z))}$, as a rational function in $r_j(z)$ where the denominator is the derivative of a specific polynomial of which $r_j(z)$ is a root. Since the root $r_j(z)$ depends on z one or several coefficients of this polynomials must depend on z. Then, by Lemma 1(ii) (providing the conditions of Lemma 1(ii) are met), the sum of the integrands becomes zero, which implies that the sum of the corresponding integrals in (54) becomes a constant. The auxiliary variable z is only needed in the proof and does not have to be included in the statement of Theorem 1.

In the following, the exposition of the proof is broken up in a number of additional Lemmas.

Lemma 2
Let $u = r_j(z), j = 1, 2, \ldots, K$ be differentiable functions such that $x_j = r_j(v)$ where $\{x_j\}$ are given. Then

$$\sum_{j=1}^{K} \varepsilon_j F(x_j) = \sum_{j=1}^{K} \varepsilon_j \int_c^{x_j} \frac{du}{\sqrt{p(u)}} = \int_{v_0}^{v} \xi(z)dz + \widetilde{C} \tag{55}$$

where

$$\xi(z) \equiv \sum_{j=1}^{K} \frac{\varepsilon_j r'_j(z)}{\sqrt{p(r_j(z))}} \tag{56}$$

and $v_0, \widetilde{C}$ are constants.

The proof of Lemma 2 follows readily by the change of variable, $u = r_j(z)$, where the function $r_j(z)$ is unspecified (but differentiable) apart from the condition that $x_j = r_j(v)$. The constant $\widetilde{C}$ is given by

$$\widetilde{C} = \sum_{j=1}^{K} \int_{v_{0,j}}^{v_0} \frac{r'_j(z)dz}{\sqrt{p(r_j(z))}}$$

where $c = r_j(v_{0j})$.

Although the integral on the right side of (55) does not immediately appear to represent a simplification of the original integral, it has the advantage that the original sum of integrals in (54) can be expressed as a single integral.

The next lemma shows that after a suitable choice of the function $\{r_j(z)\}$, one can get rid of the square root in the integrand of the elliptic integral and express $\sqrt{p(r_j(z))}$ as a rational function of $r_j(z)$.

Lemma 3
Let $f(x, z)$ and $\varphi(x, z)$ be polynomials in x of degrees m_1 and m_2, respectively, where at least one of the coefficients of $f(x, z)$ and/or $\phi(x, z)$ depend on a free variable z. If $r_j(z), j = 1, 2, \ldots, K$ are the roots (distinct and real) of the polynomial $\psi(x, z) = f(x, z)^2 - \varphi(x, z)^2 p(x)$ where $K = \max(2m_1, 2m_2 + q)$, then

$$\frac{\varepsilon_j}{\sqrt{p(r_j(z))}} = \frac{\varphi(r_j(z))}{f(r_j(z))}. \tag{57}$$

If q is the degree of $p(x)$, then the first $m_1 + m_2 + 1$ roots of $\psi(x)$ can be chosen freely whereas the remaining $K - m_1 - m_2 - 1 = \max(m_1 - m_2, m_2 - m_1 + q) - 1$ roots become functions of the first $m_1 + m_2 + 1$ roots.

Proof of Lemma 3:

For simplicity, the variable z is suppressed in the proof. Note that (57) is equivalent to $f(r_j)^2 - p(r_j)\varphi(r_j)^2 = 0$ which is equivalent to stating that (57) implies that $\{r_j\}$ are the roots of $\psi(x)$. A key question is which constraints (if any) the roots of the polynomial $\psi(x)$ must satisfy. Let $f = a_0 + a_1 x + \ldots + x^{m_1}$ and $\varphi = b_0 + b_1 x + \ldots b_{m_2} x^{m_2}$ where the coefficient a_{m_1} is normalized such that $a_{m_1} = 1$. Furthermore, (55) implies that

$$a_0 + a_1 r_j + \ldots + r_j^{m_1} = \varepsilon_j \left(b_0 + b_1 r_j + \ldots + b_{m_2} r_j^{m_2} \right) \sqrt{p(r_j)} \tag{58}$$

for $j = 1, 2, \ldots, K$. In (58), there are $m_1 + m_2 + 1$ unknown coefficients, namely the coefficients of the polynomials f and φ. Hence, the proof of Lemma 3 is complete.

In the following, the notation $g'_k(x, z)$ means the partial derivative of the function g with respect to argument k, $k = 1, 2$.

The next lemma is of crucial importance because it shows that after the change of variable and a suitable choice of the functions $\{r_j(z)\}$, it is possible to express the integrand of $F(x)$ as a rational function of $\{r_j(z)\}$ such that Lemma 1 can be applied.

Lemma 4

Letf (x, z) and $\varphi(x, z)$ be polynomials in x where one or several coefficients depend on a free variable z and let $\psi(x, z) = f(x, z)^2 - \varphi(x, z)^2 p(x)$. Then

$$\frac{\varepsilon_j r'_j(z)}{\sqrt{p(r_j(z))}} = \frac{2\theta(r_j(z), z)}{\psi'_1(r_j(z), z)} \tag{59}$$

where

$$\theta(x, z) = -f(x, z)\varphi'_2(x, z) + \varphi(x, z) f'_2(x, z). \tag{60}$$

Proof of Lemma 4:

By differentiating the equation $\psi(r_j(z), z) = 0$ with respect to z, it follows that

$$\psi'_1(r_j(z), z) r'_j(z) + \psi'_2(r_j(z), z) = 0. \tag{61}$$

From (59) and (60), and the fact that Lemma 3 yields

$$p(r_j(z)) = f(r, (z), z)^2 / \varphi(r_j(z), z)^2,$$

it follows that

$$= \frac{2\theta\left(r_j(z),z\right)}{\psi_1'\left(r_j(z),z\right)r_j'(z)} = -\frac{2\theta\left(r_j(z),z\right)}{\psi_2'\left(r_j(z),z\right)}$$

$$= -\frac{2\theta\left(r_j(z),z\right)}{2f\left(r_j(z),z\right)f_2'\left(r_j(z),z\right) - 2p\left(r_j(z)\right)\varphi\left(r_j(z),z\right)\varphi_2'\left(r_j(z),z\right)}$$

$$= \frac{\varphi\left(r_j(z),z\right)f_2'\left(r_j(z),z\right) - f\left(r_j(z),z\right)\varphi_2'\left(r_j(z),z\right)}{f\left(r_j(z),z\right)f_2'\left(r_j(z),z\right) - f\left(r_j(z),z\right)^2\varphi\left(r_j(z),z\right)^{-2}\varphi\left(r_j(z),z\right)\varphi_2'\left(r_j(z),z\right)}$$

$$= \frac{\varphi\left(r_j(z),z\right)}{f\left(r_j(z),z\right)} = \frac{\varepsilon_j}{\sqrt{p\left(r_j(z)\right)}}.$$

Hence, the proof of Lemma 4 is complete.

The results obtained above shows that it is possible to express $\xi(z)$, given in (56), as

$$\xi(z) \equiv \sum_{j=1}^{K} \frac{\varepsilon_j r_j'(z)}{\sqrt{p\left(r_j(z)\right)}} = \sum_{j=1}^{K} \frac{2\theta\left(r_j(z),z\right)}{\psi_1'\left(r_j(z),z\right)}. \tag{62}$$

The right side of (62) is particularly interesting because it has the structure that enables the use of Lemma 1(ii). However, to be able to do so, it must be the case that $\psi(0, z) \neq 0$ which means that zero cannot be a root of $\psi(x, z)$. Furthermore, the degree of $\theta(x, z)$ (as a polynomial in x) must be less than or equal to the degree of $\psi(x, z)$ (as a polynomial in x) minus 2. The next lemma establishes the properties needed for the application of Lemma 1(ii).

Lemma 5
The degree of θ is lower than or equal to the degree of ψ minus 2.

Proof of Lemma 5
Since both $f_2'(x, z)$ and $\phi_2'(x, z)$ do not have constant terms, it follows from (60) that $\theta(0, z) = 0$. Recall that q is the degree of $p(x)$, m_1 the degree of $f(x, z)$, and m_2 the degree of $\varphi(x, z)$ (as polynomials in x). Then the degree of ψ becomes equal to $\max(2m_1, 2m_1 + q)$. For the condition of Lemma 1(ii) to hold the inequality, $\max(2m_1, 2m_2 + q) \geq m_1 + m_2 + 2$ must hold which is equivalent tomax $(m_1 - m_2, q - (m_1 - m_2)) \geq 2$. Since $q \geq 3$, this inequality is evidently fulfilled for all $m_1 - m_2 \leq 0$ and for $m_1 - m_2 \geq 2$. Furthermore, it also holds when $m_1 - m_2 = 1$. This means that the inequality $\max(2m_1, 2m_2 + q) \geq m_1 + m_2 + 2$ holds for all integer values of m_1 and m_2.

Thus, we have proved that (54) also holds in the hyperelliptic case, that is, when $q > 4$. This completes the proof of Lemma 5.

From the Lemmas 1, 2, 3, 4, and 5, it follows that $\xi(z) = 0$. It thus follows from (57) and (59) that

$$\sum_{j=1}^{k} \varepsilon_j F(x_j) = \widetilde{C} + \sum_{j=1}^{K} \varepsilon_j F(c_j) \equiv C.$$

Hence, the proof of Theorem 1 is complete.

Let x_K be a root that corresponds to $\varepsilon_K = -1$. From Theorem 1, the next Corollary follows.

Corollary 1

Assume that the degree of $p(x)$ is equal to 3 or 4. Then

$$F(x_K) = \sum\nolimits_{j=1}^{K-1} \varepsilon_j F(x_j) + C$$

where C is a constant, $\varepsilon_j = \pm 1$ is determined such that $f(x_j) = \varepsilon_j \varphi(x_j)\sqrt{p(x_j)}, j = 1, 2, \ldots, K-1$ and x_K is determined by $f(x_K) + \varphi(x_K)\sqrt{p(x_K)} = 0$.

Next, consider the addition theorem for the integral $F(x, a)$. To this end, the following lemma is needed.

Lemma 6

Let $h_1(z)$ and $h_2(z)$ be two differentiable functions and κ a positive constant. Then

$$\int \frac{(h_2(z)h_1'(z) - h_1(z)h_2'(z))dz}{h_2(z)^2 - h_1(z)^2\kappa} = C + \frac{1}{2\sqrt{\kappa}} \log\left(\left|\frac{h_2(z) + h_1(z)\sqrt{\kappa}}{h_2(z) - h_1(z)\sqrt{\kappa}}\right|\right)$$

where C is a constant.

The proof of Lemma 6 follows immediately by differentiation of the right side of the equation stated in Lemma 6. By means of Lemma 1(i) and Lemma 2, Abel (1828d) used the same approach as in the proof of Theorem 1 to prove the following result.

Theorem 2

Let $f(x)$, $\varphi(x)$, and $p(x)$ be polynomials, let x_j, $j = 1, 2, \ldots, K$, be the roots of the polynomial $f(x)^2 - \varphi(x)^2 p(x)$, with $p(x_j) > 0$, and let $\varepsilon_j = \pm 1$ be determined by $f(x_j) = \varepsilon_j \varphi(x_j)\sqrt{p(x_j)}$. If all the roots $\{x_j\}$ are distinct, real, and different from zero, then

$$\sum_{j=1}^{m} \varepsilon_j F(x_j, a) = C - \frac{\mu(a)}{\sqrt{p(a)}} \log\left(\left|\frac{f(a) + \varphi(a)\sqrt{p(a)}}{f(a) - \varphi(a)\sqrt{p(a)}}\right|\right) \tag{63}$$

where C is a constant.

Proof of Theorem 2

By proceeding in the same way as in the proof of Theorem 1, it turns out that similarly to (59), it is possible to express the integrand of the left side of (63) as

$$\sum_{j=1}^{K} \frac{\varepsilon_j \mu(r_j(z)) r_j'(z)}{(a - r_j(z))\sqrt{p(r_j(z))}} = \sum_{j=1}^{K} \frac{2\mu(r_j(z))\theta(r_j(z), z)}{(a - r_j(z))\psi_1'(r_j(z), z)}. \tag{64}$$

Due to Lemma 1(i), (59), and (60), the right side of (64) equals

$$\frac{2\mu(a)\theta(a,z)}{\psi(a,z)} = 2\mu(a)\left(f(x,z)\varphi_2'(x,z) - \varphi\left(\frac{x,z) f_2'(x,z))dz}{f(a,z)^2 - p(a)\varphi(a,z)^2}.\right.$$

Using Lemma 6, with $\kappa = \sqrt{p(a)}$, $h_2(z) = f(a, z)$, and $h_1(z) = \varphi(a, z)$, we therefore obtain that

$$\int \frac{2\mu(a)\theta(a,z)dz}{\psi(a,z)} = C - \frac{\mu(a)}{\sqrt{p(a)}} \log\left(\left|\frac{f(a,z) + \varphi(a,z)\sqrt{p(a)}}{f(a,z) - \varphi(a,z)\sqrt{p(a)}}\right|\right)$$

from which (63) follows. Note that in Theorem 2, the dependence of the functions f and φ, on the free variable z is suppressed. This completes the proof.

4.2 A More General Version of Abel's Addition Theorem

A great advantage of Abel's approach is that it applies to much more general cases than the ones discussed.

Consider now the addition theorem for Abelian integrals of the form

$$G(x) = \int_c^x \frac{\mu(u)du}{y(u)}$$

where $\mu(x)$ is a polynomial of degree t and $y(x)$ is a root of the equation $Q(y) - p(x) = 0$, where $Q(y)$ is a polynomial of degree m that has constant coefficients and does not contain a constant term. Moreover, $p(x)$ is a polynomial in x with degree q which is equal to or greater than 3. The question to be investigated here is whether one can, similarly to the results obtained in the previous section, write

$$\sum\nolimits_{j=1}^{K} G(x_j) = C \tag{65}$$

where C is a constant and where some of the distinct real numbers $x_1, x_2, \ldots, x_K$ can be chosen freely whereas the remaining numbers are algebraic functions of the former ones. When $Q(y) = y^2$, the problem reduces to the case considered in Sect. 4.1. In the same way as the approach used in Sect. 4.1, one attempts to express $y(r_j(z))$ on the following form

$$y\big(r_j(z)\big) = f\big(r_j(z), z\big)/\varphi\big(r_j(z), z\big) \tag{66}$$

where $r_j(z)$, $j = 1, 2, \ldots, K$, are the roots of a suitable polynomial, $\phi(x, z)$ and $f(x, z)$ are polynomials in x of degrees m_2 and m_1. As above, it is understood that at least one of the coefficients of these polynomial depends on a free variable z. This means that the polynomials φ and f must satisfy the equation

$$p\big(r_j(z)\big) = Q\left(\frac{f\big(r_j(z), z\big)}{\varphi\big(r_j(z), z\big)}\right). \tag{67}$$

Let $\psi(x, z)$ be defined as

$$\psi(x, z) = \varphi(x, z)^m \left(Q\left(\frac{f(x, z)}{\varphi(x, z)}\right) - p(x)\right). \tag{68}$$

Since Q has degree m, $\psi(x, z)$ becomes a polynomial in x. Let $r_j(z)$, $j = 1, 2, \ldots, K$, be the roots of $\psi(x, z)$ where, as above, K is the degree of $\psi(x, z)$ which is given by $K \equiv \max(mm_1, mm_2 + q)$. By differentiation with respect to z and subsequently inserting the root $r_j(z)$ into (68), (67) and (68) imply that

$$\psi_2'\big(r_j(z), z\big) = \varphi\big(r_j(z), z\big)^{m-2} Q'\left(\frac{f\big(r_j(z), z\big)}{\varphi\big(r_j(z), z\big)}\right)\theta\big(r_j(z), z\big) \tag{69}$$

where, as above,

$$\theta(x, z) = \varphi(x, z) f_2'(x, z) - \varphi_2'(x, z) f(x, z).$$

It thus follows from (69) that

$$\frac{r_j'(z)}{y\big(r_j(z)\big)} = r_j'(z)\varphi\big(r_j(z), z\big)/f\big(r_j(z), z\big) = \frac{r_j'(z)\Lambda\big(r_j(z), z\big)}{\psi_2'\big(r_j(z), z\big)} = -\frac{\Lambda\big(r_j(z), z\big)}{\psi_1'\big(r_j(z), z\big)} \tag{70}$$

where

$$\Lambda\left(r_j(z),z\right)=\frac{\varphi\left(r_j(z),z\right)\psi_2'\left(r_j(z),z\right)}{f\left(r_j(z),z\right)}$$
$$=\varphi\left(r_j(z),z\right)^{m-1}Q'\left(\frac{f\left(r_j(z),z\right)}{\varphi\left(r_j(z),z\right)}\right)\theta\left(r_j(z),z\right)/f\left(r_j(z),z\right).$$

Since the polynomial Q does not contain a linear term, $\Lambda(x, z)$ becomes a polynomial in x. Furthermore, it is easily verified that the degree of $\mu(x)\Lambda(x, z)$, which may be denoted by T, is given by

$T = t + m_2 + (m - 1)m_1$. As discussed in Sect. 4.1, one must have that $T \leq K - 2$ in order for Lemma 1 to apply. This requirement implies that t, m_1, and m_2 must satisfy the inequality

$$2 + t + m_2 + (m - 1)m_1 \leq \max\left(mm_1, mm_2 + q\right) \tag{71}$$

for given values $q \geq 3$, $t \geq 0$ and $m \geq 4$. Equation (71) will not always hold, in contrast to the corresponding case analyzed in Sect. 4.1. When (71) holds, then Lemma 1(ii) implies that

$$\sum_{j=1}^{K}\frac{\Lambda\left(r_j(z),z\right)}{\psi_1'\left(r_j(z),z\right)}=0$$

which, according to (70), implies that

$$\sum_{j=1}^{K}\frac{\mu\left(r_j(z)\right)r_j'(z)}{y\left(r_j(z)\right)}=0. \tag{72}$$

In the same way as in the derivation of Theorem 1, we obtain the equation in (65) where $x_1, x_2, \ldots, x_K$ satisfy the equation

$$Q\left(\frac{\mu\left(x_j\right)f\left(x_j\right)}{\varphi\left(x_j\right)}\right)=p\left(x_j\right) \tag{73}$$

for $j = 1, 2, \ldots, K$. Thus, it has been demonstrated that the more general version of Abel's addition theorem, expressed in (65), can be obtained by the same approach as the one applied for the case of elliptic and hyperelliptic integrals. It is, however, beyond the scope of this chapter to investigate how many of the roots $x_1, x_2, \ldots, x_K$ can be chosen freely and how the remaining roots depend on the polynomials Q and p.

The version of the addition theorem discussed in this section is, however, only a special case of Abel's most general version of his theorem.

4.3 Abel's Addition Theorem in Its Most General Form

A summary of Abel's addition theorem in its most general form will be presented in this section, but without going through the proof. Let $y = y(x)$ be the algebraic function determined by

$$Q(x, y) = 0 \tag{74}$$

where $Q(x, y)$ is a polynomial in y of degree n with coefficients that are polynomials in x. Furthermore, let

$$V(x, y) = r_0(x) + r_1(x)y + \ldots + r_{n-1}(x)y^{n-1}$$

be another polynomial in y of degree $n - 1$ where the coefficients $\{r_j(x)\}$ are polynomials in x and let $\{\beta_k\}$ denote the coefficients of $\{r_j(x)\}$. Assume that

$$V(x, y) = 0. \tag{75}$$

When both (74) and (75) hold, y can be eliminated and it follows that x must satisfy an equation $\rho(x) = 0$ (say) where $\rho(x)$ is a polynomial in x of degree K with constant coefficients.

Theorem 3 *Let $x_j, j = 1, 2. \ldots, K$, be the roots of $\rho(x)$ and define*

$$M(x) = \int_0^x R(u, y(u))du$$

where $R(x, y)$ is a rational function in (x, y). Then

$$\sum_{j=1}^{\mu} M(x_j) = v_0 + \sum_{j=1}^{\eta} c_j \log w_j$$

where $\{w_j\}$ are rational functions of $\{\beta_s\}$ and $\{c_j\}$ are constants.

When $Q(x, y) = Q(y) - p(x)$ where $Q(y)$ is a polynomial in y with constant coefficients, $p(x)$ is a polynomial in x with constant coefficients and $R(x, y) = \mu(x)/y$, then Theorem 3 reduces to the result discussed in Sect. 4.2.

The proof of Theorem 3 is given in the article "Mémoire sur une propriété générale d'une classe très étendue des fonctions transcendantes" (Abel 1841). An extremely compact version of the proof is given in the article "Démonstration d'une propriété général d'une certaine classe de fonctions transcendantes" (Abel 1829b).

4.4 Elliptic Integrals and Multiplication

From trigonometry, it is well known that $\sin(my)$ can be expressed as a function of $\sin(y)$ by a simple algebraic formula where m is a natural number. Abel proved that similar formulas hold for elliptic integrals. As reviewed above, Fagnano obtained multiplication formulas, such as the one given in (17).

He also considered cases of complex multiplication, see Ayoub (1984).

Let $H(x)$ denote the elliptic integral of the first kind, namely

$$H(x) = \int_0^x \frac{du}{\Delta(u)}$$

where $\Delta(u) = \sqrt{(1-u^2)(1-k^2u^2)}$, and k is a real constant, $|k| < 1$. In the addition theorems above the roots $\{x_f\}$ were assumed to be distinct. However, by letting $x_1 \to x_2$, (say) it becomes clear that the corresponding limiting version of Theorem 1 will hold when two or several roots are equal, because the functions involved are continuous.

Corollary 2
Let $m > 1$ be an integer. Then

$$mH(x) = H(y)$$

where the relation between x and y is determined by

$$f(u)^2 - \varphi(u)^2\Delta(u)^2 = \left(u^2 - x^2\right)^m\left(u^2 - y^2\right). \tag{76}$$

One can determine the coefficients in the polynomial on the left-hand side of (76) and subsequently determine y as function of x. However, Abel used an alternative approach that is considerably simpler (Abel 1929a, pp. 256–258).

Corollary 3
Let g_m be defined by

$$H(g_m(x)) = mH(x).$$

Then g_m is given recursively by

$$g_{m+1}(x) = -g_{m-1}(x) + \frac{2g_m(x)\Delta(x)}{1 - k^2x^2g_m(x)^2}, \tag{77}$$

$$g_{2m-1}(x) = \frac{g_m(x)^2 - g_{m-1}(x)^2}{x\left(1 - k^2 g_m(x)^2 g_{m-1}(x)^2\right)} \tag{78}$$

and

$$g_{2m}(x) = \frac{2g_m(x)\Delta(g_m(x))}{1 - k^2 g_m(x)^4} \tag{79}$$

where $g_0(x) = 0$, $g_1(x) = x$. *The function* g_m *is rational when* m *is odd and has the form* $\xi(x)\Delta(x)$ *when* m *is even, where* $\xi(x)$ *is rational.*

Proof of Corollary 3

From Corollary 2, it follows, for natural numbers m and r that

$$H\left(g_{m+r}\right) = (m+r)H(x) = mH(x) + rH(x) = H(g_m) + H(g_r)$$

which implies that

$$H\left(g_{m+r}\right) = H(g_m) + H(g_r). \tag{80}$$

From (80) and the addition theorem, it follows that

$$g_{m+r} = \frac{g_m\Delta(g_r) + g_r\Delta(g_m)}{1 - k^2 g_r^2 g_m^2} \tag{81}$$

Note furthermore that $H(-x) = -H(x)$. From (80), it also follows that

$$H(g_m) = H(g_{m-r}) + H(g_r)$$

which is equivalent to

$$H(g_{m-r}) = H(g_m) - H(g_r) = H(g_m) + H(-g_r). \tag{82}$$

Hence, (81) and (82), with g_r replaced by $-g_r$, yields

$$g_{m-r} = \frac{g_m\Delta(g_r) - g_r\Delta(g_m)}{1 - k^2 g_r^2 g_m^2} \tag{83}$$

for $m > r$. By letting $r = 1$ in (82) and (86), one obtains

$$g_{m+1} = \frac{g_m\Delta(x) + x\Delta(g_m)}{1 - k^2 x^2 g_m^2} \tag{84}$$

and

$$g_{m-1} = \frac{g_m \Delta(x) - x\Delta(g_m)}{1 - k^2 x^2 g_m^2}. \tag{85}$$

When adding together the expressions in (84) and (85), the following recursion formula for $\{g_m\}$ is obtained, namely

$$g_{m+1} = -g_{m-1} + \frac{2g_m \Delta(x)}{1 - k^2 x^2 g_m^2}. \tag{86}$$

By calculating $g_2, g_3, \ldots$, it becomes clear that g_m is rational when m is odd and has the form $\xi(x)\Delta(x)$ when m is even, where $\xi(x)$ is rational. Furthermore, by multiplying the expressions in (81) and (85), one obtains

$$g_{m+r} g_{m-r} = \frac{g_m^2 - g_r^2}{1 - k^2 g_m^2 g_r^2}. \tag{87}$$

By letting $r = m - 1$ in (85), one gets $g_{m-r} = g_1 = x$, so that (87) gives

$$g_{2m-1} = \frac{g_m^2 - g_{m-1}^2}{x\left(1 - k^2 g_m^2 g_{m-1}^2\right)}.$$

Furthermore, by letting $r = m$ in (81) follows, due to $g_1 = x$, that

$$g_{2m} = \frac{2g_m \Delta(g_m)}{1 - k^2 g_m^4}.$$

Hence, the proof is complete.

4.5 Examples

To fix ideas, we shall discuss a few examples on elliptic and hyperelliptic integrals in this section.

Example 4.5.1
Let $f(x) = bx + x^3$, $\varphi(x) = d$, and $p(x) = 1 + \gamma_1 x^2 + \gamma_2 x^4 + \gamma_3 x^6$, where a, γ_1, γ_2, γ_3, b, and d are constants. The results achieved in Sect. 4.1 and 4.2 will now be applied to work out the details of the addition theorem for the hyperelliptic integral

$$F(x) = \int_0^x \frac{du}{\sqrt{p(u)}}.$$

Notice that $f(x)^2 - \varphi(x)^2 p(x)$ is of degree 6 implying that this polynomial has 6 roots. Note also that this polynomial contains solely odd powers so that it is a cubic

polynomial in x^2. According to Theorem 1, the roots of $\psi(x)$ are determined by the equations

$$f(x_j) - \varepsilon_j \varphi(x_j)\sqrt{p(x_j)} = bx_j + x_j^3 - \varepsilon_j d\sqrt{p(x_j)} = 0, \tag{88}$$

$j = 1, 2, \ldots, 6$, and these equations link the given polynomial $p(x)$, the roots $x_j, j = 1, 2, \ldots, 6$, and the coefficients b and d together. With no loss of generality, let $\varepsilon_1 = \varepsilon_2 = \varepsilon_6 = 1$ and $\varepsilon_3 = \varepsilon_4 = \varepsilon_5 = -1$. Moreover, since the roots are determined by three equations in x^2, we can put $x_4 = -x_1, x_5 = -x_2$, and $x_6 = -x_3$. Let x_1 and x_2 be given values. From (88) for $j = 1, 2$, it follows that b and d are determined by

$$b = b(x_1, x_2) = \frac{x_1^3\sqrt{p(x_2)} - x_2^3\sqrt{p(x_1)}}{x_2\sqrt{p(x_1)} - x_1\sqrt{p(x_2)}} \tag{89}$$

and

$$d = d(x_1, x_2) = \frac{x_2x_1^3 - x_1x_2^3}{x_2\sqrt{p(x_1)} - x_1\sqrt{p(x_2)}}. \tag{90}$$

From (88) and (89) and (90), it follows that the root x_3 is determined by the equation

$$b(x_1, x_2)x_3 + x_3^3 = -d(x_1, x_2)\sqrt{p(x_3)}$$

where it is understood that the expressions for b and d given in (90) and (91) are inserted. Consider next how to obtain a simplified explicit formula for x_3 as a function of x_1 and x_2. Specifically, an alternative way to determine x_3 is to use that the constant term of the polynomial $f(x)^2 - \varphi(x)^2p(x)$ equals the product of the roots of the polynomial. Since the products of the roots equals $x_1x_2x_3x_4x_5x_6 = -(x_1x_2x_3)^2$ and the constant term of the polynomial is equal to d^2, it follows that

$$x_3 = \frac{d(x_1, x_2)}{x_1x_2} = \frac{x_1^2 - x_2^2}{x_1\sqrt{p(x_2)} - x_2\sqrt{p(x_1)}} = \frac{x_1\sqrt{p(x_2)} + x_2\sqrt{p(x_1)}}{1 - \gamma_2x_1^2x_2^2 - \gamma_3x_1^2x_2^2(x_1^2 + x_2^2)}. \tag{91}$$

The third equality in (91) follows by multiplying the numerator and the denominator of the expression in the middle above by $x_1\sqrt{p(x_2)} + x_2\sqrt{p(x_1)}$. By making the change of variable $u \to -w$ yields that $F(-x) = -F(x)$. Hence, the addition theorem implies that

$$C = \sum_{j=1}^{6} \varepsilon_j F(x_j) = F(x_1) + F(x_2) - F(x_3) - F(x_4) - F(x_5) + F(x_6). \tag{92}$$

The constant C becomes equal to zero due to the property that $F(0) = 0$. Hence, from (92) it follows that

$$\begin{aligned} 0 &= F(x_1) + F(x_2) - F(x_3) - F(x_4) - F(x_5) + F(x_6) \\ &= F(x_1) + F(x_2) - F(x_3) - F(-x_1) - F(-x_2) + F(-x_3) \\ &= 2F(x_1) + 2F(x_2) - 2F(x_3) \end{aligned}$$

which yields

$$F(x_1) + F(x_2) = F(x_3) \tag{93}$$

where x_3 is given by (92). By letting $x_1 = x_2 = x$ and $y = x_3$, it follows from (91) by using l'Hopital's rule and (93) that if y is given by

$$y = \frac{2x\sqrt{p(x)}}{1 - \gamma_2 x^4 - 2\gamma_3 x^6}.$$

and

$$F(y) = 2F(x).$$

Example 4.5.2
When $\gamma_1 = -1 - k^2$, $\gamma_2 = k^2$, and $\gamma_3 = 0$, it follows that $p(x) = (1 - x^2)(1 - k^2x^2)$ where k is a constant, $|k| < 1$, and $F(x)$ becomes an elliptic integral of the first kind. In this case, the equation in (91) reduces to

$$x_3 = \frac{x_1\sqrt{p(x_2)} + x_2\sqrt{p(x_1)}}{1 - k^2x_1^2x_2^2}$$

which, combined with (93), yields the addition theorem for elliptic integrals of the first kind.

Example 4.5.3
This example is a slightly modified version of the example discussed in a letter Abel wrote on 6 August 1826 to August Leopold Crelle (1780–1855), editor and founder of the *Journal für die reine und angewandte Mathematik*, see Abel (2012). In this letter, Abel mentions an application of his addition theorem to the hyperelliptic integral

$$F(x) = \int_b^x \frac{du}{\sqrt{p(u)}}$$

where $p(x) = a + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + x^6$ and $a, a_1, \ldots, a_5$ are given numbers such that $p(x)$ is positive on a suitable set. Thus, the integral above is a hyperelliptic integral. However, Abel provides no details in his letter of how he achieved his results; this task is left to the reader. Here, a proof is provided of the assertion put forward in Abel's letter by applying the addition theorem.

Let

$$\begin{aligned}\psi(x) &= \left(c + c_1x + c_2x^2 + x^3\right)^2 - p(x)\\ &= c^2 - a + (2cc_1 - a_1)x + \left(c_1^2 + 2cc_2 - a_2\right)x^2 + (2c + 2c_1c_2 - a_3)x^3\\ &\quad + \left(2c_1 + c_2^2 - a_4\right)x^4 + (2c_2 - a_5)x^5 \end{aligned} \tag{94}$$

where c, c_1, and c_2 are free coefficients (variables). From this it follows that $\psi(x)$ has only five roots ($K = 5$) because the term x^6 cancels. Let x_1, x_2, x_3, y_1, and y_2 denote the roots of $\psi(x)$. As in the examples given above, the free coefficients c, c_1, and c_2 can be expressed as functions of x_1, x_2, and x_3 by solving the equations $\psi(x_j) = 0$, for $j = 1, 2, 3$, which implies that the free coefficients must satisfy the linear equations

$$c + c_1x_j + c_2x_j^2 = \sqrt{p(x_j)} - x_j^3$$

for $j = 1, 2, 3$. The remaining roots y_1 and y_2 are not free variables but determined by $\psi(y_j) = 0$ for $j = 1, 2$, after the coefficients c, c_1, and c_2 have been determined as functions of x_1, x_2, and x_3. From (94) it follows by using the relationship between roots and coefficients in a polynomial that

$y_1 + y_2 = \frac{2c_1 + c_2^2 - a_4}{a_5 - 2c_2} - x_1 - x_2 - x_3$ and $y_1y_2 = \frac{c^2 - a}{(a_5 - 2c_2)x_1x_2x_3}$ which imply that y_1 and y_2 are roots of the equation

$$y^2 - \left(\frac{2c_1 + c_2^2 - a_4}{a_5 - 2c_2} - x_1 - x_2 - x_3\right)y + \frac{c^2 - a}{(a_5 - 2c_2)x_1x_2x_3} = 0. \tag{95}$$

Hence y_1 and y_2 can be determined by solving the quadratic equation in (95). As a result, Theorem 1 gives

$$F(x_1) + F(x_2) + F(x_3) = C - F(y_1) - F(y_2)$$

where C is a constant. Thus, the claim expressed in Abel's letter is verified.

Recall that in the hyperelliptic case where $p(x)$ has degree q, $q \geq 5$, it is, in contrast to the elliptic case, not generally possible to specify $K - 1$ independent roots with the K-th root being an algebraic function of the independent roots, as we have seen in the current example. However, in special cases of hyperelliptic integrals, such as in Example 4.5.1, it is still possible to obtain a similar addition theorem as in the case with elliptic integrals.

Example 4.5.4
Let $Q(y) = y^3, f(x) = b + x, \varphi(x) = d, \mu(x) = 1$, and $p(x) = 1 + \alpha x^2 + \beta x^4$. Consider the Abelian integral

$$L(x) = \int_c^x \frac{du}{\sqrt[3]{p(u)}}.$$

Note that $L(x)$ is not an elliptic integral because the integrand contains the cubic root of a polynomial $p(x)$. In this case

$$\psi(x) = f(x)^3 - \varphi(x)^3 p(x) \tag{96}$$

which is a polynomial of degree 4. In this case, $K = 4$ and $T = 2$ so that (71) is fulfilled. From (73), it follows that the roots must satisfy the equation

$$b + x_j = d\sqrt[3]{p(x_j)} \tag{97}$$

for $j = 1, 2, 3, 4$. From (97), it follows that the coefficients b and d are determined by

$$b = b(x_1, x_2) = \frac{x_1\sqrt[3]{p(x_2)} - x_2\sqrt[3]{p(x_1)}}{\sqrt[3]{p(x_1)} - \sqrt[3]{p(x_2)}} \text{ and } d = d(x_1, x_2)$$
$$= \frac{x_1 - x_2}{\sqrt[3]{p(x_1)} - \sqrt[3]{p(x_2)}}. \tag{98}$$

Moreover, the constant term in the polynomial $-\psi(x)/\beta d^3$ turns out to be equal to $(d(x_1, x_2)^3 - b(x_1, x_2)^3)/\beta d(x_1, x_2)^3$ and the coefficient associated with the x^3 is equal to $-1/\beta d(x_1, x_2)^3$.

By utilizing the relationship between the roots and the coefficients of a polynomial, it follows that $x_3 + x_4 = \frac{1}{\beta d(x_1,x_2)^3} - x_1 - x_2$ and $x_3 x_4 = \frac{d(x_1,x_2)^3 - b(x_1,x_2)^3}{x_1 x_2 \beta d(x_1,x_2)^3}$ which implies that x_3 and x_4 are functions of (x_1, x_2), and they are determined by the quadratic equation

$$x_j^2 + \left(x_1 + x_2 - 1/\beta d(x_1, x_2)^3\right)x_j - \frac{b(x_1, x_2)^3 - d(x_1, x_2)^3}{x_1 x_2 \beta d(x_1, x_2)^3} = 0 \tag{99}$$

for $j = 3, 4$, where b and d are given in (98). Consequently, the following addition theorem

$$\sum_{j=1}^{4} L(x_j) = C$$

has been established where C is a constant and the roots x_1 and x_2 can be selected freely whereas the remaining two roots x_3 and x_4 are determined by (99). Consider next the special case where $x_1 = x_2 = x$. To examine this case, it is convenient to let $x_1 \to x_2$ in the formulas in (98) and then apply l'Hopital's rule. Hence $b(x,x) = \frac{3+\alpha x^2-\beta x^4}{2\alpha x+4\beta x^3}$ and $d(x,x) = \frac{3p(x)^{2/3}}{2\alpha x+4\beta x^3}$.

Furthermore, (99) implies that x_3 and x_4 are given by the roots of the quadratic polynomial

$$x_j^2 + \left(2x - 1/\beta d(x,x)^3\right)x_j - \frac{b(x,x)^3 - d(x,x)^3}{x^2\beta d(x,x)^3} = 0$$

for $j = 3, 4$.

Example 4.5.5
From Corollary 3 with $m = 2$ and $m = 3$, respectively, it follows that

$$H(y_2) = 2H(x) \text{ and } H(y_3) = 3H(x)$$

where

$$y_2 = \frac{2x\Delta(x)}{1-k^2x^4} \text{ and } y_3 = \frac{3x - \left(4+4k^2\right)x^2 + 6k^2x^5 - k^4x^9}{1 - 6k^2x^4 + 4k^2\left(1+k^2\right)x^6 - 3k^4x^8}.$$

5 A Short Review of Abel's Work on Elliptic Integrals and the Addition Theorem

Several biographers have written about Niels Henrik Abel: see, for example, Ore (1957), Stubhaug (2000), and Sørensen (2004). Sørensen (2004) also provides an overview of the state of mathematics at the turn of the eighteenth century. Here, a brief review of selected parts of Abel's life and work is provided which have relevance to his contribution to the theory of elliptic integrals and elliptic functions.

Abel's life was short and dramatic. He was born in 1802 near the town of Stavanger in Norway. In 1804, his father was appointed pastor at Gjerstad, near the town of Risør, and the family moved there. In 1815, he entered the Cathedral School in Christiania at the age of 13. (Oslo was called Christiania at that time). A new mathematics teacher, Bernt Michael Holmboe (1795–1850), was appointed in 1818 and he set the students' mathematical tasks to do at home. He quickly became aware of Abel's talent and encouraged him to study mathematics at an advanced level. Abel entered the university in Christiania in 1821. At that time, he was already the foremost mathematician in Norway. He had studied all the latest mathematical literature in the university library and had also started work on his own mathematical

problems, such as the solution to the general quintic equation in radicals. In a letter dated 21 May 1821 from Carl Ferdinand Degen (1766–1825), professor of mathematics at the University of Copenhagen to Christopher Hansteen (1784–1873), professor of astronomy and geophysics at the University of Christiania, the content of which was made known to Abel, Degen wrote that such a talented young manas Abel should devote himself to the theory of "elliptic transcendents" (elliptic integrals), adding that he would "discover Magellanian thoroughfares to large portions of a vast analytical ocean." This almost prophetic statement might well have been influential for the direction of Abel's work.

From 1823 to 1826, Abel wrote three papers where one was published in 1826 in *Magazin for Naturvidenskaberne* (Journal of the Natural Sciences), Norway's first scientific journal, which had been co-founded by Hansteen at the University of Christiania. This paper was later translated into French and published in Abel (2012a). The two other papers where not published before Abel's death (Abel 2012b, c). In the summer of 1823, he was awarded a scholarship that allowed him to make his first trip abroad, to Copenhagen. From the letters he wrote to his former teacher Holmboe, it is clear that he had started to work on the theory of elliptic integrals and had already achieved new results. It would appear that the article "Petite contribution à la théorie de quelques fonctions transcendantes"(Abel 2012a) is the first article he published that, among other topics, concerns elliptic and hyperelliptic integrals. In a letter to Degen dated 2 March 1824, Abel writes about the progress he has made since the last time they met in Copenhagen, in the summer of 1823. From this letter and a paper that Abel wrote a short time afterwards, it seems that he had already by that time formulated the most general version of his addition theorem (Skau 2011).

In his article "Petite contribution à la théorie de quelques fonctions transcendantes" (2012a), Abel starts by discussing the integral

$$\int \frac{\varphi(x)\exp(\ f(x))dx}{x-a}$$

where a is a constant, $f(x)$ is a rational function, and $\varphi(x)$ has the form

$$\varphi(x) = \kappa(x+\alpha_1)^{\beta_1}(x+\alpha_2)^{\beta_2}\cdot\cdot(x+\alpha_r)^{\beta_r},$$

where $\alpha_1, \alpha_2, \ldots, \beta_1, \beta_2, \ldots, \boldsymbol{\kappa}$, are given coefficients and neither of the functions f and φ depends on a. By ingenious manipulations, including suitable operations (differentiation and integration), Abel obtains a set of surprising equations that can be applied to produce a number of amazing results. For example, when $p(x) = \gamma_0 + \gamma_1 x + \cdots + \gamma_s x^s$, where $\gamma_1, \gamma_2, \ldots$ are given coefficients and s is an integer, he obtains as a special case

$$\int \frac{dx}{(x-a)\sqrt{p(x)}} = \frac{1}{2\sqrt{p(a)}}\sum_n\sum_m (m-n)\gamma_{n+m+2}\int \frac{a^m da}{\sqrt{p(a)}}\cdot\int \frac{x^n dx}{\sqrt{p(x)}} \qquad (100)$$

for s greater than 2 and x being a root of p. Consider, for example, the case where $p(x) = (1 - x^2)(1 - k^2x^2) = \Delta(x)^2$. Then $\gamma_0 = 1$, $\gamma_1 = \gamma_3 = 0$, $\gamma_2 = -1 - k^2$ and $\gamma_4 = k^2$, which by (100) yields

$$\int \frac{dx}{(x-a)\Delta(x)} = \frac{k^2}{2\Delta(a)} \int \frac{da}{\Delta(a)} \cdot \int \frac{x^2 dx}{\Delta(x)} - \frac{k^2}{2\Delta(a)} \int \frac{a^2 da}{\Delta(a)} \cdot \int \frac{dx}{\Delta(x)}. \tag{101}$$

Since

$$\int \frac{dx}{(x^2 - a^2)\Delta(x)} = \int \frac{dx}{2a(x-a)\Delta(x)} - \int \frac{dx}{2a(x+a)\Delta(x)},$$

it follows immediately from (101) (by successively changing the sign of a) that

$$\int \frac{dx}{(x^2 - a^2)\Delta(x)} = \frac{k^2}{2a\Delta(a)} \int \frac{da}{\Delta(a)} \cdot \int \frac{x^2 dx}{\Delta(x)} - \frac{k^2}{2a\Delta(a)} \int \frac{a^2 da}{\Delta(a)} \cdot \int \frac{dx}{\Delta(x)}. \tag{102}$$

The equation in (102) shows that Legendre's elliptic integral of the third kind can be expressed in a simple way by elliptic integrals of the first and second kind when x is a root of p.

In 1825, Abel was awarded a second scholarship to travel abroad. The plan was to visit Carl Friedrich Gauss (1777–1855) in Gottingen and then continue on to Paris, which at that time was the leading center for mathematics. He initially intended to travel with four of his friends. However, when he got as far as Copenhagen, he changed his mind, following his friends to Berlin instead, having decided to visit Gauss and go on to Paris afterwards. For some unknown reason, though, he never met with Gauss. On the way, he visited the astronomer Heinrich Christian Schumacher (1780–1850) in Altona, now a district of Hamburg. Schumacher was at that time the editor of *Astronomische Nachrichten*, where Abel and Jacobi published important papers. He then spent 4 months in Berlin, where he became well acquainted with August Leopold Crelle (1780–1885). Crelle was about to establish a new mathematical journal, *Journal für die reine und angewandte Mathematik*, which would become very important for the dissemination of Abel's work. Crelle also became a close personal friend. In the journal's first year, Abel contributed seven articles. However, he saved what he regarded as his most important work, the Paris memoir, for the French Academy of Sciences.

Abel revolutionized the theory of elliptic integrals by proposing to study elliptic functions instead of elliptic integrals. The elliptic functions may be defined as the inverse of corresponding elliptic integrals. More precisely, let

$$H(x,k) = \int_0^x \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}} \tag{103}$$

denote the elliptic integral of the first kind with k as parameter. Then the corresponding elliptic function, say $g(y, k)$, is defined as the inverse of the integral $H(x, k)$: that is $H(g(y, k), k) = y$ for $y \in [0, 1]$. The elliptic functions can be viewed as a generalization of the trigonometric functions. To realize this, note that when $k = 0$, then $H(x, 0)$ reduces to arcsin$x$. The so-called Jacobian elliptic functions were introduced by Jacobi and the notation *sny*, *cny*, *dny*, and *tny* proposed by Christoph Gudermann (1798–1852) is nowadays common for elliptic functions. Specifically, $sny = g(y, k)$, $cn^2y = 1 - sn^2y$, $dn^2y = 1 - k^2sn^2y$, and $tny = sny/cny$. When $k = 0$, *sny*, cny, and *tny* reduce to the trigonometric functions sin$y$, cos$y$, and tg$y$, respectively. An addition theorem for elliptic functions that is dual to Euler and Legendre's addition theorem for elliptic integrals can be expressed as

$$sn(y + v) = \frac{snycnvdnv + snvcnydny}{1 - k^2sn^2ysn^2v}.$$

Note that $sn(y + v)$ reduces to the usual addition formula for $\sin(y + v)$ when $k = 0$.

Another area in which fierce competition between Abel and Jacobi took place was the theory of transformations. This theory includes the study of solutions (transformations) $\omega(x)$, say, satisfying

$$H(\omega(x), k) = bH(x, k') + C$$

where H is given in (103), k and k' are the respective parameters, which may be equal or different, and b and C are constants. The transformation omega(x) will of course depend on b, k, and k', but this is suppressed in the notation here. Note that Corollaries 2 and 3 can be viewed as part of the theory of transformations in the case where $k = k'$. See Stubhaug (2000) and Sørensen (2004) for more details about the rivalry between Abel and Jacobi.

In Paris, Abel continued his work on his memoir, which he finished in October 1826 and submitted to the academy on 30 October. It was to be reviewed by Augustin-Louis Cauchy (1789–1857) and Legendre. Unfortunately, the Paris memoir, regarded as Abel's masterpiece, was put aside among Cauchy's papers and forgotten. In early 1827, Abel returned to Berlin, where he stayed several months before returning to Norway. In Berlin, he completed an extensive paper which was published in two parts (Abel 1827, 1828a). In it he does not, however, apply his addition theorem but rather the one established by Euler and Legendre. In this paper, Abel develops the theory of elliptic functions and obtains many new results, and he also discusses the theory of transformations and complex multiplication. He also considered ways to solve for *sny* when $sn(my)$ is given, where m is an integer. As an example of multiplication of elliptic functions, one can show that

$$sn(3y) = \frac{3sny - \left(4 + 4k^2\right)sn^2y + 6k^2sn^5y - k^4sn^9y}{1 - 6k^2sn^4y + 4k^2\left(1 + k^2\right)sn^6y - 3k^4sn^8y}.$$

The continuation of the article "Recherches sur les fonctions elliptiques" (Abel 1827) was finished and published in the following year (Abel 1828a). By May 1827, Abel was back in Norway. Jacobi wrote a paper on the theory of transformations of elliptic integrals in which he also studies elliptic functions (Jacobi 1827). Although Abel (1827) had discussed the general theory of transformation and the properties of elliptic functions, and his paper had been published several months before Jacobi sent his paper to the journal, Jacobi (1827) did not publicly refer to the work of Abel. Therefore, Abel wrote a second paper and then a third one with special emphasis on the theory of transformations (Abel 1828b, 1828c). In 1828, he published a paper (Abel 1828d) in which he demonstrates his addition theorem for hyperelliptic integrals by applying the ideas from his Paris memoir (Abel 1841). Subsequently, he wrote another and very long paper (Abel 1829a) in which he presented his addition theorem, the elliptic functions, and the theory of transformations in a way that differed from his earlier work. Due to a combination of unfortunate circumstances (Stubhaug 2000), the Paris memoir was not published before 1841, 12 years after Abel's early death from tuberculosis. In the last paper Abel wrote before he died (Abel 1829b), on 6 April 1829, he gave a very compact proof of the addition theorem in its most general form (presented in Sect. 4.3), upon which his Paris memoir is based. It consists of only two pages, the first containing the statement of the theorem and the second containing the proof.

The French mathematician Emile Picard (1856–1941) wrote the following about Abel's theorem at the end of the nineteenth century: "*Le théorème parait tout à fait élémentaire, et il n'y a peut-etre pas, dans l 'histoire de la Science, de proposition aussi importante obtenue á l 'aide de considérations aussi simples*" ("The theorem appears utterly elementary, and perhaps there has never been in the history of science a proposition so important which can be reached by such simple considerations.") The Norwegian mathematician Atle Selberg (1917–2007) wrote in 2002: "*Det har alltid stått for meg som den rene magi. Hverken Gauss eller Riemann, eller noen annen, har noe som riktig kan måle seg meddette.*" ("For me this has always appeared as pure magic. Neither Gauss nor Riemann nor anyone else have anything that really measures up to this.")

After Abel's death, Jacobi's book on elliptic integrals and functions (Jacobi 1829) was published. It became a key reference for mathematicians working on the theory of elliptic integrals and functions in the following years. However, it was later criticized for making insufficient reference to Abel's work.

After Abel had published some of his groundbreaking articles, it became known that Gauss had also been working on the theory of elliptic integrals and functions and that as early as in 1798 he had achieved some of the same results as Abel (Sørensen 2004). However, Gauss never published his results and thus the mathematical community had no knowledge of his work before the publication of Abel's articles.

Whereas Abel's theory solves the problem of defining and analyzing the properties of elliptic functions (and more generally abelian functions) with elementary means, it does not immediately extend to the complex field, such as in the case where the polynomial $p(x)$ in Theorems 1 and 2 is negative, has complex roots or has

complex coefficients. Still there is no essential difference between the functions which correspond to real and to complex values of the coefficients and constants. It became the fate of Karl Theodor Wilhelm Weierstrass (1815–1897) to remedy this deficiency and provide a rigorous foundation of Abel's theory, outlined in Weierstrass (1856). Unfortunately, the theory then loses its elementary character.

6 Conclusion

To learn about ideas and approaches in mathematics is a highly demanding task for the nonspecialist. Often the background knowledge necessary of an understanding beyond a fragmentary and superficial level is formidable. However, as regard the development of the theory of elliptic integrals and functions some 200 years ago, it is possible to understand key results, including proofs, solely with background in elementary calculus, as demonstrated in this chapter. As noted above, Abel's approach was fundamentally different from earlier contributions, and, as has been mentioned by many mathematicians in the past, it is astonishing that such important results can be achieved by such simple means. Abel's approach is so powerful that it extends to the case hyperelliptic integrals of any degree and more generally to Abelian integrals and functions. Abel's work in this field inspired a great number of mathematicians in the years that followed and eventually led to Riemann and Weierstrass developing the theory of algebraic functions and algebraic curves further.

References

Abel NH (1827) Recherches sur les fonctions elliptiques. J Reine Angew Math 2:101–181

Abel NH (1828a) Recherches sur les fonctions elliptiques (second part). J Reine Angew Math 3: 160–190

Abel NH (1828b) Solution d'un problème général concernant la transformations des fonctions elliptiques. Astron Nachr 6:365–380

Abel NH (1828c) Addition au mémoire sur les fonctions elliptiques. Astron Nactr 7:33–34

Abel NH (1828d) Remarques sur quelques propriété générales d'une certaine sorte de fonctions transcendantes. J Reine Angew Math 3:313–323

Abel NH (1829a) Précis d'une théorie des fonctions elliptiques. J Reine Angew Math 4:236–277, 309–348

Abel NH (1829b) Démonstration d'une propriété générale d'une certaine classe de fonctions transcendantes. J Reine Angew Math 4:200–201

Abel NH (1841) Mémoire sur une propriété générale d'une classe très étendue de fonctions transcendantes. Mémoires présentés par divers savants, vol VII, Paris

Abel N H (2012a) Petite contribution à la théorie de quelques fonctions transcendantes. In: Sylow L, Lie S (eds) Oeuvres complètes de Niels Henrik Abel, vol I. Cambridge University Press, London (First published in Norwegian (1824–7): Et lidet bidrag til læren om adskillige transcendente functioner. Det Kgl. Norske Videnskabers Selskap i Throndhjem, 1824–1827, 177–207

Abel NH (2012b) Sur une propriété remarquable d'une classe très étendue de fonctions transcendantes. In: Sylow L, Lie S (eds) Oeuvres complètes de Niels Henrik Abel, vol II. Cambridge University Press, London

Abel NH (2012c) Extension de la théorie précédente. In: Sylow L, Lie S (eds) Oeuvres complètes de Niels Henrik Abel, vol II. Cambridge University Press, London

Ayoub R (1984) The lemniscate and Fagnano's contributions to elliptic integrals. Arch Hist Exact Sci 29:131–149

Cayley A (1961) An elementary treatise on elliptic functions. Constable and Company Ltd., London. (Originally published by Bells G & Sons 1876)

Euler L (1761) On the integration of the differential equation) $mdx/(1-x^4)^{1/2} = ndy/(1-y^4)^{1/2}$. Article 251. In: Enestrom index (*Opera Omnia* Z_{20}, 58–79). (Originally published as: De integrationea equation is differential is In: Novi commentarii academiae scientiarum Petropolitanae 6: 37–57)

Fagnano GC (1718) Metodo per misurare la lemniscata (Method for measuring the lemniscate). Opere matematiche 2:293–313

Huntington Barnum H (1910) Abel's theorem and the addition formulae for elliptic integrals. Ann Math Second Ser 11:103–114

Hoffmann JE (1949) Die entwicklungsgeschichte der Leibnizschen mathematikwährend des Aufenthaltes in Paris (1672–1676). Leibniz Verlag, Munich

Houzel C (1986) Fonctions elliptiques et intégrales abeliennes. In: Dieudonne J (ed) Abrégé d'histoire des mathématiques, Ch. 7. Hermann, Paris

Houzel C (2004) The work of Niels Henrik Abel. In: Laudal AO, Piene R (eds) The legacy of Niels Henrik Abel – The Abel bicentennial, Oslo, 2002. Springer, Berlin, pp 21–178

Jacobi CGJ (1827) Demonstratio theorematisad theoriam functionumellipticarum spectantis. Astron Nachr 6:133–141

Jacobi CGJ (1829) Fundamenta nova theoriae functionumeli picarum. Borntreger Regiomonti. Also published in C.G.J. Jacobi's Gesammelte Werke (1881), vol I Remer G. Berlin

Lagrange JL (1766–9) Sur quelques équations différentielles dont les indéterminées sont séparés mais dont chaque member en particulier nést point intégrable. In: Serret JA (ed) Oeuves de Lagrange Vol. II, 1867–92. Gauthiers-Villars:Paris. First published in Miscellanea Taurinensia, IV. Mémoires de l 'Académie Royale de Sciences des Turins, Turin

Landen J (1775) An investigation of a general theorem for finding the length of an arc of any conic hyperbola, by means of two elliptic arcs, with some other new and useful theorems deduced therefrom. Philos Trans 65: 283–289, The Royal Society of London. Also published in Mathematical Memoirs (1780), Nourse J: Bookseller to His Majesty, London

Legendre AM (1793) Mémoire sur les transcendantes elliptiques. Du Pont et Firmin Didot, Paris

Legendre AM (1811–17) Exercices de calcul intégral sur divers orders de transcendantes, et sur les quadratures (1811, 1817, 1816). Courcier, Paris

Legendre AM (1825–8) Traité des fonctions elliptiques et des intégrales eulèriens (1825, 1826, 1828). Huzard-Courcier, Paris.

Miller CD, Lial ML, Schneider DL (1990) Fundamentals of college algebra, 3rd edn. New York, Addison-Wesley Educational Publishers

Mittag-Leffler G (1923) An introduction to the theory of elliptic functions. Ann Math 24:271–351

Ore Ø (1957) Niels Henrik Abel. Mathematician extraordinary. Minneapolis, University of Minnesota Press

Richelot FJ (1842) Über die integration eines merkwürdigen systems differentialgleichungen. J Reine Angew Math 23:354–369

Skau C (2011) Om Niels Henrik Abel (1802–1829) sin avhandling, 'Et lidet Bidrag til Læren om adskillige transcendente Functioner.' Det Kgl. Norske Videnskabers Selskaps skrifter, 90-100. In Norwegian (On the article by Niels Henrik Abel (1802–1829), 'Petite contribution a la théorie de quelques fonctions transcendantes')

Stephenson RJ (1960) Mechanics and properties of matter. Wiley, London
Stubhaug A (2000) Niels Henrik Abel and his times. Called too soon by flames afar. Berlin, Springer
Sørensen HK (2004) The mathematics of Niels Henrik Abel. Ph.D. thesis, Institute for Philosophy of Science, University of Arhus
Weierstrass KTW (1856) Theorie der Abel'schen functionen. J Reine Angew Math 52:285–380

The Absolute Arithmetic Continuum and Its Geometric Counterpart

Philip Ehrlich

Contents

1 Introduction 1678
2 The Surreal Number Tree: A Prelude 1681
3 Lexicographically Ordered Binary Trees: Preliminary Definitions 1682
4 The Construction of the Ordered Surreal Number Tree $\langle No, <, <_s \rangle$ 1684
4.1 Construction of Games 1685
4.2 Preliminary Definitions 1685
4.3 Construction of Surreal Numbers 1685
4.4 Isolation of the Ordinals 1686
4.5 The Rule of Order 1686
5 The Structure of the Ordered Surreal Number Tree $\langle No, <, <_s \rangle$ 1687
6 Ordered Fields with Simplicity Hierarchies 1689
7 S-Hierarchical Mappings and S-Hierarchical Morphisms 1692
8 The S-Hierarchical Ordered Field of Real Numbers 1695
9 Every Surreal Number Has its Own Proper Name 1696
10 The Ordered Field of Surreal Numbers 1700
11 Arithmetic Continua: Archimedean and Absolute 1703
12 S-Hierarchical Arithmetic Continua: Archimedean and Absolute 1705
13 Euclidean Geometric Continua: Archimedean and Absolute 1707
14 Bridging the Gap between the Domains of Discreteness and of Continuity 1712
15 Postscript 1714
References 1715

Abstract

In a number of works, we have suggested that whereas the ordered field **R** of real numbers should merely be regarded as constituting an *arithmetic continuum (modulo the Archimedean axiom)*, the ordered field **No** of *surreal numbers* may be regarded as a sort of *absolute arithmetic continuum (modulo NBG)*. In the present chapter, as part of a more general exposition of the absolute arithmetic

P. Ehrlich (✉)
Ohio University, Athens, OH, USA
e-mail: ehrlich@ohio.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_54

continuum, we will outline some of the properties of the system of surreal numbers that we believe lend credence to this mathematico-philosophical thesis. We will also provide an overview of **No**'s rich structure as a *simplicity-hierarchical (*or *s-hierarchical) ordered field* that recursively emerges from the interplay between its structure as an ordered field and its structure as a *lexicographically ordered full binary tree*. Finally, we will draw attention to how properties of the system of surreal numbers considered as an s-hierarchical ordered algebraic structure can be appealed to in conjunction with classical relations between ordered algebraic and geometric systems to help resolve, for the surreal case, one of the purported difficulties that lies at the heart of attempts to bridge the gap between the domains of number and of geometrical magnitude. In particular, we will explain how it is possible that despite the fact that every surreal number, considered as a member of an s-hierarchical ordered field, differs from every other in characteristic individual properties, the *absolute (Euclidean) geometrical continuum*, which is modeled by the Cartesian space over the ordered field **No** of surreal numbers, appears as an amorphous pulp of points that display little individuality.

Keywords

Absolute Arithmetic Continuum · Absolute Geometric Continua · Surreal Numbers · Non-Archimedean Continua · Non-Archimedean ordered algebraic systems

1 Introduction

Since the discovery that a diagonal of a square is incommensurable with its sides, the question of the possibility of bridging the gap between the domains of discreteness and of continuity, or between arithmetic and geometry, has been of fundamental concern to mathematicians, logicians, and philosophers alike. Indeed, as Abraham Fraenkel, Yehoshua Bar-Hillel, and Azriel Levy aptly observed:

> Bridging the gap between these two heterogeneous domains is not only the central but also the oldest problem in the foundations of mathematics and in the related philosophical fields. (1973, p. 212)

Moreover, as they go on to say:

> To understand the nature of the problem one should stress the fundamental difference between the *discrete, qualitative, individual* nature of *number* in the "combinatorial" domain of counting (arithmetic) and the *continuous, quantitative, homogeneous* nature of the points of *space* (or of time) in the "analytical" domain of *measuring* (geometry). Every integer differs from every other in characteristic individual properties comparable to the differences between human beings, while the continuum appears as an amorphous pulp of points which display little individuality. (1973, p. 212)

Cantor and Dedekind believed they had bridged the gap with the creation of their arithmetico-set theoretic continuum **R** of real numbers, and for more than a century now it has been one of the central tenets of standard mathematical philosophy that indeed they had. In accordance with this view, the geometric linear continuum is taken to be isomorphic with the *arithmetic continuum*, the axioms of geometry being so selected to ensure this would be the case. Given the Archimedean nature of **R** once this mathematico-philosophical thesis is adopted, we have the classic result of standard mathematical philosophy that infinitesimals are superfluous to the analysis of the structure of a continuous straight line.

During the past few decades, however, we have suggested that (Ehrlich 1987, 1989b, 1992, 1994, p. xxviii; 2002a, 2005, p. 508; 2010, 2012, 2021) whereas the Cantor-Dedekind theory succeeds in bridging the gap between the domains of arithmetic and of classical Euclidean geometry, it only reveals a glimpse of a far richer theory of continua that not only allows for infinitesimals, but also leads to a vast generalization of portions of Cantor's theory of the infinite, a generalization that also provides a setting for Abraham Robinson's nonstandard approach to analysis (1961, 1966) as well as for the important and all too often overlooked non-Cantorian theories of the infinite (and infinitesimal) pioneered by Giuseppe Veronese (1891, 1894), Tullio Levi-Civita (1892–3, 1898), David Hilbert (1899), and Hans Hahn (1907; also see Ehrlich 1995) in connection with their work on non-Archimedean ordered algebraic and geometric systems, and by Paul du Bois-Reymond (e.g., 1870–1871, 1875, 1877, 1882), Otto Stolz (1883, 1885), Felix Hausdorff (1907, 1909), and G. H. Hardy (1910, 1912) in connection with their work on the rates of growth of real functions, theories which have come to play fundamental roles in important segments of contemporary non-Archimedean mathematics. Central to the theory is J. H. Conway's ordered field of *surreal numbers* (1976), a system of numbers containing the reals and the ordinals as well as a great many less familiar numbers including $-\omega$, $\omega/2$, $1/\omega$, $\sqrt{\omega}$, e^{ω}, $\log\omega$, and $\sin(1/\omega)$ to name only a few, where ω is the least infinite ordinal. Indeed, this particular number system, which Conway calls **No**, is so remarkably inclusive that, subject to the proviso that numbers – construed here as members of ordered fields – be individually definable in terms of sets of von Neumann-Bernays-Gödel set theory with Global Choice (NBG), it may be said to contain "All Numbers Great and Small" (Conway 1976, p. 3).

In addition to its inclusive structure as an ordered field, **No** has a rich algebraico-tree-theoretic structure which was brought to the fore by Ehrlich (1994, 2001, 2002a, b) and further explored in (van den Dries and Ehrlich 2001; Ehrlich 2011, 2012; Ehrlich and Kaplan 2018, 2021; Fornasiero 2006, 2016; also see Berarducci and Mantova 2018; Kuhlmann and Matusinski 2015). This *simplicity-hierarchical* (or *s-hierarchical*) *structure* depends upon **No**'s structure as a *lexicographically ordered full binary tree* and arises from the fact that the sums and products of any two members of the tree are the simplest possible elements of the tree consistent with **No**'s recursively defined structure as an ordered group and an ordered field, respectively, it being understood that x is *simpler than* y (written $x <_s y$) just in case x is a predecessor of y in the tree.

Among the striking *s*-hierarchical features of **No** is that every surreal number can be assigned a canonical "proper name" that is a reflection of its characteristic *s*-hierarchical properties. These *Conway names* or *normal forms* are expressed as formal sums of the form

$$\sum_{\alpha<\beta} r_\alpha \omega^{y_\alpha},$$

where β is an ordinal, $(y_\alpha)_{\alpha<\beta}$ is a strictly decreasing sequence of surreals, and $(r_\alpha)_{\alpha<\beta}$ is a sequence of nonzero real numbers, the Conway name of an ordinal being just its Cantor normal form.

Figure 1 offers a glimpse of some of the early stages of the recursive unfolding of the members of the s-hierarchical ordered field of surreal numbers in terms of their Conway names, where ω is the least infinite ordinal as well as *the simplest positive infinite number*, $-\omega$ is the additive inverse of ω as well as *the simplest negative infinite number*, $1/\omega$ is the multiplicative inverse of ω as well as *the simplest positive infinitesimal number*, and so on. A complete recursive unfolding of the surreals in terms of their Conway names is given in (Ehrlich 2011).

In a number of works (*op. cit.*), we have suggested that whereas the ordered field of real numbers should merely be regarded as constituting an arithmetic continuum modulo the Archimedean axiom (see Sect. 11), the ordered field of surreal numbers may be regarded as a sort of *absolute arithmetic continuum* (modulo NBG). In the present chapter, as part of a more general exposition of the purported absolute arithmetic continuum and its s-hierarchical counterpart, we will outline some of

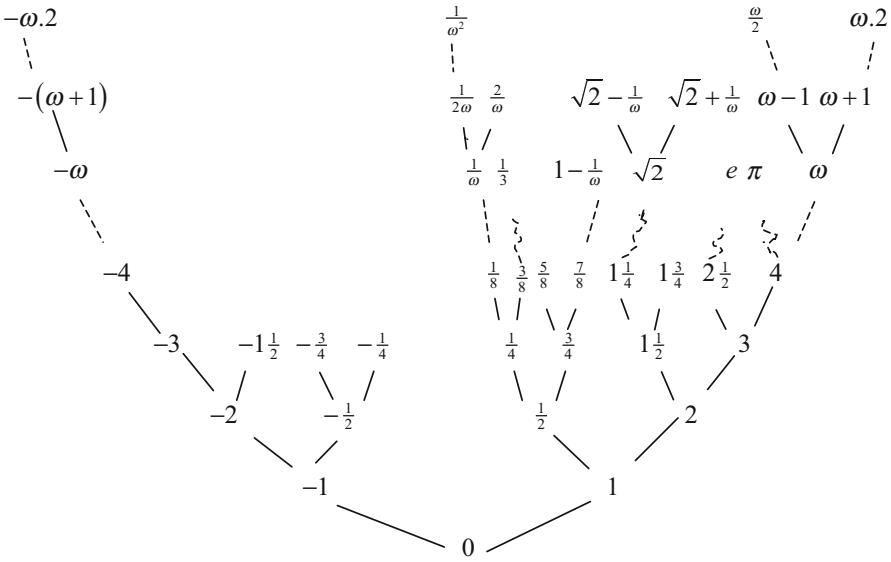


Fig. 1 Some of the early stages of the recursive unfolding of **No**

the properties of the system of surreal numbers that we believe lend credence to this mathematico-philosophical thesis. In doing so, we hope to provide a gateway to some of the leading ideas of the theory. Following the above, we will draw attention to how properties of the system of surreal numbers considered as an s-hierarchical ordered algebraic structure can be appealed to in conjunction with classical relations between ordered algebraic and geometric systems to help resolve, for the surreal case, one of the purported difficulties that lies at the heart of attempts to bridge the gap between the domains of number and of geometrical magnitude. That is, inspired by the challenge implicit in the remarks of Fraenkel, Bar-Hillel, and Levy that don the title page of our text, we will attempt to explain how it is possible that despite the fact that every surreal number differs from every other in characteristic individual properties comparable to the differences between human beings, *the absolute (Euclidean) geometrical continuum*, which is modeled by the *Cartesian space over the ordered field of surreal numbers*, appears as an amorphous pulp of points which display little individuality. Among the results that are central to our resolution is *despite the fact that* **No** *considered as an s-hierarchical ordered algebraic structure admits only the identity automorphism, the Cartesian space over the ordered field of surreal numbers admits a vast array of nontrivial automorphisms in which any given point may be associated with any other given point.* While our discussion will be restricted to plane Euclidean geometry, our analysis generalizes (with suitable modifications) to n-dimensional Euclidean, hyperbolic, and elliptic geometry for all $n \geq 2$ (Ehrlich 1987).

Throughout the chapter, the underlying set theory is assumed to be NBG, where a sentence in the language of sets is true if and only if it is true in ZFC and where all proper classes (i.e., classes that are not sets) are in bijective correspondence with the class of all ordinals (e.g., Mendelson 2010). Since in the language of sets, the sets of NBG have precisely the same properties as the sets of ZFC, the above assertion that **No** may be regarded as containing all numbers great and small that are individually definable in terms of the sets of NBG may be understood as asserting that **No** may be regarded as containing all numbers great and small that are individually definable in terms of sets of standard set theory.

2 The Surreal Number Tree: A Prelude

In Conway's development of the theory of surreal numbers, the system of surreal numbers has the structure of an ordered field, the field operations like the surreal numbers themselves being defined by transfinite recursion. Conway shows that the recursively defined ordered class of surreal numbers can be given the structure of a lexicographically ordered full binary tree, but this structure plays a limited role in his treatment. What does play a more significant role in Conway's development is a *birthday* function that maps each surreal number to the level of recursion at which it is created (along with a related though weaker notion of simplicity that is applicable to Conway's more general theory of games). The birthday of a surreal number in

Conway's treatment corresponds to the tree-rank of a surreal number in the author's development of surreal numbers presented below.

As in several of the classical constructions of real numbers, surreal numbers for Conway are formally construed as equivalence classes. In accordance with Conway's construction, each surreal number is in fact an equivalence class consisting of an entire proper class of equivalent representations, which makes it impossible to collect the proper class of all surreal numbers thus construed into a legal class of NBG. If one wishes, one can carry out Conway's construction in the equiconsistent set theory of Ackermann (Ehrlich 1994, p. 246; 2006, p. 9). In NBG, on the other hand, the numbers must be "cut down to size"; this can be done most easily by identifying each number with a canonical representative of each of Conway's equivalence classes, or by identifying each number with a canonical subset of each class. For the purpose of formalization, Conway (1976, p. 65) proposes the latter approach, making use of the set-theoretic rank function. By contrast, we have introduced two alternative constructions that employ canonical representatives: one based on cuts due to Cuesta Dutari (1954) that generalize the familiar cuts of Dedekind (Ehrlich 1988; Alling and Ehrlich 1986, 1987), and the other being a generalization of the von Neumann ordinal construction (Ehrlich 1994, 2002a, b, 2012). In each case, the surreal numbers can be defined recursively or by means of an explicit definition (Ehrlich 1994, p. 242; also see Appendix I). One can also introduce the system of surreal numbers using *sign-expansions* (see below) of surreal numbers. However, this technically simpler approach, which was introduced by Conway (1976, p. 65), and made popular by Gonshor (1986), divorces the theory of surreal numbers from Conway's more general theory of games, and it presupposes the existence of the ordinals, something neither Conway's construction nor the generalization of the von Neumann construction do. In the present discussion, following a review of some of the basic definitions from the theory of lexicographically ordered binary trees, we will introduce the surreal numbers recursively via the the above-mentioned generalization of the von Neumann ordinal construction.

3 Lexicographically Ordered Binary Trees: Preliminary Definitions

A *tree* $\langle A,<_A\rangle$ is a partially ordered class such that for each $x \in A$, the class $\{y \in A : y <_A x\}$ of *predecessors* of x, written $\mathrm{pr}_A(x)$, is a set well ordered by $<_A$. A maximal subclass of A well ordered by $<_A$ is called a *branch* of the tree. Given any two distinct elements x and y of A, precisely one of the following is the case: Either x is a predecessor of y (i.e., $x <_A y$), y is a predecessor of x (i.e., $y <_A x$), or neither $x <_A y$ nor $y <_A x$. In the latter case, x is said to be *incomparable* with y. Thus, for distinct elements x and y of A, x is incomparable with y if and only if x and y lie on different branches of the tree. An *initial subtree* of $\langle A,<_A\rangle$ is a subclass A' of A with the order inherited from $<_A$ such that for each $x \in A'$, the set of predecessors of x in A' coincides with the set of predecessors of x in A. The *tree-rank* of $x \in A$, written

$\rho_A(x)$, is the ordinal corresponding to the well-ordered set of predecessors of x; the α *th level* of A is the set of all members of the tree having tree-rank α; and a *root* of A is a member of the zeroth level. If $x, y \in A$, then y is said to be an *immediate successor* of x if $x <_A y$ and $\rho_A(y) = \rho_A(x) + 1$, and if $(x_\alpha)_{\alpha<\beta}$ is a *chain* in A (i.e., a subclass of A well ordered by $<_A$), then y is said to be *an immediate successor of the chain* if $x_\alpha <_A y$ for all $\alpha < \beta$ and $\rho_A(y)$ is the least ordinal greater than the tree-ranks of the members of the chain. The *length of a chain* $(x_\alpha)_{\alpha<\beta}$ in A is the ordinal β. A tree $\langle A, <_A\rangle$ is said to be *binary* if each member of A has at most two immediate successors and every chain in A of limit length (including the empty chain) has at most one immediate successor. If every member of A has two immediate successors and every chain in A of limit length has an immediate successor, then the binary tree $\langle A, <_A\rangle$ is said to be *full*. Since a full binary tree has a level for each ordinal, the universe of a full binary tree is a proper class.

Using the Axiom of Global Choice (or simply the Axiom of Choice if A is a set), a tree may be shown to be binary if and only if it is isomorphic to an initial subtree of the *canonical full binary tree* $\langle B, <_B\rangle$, where B is the class of all sequences of +s and -s indexed over some ordinal and $x <_B y$ signifies that x is a proper initial subsequence of y (e.g., Drake 1974, p. 216).

By an ordered tree $\langle A, <, <_A\rangle$, we mean a tree $\langle A, <_A\rangle$ whose universe is totally ordered by $<$. Thus, an ordered tree $\langle A, <, <_A\rangle$ has two orderings: a partial ordering $<_A$ in virtue of which $\langle A, <_A\rangle$ is a tree and a total ordering $<$ in virtue of which $\langle A, <\rangle$ is an ordered class. As is well known, $\langle B, <_B\rangle$ can be totally ordered lexicographically in accordance with the definition: $(x_\alpha)_{\alpha<\mu} <_{lex(B)} (y_\alpha)_{\alpha<\sigma}$ if and only if $x_\beta = y_\beta$ for all $\beta <$ some δ, but $x_\delta < y_\delta$, it being understood that $-<$ *undefined* $<+$. The resulting structure $\langle B, <_{lex(B)}, <_B\rangle$ is called the *lexicographically ordered canonical full binary tree*.

Figures 2, 3, and 4 exhibit levels 0–3 of $\langle B, <_{lex(B)}, <_B\rangle$, the level of the tree being indicated in Fig. 2 on the right. In Fig. 3, the members of the tree that are connected to a given member by strictly descending sequences of dots are the predecessors of the given member. And in Fig. 4, the ordering of the positions of the projected sequences on the line indicates the total ordering of the members of levels 0–3 in accordance with the lexicographical ordering.

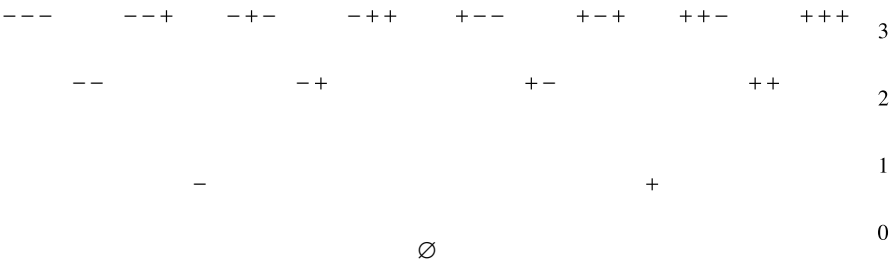


Fig. 2 Levels 0–3 of $\langle B, <_{lex(B)}, <_B\rangle$

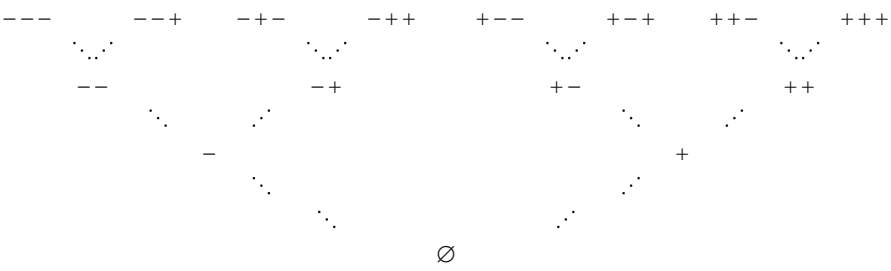


Fig. 3 The tree structure of levels 0–3 of $\langle B, <_{lex(B)}, <_B \rangle$

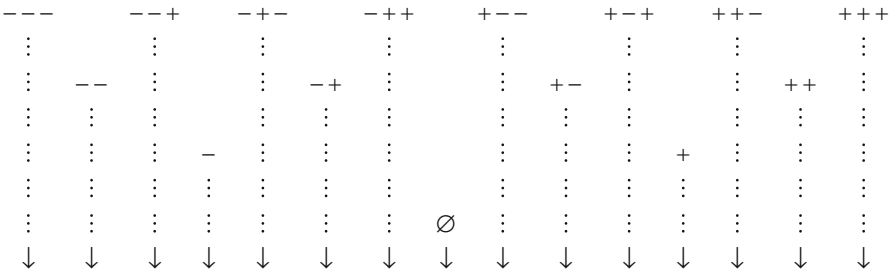


Fig. 4 The total ordering of levels 0–3 of $\langle B, <_{lex(B)}, <_B \rangle$

In a subsequent section, we will explore the relation between the ordered surreal number tree and $\langle B, <_{lex(B)}, <_B \rangle$. For the time being, we turn our attention to the ordered surreal number tree itself.

4 The Construction of the Ordered Surreal Number Tree $\langle$No, <, $<_s\rangle$

Von Neumann defines an ordinal as a transitive set that is well ordered by the membership relation. As a result, for von Neumann, an ordinal emerges as the set of all of its predecessors in the long though rather trivial binary tree $\langle Ord, \in \rangle$ of all ordinals. So, for example, 0 is identified with Ø, 1 is identified with $\{\emptyset\} = \{0\}$, 2 is identified with $\{\emptyset, \{\emptyset\}\} = \{0, 1\}$, and so on. In the following generalization of the von Neumann ordinal construction, each surreal number x emerges as an ordered pair (L_x, R_x) of sets of surreal numbers that are predecessors of x. Although L_x and R_x are defined independently of the total ordering that is imposed on the surreal number tree, they ultimately coincide with the sets of all predecessors of x that are less than x and greater than x, respectively. $\langle Ord, \in \rangle$ is not usually described as a binary tree since its familiar structure as a well-ordered class is indistinguishable from its structure as a tree. In the theory of surreal numbers, however, one must distinguish

between its binary tree structure, on the one hand, and its structure as a totally ordered class, on the other. The particular approach we will employ requires one to have a surreal number (L, R) at hand before it can be recusively identified as a surreal number. To ensure the availability of the requisite class of set-theoretical entities, we follow Conway in first recursively defining the class of *games*.

4.1 Construction of Games

If L and R are any two sets of games, then there is a game (L, R). All games are constructed in this way.

It is not difficult to see that this construction leads to an entire proper class of games beginning with the game $(\emptyset, \emptyset)$, where $\emptyset$ is the empty set (of games). The closure condition "All games are constructed in this way" simply ensures that nothing is a game that does not arise in the specified way. To recursively extract the surreal numbers from the class of games, we make use of the following terminology from (Ehrlich 2002a, b).

4.2 Preliminary Definitions

A game x is said to be *simpler than* a game $y = (L_y, R_y)$, written $x <_s y$, if $x \in L_y$ or $x \in R_y$; a chain of games (i.e., a class of games totally ordered by $<_s$) is said to be *ancestral* if it is closed under the simpler than relation, i.e., x is a member of the chain whenever y is a member of the chain and $x <_s y$; and by an *orderly partition* L, R of an ancestral chain C of games we mean an ordered pair (L, R) of disjoint subsets of C whose union is C and for which $L \supseteq L_x$ and $R_x \subseteq R$ for each element $x = (L_x, R_x)$ of the chain.

The members of the class **No** of surreal numbers are now identified by means of the following simple construction in which the clause "there is a surreal number (L, R)" signifies that the game (L, R) is a surreal number.

4.3 Construction of Surreal Numbers

If (L, R) is an orderly partition of an ancestral chain of surreal numbers, then (L, R) is a surreal number. All surreal numbers are constructed in this way.

At this point (employing much the same argument from (Ehrlich 1994, pp. 242–144)), it is not difficult to show that $\langle \mathbf{No}, <_s \rangle$ is a full binary tree where the definition of the simpler than relation for surreal numbers is inherited from the definition for games. For this purpose, however, one needs the ordinals. If one wishes, one could avail oneself of the von Neumann ordinals, which are already at hand. On the other hand, if one wants to rely solely on the ordinals developed within the theory of surreal numbers, one must first identify "our" ordinals.

4.4 Isolation of the Ordinals

A surreal number $x = (L_x, R_x)$ will be said to be an ordinal if $R_x = \emptyset$. By *On*, we mean the class of ordinals so defined. For all ordinals x and y, x will be said to be less than y, written $x <_{On} y$, if $L_x \subset L_y$.

In accordance with the above definition, $(\emptyset, \emptyset)$ is an ordinal, $(\{(\emptyset, \emptyset)\}, \emptyset)$ is an ordinal, $(\{(\emptyset, \emptyset), (\{(\emptyset, \emptyset)\}, \emptyset)\}, \emptyset)$ is an ordinal, and so on. In fact, the just-cited ordinals are eventually identified with the finite ordinals $0 = (\emptyset, \emptyset)$, $1 = (\{0\}, \emptyset)$, and $2 = (\{0, 1\}, \emptyset)$, respectively.

That the ordered class $\langle On, <_{On} \rangle$ of ordinals has all of the requisite properties possessed by any of the more familiar constructs so-called follows from

Propostion 1 *There is a one-to-one order preserving correspondence between the ordered class* $\langle On, <_{On} \rangle$ *and the ordered class* $\langle Ord, \in \rangle$ *of von Neumann ordinals.*

Using the ordered class of ordinals so defined, the structure of the surreal numbers that emerge vis-à-vis the above construction may be characterized as follows.

Theorem 1 (Ehrlich 1994) $(\emptyset, \emptyset)$ *is a surreal number; if* $x = (L_x, R_x)$ *is a surreal number, then* $(L_x, \{x\} \cup R_x)$ *and* $(L_x \cup \{x\}, R_x)$ *are surreal numbers; moreover, if* $(x_\alpha)_{\alpha < \beta}$ *is a chain of surreal numbers (ordered by* $<_s$*) of infinite limit length, then* $\left(\cup_{\alpha<\beta} L_{x_\alpha}, \cup_{\alpha<\beta} R_{x_\alpha}\right)$ *is a surreal number. Nothing is a surreal number except in virtue of the above.*

The reader will note that $(L_x, \{x\} \cup R_x)$ and $(L_x \cup \{x\}, R_x)$ are the immediate successors in $\langle \mathbf{No}, <_s \rangle$ of the surreal number x, and that $\left(\cup_{\alpha<\beta} L_{x_\alpha}, \cup_{\alpha<\beta} R_{x_\alpha}\right)$ is the immediate successor in $\langle \mathbf{No}, <_s \rangle$ of the chain $(x_\alpha)_{\alpha < \beta}$ of surreal numbers indexed over the infinite limit ordinal β. In fact, as it turns out, $(L_x, \{x\} \cup R_x)$ is the immediate successor of x less than x, $(L_x \cup \{x\}, R_x)$ is the immediate successor of x greater than x, and $\left(\cup_{\alpha<\beta} L_{x_\alpha}, \cup_{\alpha<\beta} R_{x_\alpha}\right)$ is always greater than the members of $\cup_{\alpha<\beta} L_{x_\alpha}$ and less than the members of $\cup_{\alpha<\beta} R_{x_\alpha}$. To lend precision to these ideas and to ideas regarding **No** considered as a totally ordered class more generally, we require the following definition.

4.5 The Rule of Order

For all surreal numbers $x = (L_x, R_x)$ and $y = (L_y, R_y)$, $x < y$ if and only if $x \in L_y$ or $y \in R_x$ or $R_x \cap L_y \neq \emptyset$.

The three cases that comprise the rule of order correspond to the three mutually exclusive and collectively exhaustive relations that two surreal numbers x and y may bear to one another when x is less than y; namely, x is simpler than y, y is simpler than x, and x is incomparable with y, respectively. Using this in conjunction with Theorem 1, one may prove

Corollary 1 (Ehrlich 1994) *For each surreal number x,*

$$x = \left(L_{s(x)}, R_{s(x)}\right),$$

where

$L_{s(x)} = \{a \in \mathbf{No} : a <_s x \text{ and } a < x\}$ and $R_{s(x)} = \{a \in \mathbf{No} : a <_s x \text{ and } x < a\}$.

Moreover, by appealing to Corollary 1 together with the definition of **No**'s ordinals and the definition of the corresponding ordering thereof, we also have

Corollary 2 (Ehrlich 1994) **No***'s ordinals are the members of the rightmost branch of* $\langle \mathbf{No}, <, <_s \rangle$*, the ordinal* α *being the member of tree-rank* α*. As such, for all ordinals* α *and* β*,* $\alpha <_{On} \beta$ *if and only if* $\alpha <_s \beta$ *if and only if* $\alpha < \beta$*. Moreover, if* α *is an ordinal, then*

$$\alpha = (L_{s(\alpha)}, \emptyset).$$

Remark The above construction of the surreals makes no use of the ordinals. Like Conway, we believe this is befitting a construction of "all numbers great and small." It is worth noting, however, that if one permits the use of the von Neumann ordinals in the construction, one can introduce $\langle \mathbf{No}, <, <_s \rangle$ more simply employing the following definition along with $<_s$ and $<$ as defined above.

Definition (Ehrlich 2010, p. 248) $(\emptyset, \emptyset)$ is a surreal number; if $x = (L_x, R_x)$ is a surreal number, then $(L_x, \{x\} \cup R_x)$ and $(L_x \cup \{x\}, R_x)$ are surreal numbers; moreover, if $(x_\alpha)_{\alpha<\beta}$ is a sequence of surreal numbers of infinite limit length β in which $x_\mu \in L_{x_\nu} \cup R_{x_\nu}$ whenever $\mu < \nu < \beta$, then $\left(\cup_{\alpha<\beta} L_{x_\alpha}, \cup_{\alpha<\beta} R_{x_\alpha}\right)$ is a surreal number. Nothing is a surreal number except in virtue of the above.

5 The Structure of the Ordered Surreal Number Tree $\langle \mathbf{No}, <, <_s \rangle$

For each surreal number x, there is a unique enumeration $(x_\alpha)_{\alpha<\rho_{\mathbf{No}}(x)}$ of the predecessors of x well ordered by the simpler than relation. Following (Ehrlich 1994), we refer to this enumeration as the *genealogy* of x. By the *sign-expansion* of x, we mean the sequence $(g_\alpha(x))_{\alpha<\rho_{\mathbf{No}}(x)}$ defined by the condition: for all $\alpha < \rho_{\mathbf{No}}(x)$,

$$g_\alpha(x) = \begin{cases} +, & \text{if } x_\alpha \in L_{s(x)} \\ -, & \text{if } x_\alpha \in R_{s(x),} \end{cases}$$

where x_α is the α *th* term in the genealogy of x (i.e., the predecessor of x having tree-rank α). This definition of the sign-expansion of a surreal number differs

considerably from Conway's (1976, pp. 29–30) which is defined using the ordered additive structure of **No**.

It is not difficult to show that every surreal number has a sign-expansion; distinct surreal numbers have distinct sign-expansions; and every sequence of elements of $\{+, -\}$ of length $\beta \in On$ is the sign-expansion of some surreal number (Ehrlich 1994, Theorem 1.4). Moreover, and more importantly, however, is

Theorem 2 (Ehrlich 1994) $\langle \mathbf{No}, <, <_s \rangle$ and $\langle B, <_{lex(B)}, <_B \rangle$ *are isomorphic as ordered trees, the map that sends each surreal number to its sign-expansion being the unique such isomorphism*.

While $\langle B, <_{lex(B)}, <_B \rangle$ sheds important light on the structure of $\langle \mathbf{No}, <, <_s \rangle$, in the general theory of surreal numbers it is advantageous to have available a representation independent characterization of those ordered trees that are isomorphic to initial ordered subtrees of $\langle B, <_{lex(B)}, <_B \rangle$ and, hence, of $\langle \mathbf{No}, <, <_s \rangle$. There are two particularly useful such characterizations that will be appealed to in the subsequent discussion, the first of which we take as the standard definition.

Definition 1 (Ehrlich 2001) *A lexicographically ordered binary tree is an ordered binary tree in which distinct members are incomparable if and only if they have a common predecessor lying between them*.

A subclass I of an ordered class $\langle A, < \rangle$ is said to be *convex* if every member of A that lies between two members of I is likewise a member of I. The familiar intervals of an ordered class $\langle A, < \rangle$ of the forms $\{x \in A : a \leq x \leq b\}$, $\{x \in A : a < x < b\}$, $\{x \in A : a \leq x < b\}$, and $\{x \in A : a < x \leq b\}$ are all convex subclasses. Besides these, however, there may be convex subclasses of $\langle A, < \rangle$ without a greatest lower bound a or without a least upper bound b. The ordered set of rational numbers of course contains many such convex subclasses, and non-Archimedean ordered fields invariably contain many such convex subclasses as well.

The idea of a convex subclass of an ordered class plays a leading role in the formulation of the second of the two representation independent characterizations of a lexicographically ordered binary tree alluded to above. Also important in this regard are the following notational and terminological conventions that will be employed throughout the remainder of the chapter.

If L and R are subclasses of an ordered class $\langle A, < \rangle$ (where L or R may be empty), then "$L < R$" signifies that every member of L precedes every member of R. In addition, if x and y are members of an ordered tree $\langle A, <, <_s \rangle$, then (as in the case of surreal numbers) $x <_s y$ will be read *x is simpler than y*; moreover, an element x of a nonempty subclass I of A will be said to be "the simplest member" of I if $x <_s y$ for all $y \in I - \{x\}$. Finally, mimicking the stipulations introduced in Corollary 1, by "$L_{s(x)}$" we mean $\{a \in A : a <_s x \text{ and } a < x\}$ and by "$R_{s(x)}$" we mean $\{a \in A : a <_s x \text{ and } x < a\}$.

The second of our representation independent characterizations of a lexicographically ordered binary tree is given by

Theorem 3 (Ehrlich 2001: Theorems 1 and 2) *$\langle A, <, <_s\rangle$ is a lexicographically ordered binary tree if and only if $\langle A, <, <_s\rangle$ is an ordered tree in which every nonempty convex subclass contains a simplest member, and for all $x, y \in A$, $x <_s y$ if and only if $L_{s(x)} < \{y\} < R_{s(x)}$ and $x \neq y$.*

Now suppose L and R are subsets of a lexicographically ordered binary tree $\langle A, <, <_s\rangle$ where $L < R$. It follows from the just-stated theorem that the convex subclass of A consisting of all the members of A lying between L and R is either empty or contains a simplest member. If a simplest such member always exists, $\langle A, <, <_s\rangle$ will be said to be *complete*. As the following theorem makes clear, the concept of a complete lexicographically ordered binary tree and the concept of a full lexicographically ordered binary tree provide us with complementary perspectives on the nature of **No**.

Theorem 4 (Ehrlich 1994, 2001) *A lexicographically ordered binary tree is complete if and only if it is full and if and only if it is isomorphic to $\langle \mathbf{No}, <, <_s\rangle$.*

An automorphism is an isomorphism of a structure $\langle A, \ldots\rangle$ onto itself. Roughly speaking, if f is an automorphism of $\langle A, \ldots\rangle$ in which $f(x) \neq x$ for some $x \in A$, then $f(x)$ and x can play exactly the same roles as one another in isomorphic structures having the common universe A. As is evident from (Ehrlich 1989a, Sect. 2), $\langle \mathbf{No}, <\rangle$ has a vast array of automorphisms. However, as the following result makes clear, all but one fail to take into account from whence the numbers come.

Theorem 5 (Ehrlich 1994) *The identity map is the sole automorphism of $\langle \mathbf{No}, <\rangle$ that preserves predecessors. That is, an automorphism f of $\langle \mathbf{No}, <\rangle$ satisfies the condition $\mathrm{pr}_{\mathbf{No}}(f(x)) = \mathrm{pr}_{\mathbf{No}}(x) = L_{s(x)} \cup R_{s(x)}$ for all $x \in \mathbf{No}$ if and only if $f(x) = x$ for all $x \in \mathbf{No}$.*

The identity map is likewise the only automorphism of $\langle \mathbf{No}, <\rangle$ that preserves tree-rank. Moreover, the following consequence of Theorem 5 provides still another way of characterizing this map.

Corollary 3 (Ehrlich 1994) *$\langle \mathbf{No}, <, <_s\rangle$ has no automorphism other than the identity automorphism.*

6 Ordered Fields with Simplicity Hierarchies

Let $\langle A, <, <_s\rangle$ be a lexicographically ordered binary tree. If x is the simplest member of A lying between two subsets L and R of A where $L < R$, then we will denote x by "$\{L|\ R\}^A$" or simply by "$\{L|\ R\}$," when no ambiguity will arise. Moreover, applying

Conway's conventions (Conway 1976, p. 4) to our definition of "$\{L \mid R\}$," if $x = \{L \mid R\}$ we write "x^L" for a typical member of L and "x^R" for a typical member of R; for x itself, we write "$\{x^L \mid x^R\}$"; and $x = \{a, b, c, \ldots | d, e, f, \ldots\}$ means that $x = \{L \mid R\}$ where $a, b, c, \ldots$ are typical members of L and $d, e, f, \ldots$ are typical members of R.

The following consequence of Theorem 3 is central to the subsequent development.

Proposition 2 *If $\langle A, <, <_s\rangle$ is a lexicographically ordered binary tree and $x \in A$, then there are (possibly empty) subsets L and R of A such that $L < R$ and for which $x = \{L \mid R\}$; in particular, x is the simplest member of A lying between its predecessors on the left and its predecessors on the right, that is, $x = \{L_{s(x)} \mid R_{s(x)}\}$.*

In the following definition, x^L, x^R, y^L, and y^R are understood to be typical members of arbitrary sets L_x, R_x, L_y, and R_y, respectively, for which $x = \{L_x \mid R_x\}$ and $y = \{L_y \mid R_y\}$.

Definition 2 (Ehrlich 2001) *$\langle A, +, \cdot, <, <_s, 0, 1\rangle$ is said to be an s-hierarchical ordered field if (i) $\langle A, +, \cdot, <, 0, 1\rangle$ is an ordered field; (ii) $\langle A, <, <_s\rangle$ is a lexicographically ordered binary tree; and (iii) for all $x, y \in A$,*

$$x + y = \left\{x^L + y, x + y^L | x^R + y, x + y^R\right\}$$

and

$$xy = \{x^L y + xy^L - x^L y^L, x^R y + xy^R - x^R y^R \mid x^L y + xy^R - x^L y^R, x^R y + xy^L - x^R y^L\}.$$

The reader will note that in virtue of the conventions introduced at the beginning of this section, the s-hierarchical constraint on sums in the above definition asserts that $x + y = \{L_{x+y} \mid R_{x+y}\}$ where L_{x+y} is the set consisting of all sums of the form $x^L + y$ where $x^L \in L_x$ as well as all sums of the form $x + y^L$ where $y^L \in L_y$, and R_{x+y} is the set consisting of all sums of the form $x^R + y$ where $x^R \in R_x$ as well as all sums of the form $x + y^R$ where $y^R \in R_y$. The s-hierarchical constraint on products is analogously understood.

Using the s-hierarchical constraints on sums and products, respectively, it is a simple matter to show that $0 = \{\emptyset \mid \emptyset\} = \{|\}$ is the simplest element of, as well as the unique root in, an s-hierarchical ordered field A, and that $1 = \{\{\{\emptyset \mid \emptyset\}\} \mid \emptyset\} = \{0|\}$ is the simplest positive element of A (Ehrlich 2001). Moreover, it is a consequence of the following result that except for the signs of elements, the portion of the tree less than 0 is a mirror image of the portion of the tree greater than 0.

Proposition 3 (Ehrlich 2001) For each element x of an s-hierarchical ordered field

$$-x = \{-x^R | -x^L\}.$$

Like the just-stated algebraico-tree-theoretic condition governing additive inverses, the algebraico-tree-theoretic constraints placed on sums and products of elements of an s-hierarchical ordered field are apt to appear rather cryptic to readers unfamiliar with the theory of surreal numbers. We will decipher their significance shortly. Before doing so, however, we note that a good deal of their motivation arises from the following straightforward tree-theoretic variation of a remarkable result due to Conway (1976).

Theorem 6 (Ehrlich 2001) $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$ *is an s-hierarchical ordered field when* $+$, $-$, *and* $\cdot$ *are defined by recursion as follows where* x^L, x^R, y^L, *and* y^R *are understood to range over the members of* $L_{s(x)}$, $R_{s(x)}$, $L_{s(y)}$, *and* $R_{s(y)}$, *respectively.*

Definition of $x + y$.

$$x + y = \{x^L + y, x + y^L | x^R + y, x + y^R\}.$$

Definition of $-x$.

$$-x = \{-x^R | -x^L\}.$$

Definition of xy.

$$xy = \{x^L y + xy^L - x^L y^L, x^R y + xy^R - x^R y^R \mid x^L y + xy^R - x^L y^R, x^R y + xy^L - x^R y^L\}.$$

As our remarks preceding the statement of Theorem 6 suggest, the constraints on sums and products found in Definition 2 are tree-theoretic adaptations of the recursive definitions of $+$ and $\cdot$ employed by Conway in his construction of the ordered field of surreal numbers (1976, p. 5). Appealing to ideas alluded to in the Introduction, the import of these constraints may be encapsulated thus.

Let $x, y \in A$. Since $x = \{x^L \mid x^R\}$ and $y = \{y^L \mid y^R\}$, it follows that for all x^L, x^R, y^L, and y^R:

$$(*) \qquad x^L < x < x^R \text{ and } y^L < y < y^R.$$

Furthermore, combining $(*)$ with the fact that A is an ordered field (and, hence, an ordered additive group), it follows that $x^L + y < x + y < x^R + y$ and $x + y^L < x + y < x + y^R$, and, hence, that

$$(**) \qquad \{x^L + y, x + y^L\} < \{x + y\} < \{x^R + y, x + y^R\}.$$

Therefore, since $x + y$ must lie between the two sets of members of A specified in $(**)$, clause (iii) of Definition 2 requires that $x + y$ be *the simplest member of A lying between those sets*.

Similarly, since, by clause (i) of Definition 2, A is an ordered field, it follows from $(*)$ that each of the differences $x - x^L, x^R - x, y - y^L, y^R - y$ is positive and, hence, so is each of the products $(x - x^L)(y - y^L)$, $(x^R - x)(y^R - y)$, $(x - x^L)(y^R - y)$, and $(x^R - x)(y - y^L)$. But by applying this together with routine high school algebra of ordered fields to each of these products, one obtains

$$(***) \qquad \begin{aligned} &\left\{x^L y + xy^L - x^L y^L, x^R y + xy^R - x^R y^R\right\} < \\ &\qquad < \{xy\} < \left\{x^L y + xy^R - x^L y^R, x^R y + xy^L - x^R y^L\right\}. \end{aligned}$$

Therefore, since xy must lie between the two sets of members of A specified in $(***)$, clause (iii) of Definition 2 requires that xy be *the simplest member of A lying between those sets*.

Given an ordered field $\langle A, +, \cdot, <, 0, 1\rangle$, it is possible to define a lexicographically ordered binary tree structure on its ordered class $\langle A, <\rangle$ of elements in infinitely many ways. That is, there are infinitely many distinct predecessor relations $<_A$ that can be defined on the members of A in accordance with which $\langle A, <, <_A\rangle$ is a lexicographically ordered binary tree. In the case of $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$, in particular, a lexicographically ordered binary tree structure can be imposed on its ordered class of elements in as many ways as there are ordinals. Moreover, while the constraints on sums and products dictated by $(**)$ and $(***)$, respectively, will be satisfied in any of the resulting systems, one is largely free to assign the sum and product of any two members of the system to any position in the tree one pleases. If the resulting structure is to be an s-hierarchical ordered field, however, this is not the case. Indeed, requiring that the sums and products of members of A be the simplest possible elements of the tree consistent with $(**)$ and $(***)$ is tantamount to requiring that sums and products of members of A have the lowest possible tree rank consistent with those constraints. In the case of an s-hierarchical ordered field like **No**, whose sums and products are defined by recursion, this implies that the sums and products of elements get defined just as soon as there is sufficient previously defined ordered algebraico-tree-theoretic information to do so. As we mentioned in the Introduction and the subsequent sections of the chapter will show, this has profound implications for the s-hierarchical ordered field of surreal numbers.

7 S-Hierarchical Mappings and S-Hierarchical Morphisms

There is a uniquely defined image of each rational number r in every ordered field. Moreover, if A is an ordered subfield of the ordered field $\mathbf{R}_{alg}$ of *real algebraic numbers* – i.e., the ordered subfield of $\mathbf{R}$ whose members are solutions to polynomial equations with integer coefficients – and r is an irrational member of A, such as $\sqrt{2}$, then every ordered field containing an isomorphic copy of A will have a uniquely

defined image of r as well. However, there are no other members of ordered fields for which this is true without qualification. Of course, every member r of an ordered subfield of the reals has a unique image in the ordered field of real numbers, and, as such, within the confines of an ordered subfield of real numbers one can speak unambiguously of the real number r whether it be a rational number, a real algebraic number that is not rational, or a transcendental number, i.e., an irrational number, such as π or e, that is not a real algebraic number. However, if A is an Archimedean ordered field containing a *transcendental number r*, then *no non-Archimedean ordered field containing an isomorphic copy of A will have a uniquely defined image of r.* That is, there will always be isomorphisms f and g from A into the non-Archimedean ordered field in question for which $f(r) \neq g(r)$. In fact, *if A is an arbitrary ordered field, whose universe is a set, and r a member of A that is not one of A's real algebraic numbers, there is an entire proper class of isomorphisms of A into* **No** *for which no two of them send r to the same surreal number!* Happily, however, as the remainder of this and the subsequent section will show, not only does the absence of uniquely defined images fail to arise for members of s-hierarchical ordered fields under suitably defined morphisms, but every such element has a uniquely defined image in $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$. In saying as much, we are alluding to the fact that the type of isomorphisms appropriate for s-hierarchical ordered fields are more stringent than those appropriate for ordered fields without s-hierarchical structures. It is to these *s-hierarchical morphisms*, as we call them, to which we now turn.

Definition 3 (Ehrlich 2001) *A mapping $f: A \to A'$ between two lexicographically ordered binary trees $\langle A, <, <_A\rangle$ and $\langle A', <', <_{A'}\rangle$ is said to be an s-hierarchical mapping if for all $x \in A$, $f(x) = \{ f(L) \mid f(R)\}^{A'}$ whenever $x = \{L \mid R\}^A$. If, in addition, A and A' are s-hierarchical ordered fields and the mapping is also an embedding of ordered fields, it is said to be an s-hierarchical embedding or an s-hierarchical isomorphism.*

It is not difficult to show that a mapping $f: A \to A'$ between two lexicographically ordered binary trees $\langle A, <, <_A\rangle$ and $\langle A', <', <_{A'}\rangle$ is s-hierarchical if and only if the mapping preserves order, the predecessor relation, and tree-rank (Ehrlich 2001, Lemma 1): that is, for all $x, y \in A$: (i) $f(x) <' f(y)$ whenever $x < y$; (ii) $f(x) <_{A'} f(y)$ whenever $x <_A y$; and (iii) $\rho_A(x) = \rho_{A'}(f(x))$. This in conjunction with the structure of s-hierarchical ordered fields underlies the following result that brings to the fore just how distinguished s-hierarchical mappings of s-hierarchical ordered fields really are.

Theorem 7 (Ehrlich 2001: Lemma 2, Theorem 5) *Every s-hierarchical mapping between s-hierarchical ordered fields is an s-hierarchical embedding; moreover, if $f: A \to S$ and $g: A \to S$ are s-hierarchical mappings, then $f = g$ and $f(A)$ is an initial subtree of S. There is an s-hierarchical mapping of every s-hierarchical ordered field into $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$.*

In virtue of Corollary 3, it is clear that the sole automorphism of $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$ is the identity automorphism. However, in virtue of Theorem 7, we have more.

Corollary 4 *The identity automorphism is the sole s-hierarchical embedding of an s-hierarchical ordered field into itself. Accordingly, if* $\langle A, +, \cdot, <, <_s, 0, 1\rangle$ *is an s-hierarchical ordered field, then the identity automorphism is the sole embedding of* $\langle A, +, \cdot, <, 0, 1\rangle$ *into itself that preserves* $<_s$.

In particular:

Corollary 5 *The identity automorphism is the sole s-hierarchical embedding of* $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$ *into itself. Thus, the identity automorphism is the sole embedding of* $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ *into itself that preserves* $<_s$.

Structures whose automorphism groups consist solely of the identity automorphism are said to be *rigid*. If the identity automorphism is the sole embedding of a structure into itself (i.e., is the sole endomorphism of a structure), the structure is said to be *strongly rigid*. While strongly rigid structures are rigid, rigid structures need not be strongly rigid. For example, whereas the ordered set **N** of natural numbers is rigid, it is not strongly rigid since it is isomorphic to the ordered set $\mathbf{N} - \{0\}$ of natural numbers greater than 0. However, in accordance with Corollary 4, s-hierarchical ordered fields are strongly rigid, and in this regard they behave like **R** and the ordered ring **Z** of integers. We do not know whether it is the rigidity of **Z**, the strong rigidity of **Z**, or something else entirely that Fraenkel, Bar-Hillel, and Levy had in mind when they maintained (see Introduction) that "[e]very integer differs from every other in characteristic individual properties comparable to the differences between human beings." However, while the rigidity of **Z** certainly suffices to underwrite their contention, it does not provide the full story. The strong rigidity of **Z** ensures that neither a permutation of the members of **Z** nor the members of a proper substructure of **Z** could exhibit all the properties of all the members of **Z** expressed in the language $L = \{+, \cdot, <, 0, 1\}$ of ordered rings, whereas the rigidity of **Z** only ensures the former. In sharp contrast with **Z**, the ordered field of surreal numbers has a vast array of endomorphisms that are not automorphisms as well as a vast array of automorphisms other than the identity automorphism. It is only when the s-hierarchical nature of **No** is taken into account that its strong rigidity emerges.

A subfield A of an s-hierarchical ordered field S is said to be an *initial subfield* if the universe of A is an initial subtree of S.

Proposition 4 (**Ehrlich 2001**) *An initial subfield of an s-hierarchical ordered field is itself an s-hierarchical ordered field; furthermore, an ordered field admits a*

relational expansion to an s-hierarchical ordered field if and only if it is isomorphic as an ordered field to an initial subfield of an s-hierarchical ordered field.

In a number of subsequent sections, the closely related notions of s-hierarchical embedding and initial subfield play important roles in shedding further light on **No**'s structure as an s-hierarchical ordered field. We begin with **No**'s s-hierarchical ordered subfield of real numbers.

8 The S-Hierarchical Ordered Field of Real Numbers

Thus far, the only s-hierarchical ordered fields whose existence we have discussed are complete s-hierarchical ordered fields. However, there is a vast array of s-hierarchical ordered fields that fall short of completeness, the full spectrum of which was characterized by Ehrlich (2001, §4). It is a consequence of Theorem 7 and Proposition 4 that revealing the spectrum of s-hierarchical ordered fields reduces to revealing the spectrum of initial subfields of **No**. While we will make no attempt in this chapter to spell out the entire spectrum, in this and a subsequent section we will draw attention to components of the spectrum that are of particular importance from the standpoints of the classical, elementary (i.e., first-order), and absolute theories of continua. In this section, we introduce the s-hierarchical ordered field of real numbers, a structure that, as we shall later see, plays a critical role in assigning canonical appellations to arbitrary surreal numbers.

Definition 4 Let **D** be the set of all surreal numbers having finite tree-rank, and further let

$$\mathbf{R} = \mathbf{D} \cup \{\{X \mid Y\} : (X, Y) \text{ is a Dedekind gap in } \mathbf{D}\},$$

where, as the reader will recall, a Dedekind gap (X, Y) *in* **D** is a partition of **D** into two nonempty subsets X and Y such that $X < Y$, X has no greatest member, and Y has no least member.

The following result is essentially due to Conway (1976, pp. 23–25).

Proposition 5 **R** *(with* $+, -, \cdot$ *and* $<$ *defined à la* **No***) is isomorphic to the ordered field of reals defined in any of the more familiar ways,* **D** *being* **No***'s ring of dyadic rationals (*i.e.*, rationals of the form* $m/2^n$ *where* m *and* n *are integers);* $n = \{0, \ldots, n-1|\}$ *for each positive integer* n, $-n = \{|-(n-1), \ldots, 0\}$ *for each positive integer* n, $0 = \{|\}$, *and the remainder of the dyadics are the arithmetic means of their left and right predecessors of greatest tree-rank;* e.g., $\frac{1}{2} = \{0|1\}$.

Being an ordered field, **No** contains exactly one subring of integers, one subring of dyadic rationals, and one subfield of rationals. On the other hand, as we alluded to

above, **No** contains an entire proper class of distinct isomorphic copies of the ordered field of reals. What distinguishes **R** from the rest is captured by the first part of

Theorem 8 (**Ehrlich 2001**) **R** *is the unique isomorphic copy of the ordered field of reals that is an initial subfield of* **No**. *In fact, every Archimedean ordered field is isomorphic to exactly one initial subfield of* **No**, *the latter being an initial subfield of* **R**.

The first part of Theorem 8 being its requisite justification, henceforth we shall refer to **R** as the ordered field of real numbers. In accordance with this convention, **R** is the set of all surreal numbers whose sign-expansions have length $\leq \omega$ that end neither in an infinite sequence of $-$ s nor in an infinite sequence of $+$ s. The real numbers whose sign-expansions have length $< \omega$ are the dyadic rationals, and the remaining rational numbers are the members of **R** having sign-expansions that end with finite sequences of signs, containing at least one $+$ and at least one $-$, that recurs ω times. For instance, $\frac{1}{3}$ is the surreal number having the sign-expansion the begins with $+-$ and is followed by the sequence $-+$ repeated ω times. For a simple rule, due to Elwyn Berlekamp, for obtaining the value of a real number from its sign-expansion, see (Conway 1976, p. 31).

9 Every Surreal Number Has its Own Proper Name

As we mentioned in the Introduction, each surreal number can be assigned a canonical proper name that is a reflection of its characteristic individual properties. In order to define these *Conway names*, or *normal forms* as Conway (1976, p. 33) calls them, we require a number of definitions beginning with the following classical ones from the theory of ordered algebraic systems applied to **No**.

Two elements a and b of **No** are said to be *Archimedean equivalent*, written $a \approx b$, if there are positive integers m and n such that $m|a| > |b|$ and $n|b| > |a|$; if $a \approx b$ and $|a| < |b|$, then we write $|a| \ll |b|$ and a is said to be *infinitesimal (in absolute value) relative to* b and b is said to be *infinite (in absolute value) relative to* a; the class of all members of **No** that are Archimedean equivalent to some member of **No** is said to constitute an *Archimedean class* of **No**. 0, which is infinitesimal (in absolute value) relative to every other surreal number, is the sole surreal number that is not a member of an Archimedean class.

Following Conway (1976, p. 31), an element of **No** is said to be a *leader* if it is the simplest member of the class of positive elements of an Archimedean class of **No**. Since the class of positive elements of an Archimedean class of **No** is a convex subclass of **No**, the concept of a leader is well defined. Henceforth, by Lead(**No**) we mean the class of all leaders of **No**.

One of the salient features of Lead(**No**) is given by the following result, where $<_{\mathbf{No}}|\text{Lead}(\mathbf{No})$ and $<_s|\text{Lead}(\mathbf{No})$ denote the restrictions to Lead(**No**) of the order and simpler than relations in **No**.

Theorem 9 (**Ehrlich 2001**) $\langle \text{Lead}(\mathbf{No}), <_{\mathbf{No}} | \text{Lead}(\mathbf{No}), <_s | \text{Lead}(\mathbf{No}) \rangle$ is a lexicographically ordered full binary tree, and, as such, the unique s-hierarchical mapping from Lead(**No**) to **No** (henceforth, $\Phi_{\mathbf{No}}$) is a bijection.

The following recursive definition, which is a variation on an idea due to Conway (1976, p. 31), provides a unique appellation for each leader in **No**. In effect, it assigns the appellation "ω^0" to 1, which is the simplest leader in **No**, the appellation "ω^{-1}" to the simplest leader in **No** less than ω^0, the appellation "ω^1" to the simplest leader in **No** greater than ω^0, and so on.

Definition 5 *For each* $y \in \mathbf{No}$, $\omega^y = \left\{0, n\omega^{L_{S(y)}} | \frac{1}{2^n}\omega^{R_{S(y)}}\right\}^{\mathbf{No}}$, *where n is understood to range over all positive integers and* $\omega^{L_{S(y)}}$ *and* $\omega^{R_{S(y)}}$ *are understood to range over the members of* $\{\omega^x : x \in L_{s(y)}\}$ *and* $\{\omega^x : x \in R_{s(y)}\}$, *respectively.*

Justification for the contention that Definition 5 provides a unique appellation to each leader in **No** accrues from

Theorem 10 (**Ehrlich 2001**) $\omega^y = \Phi_{\mathbf{No}}{}^{-1}(y)$, *for each* $y \in \mathbf{No}$.

Using a classical argument, one may show that if $x \in \mathbf{No} - \{0\}$, there is a unique $r \in \mathbf{R} - \{0\}$, a unique $y \in \mathbf{No}$, and a unique $a \in \mathbf{No}$ such that $x = r\omega^y + a$ where a is infinitesimal (in absolute value) relative to ω^y. To put this somewhat differently, if $x \in \mathbf{No} - \{0\}$, there is a unique $r \in \mathbf{R} - \{0\}$ and a unique $y \in \mathbf{No}$ for which $|x - r\omega^y| \ll \omega^y$. This elementary observation plays pivotal roles in both the formulation and proof of the following s-hierarchical recasting of an argument sketched by Conway (1976, pp. 32–33).

Theorem 11 (**Ehrlich 2001**) *For each* $x \in \mathbf{No} - \{0\}$, *there is a unique descending chain* I^x_α, $\alpha < \beta \in On$, *of convex subclasses of* **No** *whose intersection contains x as its simplest member, and whose components are defined by recursion as follows:* I^x_α *is the subclass of all members of* **No** *of the form* $s_\alpha + r_\alpha\omega^{y_\alpha} + a_\alpha$, *where*

(i) $s_\alpha = 0$, if $\alpha = 0$, and s_α is the simplest member of $\cap_{\gamma<\alpha} I^x_\gamma$, otherwise.
(ii) $r_\alpha\omega^{y_\alpha}$– henceforth called the α-term of x– is the unique member of **No** for which $r_\alpha \in \mathbf{R} - \{0\}$, $y_\alpha \in \mathbf{No}$ and $|x - (s_\alpha + r_\alpha\omega^{y_\alpha})| \ll \omega^{y_\alpha}$.
(iii) $|a_\alpha| \ll \omega^{y_\alpha}$.

Since the intersection of a chain of convex subclasses is itself a convex subclass, it follows from Theorem 11 that, for each $x \in \mathbf{No} - \{0\}$ there is a unique convex subclass of **No** containing x as its simplest member consisting of all and only those surreal numbers whose α-term is $r_\alpha\omega^{y_\alpha}$ for all $\alpha <$ some ordinal β that depends on x. With this as its underlying justification, the completion of the assignment of appellations to surreal numbers begun with Definition 5 is carried out via

Definiton 6 We will refer to the formal expression

$$\sum_{\alpha<\beta}\omega^{y_\alpha}.r_\alpha$$

as the Conway name of a surreal number x, treat the Conway name of x as a proper name of x, and, accordingly, write "$x = \sum_{\alpha<\beta}\omega^{y_\alpha}.r_\alpha$" if and only if either $x = 0$ and $\beta = 0$ or x is the simplest surreal number whose α-term is $r_\alpha\omega^{y_\alpha}$ for all $\alpha < \beta$. (Moreover, we will delete the summation sign or insert sum signs between successive terms when it is convenient to do so, as well as write $\omega^0.r = r$, $\omega^y.-r = -\omega^y.r$, $\omega^1.1 = \omega$, $\omega.\frac{1}{2} = \frac{\omega}{2}$, $\omega^{-y} = \frac{1}{\omega^y}$, and so on, in accordance with standard practice.)

As is well known, in virtue of a celebrated theorem due to Cantor (1897) and Hessenberg (1906), every ordinal has a unique *Cantor normal form*, i.e., a unique representation of the form

$$\sum_{\alpha<n}^{c}\omega^{\cdot_c y_\alpha}.{}_c a_\alpha$$

where $\{y_\alpha : \alpha < n\}$ is a (possibly empty) finite descending sequence of ordinals, $\{a_\alpha : \alpha < n\}$ is a sequence of nonzero finite ordinals, and the "c" affixed to the various operations indicates that the operations are the familiar Cantorian operations on ordinals. The following result shows, among other things, that Conway names of surreal numbers generalize the Cantor normal forms of ordinals.

Theorem 12 (Conway 1976; Ehrlich 2001) *Every surreal number has a Conway name; distinct surreal numbers have distinct Conway names; moreover, the formal expression*

$$\sum_{\alpha<\beta}\omega^{y_\alpha}.r_\alpha$$

is the Conway name of some surreal number if and only if $\{y_\alpha : \alpha < \beta \in On\}$ is a (possibly empty) descending sequence of members of **No** *and $\{r_\alpha : \alpha < \beta\}$ is a sequence of nonzero real numbers, the Conway name of an ordinal being its Cantor normal form.*

Conway names provide highly perspicuous representations of surreal numbers. $\omega - \pi$ is indeed the surreal number obtained by subtracting π from ω, $\frac{\omega}{2}$ is ω divided by 2, $\frac{1}{\omega}$ is the multiplicative inverse of ω as well as the cube root of $\frac{1}{\omega^3}$, $\omega^\omega + \omega.2 + 1$ and $\sqrt{2}$ are the ordinal and real number, respectively, so named, and so on. Even Conway names indexed over all $\alpha <$ some infinite limit ordinal β encode significant

algebraico-tree-theoretic information about the surreal numbers they denote (e.g., Ehrlich 2001, Theorem 15 (iii)).

Besides providing perspicuous representations, Conway names make surreal numbers more tractable from an algebraic point of view. Indeed, when surreal numbers are denoted by their respective Conway names, the basic field operations can be performed on them analogously to the familiar operations on polynomials as can the familiar ordering of elements by first-differences, the identification of multiplicative inverses, and the extraction of n-th roots. In the case of addition and multiplication, in particular, the operations are the familiar polynomial operations where terms having a common exponent are added by the rule

$$\omega^x.a +_H \omega^x.b = \omega^x.(a+b)$$

where $a + b$ is the ordinary sum of real numbers, and the product of an arbitrary pair of terms is given by

$$\omega^x.a \cdot_H \omega^y.b = \omega^{x +_H y}.ab$$

where ab is the ordinary product of real numbers and $x +_H y$ is the aforementioned sum. So, for example, given two surreal numbers $\omega - \frac{1}{2}$ and $1 + \frac{1}{\omega}$, their sum is given by $\omega + \frac{1}{2} + \frac{1}{\omega}$ and their product by $\omega - \frac{1}{2} + \frac{1}{\omega.2}$. Moreover, $1 + \frac{1}{\omega} < \omega - \frac{1}{2}$ since at the first corresponding term in which $\omega^1.0 + \omega^0.1 + \omega^{-1}.1$ and $\omega^1.1 + \omega^0.-\frac{1}{2} + \omega^{-1}.0$ differ (i.e., $\omega^1.0$ and $\omega^1.1$) $0 < 1$.

Underwriting the preceding observations regarding the representational and algebraic virtues of Conway names is the following result where the operations $+_H$ and $\cdot_H$ are defined on, and the relation $<_H$ is defined between, surreal numbers denoted by their respective Conway names (supplemented with "dummy" terms with zeros for coefficients to permit a uniform representation of all surreal numbers).

Theorem 13 (Conway 1976; Ehrlich 2001) $\langle \mathbf{No}, +_H, \cdot_H, <_H, <_s, 0, 1\rangle = \langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$, *when* $+_H$, $\cdot_H$, *and* $<_H$ *are defined as follows:*

$$\sum_{y \in \mathbf{No}} \omega^y.a_y +_H \sum_{y \in \mathbf{No}} \omega^y.b_y = \sum_{y \in \mathbf{No}} \omega^y.(a_y + b_y),$$

$$\sum_{y \in \mathbf{No}} \omega^y.a_y \cdot_H \sum_{y \in \mathbf{No}} \omega^y.b_y = \sum_{y \in \mathbf{No}} \omega^y.\left[\sum_{\mu +_H \nu = y} a_\mu b_\nu\right],$$

$$\sum_{y \in \mathbf{No}} \omega^y.a_y <_H \sum_{y \in \mathbf{No}} \omega^y.b_y, \text{ if } a_y = b_y \text{ for all } y > \text{some } x \in \mathbf{No} \text{ and } a_x < b_x.$$

When restricted to ordinals, $+_H$ and $\cdot_H$ are not the familiar noncommutative sums and products of Cantor, but rather the so-called *natural sums* and *natural products* of

ordinals due to Hessenberg (1906) and Hausdorff (1927), respectively, and $<_H$ is the standard ordering of ordinals (written in *Cantor normal form*) by first-differences (e.g., Ehrlich 2006, pp. 24–25). The subscript "*H*" affixed to the symbols denoting the field operations and the ordering relation, however, is not a tribute to Hessenberg and Hausdorff. After all, Cantor normal forms of ordinals are always indexed over finite ordinals. Rather, the subscripts are in honor of Hans Hahn (1907), whose corresponding field operations and ordering relation defined on transfinite power series are the *loci classici* for the types of operations and ordering relation defined in Theorem 13. In fact, Theorems 12 and 13 essentially show that **No**, considered as an ordered field, is isomorphic to the *On*-restricted Hahn field $\mathbf{R}((\mathbf{No}))_{On}$ of formal power series with coefficients in the ordered field **R** and exponents in **No** considered as an ordered additive group (Alling 1987; Ehrlich 1988, 2001).

10 The Ordered Field of Surreal Numbers

Following Emil Artin and Otto Schreier (1926), an ordered field *A* may be said to be *real-closed* if it admits no extension to a more inclusive ordered field that results from supplementing *A* with solutions to polynomial equations with coefficients in *A*. Intuitively speaking, real-closed ordered fields are precisely those ordered fields having no "holes" that can be filled by algebraic means alone. The ordered field of real numbers is the paradigmatic real-closed ordered field; indeed, it was the desire to capture the algebraic theory of the reals that motivated the development of Artin and Schreier's theory (Sinaceur 1994). Among a number of important properties of **R** they showed are satisfied by real-closed ordered fields more generally is the *intermediate value theorem for polynomials of a single variable*. In fact, as soon became clear, for an ordered field *A*, the satisfaction of the intermediate value theorem for polynomials of a single variable over *A* is equivalent to *A* being real-closed (e.g., Tarski 1939; Warner 1965, pp. 492–493). Accordingly, the idea of a real-closed ordered field is equivalent for the case of polynomials of a single variable to satisfying what has traditionally been regarded as one of the quintessential properties of a continuous function – the formalization of the intuitive idea that a continuous curve connecting points on the opposite sides of a straight line intersects the given line. Tarski (1948) also shed important model-theoretic light on Artin and Schreier's algebraic conception by showing that real-closed ordered fields are precisely the ordered fields that are first-order indistinguishable from the ordered field of reals. For this reason, they are sometimes called *elementary arithmetic continua*.

Like **R**, **No** is real-closed (Conway 1976). However, since there is a multitude of real-closed ordered fields, it is natural to inquire if, like **R**, it is possible to distinguish **No** (to within isomorphism) from the remaining real-closed ordered fields by appealing solely to its order. As we shall now see, the following definition, which

embodies the order-theoretic content of the above-mentioned idea of a complete lexicographically ordered binary tree, enables one to do just that.

Definition 7 (Ehrlich 1988) An ordered class $\langle A, < \rangle$ is said to be an *absolute linear continuum* if for all subsets L and R of A where $L < R$ there is a $y \in A$ such that $L < \{y\} < R$.

In (Ehrlich 1988, 1992, 1994) and a number of other early works, we refer to absolute linear continua as η_{On}*-orderings* because they extend to proper classes Hausdorff's (1907, 1909) idea of an η_α*-ordering of power* $\aleph_\alpha$, i.e., an ordered set A of power $\aleph_\alpha$ such that for all subsets L and R of A where $L < R$ and $|L|, |R| < \aleph_\alpha$ there is a member of A lying strictly between those of L and those of R. In a number of earlier works (Ehrlich 1987, p. 243; 1992, p. 172; 1994, p. xxvii), we also described the property expressed in the definition of an absolute linear continuum as a condition of absolute density. However, we subsequently came to believe it is more appropriate and more revealing to split the condition into two conditions in the following fashion, where one of the conditions, which is a more restricted sense of absolute density, is now called "absolute density."

An absolute linear continuum $\langle A, < \rangle$ is both *absolutely dense* in the sense that for each pair of *nonempty* subsets L and R of A where $L < R$, there is a $y \in A$ such that $L < \{y\} < R$, and *absolutely extensive* in the sense that given any subset L of A there are $a, b \in A$ such that $\{a\} < L < \{b\}$. In fact, since in the definition of an absolute linear continuum L or R may be empty, one can readily show that an ordered class is an absolute linear continuum if and only if it has both of the just-stated properties. Thus, since every element of an ordered class must either lie between two of its nonempty subclasses or lie to the right or to the left of some (possibly empty) subclass, absolute linear continua are ordered classes having *no order-theoretic limitations that are definable in terms of sets of standard set theory.*

In the first installment of his *Contributions to the Founding of the Theory of Transfinite Numbers*, Cantor (1895, Sect. 11) provided an order-theoretic characterization of the classical linear continuum. The following is the analog for absolute linear continua of Cantor's result.

Theorem 14 (Ehrlich 1988) $\langle \mathbf{No}, < \rangle$ *is (up to isomorphism) the unique absolute linear continuum.*

Unlike the ordered field of reals, however, the ordered field of surreal numbers is not distinguished (up to isomorphism) from other ordered fields by its structure as an ordered class. Indeed, there are infinitely many pairwise nonisomorphic ordered fields that are absolute linear continua. In fact, since every α-saturated ordered field is an η_α-ordering, to obtain such an ordered field one need only form the union of a continuous chain A_α ($\alpha < On$) of ordered fields where A_0 is not real-closed and $A_{\alpha+1}$ is an $\alpha + 1$-saturated elementary extension of A_α for each α (e.g., Chang and Keisler 1990, ch. 5). Happily, however, what one can prove is:

Theorem 15 (Ehrlich 1988, Lemma 2) $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ *is (up to isomorphism) the unique real-closed ordered field that is an absolute linear continuum.*

In virtue of Theorem 15, **No** is not only devoid of set-theoretically defined order-theoretic limitations, but also it is devoid of algebraic limitations; moreover, to within isomorphism, it is the unique ordered field that is devoid of both types of limitations. That is, **No** not only exhibits *all possible algebraic and set-theoretically defined order-theoretic gradations consistent with its structure as an ordered field*, it is to within isomorphism the unique structure that does. It is ultimately this together with a number of closely related results that underlies our contention that **No** may be naturally regarded as an absolute arithmetic continuum (modulo NBG).

Whereas every absolute linear continuum is absolutely dense, there is a vast array of absolutely dense ordered classes that are not absolute linear continua. Simple examples are the semiopen and nondegenerate closed intervals of absolute continua. While these ordered classes are absolutely dense, they are not absolutely extensive. More distinguished examples are the convex proper subrings of **No** (Ehrlich and Kaplan 2018, Theorem 7.2 and pp. 15–16). On the other hand, there is no absolutely dense ordered field that is not an absolute linear continuum. To demonstrate this, we make use of the following definitions which are straightforward extensions for ordered classes of definitions for ordered sets due to Hausdorff (1914, ch VI; also see Harzheim 2005, pp. 77–79).

A subclass Y of an ordered class X is said to be *cofinal* (*coinitial*) with X if for each $x \in X$ there is a $y \in Y$ such that $y \geq x$ ($y \leq x$). Moreover, an ordered class X is said to have *cofinal character* α (*coinitial character* α) if α is the least ordinal $\leq On$ such that there is a cofinal (coinitial) subclass of X that is isomorphic with α (α^*), where α^* is the inverse of α. In addition, if x is a member of an ordered class X, then x is said to have *character* (α, β) if $\{y \in X : y < x\}$ has cofinal character α and $\{y \in X : y > x\}$ has coinitial character β.

Lemma 1 *An ordered field is an absolute linear continuum if and only if it is absolutely dense.*

Proof Let A be an absolutely dense ordered field. By results of Roman Sikorski (1948, pp. 74–75: (ii), (iii), (v)), which readily extend from sets to classes, if x is a member of an ordered field A, then x has character (α, α), where α is both the cofinal and coinital character of A. Accordingly, to prove the nontrivial portion of the lemma, it suffices to show that the coinital character of the additive identity 0_A of A is On. For this purpose, suppose the coinital character of 0_A is $<On$. Then there is a cointial subset B of $A^+ = \{x \in A : x > 0_A\}$. And so, by absolute density, there is an $x \in A$ such that $\{0_A\} < \{x\} < B$, and hence since B is a cointial subset of A^+, $\{0_A\} < \{x\} < A^+$, which is impossible.

Accordingly, in light of Theorem 15 and Lemma 1, we have:

Theorem 16 $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ *is (up to isomorphism) the unique absolutely dense real-closed ordered field.*

11 Arithmetic Continua: Archimedean and Absolute

An ordered field A is said to be *Archimedean* if whenever a and b are positive members of A where $a < b$, there is a positive integer n such that $na > b$; if A is not Archimedean, it is said to be *non-Archimedean*. Since every ordered field A contains a unique isomorphic copy $\mathbf{Q}_A$ of the ordered field of rational numbers, an element a of A may be said to be infinitesimal if $|a|$ is less than every positive member of $\mathbf{Q}_A$ and it may be said to be infinite if $|a|$ is greater than every positive member of $\mathbf{Q}_A$. An ordered field is Archimedean if and only if it contains neither infinite nor nonzero infinitesimal elements. Non-Archimedean ordered fields, by contrast, contain infinite as well as nonzero infinitesimal elements, the latter being the multiplicative inverses of the former. The zero of the field is the sole infinitesimal member of an Archimedean ordered field.

Before concluding our discussion of $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$, we wish to shed a bit more light on its inclusiveness as an ordered field, its relation to the classical *Archimedean arithmetic continuum* $\langle \mathbf{R}, +, \cdot, <, 0, 1\rangle$, and its amorphousness as a system of numbers great and small.

To begin with, every categorical characterization of the ordered field of surreal numbers expresses in one way or another just how rich in types of numbers great and small **No** really is. Unlike the ones offered in Theorems 15 and 16, the two provided in Theorem 17 below are of a model-theoretic nature. The first is based upon

Definition 8 (**Ehrlich 1987, 1988**) An ordered field A is said to be universally extending if for each ordered subfield B of A of power $<On$ and each ordered field A' extending B, there is an isomorphism $f: A' \rightarrow A$ that is an extension of the identity map on B.

Intuitively speaking, A is a universally extending ordered field if every possible way of enriching the ordered field-theoretic gradations of an ordered subfield B of A that is consistent with NBG is already (isomorphically) realized as an extension of B in A, if the universe of B is a set. The reader will notice that insofar as every ordered field is an extension of the ordered field of rationals, every universally extending ordered field contains an isomorphic copy of every ordered field. There is, however, a plethora of ordered fields that are inclusive in this sense but are not universally extending. However, as we shall shortly see, the following definition provides the means to distinguish between the universally extending ordered fields and those that are "merely" universal.

Definition 9 (**Ehrlich 1987, 1988**) *An ordered field A is said to be homogeneous-universal if it is universal – every ordered field can be embedded in A, and it is homogeneous – every isomorphism between ordered subfields B and B′ of A of power <On can be extended to an automorphism of A.*

Since an automorphism of A is an isomorphism of A onto itself, the homogeneity condition essentially ensures that any pair of ordered subfields of A whose universes

are sets that are structurally indistinguishable have structurally indistinguishable surroundings as well. As a result, any role that can be played in A by a given ordered subfield whose universe is a set can be played equally well by any of its other isomorphic counterparts in A.

Mimicking the corresponding definitions for ordered fields, an Archimedean ordered field A is said to be *homogeneous-universal* if every Archimedean ordered field can be embedded in A and every isomorphism between ordered subfields of A can be extended to an automorphism of A; and an Archimedean ordered field A is said to be *universally extending* if for each ordered subfield B of A and each Archimedean ordered field A' extending B, there is an isomorphism $f : A' \to A$ that is an extension of the identity map on B. The reader will notice that in the above two definitions implicit use is made of the fact that every subfield of an Archimedean ordered field is itself an Archimedean ordered field whose universe is a set.

The following theorem helps to bring to the fore the intimate relation between the classical arithmetic continuum and the purported absolute arithmetic continuum considered as inclusive ordered fields.

Theorem 17 (**Ehrlich 1987, 1989a, 1992**) *(I) The following sets of axioms constitute (categorical) axiomatizations of* $\langle \mathbf{R}, +, \cdot, <, 0, 1 \rangle$*; (II) by deleting the Archimedean axiom from the following axiomatizations, one obtains categorical axiomatizations of* $\langle \mathbf{No}, +, \cdot, <, 0, 1 \rangle$.

Axioms for ordered fields.

Archimedean axiom.

Axiom of (arithmetic) homogeneous-universality,

or, alternatively,

Axiom of (arithmetic) universal extensibility,

whereby the axiom of arithmetic homogeneous-universality (universal extensibility), we mean the assertion: The collection of numbers together with the corresponding relations defined on it constitutes a homogeneous-universal (universally extending) model of the above stated axioms.

Intuitively speaking, Theorem 17 asserts that whereas $\langle \mathbf{R}, +, \cdot, <, 0, 1 \rangle$ is (up to isomorphism) the unique ordered field containing all possible types of numbers great and small modulo the Archimedean axiom, $\langle \mathbf{No}, +, \cdot, <, 0, 1 \rangle$ is (up to isomorphism) the unique ordered field containing all possible types of numbers great and small modulo NBG. We believe the axioms of arithmetic homogeneous-universality and universal extensibility may be naturally regarded as *continuity axioms* since in the context of the above axiomatizations they are equivalent to any of the more familiar continuity conditions including those due to Cantor (1872), Dedekind (1872), and Hilbert (1900, 1971). An analogous remark likewise applies to their geometric counterparts expressed in Theorem 22 below.

We hasten to emphasize, however, that the import of the homogeneity portion of the first continuity condition differs substantially from the Archimedean case to the non-Archimedean case. This is a consequence of the fact that whereas $\langle \mathbf{R}, +, \cdot, <, 0, 1 \rangle$ contains exactly one isomorphic copy of each Archimedean ordered field,

$\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ contains a plethora of isomorphic copies of most ordered fields whose universes are sets. To be more precise, if A is an ordered field whose universe is a set and A is not isomorphic to a subfield of $\mathbf{R}_{alg}$, then $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ contains a proper class of pairwise distinct isomorphic copies of A. Thus, whereas the homogeneity condition holds vacuously in $\langle \mathbf{R}, +, \cdot, <, 0, 1\rangle$ (since every isomorphism between subfields of **R** extends solely to the identity automorphism of **R**), it points to a genuine amorphousness in $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ that applies not only to the members of **No** outside of $\mathbf{R}_{alg}$ but also to a vast array of subfields of **No** as well.

In the following section, we will see that as striking as the relation between the Archimedean and absolute arithmetic continua is, the relation between the *s-hierarchical Archimedean arithmetic continuum* and the *s-hierarchical absolute arithmetic continua* is more striking still.

12 S-Hierarchical Arithmetic Continua: Archimedean and Absolute

As we alluded to above, an ordered field is Archimedean if and only if it is isomorphic to a subfield of **R**. That is, **R** is up to isomorphism the unique universal Archimedean ordered field. **R** is also (up to isomorphism) the unique *nonextensible* Archimedean ordered field, the condition of maximality or of nonextensibility being Hilbert's (1900, 1971, p. 45) classical continuity condition. Analogs of these results, however, fail to hold for **No** considered as an ordered field (Ehrlich 2001, Proposition 3). On the other hand, as Theorems 6, 7, and 8 show, analogous results do in fact hold for the s-hierarchical counterparts of $\langle \mathbf{R}, +, \cdot, <, 0, 1\rangle$ and $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$. This being the case, the latter results may be encapsulated as follows, where if A is an s-hierarchical Archimedean ordered field (resp. s-hierarchical ordered field), then A is said to be *universal* if every s-hierarchical Archimedean ordered field (resp. s-hierarchical ordered field) is isomorphic to an initial subfield of A, and A is said to be *maximal* (or *nonextensible*) if there is no s-hierarchical Archimedean ordered field (resp. s-hierarchical ordered field) that properly contains A as an initial subfield.

Theorem 18 (Ehrlich 2001) *(I) The following sets of axioms constitute (categorical) axiomatizations of* $\langle \mathbf{R}, +, \cdot, <, <_s, 0, 1\rangle$*; (II) by deleting the Archimedean axiom from the following axiomatizations, one obtains categorical axiomatizations of* $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$.

Axioms for s-hierarchical ordered fields.

Archimedean axiom.

Axiom of s-hierarchical universality,

or, alternatively,

Axiom s-hierarchical maximality,

whereby the axiom of s-hierarchical universality (axiom of s-hierarchical maximality), we mean the assertion: The collection of numbers together with the

corresponding relations defined on it constitutes a universal (maximal) model of the above-stated axioms.

Unlike Archimedean ordered fields, not every ordered field is isomorphic to an initial subfield of **No**. On the other hand, every ordered field has a theoretically distinguished extension to an ordered field that is. More specifically, by a classical result of Artin and Schreier (1926), every ordered field A has (up to isomorphism) a unique *real-closure*, i.e., a smallest real-closed ordered field containing A. If A is itself real-closed, then A is its own real-closure, otherwise the real-closure of A is a (cardinality preserving) proper extension of A. Accordingly, while the following result does not expose the full range of initial subfields of **No**, it shows that the spectrum of initial subfields of **No** encompasses a class that is as theoretically significant as it is robust.

Theorem 19 (**Ehrlich 2001**) *Every real-closed ordered field (*i.e.*, elementary arithmetic continuum) is isomorphic to an initial subfield of* **No**.

Since the initial subfields of **No** emerge by recursion from the empty set of surreal numbers, Theorem 19 tells us that every real-closed ordered field, beginning with $\mathbf{R}_{alg}$ and terminating with **No**, unfolds by recursion as the surreal numbers themselves unfold. While some of these elementary continua, including $\mathbf{R}_{alg}$, **R**, and **No**, recursively unfold in precisely one way, there are initial real-closed subfields of **No** that are isomorphic as ordered fields though not isomorphic as s-hierarchical ordered fields.

For the full spectrum of ordered fields that are isomorphic to initial subfields of **No**, see (Ehrlich 2001, Theorem 18; 2012, Theorem 14; Ehrlich and Kaplan 2021, Theorem 8.1).

As we have already seen, $\langle \mathbf{No}, +, \cdot, <, 0, 1\rangle$ is up to isomorphism the unique absolutely dense real-closed ordered field. This result does not take into account the s-hierarchical nature of **No**. We draw this section to a close by observing that when this aspect of **No** is taken into account the elementary arithmetic nature of **No** can be omitted from the characterization.

Theorem 20 $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$ *is up to isomorphism the unique absolutely dense s-hierarchical ordered field.*

Proof In virtue of Theorems 6 and 16, $\langle \mathbf{No}, +, \cdot, <, <_s, 0, 1\rangle$ is an absolutely dense s-hierarchical ordered field. For the converse, first observe that by Lemma 1 an absolutely dense s-hierarchical ordered field is an absolute linear continuum. Furthermore, by Theorem 4 and the definition of a complete lexicographically ordered binary tree, a lexicographically ordered binary tree is an absolute linear continuum if and only if it is complete if and only if it is isomorphic to $\langle \mathbf{No}, <, <_s\rangle$. Accordingly, to complete the proof, it suffices to note that in virtue of Theorem 7: (i) The only s-hierarchical mapping of $\langle \mathbf{No}, <, <_s\rangle$, and, hence, an isomorphic copy thereof, into

$\langle \mathbf{No}, <, <_s \rangle$ is a surjection, and (ii) the s-hierarchical mapping is an embedding of ordered fields.

13 Euclidean Geometric Continua: Archimedean and Absolute

Whereas Theorem 17 may be regarded as providing comparative categorial characterizations of **R** and **No** considered as arithmetic continua, Theorem 18 may be regarded as providing comparative categorial characterizations of **R** and **No** considered as s-hierarchical arithmetic continua. In this section, we provide the Euclidean analog of Theorem 17.

The relation between **No** and Euclidean geometry can be fleshed out using any of the standard Euclidean geometrical frameworks. For our purpose here, however, Tarski's system $\mathcal{P}$ (Tarski 1959; Schwabhäuser et al. 1983) provides an especially convenient formulation.

In Tarski's system $\mathcal{P}$, which is formulated in the first-order predicate calculus with "=" treated as logical identity, only points are treated as individuals and the only nonlogical predicates employed in the axioms are a ternary predicate B (where "$Bxyz$" is read *y lies between x and z* – the case when y coincides with x or z not being excluded) and a quarternary predicate $\equiv$ (where "$xy \equiv zu$" is read *x is as distant from y as z is from u*). Over the years, Tarski's system of axioms for $\mathcal{P}$ has undergone substantial evolution (Tarski and Givant 1999), culminating in a set of ten axioms (e.g., Schwabhäuser et al. 1983; Ehrlich 1997, p. 61) that are collectively equivalent to Hilbert's classical axioms of order, incidence, congruence, and parallelism (1899, 1971). That is, Tarski's system of axioms for $\mathcal{P}$, henceforth A_1–A_{10}, is equivalent to Hilbert's system of axioms for Euclidean geometry save Hilbert's continuity axioms – the axioms that limit the models of Hilbert's other axioms to isomorphic copies of classical Cartesian geometry over **R**. In place of the latter axioms, Tarski supplements A_1–A_{10} with an *elementary continuity scheme*, henceforth A_{11}, that restricts the models of A_1–A_{10} to those that are elementary equivalent to classical Cartesian geometry over **R**. The resulting system of geometry based on A_1–A_{11}, henceforth $\mathcal{E}$, is referred to as *complete elementary Euclidean geometry* in virtue of the completeness of $\mathcal{E}$.

To simplify readability, the strings of universal quantifiers that precede A_1–A_{11} are deleted below.

A_1. *Reflexivity axiom for equidistance.*

$$ab \equiv ba.$$

A_2. *Transitivity axiom for equidistance.*

$$ab \equiv pq \wedge ab \equiv rs \rightarrow pq \equiv rs.$$

A_3. *Identity axiom for equidistance.*

$$ab \equiv cc \rightarrow a = b.$$

A_4. *Axiom of segment construction.*

$$\exists x(Bqax \wedge ax \equiv bc).$$

A_5. *Five-segment axiom.*

$$a \neq b \wedge Babc \wedge Ba'b'c' \wedge ab \equiv a'b' \wedge bc \equiv b'c' \wedge ad \equiv a'd' \wedge bd \equiv b'd' \rightarrow cd \equiv c'd'.$$

A_6. *Identity axiom for betweenness.*

$$Baba \rightarrow a = b.$$

A_7. *Pasch's axiom.*

$$Bapc \wedge Bbqc \rightarrow \exists x(Bpxb \wedge Bqxa).$$

A_8. *Lower dimension axiom.*

$$\exists a \exists b \exists c(\neg Babc \wedge \neg Bbca \wedge \neg Bcab).$$

A_9. *Upper dimension axiom.*

$$p \neq q \wedge ap \equiv aq \wedge bp \equiv bq \wedge cp \equiv cq \rightarrow Babc \vee Bbca \vee Bcab.$$

A_{10}. *Euclid's parallel axiom.*

$$Badt \wedge Bbdc \wedge a \neq d \rightarrow \exists x \exists y(Babx \wedge Bacy \wedge Bxty).$$

A_{11}. *Elementary continuity schema.* All sentences in the language $\equiv$, B of the form:

$$\exists a \forall x \forall y(\phi(x) \wedge \psi(y) \rightarrow Baxy) \rightarrow \exists b \forall x \forall y(\phi(x) \wedge \psi(y) \rightarrow Bxby),$$

where φ and ψ are first order formulas in which the variables a, b, y do not occur in φ and the variables a, b, x do not occur in ψ.

Unlike $A_1 - A_{10}$, A_{11} is entirely modern in conception; roughly speaking, it is the infinite set of assertions obtained from the following formulation of the Dedekind continuity condition due to Tarski that results by replacing the explicit references to arbitrary sets of points X and Y with implicit references to those sets which are definable in the language $\equiv$, B by first order means.

Dedekind Continuity Axiom

$$\forall X \forall Y(\exists a \forall x \forall y(x \in X \wedge y \in Y \rightarrow Baxy) \rightarrow \exists b \forall x \forall y(x \in X \wedge y \in Y \rightarrow Bxby)).$$

The version of the Archimedean axiom appropriate for $\mathcal{E}$ is given by the following assertion from (Ehrlich 1997):

Archimedean Axiom.

Whenever Ba_0a_1c, *there is a finite set of points* $a_0, \ldots, a_n$ *such that* $a_0a_1 \equiv a_ia_{i+1}$ *and* $Ba_{i-1}a_ia_{i+1}$ *for all* $1 \leq i < n$ *and* Ba_1ca_n.

The conception that bridges the gap between the domains of number and Euclidean magnitude is the classical notion of a Cartesian space over an ordered field. In the most recent version of Tarski's framework (Schwabhäuser et al. 1983), this familiar concept assumes the following simple form.

A *Cartesian space* $C_2(F)$ *over an ordered field* $\langle F, +, \cdot, \leq \rangle$ is a structure $\langle A_F, B_F, \equiv_F \rangle$ where the relations B_F and $\equiv_F$ are defined on $A_F = \{(x_1, x_2) : x_1, x_2 \in F\}$ by the stipulations:

$$B_F\, xyz \text{ if and only if there is a } \lambda \in F \text{ for which}$$

$$0 \leq \lambda \leq 1 \text{ and } y_\alpha - x_\alpha = \lambda(z_\alpha - x_\alpha) \text{ for } \alpha = 1, 2;$$

$$xy \equiv_F uv \text{ if and only if } \sum_{\alpha=1}^{2} (x_\alpha - y_\alpha)^2 = \sum_{\alpha=1}^{2} (u_\alpha - v_\alpha)^2.$$

The relation between Cartesian spaces and models of $\mathcal{E}$ is given by the following classical result due to Tarski (1948, 1959).

Representation Theorem. M is a model of $\mathcal{E}$ if and only if M is isomorphic to $C_2(F)$ for some real-closed ordered field F.

To prove the representation theorem, Tarski first shows that a Cartesian space over a real-closed ordered field is elementary equivalent to $C_2(\mathbf{R})$, which is known to be a model of $\mathcal{E}$. And his proof of the converse essentially consists of two components, the first of which is an appeal to what may be called *the fundamental lemma of elementary Cartesian geometry*: If o and e are distinct points in a model $M = \langle A, B, \equiv \rangle$ of $\mathcal{P}$ and l is the line in M through o and e, then by letting o and e serve as the zero and unit, respectively, and by appealing to the Euclidean theory of proportions which can be derived from $\mathcal{P}$, one can define (in terms of the relations B and $\equiv$) operations $+_l$ and $\cdot_l$ on, and a relation $\leq_l$ between, any two points of l so that the resulting structure $F^l_{oe}(M) = \langle l, +_l, \cdot_l, \leq_l, o, e \rangle_M$ is an ordered field (that is, to within isomorphism, independent of the choices of l, o, and $e \neq o$). Employing $F^l_{oe}(M)$ together with lines l and l' and point e', where l and l' are perpendicular and have the point o in common and e' is the point on l' such that $oe \equiv oe'$, one introduces a system of Cartesian coordinates and determines the distance between pairs of points in accordance with the Pythagorean Theorem, and shows more generally: There is an isomorphism – henceforth *the coordinatization mapping* – from M onto $C_2\left(F^l_{oe}(M)\right)$. To complete the proof, it is shown that since the elementary continuity scheme holds in

M, $F^l_{oe}(M)$ is real-closed (Tarski 1959, pp. 21–22; Schwabhäuser et al. 1983, pp. 144–145). Without the elementary continuity scheme, one can only show that $F^l_{oe}(M)$ is a *Pythagorean ordered field*, i.e., an ordered field in which $\sqrt{a^2+b^2}$ is in the field whenever a, b are in the field.

The ordered field $F^l_{oe}(M)$, or any isomorphic copy thereof, may be said to be the *characteristic field* of M. Since every ordered field has (up to isomorphism) a unique real-closure, if M is a model of $\mathcal{P}$ then there is (up to isomorphism) a least model M' of $\mathcal{E}$ containing M, which may be said to be the $\mathcal{E}$-*closure* of M; moreover, the characteristic field of M' is a real-closure of the characteristic field of M. In particular, if $C_2\left(F^l_{oe}(M)\right)$ is a model of $\mathcal{P}$, there is (up to isomorphism) a unique $\mathcal{E}$-closure $C_2\left(F^{l'}_{oe}(M')\right)$ of $C_2\left(F^l_{oe}(M)\right)$ where $F^{l'}_{oe}(M')$ is a real closure of $F^l_{oe}(M)$, insofar as the two points o and e determine l' as well as l, since $l \subseteq l'$.

Since **No** is (up to isomorphism) the unique real-closed ordered field that is an absolute linear continuum, it is evident from the above and Lemma 1 that if M is a model of $\mathcal{E}$, then to ensure that its characteristic field is isomorphic to **No** it suffices to ensure that its characteristic field is itself an absolute linear continuum. This may be accomplished by supplementing the axioms for $\mathcal{E}$ with the following variation of Tarski's version of the Dedekind continuity axiom in which X and Y are understood to range over sets.

A_{12}. *Absolute Continuity Axiom* (Ehrlich 1987).

$$\forall X \forall Y[\exists a \forall x \forall y(x \in X \wedge y \in Y \rightarrow Baxy \wedge x \neq y) \rightarrow$$
$$\exists b \forall x \forall y(x \in X \wedge y \in Y \rightarrow Bxby \wedge b \neq x \wedge b \neq y)].$$

Indeed, using A_{12} in conjunction with A_1–A_{11} and Tarski's aforementioned representation theorem, one may prove the following result which the author stated without proof in (Ehrlich 1987).

Theorem 21 *A model M of $\mathcal{E}$ satisfies the absolutely continuity axiom if and only if M is isomorphic to C_2(**No**).*

Proof Let M be a model of $\mathcal{E}$, l be a line in M with distinct points o and e, and $F^l_{oe}(M)$ be the characteristic field of M. Further suppose X and Y are nonempty subsets of l for which $X < Y$, and let X' be a cofinal subset of X having a least member, say p. Such an X' plainly exists and $\{p\} \leq X' \leq Y'$, which implies $\exists a \forall x \forall y(x \in X' \wedge y \in Y \rightarrow Baxy \wedge x \neq y)$. Now suppose M satisfies A_{12}. Then $\exists b\, \forall x\, \forall y(x \in X' \wedge y \in Y \rightarrow Bxby \wedge b \neq x \wedge b \neq y)$. But then $X' < \{p'\} < Y$ for some $p' \in l$, which implies $X < \{p'\} < Y$ since X' is cofinal with X. This implies l is absolutely dense with respect to $<$, and, hence by Theorem 16, $F^l_{oe}(M)$ is isomorphic to **No**, which implies M is isomorphic to C_2(**No**). For the converse, suppose M is isomorphic to C_2(**No**), which implies l is an absolute linear continuum with respect to $<$. Then l is absolutely extensive with respect to $<$, which implies $\{p'\} < X < Y$ for some $p' \in l$ and, hence, $\exists a\, \forall x\, \forall y(x \in X \wedge y \in Y \rightarrow Baxy \wedge x \neq y)$. But this combined with absolute

density implies $\exists b\, \forall x\, \forall y(x \in X \wedge y \in Y \rightarrow Bxby \wedge b \neq x \wedge b \neq y)$ and, hence, the absolutely continuity axiom.

The following result, which was announced in (Ehrlich 1987) in anticipation of (Ehrlich 1989a, 1992), is the above-mentioned Euclidean analog of Theorem 17, where the Euclidean analogs of the concepts of a homogeneous-universal ordered field (resp. Archimedean ordered field) and of a universally extending ordered field (resp. Archimedean ordered field) are defined in the expected manners.

Theorem 22 *(I) The following sets of axioms constitute (categorical) axiomatizations of* $C_2(\mathbf{R})$*; (II) by deleting the Archimedean axiom from the following axiomatizations, one obtains categorical axiomatizations of* $C_2(\mathbf{No})$.

Axioms for complete elementary Euclidean geometry ($\mathcal{E}$*).*

Archimedean axiom.

Axiom of (geometric) homogeneous-universality,

or, alternatively,

Axiom of (geometric) universal extensibility,

whereby the axiom of geometric homogeneous-universality (geometric universal extensibility), we mean the assertion: The collection of points together with the corresponding relations defined on it constitutes a homogeneous-universal (universally extending) model of the above-stated axioms.

Proof (*Archimedean case*). Let M be an Archimedean model of $\mathcal{E}$, o and e be distinct points on a line l of M, and $F^l_{oe}(M)$ be the characteristic field of M generated by the choice of o and e. Since M is Archimedean, so is $F^l_{oe}(M)$, and hence there is a unique isomorphism f of $F^l_{oe}(M)$ onto a subfield of $\mathbf{R}$, say, A. But then f induces an isomorphism g of M onto $C_2(A) \subseteq C_2(\mathbf{R})$, which suffices to show $C_2(\mathbf{R})$ is a universal Archimedean model of $\mathcal{E}$. Now suppose M and M' are models of $\mathcal{E}$ such that $M, M' \subseteq C_2(\mathbf{R})$ and let f be an isomorphism from M onto M'. Further let l and l' be lines in M and M', respectively, where l' is the image of g, g being the restriction of f to l. Moreover, let $l_{\mathbf{R}}$ and $l'_{\mathbf{R}}$ be the lines in $C_2(\mathbf{R})$ that respectively extend l and l', let o and e be distinct points on l, and let o' and e' be the images of o and e under g. Then g is an isomorphism from $F^l_{oe}(M)$ onto $F^{l'}_{o'e'}(M')$, and there is a unique g' extending g that is an isomorphism from $F^{l_{\mathbf{R}}}_{oe}(M)$ onto $F^{l'_{\mathbf{R}}}_{o'e'}(M)$, the uniqueness following from the Archimedean nature of the structures. But g' induces an automorphism f' of $C_2(\mathbf{R})$ that is an extension of f, which establishes $C_2(\mathbf{R})$ is homogeneous. To complete the Archimedean portion of the proof, it suffices to observe: If a theory T has a homogeneous-universal model whose universe is a set, then A is a homogeneous-universal model of T if and only if A is a universally extending model of T, and any two universal models of T are isomorphic (Ehrlich 1992, p. 170: (iv)).

(*Absolute case*). By results of the author (Ehrlich 1989a; also see 1992, p. 171: (v)): If T is an $\forall\exists$-theory with an infinite model that satisfies the joint embedding and

amalgamation properties for models whose universes are sets, then T has up to isomorphism a unique homogeneous-universal model; moreover, a model of T is homogeneous-universal if and only if it is universally extending. Accordingly, since $\mathcal{E}$ is an $\forall\exists$-theory (Tarski 1948, 1959) with infinite model $C_2(\mathbf{No})$, to complete the proof it suffices to show that $C_2(\mathbf{No})$ has the following two properties:

Universality for set-models: *For any set-model M of $\mathcal{E}$ (in the language B, $\equiv$), there is an embedding $f : M \to C_2(\mathbf{No})$.*

Amalgamation for set-models: *For any set-models of A, M, and M' of $\mathcal{E}$ in the language B, $\equiv$ and embeddings $f : A \to M$ and $f' : A \to M'$, there are embeddings $g : M \to C_2(\mathbf{No})$ and $g' : M' \to C_2(\mathbf{No})$ such that $gf = g'f'$.*

The proof essentially shows that universality and amalgamation, which hold for **No** with respect to real-closed fields and ordered fields more generally (Ehrlich 1989a, 1992), transfer to $C_2(\mathbf{No})$.

(*Universality*). Let M be a set-model of $\mathcal{E}$, l be a line in M, and o, e be distinct points on l. To prove universality for set-models, it suffices to note that since **No** is a universal ordered field there is an embedding of $f : F^l_{oe}(M) \to \mathbf{No}$ which induces a canonical coordinatization mapping $g : C_2\left(F^l_{oe}(M)\right) \to C_2(\mathbf{No})$.

(*Amalgamation*). Let A, M, M', f, f' be as specified above, let a, b be distinct points in A, and let $F_{ab}(A)$ be the characteristic field of A. Further, let $F^l_{oe}(M)$ and $F^{l'}_{o'e'}(M')$ be characteristic fields of M and M' where o and e are the images under f of a and b and o' and e' are the images under f' of a and b, and let k and k' be the restrictions of f to $F_{ab}(A)$ and f' to $F_{ab}(A)$, respectively. Notice that since isomorphisms of models of $\mathcal{P}$ preserve collinearity, $k(F_{ab}(A)) \subseteq F^l_{oe}(M)$ and $k'(F_{ab}(A)) \subseteq F^{l'}_{o'e'}(M')$. Moreover, since **No** has the analogous amalgamation property for real-closed ordered fields, there are ordered field embeddings $h : F^l_{oe}(M) \to \mathbf{No}$ and $h' : F^{l'}_{o'e'}(M) \to \mathbf{No}$ such that $hk = h'k'$, and these in turn induce canonical coordinatization mappings $g : M \to C_2\left(F^l_{oe}(M)\right) \subset C_2(\mathbf{No})$ and $g' : M' \to C_2\left(F^{l'}_{o'e'}(M')\right) \subset C_2(\mathbf{No})$ of models of $\mathcal{E}$ such that $gf = g'f'$.

Corollary 6 *Same as Theorem 22 except that the axioms for $\mathcal{E}$ are replaced with the axioms for $\mathcal{P}$.*

Proof This follows immediately from Theorem 22, the fact that every model of $\mathcal{P}$ has (up to isomorphism) a unique $\mathcal{E}$-closure and a result due to the author (1989a, Sect. 2; 1992, p. 171: (vi)).

14 Bridging the Gap between the Domains of Discreteness and of Continuity

For each Archimedean model of $\mathcal{E}$, there is an infinite array of pairwise distinct embeddings of it in $C_2(\mathbf{R})$. More strikingly still, for each model of $\mathcal{E}$ whose universe is a set, there is an entire proper class of pairwise distinct embeddings of it in $C_2(\mathbf{No})$.

In virtue of these observations and the homogeneity of $C_2(\mathbf{R})$ and $C_2(\mathbf{No})$, the highly amorphous nature of these geometric continua is can be made apparent. However, there are familiar geometrical considerations that lead to much the same conclusion regarding the points in models of $\mathcal{P}$ more generally. For example, it is well known that every *Euclidean rigid motion* (products of translations, rotations, and reflections) of a model of classical Euclidean geometry is an automorphism of the model (e.g., Greenberg 2008, ch. 9; Venema 2006, ch. 12) and that this theorem of classical geometry carries over to arbitrary models of $\mathcal{P}$ (Hartshorne 2000, pp. 148–158; Schwabhäuser et al. 1983, p. 36). The Euclidean rigid motions or *isometries* of a model $\langle A, B, \equiv\rangle$ of $\mathcal{P}$ coincide with the one-to-one onto mappings $f : A \to A$ that satisfy the condition: $ab \equiv f(a)f(b)$ for all $a, b \in A$; that is, they are precisely the distance preserving surjections of the plane (e.g., Hartshorne 2000, p. 155: problem 17.2). On the basis of this, it is easy to see that every such model has at least as many distinct nontrivial automorphisms as there are points of the space.

Some indication of the vast range of the isometries of a model $\langle A, B, \equiv\rangle$ of $\mathcal{P}$ is given by the following well-known result (e.g., Hartshorne 2000, p. 149):

(i) For any pair of points $a, a' \in A$, there is an isometry f such that $f(a) = a'$.
(ii) For any three points $o, a, a' \in A$, there is an isometry f such that $f(o) = o$ and f maps the ray $\overrightarrow{oa}$ onto the ray $\overrightarrow{oa'}$.
(iii) For any line l of A, there is an isometry f such that $f(a) = a$ for all $a \in l$ and f interchanges the points on the two sides of l.

As the qualifying phrase "some indication" suggests, the isometries referred to in (i)-(iii) above do not exhaust the isometries of $\langle A, B, \equiv\rangle$, nor for that matter do the isometries exhaust the automorphisms of $\langle A, B, \equiv\rangle$. A more general class of automorphisms of $\langle A, B, \equiv\rangle$ is the *similarities*, that is, the one-to-one onto mappings $f : A \to A$ that satisfy the condition: $f(a)f(b) \equiv kab$ for all $a, b \in A$ where k is a positive element of the characteristic field of $\langle A, B, \equiv\rangle$. These automorphisms are called "similarities" because they map every triangle of a plane onto a similar triangle. In the case of an Archimedean model of $\mathcal{P}$, the similarities coincide with the automorphisms of the model (Greenberg 2008, p. 409). If the Archimedean condition fails, however, this need not be the case (Greenberg 2008, p. 472: problem 69). Nevertheless, on the basis of (i)-(iii) alone, it is evident that, unlike the ordered field of real numbers, the s-hierarchical ordered field of surreal numbers, and the s-hierarchical real-closed subfields of the surreals more generally, models of $\mathcal{P}$ are highly amorphous indeed. Whereas the former structures are rigid, the latter admit vast arrays of nontrivial automorphisms.

Despite this critical disparity, the idea of a Cartesian space over an ordered field provides the requisite vehicle for bridging the gap between the domains of arithmetic and geometry. Indeed, much as the Cartesian space over the ordered field of real numbers succeeds in bridging the gap between the domains of number and of classical Euclidean geometry, the Cartesian space over the ordered field of surreal numbers bridges the gap between the domains of number and of absolutely

continuous elementary Euclidean geometry. Underlying this success is the fact that the analytic representations afforded by these structures are intrinsically amorphous, and rightfully so, even though the ordered field of reals and the s-hierarchical ordered field of surreals are not. It is this unsung feature of the classical scaffolding that we believe is central to the resolution of the age-old problem recounted by Fraenkel, Bar-Hillel, and Levy in the opening passages of our text. Moreover, this resolution applies more generally to the relation between the elementary arithmetic continua considered as rigid s-hierarchical structures and their amorphous (non-s-hierarchical) geometrical counterparts.

15 Postscript

Since the embrace of the Cantor-Dedekind thesis that the ordered field **R** of real numbers should be identified as the "arithmetic continuum," we have been accustom to regarding the arithmetic continuum as a strongly rigid system of numbers. By contrast, the ordered field of surreal numbers, which is arguably **R**'s most natural non-Archimedean counterpart, does not share this property. Indeed, like $C_2(\mathbf{No})$, **No** is remarkably rich in nontrivial automorphisms and nontrivial endomorphisms more generally. It is worth emphasizing, however, that **No**'s lack of rigidity, let alone strong rigidity, is not uncommon among non-Archimedean real-closed ordered fields. In fact, at present, there appears to be no known rigid non-Archimedean real-closed ordered field whose existence is provable in ZFC (Enayat 2018), and, to the best of our knowledge, the question of their existence in NBG for class-structures more generally is open as well. It was the apparent absence of examples that led Saltzman to ask if there exists such structures, and, while not settling the matter for ZFC or NBG, Saharon Shelah has shown that the existence of such structures is provable in ZFC supplemented with "suitable set theoretic hypotheses" (1983, p. 183). Nevertheless, even if Saltzman's query was found to have a positive answer in ZFC or NBG, the rigid non-Archimedean real-closed ordered fields that would evidence this would not include the real-closed fields that are central to present-day non-Archimedean mathematics, including those that have emerged in the theory of Hahn fields, nonstandard analysis, the theories of ordered exponential and differential fields, and the theory of surreal numbers. It is only suitable non-trivial relational expansions of the latter structures that could exhibit the strong rigidity that is characteristic of their Archimedean counterparts. On the basis of such considerations as simplicity, elegance, and the power of the results that can be derived therefrom, we believe it is the s-hierarchical nature (or natures) that the just-said elementary arithmetic continua inherit from their recursive unfolding in the s-hierarchical absolute arithmetic continuum of surreal numbers that is the natural candidate for this requisite additional structure, where and when it is desired.

Acknowledgments This chapter is an updated, revised, and expanded version of a portion of a paper entitled The Absolute Arithmetic Continuum that was written for the unconsummated proceedings of the conference *The Continuum in Mathematics and Philosophy*, The Carlsberg Academy, Copenhagen, Denmark, November 25–27, 2004.

References

Alling N (1987) Foundations of analysis over surreal number fields. North-Holland, Amsterdam

Alling N, Ehrlich P (1986) An alternative construction of Conway's surreal numbers. C R Math Acad Sci Soc R Can VIII:241–246

Alling N, Ehrlich P (1987) Sections 4.02 and 4.03 of Alling (1987)

Artin E, Schreier O (1926) Algebraische Konstruktion reeller Körper. Abh Math Semin Univ Hambg 5:85–99. English translation: Algebraic construction of real fields. In: Rosen MR (ed) (2007) Exposition by Emil Artin: a selection. American Mathematical Society and London Mathematical Society, Providence, pp 273–283

Berarducci A, Mantova V (2018) Surreal numbers, derivations and transseries. J Eur Math Soc 20: 339–390

Cantor G (1872) Über die Ausdehnung eines Satzes der Theorie der trigonometrischen Reihen. Math Ann 5:122–113

Cantor G (1895) Beiträge zur Begründung der transfiniten Mengenlehre, Part I. Math Ann 46:481–512. English translation: Contributions to the founding of transfinite numbers, Article I. Translated, and Provided with an Introduction and Notes by Philip EB Jourdain, 1915. Reprinted: Dover Publications, New York, 1955, pp 85–136

Cantor G (1897) Beiträge zur Begründung der transfiniten Mengenlehre, Part II, Math Ann 49:207–462. English translation: Contributions to the founding of transfinite numbers, Article II. Translated, and Provided with an Introduction and Notes by Philip EB Jourdain, 1915. Reprinted: Dover Publications, New York, 1955, pp 137–201

Chang CC, Keisler HJ (1990) Model theory. North-Holland Press, Amsterdam. Reprinted: Dover Publishing Company, Mineola, New York, 2012

Conway JH (1976) On numbers and games. Academic, London. 2nd edn (2001). A K. Peters, Natick

Cuesta Dutari N (1954) Ordinal Algebra. Rev Acad Ci Madrid 48:103–145

Dedekind R (1872) Stetigkeit und irrationale Zahlen. Vieweg, Braunschweig. English translation: Continuity and irrational numbers. In: Ewald W (ed) (1996) From Kant to Hilbert: a source book in the foundations of mathematics, vol II. Clarendon Press, Oxford, pp 765–779

Drake F (1974) Set theory; an introduction to large cardinals. North-Holland, Amsterdam

Du Bois-Reymond P (1870–71) Sur la grandeur relative des infinis des fonctions. Ann Mat Pura Appl 4:338–353

Du Bois-Reymond P (1875) Ueber asymptotische Werthe, infinitäre Approximationen und infinitäre Auflösung von Gleichungen. Mathematische Annalen 8:363–414; (Nachträge zur Abhandlung: ueber asymptotische Werthe etc.: 574–576)

Du Bois-Reymond P (1877) Ueber die Paradoxen des Infinitärcalcüls. Math Ann 11:149–167

Du Bois-Reymond P (1882) Die allgemeine Functionentheorie I. Verlag der H. Laupp'schen Buchhandlung, Tübingen

Ehrlich P (1987) The absolute arithmetic and geometric continua. In: Fine A, Machamer P (eds) PSA 1986, vol 2. Philosophy of Science Association, Lansing

Ehrlich P (1988) An alternative construction of Conway's ordered field No. Algebra Universalis 25: 7–16. Errata 25 (1988): 233

Ehrlich P (1989a) Absolutely saturated models. Fundam Math 133(1):39–46

Ehrlich P (1989b) Universally extending continua. Abstracts of Papers Presented to the American Mathematical Society 10(1):15

Ehrlich P (1992) Universally extending arithmetic continua. In: Sinaceur H, Salanskis J-M (eds) Le Labyrinthe du Continu, Colloque de Cerisy. Springer, Paris, pp 168–177

Ehrlich P (1994) All number great and small. In: Ehrlich P (ed) Real numbers, generalizations of the reals, and theories of continua. Kluwer Academic, Dordrecht, pp 239–258

Ehrlich P (1995) Hahn's Über die nichtarchimedischen Grössensysteme and the development of the modern theory of magnitudes and numbers to measure them. In: Hintikka J (ed) From Dedekind

to Gödel: essays on the development of the foundations of mathematics. Kluwer Academic, Dordrecht, pp 165–213
Ehrlich P (1997) From completeness to Archimedean completeness: an essay in the foundations of Euclidean geometry. In: Tauber A, Kanamori A (eds) A symposium on David Hilbert. Synthese 110(1):57–76
Ehrlich P (2001) Number systems with simplicity hierarchies: a generalization of Conways theory of surreal numbers. J Symb Log 66(3):1231–1258. Corrigendum (2005) 70(3):1022
Ehrlich P (2002a) The absolute arithmetic continuum and the unification of all numbers great and small. In: Philosophical insights into logic and mathematics (Abstracts). Université de Nancy 2, Laboratoire de Philosophie et d'Historie des Sciences, Archive Henri Poincaré, Nancy, France, pp 41–43
Ehrlich P (2002b) Surreal numbers: an alternative construction (abstract). Bull Symb Log 8(3):448
Ehrlich P (2005) Continuity. In: Borchert DM (editor-in-chief) Encyclopedia of philosophy, 2nd edn, vol 2. Macmillan Reference USA, Farmington Hills, pp 489–517
Ehrlich P (2006) The rise of non-Archimedean mathematics and the roots of a misconception I: the emergence of non-Archimedean systems of magnitudes. Arch Hist Exact Sci 60:1–121
Ehrlich P (2010) The absolute arithmetic continuum and its Peircean counterpart. In: Moore M (ed) New essays on Peirces mathematical philosophy. Open Court Publishing, Chicago, pp 235–282
Ehrlich P (2011) Conway names, the simplicity hierarchy and the surreal number tree. J Log Anal 3 (1):1–26
Ehrlich P (2012) The absolute arithmetic continuum and the unification of all numbers great and small. Bull Symb Log 18(1):1–45
Ehrlich P (2021) Contemporary infinitesimalist theories of continua and their late nineteenth- and early twentieth-century forerunners. In: Shapiro S, Hellman G (eds) History of continua: philosophical and mathematical perspectives. Oxford University Press, Oxford, pp 502–570
Ehrlich P, Kaplan E (2018) Number systems with simplicity hierarchies: a generalization of Conway's theory of surreal numbers II. J Symb Log 83:617–633
Ehrlich P, Kaplan E (2021) Surreal ordered exponential fields. J Symb Log 86:1066-1115, ArXiv Math e-prints. http://arxiv.org/abs/1905.07298/
Enayat A (2018) Ali Enayat. https://mathoverflow.net/users/9269/ali-enayat. Rigid Non-archimedean Real Closed Fields, URL (version: 2018-02-23): https://mathoverflow.net/q/293463
Fornasiero A (2006) Embedding Henselian fields into power series. J Algebra 304:112–156
Fornasiero A (2016) Initial embeddings in the surreals. In: Mini-workshop: surreal numbers, surreal analysis, Hahn fields and derivations. Abstracts from the mini-workshop held December 18–23, 2016. Organized by A. Berarducci, P. Ehrlich and S. Kuhlmann, *Oberwolfach Reports* 13, No. 4, Gerhard Huisken, Editor in Chief, European Mathematical Society Publishing House, Zürich, pp 3325–3326
Fraenkel AA, Bar-Hillel Y, Azriel L (1973) Foundations of set theory (2nd rev edn). North-Holland, Amsterdam
Gonshor H (1986) An introduction to the theory of surreal numbers. Cambridge University Press, Cambridge
Greenberg MJ (2008) Euclidean and non-Euclidean geometries: development and history, 4th edn. W H Freeman, New York
Hahn H (1907) Über die nichtarchimedischen Grössensysteme. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Wien. Mathematisch-Naturwissenschaftliche Klasse 116 (Abteilung II a): 601–655
Hardy GH (1910) Orders of infinity, the "infinitärcalcül" of Paul Du Bois-Reymond. Cambridge University Press, Cambridge. Reprinted (1971) Hafner, New York
Hardy GH (1912) Properties of logarithmico-exponential functions. Proc Lond Math Soc 10(2):54–90
Hartshorne R (2000) Geometry: Euclid and beyond. Springer, New York
Harzheim E (2005) Ordered sets. Springer, New York

Hausdorff F (1907) Untersuchungen über Ordungtypen. Berichte über die Verhandlungen der königlich sächsischen Gesellschaft der Wissenschaften zu Leipzig. Matematisch-Physische Klasse 59:84–159. English translation: Investigations into order types IV, V. In: Plotkin JM (ed) (2005) Hausdorff on ordered sets. American Mathematical Society and London Mathematical Society, Providence, pp 113–171

Hausdorff F (1909) Die Graduierung nach dem Endverlauf. Abhandlungen königlich sächsischen Gesellschaft der Wissenschaften zu Leipzig. Matematisch-Physische Klasse 31:295–335. English translation: Graduation by final behavior. In: Plotkin JM (ed) (2005) Hausdorff on ordered sets. American Mathematical Society and London Mathematical Society, Providence, pp 271–301

Hausdorff F (1914) Grundzüge der Mengenlehre, Leipzig, von Veit. Reprinted New York, Chelsea, 1949

Hausdorff F (1927) Mengenlehre. W. de Gruyter, Berlin

Hessenberg G (1906) Grundbegriffe der Mengenlehre, Abhandlungen der Friesschen Schule, (Neue Folge), Band I, Heft IV, Göttingen. Reprinted: Vandenhoeck & Ruprecht, Göttingen, 1906

Hilbert D (1899) Grundlagen der Geometrie. Teubner, Leipzig

Hilbert D (1900) Les principes Fondamentaux de la géometrie. Gautier – Villars, Paris. French translation by L. Laugel of an expanded version of (Hilbert 1899)

Hilbert D (1971) Foundations of geometry, 10th edn. Open Court Publishing, LaSalle

Kuhlmann S, Matusinski M (2015) The exponential-logarithmic equivalence classes of surreal numbers. Order 32: 53–68

Levi-Civita T (1892–93) Sugli infiniti ed infinitesimi attuali quali elementi analitici. Reprinted in: Tullio Levi-Civita (1954), opere matematiche, memorie e note, vol primo 1893–1900. Nicola Zanichelli, pp 1–39

Levi-Civita T (1898) Sui Numeri Transfiniti. Reprinted in: Tullio Levi-Civita (1954), opere matematiche, memorie e note, vol primo 1893–1900. Nicola Zanichelli, pp 315–329

Mendelson E (2010) An introduction to mathematical logic, 5th edn. CRC Press, Boca Raton

Robinson A (1961) Non-standard analysis. Indag Math 23:432–440 Reprinted in: Luxemburg WAJ, Körner S (eds) (1979) Abraham Robinson selected papers, volume 2: nonstandard analysis and philosophy. Yale University Press, New Haven/London, pp 3–11

Robinson A (1966) Non-standard analysis, 2nd edition (1974). North-Holland, Amsterdam

Schwabhäuser W, Szmielew W, Tarski A (1983) Metamathematische methoden in der Geometrie. Springer, Berlin

Shelah S (1983) Models with second order properties. IV. A general method and eliminating diamonds. Ann Pure Appl Logic 25:183–212

Sikorski R (1948) On an ordered algebraic field, Soc Sci Lett Varsovie. C R Cl III Sci Math Phys 41: 69–96

Sinaceur H (1994) Calculation, order and continuity. In: Ehrlich P (ed) Real numbers, generalizations of the reals, and theories of continua. Kluwer Academic, Dordrecht

Stolz O (1883) Zur Geometrie der Alten, insbesondere über ein Axiom des Archimedes. Math Ann 22:504–519

Stolz O (1885) Vorlesungen Über Allgemeine Arithmetik. Erster Theil: Allgemeines und Arithmetik der Reelen Zahlen. Teubner, Leipzig

Tarski A (1939) The completeness of elementary algebra and geometry. In: Centre National de la Recherche Scientifique. Institut Blaise Pascal, Paris, 1967, pp 289–346. Reprinted in Givant S, McKenzie R (eds) The collected papers of Alfred Tarski, vol IV. Birkhäuser, 1986

Tarski A (1948) A decision method for elementary algebra and geometry. Rand Corporation, Santa Monica. 2nd edn (prepared for publication with the assistance of JCC McKinsey) University of California Press, Berkeley and Los Angeles, 1951

Tarski A (1959) What is elementary geometry? In: Henkin L, Suppes P, Tarski A (eds) The axiomatic method, with special reference to geometry and physics. North-Holland, Amsterdam, pp 16–29

Tarski A, Givant S (1999) Tarski's system of geometry. Bull Symb Log 5(2):175–214

van den Dries L, Ehrlich P (2001) Fields of surreal numbers and exponentiation. Fundam Math 167: 173–188

Venema G (2006) Foundations of geometry. Pearson/Prentice Hall, Upper Saddle River

Veronese G (1891) Fondamenti di geometria a più dimensioni e a più specie di unità rettilinee esposti in forma elementare. Tipografia del Seminario, Padova

Veronese G (1894) Grundzüge der Geometrie von mehreren Dimensionen und mehreren Arten gradliniger Einheiten in elementarer Form entwickelt. Mit Genehmigung des Verfassers nach einer neuen Bearbeitung des Originals übersetzt von Adolf Schepp. Teubner, Leipzig

Warner S (1965) Modern algebra, vol. I and II. Prentice Hall, Englewood Cliffs. Reprinted by Dover Publishing Company, New York, 1990

Section VI

Practices in Euclidean and Non-Euclidean Geometry

Bharath Sriraman

Euclidean and Non-Euclidean Geometry in the History and Philosophy of Mathematical Practice

Bharath Sriraman

Contents

1 Introduction 1722
References 1726

Abstract

In the history of mathematical practice, geometry serves as an important focal point through which numerous changes in practice and associated (mathematical) worldviews can be studied. For instance, eighteenth century Kantian views on the synthetic a priori nature of geometry obfuscate practices of the likes of Menelaus of Alexandria who lived 16 centuries before him and are rarely mentioned. One could argue that Kant's philosophy on the nature of space (and geometry) was relational to human perception and cognition, as opposed to specific mathematical discoveries per se. Tracing a curated selection of advances in Euclidean and non-Euclidean geometry is useful to understand the mathematical practices of the ancient Greeks, and the remarkable contributions of Arab mathematicians leading to post-Renaissance advances that ushered in the modern era with the work of Grothendieck, and Busemann in the "return to Euclid."

Keywords

Axioms · Euclidean geometry · Grothendieck · Kant · Menelaus of Alexandria · Non-Euclidean Geometry · Poincaré · Spherics

B. Sriraman (✉)
University of Montana, Missoula, MT, USA
e-mail: sriramanb@mso.umt.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_120

1 Introduction

In Kant's (1781) *Critique of Pure Reason*, there are several references to Euclidean geometry as a means to support some of the arguments made by the philosopher. For Kant, geometry was a science which determined the properties of space a priori. It did not, like arithmetic, draw its concepts from the general concept of magnitude in general, but rather it was concerned with a concept of space as a magnitude given a priori, built on pure intuitions of space and time. Kant used Euclidean geometry as an example to illustrate the nature of synthetic a priori knowledge, which he argued was necessary and universally true but not derived from experience. To him, geometry provided a perfect example of such knowledge because it was a science that dealt with properties of space, and the necessary truths of geometry could be known through reason alone, without having to rely on empirical observation. Geometry based on intuitions of space and time and the pure concepts that were derived from them were called pure concepts of understanding. These concepts, such as those of point, line, and plane, were not derived from experience but were innate in human cognition. Thus, geometry served as a powerful example of how our a priori concepts and categories shape our understanding of the world and contribute to our knowledge of it. By using geometry to support his claims about pure reason, Kant attempted to show that certain types of knowledge were not derived from experience but were instead based on innate structures of the mind. In this way, he sought to demonstrate that the human mind had a capacity for knowledge that was independent of experience. Categories that structured our experience of the world and allowed us to make sense of space and time were consistent with the axioms and principles of Euclidean geometry. These concepts were thus necessary and universally true for all human beings, regardless of their particular experiences.

Another philosopher who used similar concepts was Gottfried Wilhelm Leibniz albeit in a completely different context. Leibniz's concept of "necessary truths" was similar to Kant's concept of synthetic a priori knowledge in that both involved knowledge that was not derived from experience but was instead based on inherent structures of the mind. Poincaré (1905) used the distinction between analytic and synthetic propositions in his philosophy of mathematics to argue that mathematical concepts and principles were grounded in intuition and experience, while also possessing a necessary and universal character. Note that Poincaré does not separate intuition and experience like Kant did. For him, mathematical concepts and principles were not simply derived from logic and definitions, as in the case of analytic propositions. Rather, he argued that mathematical knowledge is grounded in intuition and experience, and that this knowledge is both synthetic and a priori. It is synthetic because it involves the combination of simple ideas to form more complex ones, and it is a priori because it is independent of experience. Conceivably, Poincaré's view of mathematical knowledge as synthetic a priori was influenced by Kant's ideas about synthetic judgments. However, Poincaré also introduced new elements to this account, such as the role of intuition and the idea that mathematical concepts and principles are constructed through a process of mental synthesis. Poincaré in his book *Science and Hypothesis* (*La Science et l'Hypothèse*), first published in French in 1902, discussed

the nature of mathematical knowledge and its relationship to empirical science. He argued that mathematical concepts and principles are not derived from sensory experience but are instead conditions for the possibility of experience. Poincaré developed his own original ideas about the role of intuition, convention, and creativity in the formation of mathematical concepts and theories. His specific criticisms of Kant's philosophy are primarily directed at Kant's concept of space and his theory of knowledge. Poincaré argued that Kant's notion of space as purely intuitive was problematic because it relied on an implicit assumption of Euclidean geometry. Non-Euclidean geometries posed a challenge to Kant's account of space, and that there was no way to justify the choice of one geometry over another as the correct a priori structure of space. The geometrical axioms were therefore neither purely and intuitively synthetic nor experimental facts.

Poincaré discussed the implications of non-Euclidean geometry for the axiomatic method in mathematics. The axiomatic method that involves starting with a set of basic axioms and deducing further statements from them was not the only way to develop a mathematical theory. The development of non-Euclidean geometry challenged the idea that mathematics is based on absolute truths that are independent of human experience. Non-Euclidean geometries showed that different systems of geometry could be developed, depending on the assumptions or axioms that were used. The choice of axioms was not a matter of absolute truth but rather a matter of convention or convenience. Therefore, although the axiomatic method was a powerful tool for organizing and systematizing mathematical knowledge, it was not the only way to develop mathematical theories. Other methods, such as the use of analogies or the use of physical intuition, could also play a role in the development of mathematical ideas.

The axiomatic method has limitations, particularly in areas of mathematics where the axioms were not well-understood or where there were no obvious starting points for the development of a theory. In such cases, he suggested other methods, such as the use of conjectures or the use of experimental data, might be more appropriate.

> *If geometry were an experimental science, it would not be an exact science. it would be subject to continual revision ... the geometrical axioms are therefore neither synthetic* a priori *intuitions nor experimental facts. They are conventions. Our choice among all possible conventions is guided by experimental facts; but it remains free, and is only limited by the necessity of avoiding every contradiction, and thus it is that postulates may remain rigorously true even when the experimental laws which have determined their adoption are only approximate. In other words the axioms of geometry (I do not speak of those of arithmetic) are only definitions in disguise. What then are we to think of the question: Is Euclidean geometry true? It has no meaning. We might as well ask if the metric system is true and if the old weights and measures are false; if Cartesian coordinates are true and polar coordinates are false. One geometry cannot be more true than another; it can only be more convenient.* (Poincaré 1905, pp.58–59)

In this section of the handbook, several chapters address the history and context of geometry. Bardi's chapter on *Cultures of Mathematical Practice in Alexandria* focuses on the mathematical practices in ancient Alexandria, particularly the work of Claudius Ptolemy and his commentators. Bardi discusses the historical context of the

development of mathematical practices in Alexandria, including the influence of Greek and Egyptian cultures, and the significance of the Library of Alexandria as a center of intellectual activity. Bardi examines Ptolemy's mathematical works, including the Almagest and its contributions to the development of mathematical practices. The chapter then shifts to an analysis of Ptolemy's commentators, who were scholars that elaborated on and criticized Ptolemy's mathematical works. Bardi examines the works of several commentators, including Theon of Smyrna and Pappus of Alexandria, and discusses their contributions to the development of mathematical practices in Alexandria. One gets an overview of the mathematical practices in Alexandria during the second to fourth centuries CE, with a focus on the work of Claudius Ptolemy and his commentators. It highlights the historical and cultural context of the development of mathematical practices in Alexandria and the significance of Ptolemy's contributions to the field. In another chapter, Bardi discusses how Copernicus used axiomatics to establish the foundations of his heliocentric model of the solar system. Bardi notes that Copernicus was influenced by the axiomatic approach to mathematical reasoning developed by Euclid in the Elements, but that Copernicus departed from Euclid's approach in some significant ways, applying axioms to astronomy and cosmology rather than geometry and making assumptions about the physical world and the movements of celestial bodies. Bardi argues that Copernicus' use of axiomatics was a major contribution to the development of modern astronomy and cosmology.

An earlier departure from a strictly Euclidean view is explained by Papadopoulos, in his chapter on *Menelaus' Spherics.* Several misconceptions about Menelaus are clarified such as the misleading notion that this work pertains to astronomy. Such exaggerations led some scholars to believe that the Spherics was written specifically for astronomical purposes. Papdopoulos writes that the origin of this misinterpretation can be traced back to Bjørnbo and Heath, who erroneously declared that a significant part of the Spherics is dedicated to astronomy. These two scholars were likely influenced by the idea that the only "sphere of interest" (pun intended) to the Greeks was the celestial sphere, a view that mathematicians like him who have delved into original sources, find implausible. The sphere and its intersections by planes (playing the role of straight lines) are a natural object of study, just as the cone with its intersection by planes is the subject of Apollonius' Conics, or the ordinary plane with its straight lines is the subject of Euclidean geometry. Furthermore, Heath's claim that Book III of Menelaus' Spherics belongs to "trigonometry" is incorrect, as spherical trigonometry (as the study of the relations between sides and angles of a triangle) was not developed until centuries later (by the Arabs). The only use of a theorem from Menelaus' Spherics, in ancient times, was made by Ptolemy in his Almagest. The theorem was used by Ptolemy in his computations and in establishing his astronomical tables. In conversations with Papadopoulos, he mentions that Menelaus' work was translated in English only as recently as 2017 from Arabic manuscripts. Before that his name was known in the mathematical community (which excludes historians or linguists) only for a theorem in projective geometry, which carries his name, because it was used by ancient authors, including Chasles and others who worked on projective geometry, and who referred to this theorem as "Menelaus' Theorem." Incidentally, this is the result that Ptolemy used in

the Almagest. Conceivably, because of the importance of this "Theorem of Menelaus" in projective geometry, every mathematician thought that this is what he did and that is all.

The fact that the only result of Menelaus that people knew was used in astronomy contributed to the belief that Menelaus' Spherics was written for its use in astronomy. Papadopoulos' chapter reveals the vital contributions of Arab mathematicians in preserving and propagating the Spherics, its profoundness, and the difficulty of the theorems. But this knowledge did not reach the modern period. Euler, Lagrange, and Lambert, who are the three main figures of eighteenth century spherical geometry, never mention Menelaus; they were not aware of his existence. Two of the three founders of non-Euclidean geometry, Gauss and Lobachevsky, were also strongly involved in spherical geometry, but they never mention Menelaus, and they were not aware of the Spherics. Non-Euclidean geometry at the beginning of the twentieth century was already a dead subject, that is, it was not a research field; it was used in mathematical research only for some work on automorphic functions, but people there were essentially using works of Klein and Poincaré. The real introduction of non-Euclidean geometry in mathematical research started in the 1970, with W. P. Thurston, and became capital. Thurston started with both spherical and hyperbolic geometry. His first notes contain a chapter on both. It turned out that spherical 3-manifolds are very easily classified, and the hyperbolic (Lobachevsky) were the interesting ones. Research on the big problems of low-dimensional topology (like the Poincaré conjecture) started using geometry in an essential way, and hyperbolic geometry became at the forefront of research.

Brückler and Stilinovic in *From Friezes to Quasicrystals* provide a detailed history of (Euclidean) symmetry groups. They begin with earliest possible appearances and make numerous connections with art, culminating in the geometric classifications by nineteenth century crystallographers, and the related work in the theory of groups in the second half of the nineteenth century. Their chapter also provides twentieth century generalizations to symmetry groups in higher dimensions and mention that this extended to non-Euclidean spaces and applications in theoretical physics. Another notable shift in mathematical practice found in the nineteenth century is explored by Pedeferri in *Visual Representations of Euclidean Geometry: Diagrammatic Reasoning in Oliver Byrne's Work*. Pedeferri highlights several Euclidean propositions in Byrne's visual representations and the diagrammatic reasoning, which presented Euclid's Elements in a visually striking format. Pedeferri examines how Byrne's color-coded diagrams effectively convey geometric concepts and aids the reader in understanding the arguments of mathematical reasoning. Byrne's work provides an alternative means of understanding the first six books of Euclid's Elements and the significance of visual representations in mathematical practice.

Morales' chapter titled *The Notion of Space in Grothendieck* explains how he aimed to unite arithmetic and geometry into a single entity called "space-number." This notion suggests a connection between the physical aspects of space and Grothendieck's mathematical endeavors. The essay explores the evolution of Grothendieck's ideas on space, starting with his foundational work in algebraic geometry, introducing schemes and toposes as generalizations of spaces. It then

delves into topics such as tame topology and the formulation of a geometry of forms. Grothendieck's perspective places form at the core of geometry, viewing it as the branch of mathematics that studies the aspects of the universe related to form. The chapter concludes by proposing that further exploration of the space-number concept could enhance our understanding of Grothendieck's notion of space, and it presents this as a future project. Poenauru (2016) wrote that Grothendieck's vision of algebraic geometry progresses from schemes to toposes and then to motives. Schemes provide an efficient way to encode both systems of equations and their solutions, going beyond the set of points to include irreducible subvarieties. Toposes, a generalization of Riemann's multivalued functions, replace the lattice of open sets with étale sets, leading to étale cohomology and sheaves. Grothendieck's ultimate dream, motives, encompass various cohomology theories, breaking morphisms among algebraic varieties and performing other intricate tasks. This field has seen significant contributions from Voevodsky, Kontsevitch, Connes, Marcolli, and others, extending into quantum field theory and theoretical physics (Poenaru 2016). Grothendieck's unique style of mathematics involves starting with the correct level of abstraction and generality, solving problems from the top down.

The twentieth century in mathematics has seen the birth and the spectacular development of several geometric fields: algebraic topology, geometric topology, differential topology, algebraic geometry, singularity theory, and there are others. There is, however, one domain, which has advanced slowly and continuously rather than by producing spectacular results. This is the domain called "metric geometry." It is today an essential domain of the geometric landscape. Even if there are predecessors – Karl Menger (1902–1985) is among those who come to mind – the main founders of this field are two: Alexandre Danilovitch Alexandrov (1912–1999), in the Soviet Union, and Herbert Busemann (1905–1994), who was German but who practically had his career in the United States. Alexandrov founded a school of metric geometry that had an important impact. His students included Alexei Pogorelov, Victor Zalgaller, and Grigorii Perelman, and his work influenced Yuri and Dmitri Burago, Sergei Ivanov, Misha Gromov, and many others. In a tribute to Alexandrov, Semen Samsonovich Kutateladze wrote that Alexandrov used to work under the slogan: "Return to Euclid," and that he used to say that "the pathos of contemporary mathematics is the return to Ancient Greece." Busemann worked also in the same spirit, a return to Euclid, but in relative isolation. It is this important aspect of twentieth century mathematics that Papadopoulos presents in the "return to Euclid," highlighting the spirit under which Busemann used to choose the problems he worked on.

References

Kant I (1781) Kritik der reinen Vernunft (Critique of pure reason). Johann Friedrich Hartknoch, Riga

Poenaru V (2016) Alexander Grothendieck. In: Papadopoulos A (ed) Handbook of Teichmüller theory, vol VI. European Mathematical Society, EMS - Publishing House GmbH

Poincaré H (1905) Science and hypothesis. Retrieved from http://www.gutenberg.org/ebooks/37157

Cultures of Mathematical Practice in Alexandria in Egypt: Claudius Ptolemy and His Commentators (Second–Fourth Century CE)

Alberto Bardi

Contents

1 Introduction 1728
2 Claudius Ptolemy and His Time: A Brief Overview of His Life and Works 1729
3 Ptolemy's Legacy in Alexandria: Pappus and Theon 1732
4 Ptolemy's Philosophy of Mathematical Practice: Astronomy, Astrology, and Astronomical Tables 1732
5 Ptolemy's Mathematical Practice in the Alexandrian Context: Texts, Languages, and Genres 1734
6 Socio-historical and Material Aspects of Ptolemy's Mathematical Practice 1738
7 Conclusion 1740
8 Cross-References 1742
References 1742

Abstract

Claudius Ptolemy's mathematical astronomy originated in Alexandria in Egypt under Roman rule in the second century CE and held for more than a millennium, even beyond the Copernican theories (sixteenth century). To trace the flourishing of such mathematical creativity requires an understanding of Ptolemy's philosophy of mathematical practice, the ancient commentators of Ptolemaic works, and the historical context of Alexandria in Egypt, a multicultural city which became a cradle of cultures of mathematical practices and blossomed into the Ptolemaic system and its several related outputs between the second and fourth centuries CE. Ptolemy's mathematical practice will be explored under three lenses: (a) an examination of the mathematical prose of Ptolemy's *Almagest* alongside the corresponding passages in its Alexandrian commentaries; (b) the social, material, and semantic layers of Ptolemy's mathematics; and (c) the conception of true philosophy in Ptolemy's theory of knowledge. Examining the history and

A. Bardi (✉)
Department of the History of Science, Tsinghua University, Beijing, China
e-mail: brd@tsinghua.edu.cn

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_111

philosophy of Ptolemy's mathematical practice in Alexandria (1) shows that Ptolemy's ideals were usually betrayed by the practitioners of his mathematics and (2) proves that its main aspect is constituted by mathematical knowledge arranged in tabular format, that is the so-called astronomical tables, to which Ptolemy's mathematics owes its successful legacy.

Keywords

Alexandria in Egypt · Almagest · Astronomical tables · Commentaries · Handy Tables · Pappus · Ptolemy · Tables · Theon of Alexandria

1 Introduction

The mathematical astronomy of Claudius Ptolemy originated in Alexandria in Egypt under Roman rule in the second century CE and held for more than a millennium, even beyond the Copernican theories of the sixteenth century (Boll 1894; Toomer 1975; Taub 1993; Feke 2018; Juste et al. 2020). To trace the origins of such mathematical creativity calls for an examination of Ptolemy's philosophy of mathematical practice, the ancient commentators of Ptolemaic works, and the historical context of Alexandria in Egypt, a multicultural city which became a cradle of cultures of mathematical practices and flourished into the Ptolemaic system and its several related outputs (e.g., astronomical tables) between the second and the fourth centuries CE.

Alexandria was founded in 331 by Alexander the Great during his campaign in Africa and the Middle East. Due to its favorable strategic position, being located at the crossroads of Asia, Africa, and Europe, Alexander established the capital of his empire in Alexandria. The city would soon benefit from cross-cultural encounters, making it a prominent cosmopolitan center. Historical documents account for Alexandria's population as composed of native Egyptians, Greeks, Jews, Persians, Ethiopians, Syrians, Romans, and Arabs (Fraser 1972; Netz 2020, 257–305). The credit for making Alexandria an intellectual hub of the Hellenistic world goes to Ptolemy the First (305–283 BCE), the general who took over control of Egypt upon Alexander's death. Inspired by the model of famous Greek intellectual circles such as those formed around Pythagoras, Plato, and Aristotle, Ptolemy decided that Alexandria should become the center of Greek culture in this new world.

Ptolemy's patronage for the arts and sciences took shape in the *Museion* (museum), a sacred place devoted to the Muses, daughters of Zeus and Mnemosyne, the personifications of memory, and their protector Apollo. Adjacent to the Museum, Ptolemy constructed a library, which was intended not only to house and preserve important works but also to be open to the general public. This famous library contained a huge number of volumes for the time and was meant to contain all human knowledge in one place (Blum 1991, 95–123). Ptolemy invited scholars from all over the world to Alexandria and supported them with scholarship grants. Consequently, all manner of scholars gathered at this Museion, which was akin to a society of fellows or an institute of advanced studies: mathematicians, poets,

philosophers, philologists, astronomers, geographers, physicians, historians, and artists. How such exceptional circumstances came about remains a matter of historiographical discussion: two factors seem to have deeply influenced the character of the culture which grew out of the mixture of peoples and scholars and the broadened physical horizons. First, the Alexandrians' commercial interests brought geographical and navigational problems to the fore and directed attention to materials, methods of production, and the improvement of skills. Second, because commerce was carried out by free people who were not segregated socially from the scholars, the latter became aware of and involved in the problems facing the people at large. As a result, scholars were induced to unite their flourishing theoretical studies with concrete, practical, scientific, and engineering investigations. Technical fields were pursued and extended, training schools were established, and sciences were fostered and cultivated. Later, the wars between Alexander's successors saw the success of the Roman power, with Egypt becoming a Roman province in 30 BCE.

Alexandria under Roman rule was the environment in which Claudius Ptolemy worked and formed the background in which his mathematical creativity could thrive and spread. The science of that time, usually described as "Hellenistic," benefitted from encounters with several traditions that merged and mingled thanks to the cultural institutions in Alexandria, namely, the Museion and the Library (Neugebauer 1957, 145–190). The literary culture of Alexandria became so rich and elaborate that texts would reflect connections between arts and sciences; mathematics played alongside poetry in new ways (Netz 2009). The accumulated mass of data in literature and science became so vast that early systems of bibliography and cataloguing had to be invented, and the art of philology could thrive (Blum 1991). In astronomy, scholars could draw on knowledge stemming from the Egyptian, Greek, and Babylonian traditions. Besides providing a huge volume of records of observational data for astronomers, the Greek and the Babylonian traditions had conceived theoretical frameworks which were influential for Ptolemy and his generation: the Greeks conceived the cosmos geometrically, while the Babylonians dealt with celestial phenomena arithmetically (Neugebauer 1957, 97–144). Ptolemy's mathematical astronomy would produce new knowledge by combining the two traditions (Neugebauer 1957, 191–207; Goldstein 2007).

Ptolemy's mathematical practice will be explored under three lenses: (a) an examination of the mathematical prose of Ptolemy's *Almagest* alongside the corresponding passages in its Alexandrian commentaries; (b) the social, material, and semantic layers of Ptolemy's mathematics; and (c) the conception of true philosophy in Ptolemy's theory of knowledge.

2 Claudius Ptolemy and His Time: A Brief Overview of His Life and Works

Claudius Ptolemy lived in Alexandria in the second century (ca. 100–ca. 170), a period which historiography usually describes with the term *pax romana* (Roman peace). The time was characterized by bilingualism (Greek and Latin) and saw four

renowned Roman emperors in power: Trajan (ruled 98–117), Adrian (ruled 117–38), Antoninus Pius (ruled 138–61), and Marcus Aurelius (ruled 161–80) (Goldsworthy 2016). People with access to education were usually studying toward a full command of Greek and Latin because the spheres of influence between the two idioms had been shaped through cultural-historical processes, so that eventually Greek became the language of science and philosophy, Latin the language of law, administration, and trade. Any Roman of the second century who was intellectually ambitious had to learn Greek; this was achieved chiefly with the help of Greek tutors or through years of study in Athens, Alexandria, or any other city in the eastern provinces (Brown 1971; Cribiore 2001; Watts 2012).

Ptolemy wrote his works in Greek, and they are the only source we have about his life, apart from later indirect sources. For instance, the tenth-century Byzantine lexicon *Suda* mentions Ptolemy as an Alexandrian philosopher. This may sound surprising to a modern readership, but astronomy and mathematics have been considered part of philosophy for centuries, and this was a common framework in Hellenism and the Middle Ages (Taub 1993; Feke 2018; Jones 2020). Without pretending to map all the semantic possibilities of *philosophy* in Antiquity and the Middle Ages, it is worth noting that Ptolemy himself provides a philosophical statement in *Almagest* Book 1 in which he claims that mathematics is the highest form of philosophy, thus it provides the highest form of knowledge (see below).

Ptolemy's polymathy is indeed well reflected in his scholarly outputs. To mention some of his works, on the sciences of the heavens he authored the *Almagest* or *Mathematical Syntaxis* (Toomer 1984), the *Handy Tables* (Tihon and Mercier 2011), the *Tetrabiblos* (Robbins 1940), and the *Planetary Hypotheses* (Goldstein 1967; Hamm 2011). On other sciences, the *Geography* (Berggren and Jones 2000), *Optics* (Mark Smith 1996), the *Harmonics* (Swerdlow 2004), and also a philosophical work, the *Criterion*, alongside further, lesser explored works (Toomer 1975; Jones 2020).

The *Almagest* is Ptolemy's most successful work and is essential to understanding his mathematical practice (Toomer 1984; Goldstein 2007; Pedersen 2011). In the text, it is evident how Ptolemy was benefitting from Babylonian, Egyptian, and Greek astronomical traditions to generate a new mathematical system to be applied to astronomy: in sum, Babylonian sexagesimal system and astronomical parameters alongside observational records, Egyptian arrangement of calendar and division of day, and Greek geometrical devices to account for celestial motions.

The opus consists of 13 books. The first two are introductory, explaining astronomical assumptions and mathematical methods. Ptolemy proves the sphericity of the Earth and postulates the sphericity of the heavens and their revolution around the Earth immobile in the center. He discusses and redetermines the obliquity of the ecliptic by using trigonometry and relying on previous astronomers (or philosophers), Hipparchus and Menelaus of Alexandria. Every distance on the sphere is an angular one; the measurement of angles is replaced by the consideration of the chords subtending the corresponding arcs. The circle is divided into 360 parts and the diameter into 120 parts. Ptolemy used sexagesimal numbers: each of the

60 parts of the radius was divided into 60 small parts, and these again were divided into 60 smaller ones. A table of chords (sides of regular polygons) was computed for every half degree, each chord being expressed in parts of the radius, minutes, and seconds. The table of chords is followed by a geometrical argument leading to the calculation of the relations of arcs of the equator, ecliptic, horizon, and meridian. The same kind of discussion is continued in Book 2 with reference to the length of the longest day at a given latitude. Book 3 deals with the length of the year and the theory of the Sun, with the exposition of the models of eccentrics and epicycles, both probably already developed by the Greek astronomer Apollonius of Perga around the second century BCE. Book 3: length of the month and theory of the Moon. Book 4: construction of the astrolabe; theory of the Moon continued; diameters of the Sun, Moon, Earth's shadow, distance of the Sun, dimensions of the Sun, Moon, and Earth. Book 6: solar and lunar eclipses. Books 7–8: catalogue of the Stars (Graßhoff 1990); precession of the equinoxes; the Milky Way; and the construction of a celestial globe. Books 9–13: planetary motions. Book 9: order of the planets according to their distances from the Earth and periods of revolution; Mercury. Book 10: Venus; Book 11: Jupiter and Saturn; Book 12: stationary points and retrogressions; Book 13: motions of planets in latitude, inclinations, and magnitudes of their orbits.

The goal of the *Almagest* was to educate its readers to reach a contemplative status of the laws governing the heavens and thus become god-like, in the Platonic sense. Not accidentally, the collection of Greek poems from Late Antiquity known under the title *Palatine Anthology* provides an epigram ascribed to Ptolemy, in which the possibility of reaching divine realms in Platonic terms through the study of astronomy and detachment of the self from the earth is put into verse (Page 1981, 112).

Among the outcomes of the *Almagest* is to provide, to use a current word, the "know-how" to construct astronomical tables. A full command of the "know-how" of the astronomical tables enables the drawing of schematic maps of the heavens for the present, the past, and the future. Ptolemy himself produced such a set of tables, the *Handy Tables*, which is a structured set of astronomical tables that he composed after completing the *Almagest*, largely adapting them from the tables embedded in that treatise (Tihon and Mercier 2011). An astronomical table is an arrangement of numbers (plus other graphic signs), which stand for the values assumed by a quantity as a function of values of an argument. The tables can be considered a quantitative representation of the geometric model, in a tabular format. Practically, on the one hand, a table is approachable to people uninterested in the mathematical theory; on the other hand, it is also deemed the aim of tabular astronomy, for Ptolemy and his peers regard table-making as a good astronomical practice: these two aspects will prove crucial for the outcomes of Ptolemy's mathematical practice. While a table never appears on its own in the *Almagest*, but rather as a constitutive element of a mathematical argument that shows how the table is derived from the geometric model (Sidoli 2014), by contrast, the *Handy Tables* are tables on their own. The goal of Ptolemy's tables, in both *Almagest* and *Handy Tables*, is to show that apparently irregular motions are actually based on the periodic motions of circles.

Ptolemy's tables superseded other formats of astronomical knowledge, such as the tables of Babylonian tradition, which had been produced according to arithmetic models (Jones 1997a, b, 1998, 2000).

The tables acquired immense importance in antiquity and in the medieval traditions of the Western and Eastern Mediterranean and the Middle East: kings, caliphs, and emperors alike granted funding to have the best and most accurate astronomical tables (Kennedy 1956; King et al. 2001; Chabás and Goldstein 2012).

3 Ptolemy's Legacy in Alexandria: Pappus and Theon

The fourth-century Alexandrian mathematicians Pappus (Cuomo 2000) and Theon (Toomer 1976) are essential point of reference when exploring Ptolemy's mathematical practice. Between 300 and 350 CE, Pappus authored several works, the most famous being the *Synagoge* or *Collection* of mathematical problems (Jones 1986); his *Commentary on the Almagest* survives in fragments (Rome 1931–1945).

Theon was active in the second half of the fourth century. Remembered as the last attested member of the Museion, he was either a contemporary of Pappus or belonging to the generation immediately following him. He co-authored or accomplished the *Commentary on the Almagest* initiated by Pappus and he also published two commentaries on the *Handy Tables*, which are known as *The Great Commentary on the Handy Tables* (Mogenet and Tihon 1985; Tihon 1991; Tihon 1999) and *The Little Commentary on the Handy Tables* (Tihon 1978). The former is extensive, spanning five books: it explains not only how to use the tables but also the reasons behind the operations to be carried out and the theory behind the tables' construction and, accordingly, provides geometrical proofs. By contrast, *The Little Commentary* is just a collection of instructions: each chapter provides procedures on how to use the tables. Due to its practice-oriented nature, it would soon become the model for all subsequent generations of authors writing handbooks on how to use astronomical tables, from Late Antiquity till Early Modernity, in Christian, Jewish, and Islamicate contexts.

Pappus's and Theon's commentaries were chiefly meant to explain the difficult content of the *Almagest* to students and provide them with the basic knowledge to access its theorems (Jones 1999). It turned out that their practice of commenting on Ptolemy generated a new kind of mathematical practice, which, though conservative in nature, granted the tools to further develop Ptolemy's mathematics to subsequent generations – similar to what happened to other commentaries on Greek mathematics, the "deuteronomic texts," as Netz (1998) dubs them.

4 Ptolemy's Philosophy of Mathematical Practice: Astronomy, Astrology, and Astronomical Tables

As anticipated, according to Ptolemy's theory of knowledge, the highest form of philosophy is mathematics. More precisely, he attaches to this conception not just a mere theoretical dimension but a practical one too, intending the two sides to be mutually beneficial: all this is evident in the first passage of the *Almagest*:

> The true philosophers, Syrus, were, I think, quite right to distinguish the theoretical part of philosophy from the practical. For even if practical philosophy, before it is practical, turns out to be theoretical, nevertheless one can see that there is a great difference between the two: in the first place, it is possible for many people to possess some of the moral virtues even without being taught, whereas it is impossible to achieve theoretical understanding of the universe without instruction; furthermore, one derives most benefit in the first case [practical philosophy] from continuous practice in actual affairs, but in the other [theoretical philosophy] from making progress in the theory. Hence we thought it fitting to guide our actions (under the impulse of our actual ideas [of what is to be done]) in such a way as never to forget, even in ordinary affairs, to strive for a noble and disciplined disposition, but to devote most of our time to intellectual matters, in order to teach theories, which are so many and beautiful, and especially those to which the epithet 'mathematical' is particularly applied. (Toomer 1984, 35)

Furthermore, the notion of mathematics as the highest form of philosophy in Ptolemy is related to the knowledge and practice of astrology. Astrology, in Ptolemy's theory of knowledge, is the "physical" branch of the sciences of the stars, and the most influential systematization of knowledge labelable as "astrology" was redacted by Ptolemy himself in his *Tetrabiblos* (Robbins 1940). To understand astrology through the eyes of Ptolemy and the Alexandrian scholars, let us read his introduction to the *Tetrabiblos*:

> OF the means of prediction through astronomy, o Syrus, two are the most important and valid. One, which is first both in order and in effectiveness, is that whereby we apprehend the aspects of the movements of sun, moon, and stars in relation to each other and to the earth, as they occur from time to time; the second is that in which by means of the *physical* character of these aspects themselves we investigate the changes which they bring about in that which they surround. The first of these, which has its own science, desirable in itself even though it does not attain the result given by its combination with the second, has been expounded to you as best we could in its own treatise *by the method of geometrical proofs*. We shall now give an account of the second and less self-sufficient method in a properly philosophical way, so that one whose aim is the truth might never compare its perceptions with the sureness of the first, unvarying science, for he ascribes to it the weakness and unpredictability of material qualities found in individual things, nor yet refrain from such investigation as is within the bounds of possibility, when it is so evident that most events of a general nature draw their causes from the enveloping heavens. But since everything that is hard to attain is easily assailed by the generality of men, and in the case of the two before-mentioned disciplines the allegations against the first could be made only by the blind, while there are specious grounds for those leveled at the second—for its difficulty in parts has made them think it completely incomprehensible, or the difficulty of escaping what is known has disparaged even its object as useless—we shall try to examine briefly the measure of both the possibility and the usefulness of such prognostication before offering detailed instruction on the subject. (Translation Robbins 1940, 2–4; slightly reworked, my emphasis)

In sum, the mathematical branch of the science of the stars with a higher degree of certainty is the one whose certainty is granted by geometrical proofs, while there is a less self-sufficient method, which lies at the foundations of the physical branch of astronomy, that is, astrology, which hinges upon the certainty of the former branch. Astrology is knowable on the grounds of the regular occurrence of the same effects of the Sun and the Moon and of other regular occurrences observed by farmers and sailors. To these observational records, in order to be scientific, one needs to add an

accurate knowledge of the movements of the heavenly bodies. In other words, one needs to study the mathematics contained in the *Almagest*, which, in fact, as said, was meant to teach readers the mathematics needed to reach a contemplative status of the laws governing the heavens and thus become god-like, in the Platonic sense. Those who can reach the contemplative status are, by implication, able to construct astronomical tables as taught by the *Almagest*. As noted above, having a full command of the "know-how" of the astronomical tables allows one to have schematic maps (horoscopes) of the heavens for the present, the past, and the future (North 1986). They can be interpreted properly after studying the *Tetrabiblos*. On all these grounds, the true philosopher, according to Ptolemy, must learn both *Almagest* and *Tetrabiblos*, and that is the way to reach the full contemplation, the highest form of knowledge granted by the study of mathematics.

5 Ptolemy's Mathematical Practice in the Alexandrian Context: Texts, Languages, and Genres

An analysis of Ptolemy's mathematical practice can be undertaken via several approaches. For instance, (a) by recomputing the mathematical procedures and tables provided in Ptolemy's works (e.g., Van Brummelen 1993 recomputed Ptolemy's tables) or (b) by examining the *Almagest* and the corresponding passages in Alexandrian commentaries, paying attention to the language authors employed and the literary contexts in which they were working (they did not have to write "papers," their canonical genres were different). The latter has never been done in these terms so far. A case study in this concern is offered here.

Within such an approach, Ptolemaic texts must be understood in the cultural context of the production of knowledge in which Ptolemy, Pappus, and Theon were living, that is, the Museion and the Library of Alexandria, which embody and symbolize the institutionalization of the culture of mathematical practice that Ptolemy and his successors had developed. The products that have survived from that age are the texts, which act as both goals and tools for further development of the knowledge which in Alexandria was acquired, catalogued, and systematized (Renn 2020, 302–303, 308–309). In the case of astronomy, the knowledge stemming from Egyptian, Babylonian, and Greek astronomy could be systematized and used to generate the Ptolemaic system. On this account, the focus will be on the mathematical prose of Ptolemy, Pappus, and Theon.

The working questions are as follows: (1) What is the nexus between the *Almagest* and the commentaries of Pappus and Theon of Alexandria, why are the commentaries needed? (2) Are the commentaries didactic books, or how should they be conceived? (3) What is the role of the tables in Ptolemaic treatises? (4) What changes between Ptolemy and his commentators can be explained by historical circumstances, and how?

To study the Greek mathematical prose is to study a language that is basically composed of words, numbers, and diagrams. Due to historical grounds, it lacks the

level of radical formalization of contemporary mathematics, but its powerful linguistic resources were the basis for the development of Western mathematics, and its hypothetico-deductive structure shaped a solid cognitive substratum that pervaded Western mathematics till Modernity, as set out in *The Shaping of Deduction in Greek Mathematics: A Study in Cognitive History* (Netz 1999). A study of Ptolemy's mathematical discourse in terms of the function of individual units provided a classification of types of mathematical prose: theorem, problem, analysis, computation, table, and description (Sidoli 2004). A seminal study on the Greek language of Euclid has detected three different stylistic codes: the demonstrative code, the procedural code, and the algorithmic code (Acerbi 2021). The three studies mentioned in this paragraph have offered insightful interpretative tools to the present examination of Ptolemy's mathematical practice.

The present case study focuses on one of the most renowned and influential topics in the history of astronomy, namely, the assumption that celestial bodies are subjected to regular and circular motion, according to Aristotle's physical assumptions for the superlunar world. This assumption finds one of its most renowned textual witnesses in *Almagest* Book 3, in which the theory of the Sun is set out. Hinging upon the Aristotelian assumption, Ptolemy provides an epicyclic and eccentric hypothesis in order to account geometrically for the apparent variations in the motions of the Sun.

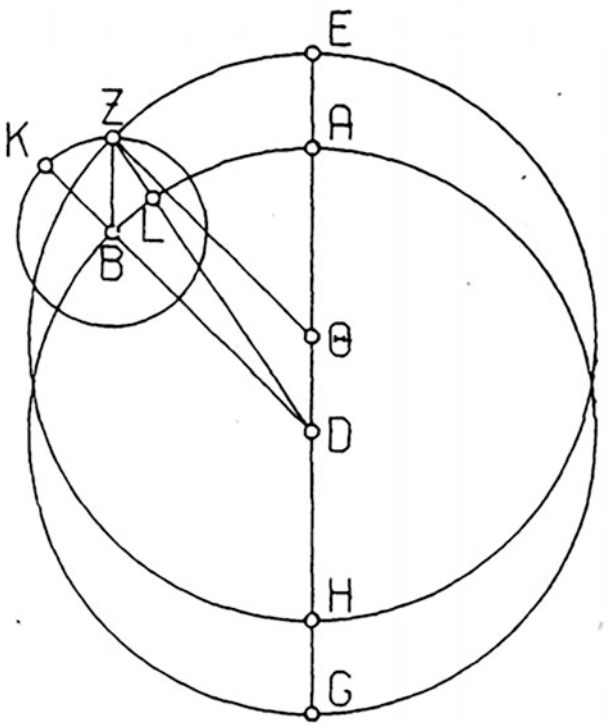
This subject, furthermore, was chosen because it is one of the most prominent in the history of astronomy. On the one hand, the assumption of regular circular motion was one of the issues for which several medieval astronomers criticized Ptolemy. On the other hand, the very same assumption was the basis for Copernicus's hypothesis of heliocentrism.

The equivalence of the eccentric and the epicyclic models is expounded, *inter alia*, in *Almagest* Book 3:

> Our next task is to demonstrate the apparent anomaly of the sun. But first we must make the general point that the rearward displacements of the planets with respect to the heavens are, in every case, just like the motion of the universe in advance, by nature uniform and circular. That is to say, if we imagine the bodies or their circles being carried around by straight lines, in absolutely every case the straight line in question describes equal angles at the centre of its revolution in equal times. The apparent irregularity [anomaly] in their motions is the result of the position and order of those circles in the sphere of each by means of which they carry out their movements, and in reality there is in essence nothing alien to their eternal nature in the 'disorder' which the phenomena are supposed to exhibit. The reason for the appearance of irregularity can be explained by two hypotheses, which are the most basic and simple. When their motion is viewed with respect to a circle imagined to be in the plane of the ecliptic, the centre of which coincides with the centre of the universe (thus its centre can be considered to coincide with our point of view), then we can suppose, either that the uniform motion of each [body] takes place on a circle which is not concentric with the universe, or that they have such a concentric circle, but their uniform motion takes place, not actually on that circle, but on another circle, which is carried by the first circle, and [hence] is known as the 'epicycle'. It will be shown that either of these hypotheses will enable [the planets] to appear, to our eyes, to traverse unequal arcs of the ecliptic (which is concentric to the universe) in equal times. (Translation Toomer 1984, 141)

This is followed by an exposition of the eccentric and the epicyclic hypotheses and their equivalence with reference to the motions of the Sun. Theorems, proofs, and diagrams are the basic elements of the prose. Theorems and proofs are written in a Greek that follows the canonic language of Greek geometry as found in Euclid's *Elements* (the demonstrative code, following Acerbi 2021).

The following diagram graphically resumes the equivalence of the two hypotheses. Imagine that we are on Earth, center of the universe, on point D, center of the circle ADG. We will end up finding the Sun at point Z on a certain time from a certain latitude either if we assume that it moves along the circle EΘH, with center Θ, *eccentric* from D, or if we take the Sun along the small circle (*epicycle*) whose center is B, which moves along the circle ADG.



(from Toomer 1984, 149)

Both hypotheses are structured according to a common scheme shared by practitioners of Greek geometry, initially developed by Euclidean scholars and later fixed into writing by Proclus (fifth century) in his *Commentary on the First Book of Euclid's Elements* (Morrow 1970). It reads as follows:

> Every problem and every theorem that is furnished with all its parts should contain the following elements: an *enunciation*, an *exposition*, a *determination*, a *construction*, a *proof*, and a *conclusion*. Of these the enunciation states what is given and what is being sought from

> it, for a perfect enunciation consists of both these parts. The *exposition* takes separately what is given and prepares it in advance for use in the investigation. The *determination* takes separately the thing that is sought and makes clear precisely what it is. The *construction* adds what is lacking in the given for finding what is sought. The *proof* draws the proposed inference by reasoning scientifically from the propositions that have been admitted. The *conclusion* reverts to the enunciation, confirming what has been proved. So many are the parts of a problem or a theorem. The most essential ones, and those which are always present, are enunciation, proof, and conclusion; for it is alike necessary to know in advance what is being sought, to prove it by middle terms, and to collect what has been proved. It is impossible that any of these three should be lacking; the other parts are often brought in but are often left out when they serve no need. (Translation Morrow 1970, 159, slightly revised; my emphasis)

Throughout the *Almagest*, Ptolemy takes Euclidean geometry for granted, and this seems to generate misunderstandings in the canonical style of Greek mathematics, especially if the reader is not trained in Euclidean geometry. Pappus and Theon intervene and provide the hidden knowledge in their commentaries (Jones 1999).

The corresponding passages of *Almagest* Book 3, in Pappus's and Theon's *Commentary on the Almagest* (Rome 1931–1945: III, 845 ff.) feature the demonstrative code and the procedural one. But, most importantly, their commentary is a paraphrase and summary of the subjects. The common elements of sentences are constituted by expressions like: "here he [meaning Ptolemy] says" + "direct quotations from the text of the Almagest." Moreover, Pappus and Theon provide passages from Euclid or other canonical works in circumstances which Ptolemy did not explain.

It is evident that the focus is on the tables, and certain keywords – such as *kanonographia* (writing tables) and *kanonopoiia* (constructing tables) – are emphasized (Rome 1931–1945: III, 891). The computations of the tables of the Sun feature the procedural code (Rome 1931–1945: III, 907 ff.), but also computations in the second person of the imperative mood to describe an operation, in a paratactic syntax (the algorithmic code, following Acerbi 2021), and these computations are aimed at summing up the operations expounded in the preceding part by applying them to a computation sample, according to the values on which tables are structured, that is, the cycle of years, months, days, hours, and latitude of Alexandria (e.g., Rome 1931–1945: III, 911).

Notably, the next chapter of the commentary (Rome 1931–1945: III, 912 ff.) is devoted to the correspondences between the computations in the *Almagest* and the *Handy Tables*. This shows that such computations were not just copied from the *Almagest* but reworked by Ptolemy himself.

Theon's *Great Commentary on the Handy Tables* treats the theory of the Sun (Mogenet and Tihon 1985, 103) by exposing how to derive the regular motions of the Sun to build tables and teaches how to read the tables. Proofs are combined with procedures. By contrast, the *Little Commentary on the Handy Tables* just provides procedures on how to use the *Handy Tables*, that is, how to work out the operations and which table to choose for a given problem, which column, which line. In the latter work, there is absolutely no requirement of understanding the theory underlying the tables.

To sum up:

1. There is a shift from complexity of syntax and morphology to a simplification (e.g., the same computations in different codes).
2. The notion of epicycle and eccentrics are progressively lost.
3. The diagram is unnecessary, thus absent, in the tables and in the commentary thereon.
4. According to Sidoli's units, while Ptolemy's *Almagest* and Pappus's and Theon's commentary thereon contain units 1 to 6 (see above), the focus on the tables leads to a deployment of just 5 (the tables) and 6 (description) in the *Handy Tables*. At the same time, there is a loss of theorems and problems but an increase in (3) analysis (of tables) and (4) computation (within the system of the tables) on the commentaries on tables.
5. The texts reflect the orientation to the practice that astronomy underwent after Ptolemy's *Almagest*:
   a. greater attention placed on the tables and how to use them (*kanonopoiia* and *kanonographia*);
   b. prevalence of the procedural code;
6. Process of systematization: in commenting on the *Almagest*, Pappus and Theon not only follow Ptolemy's division into chapters but also enumerate theorems to allow readers to orientate into the theory;
7. Commentators were likely operating when they perceived the prose as incomplete. The commentaries to the *Almagest* are therefore not to be understood as mere didactic tools or simplified versions of the work they are commenting on, but as proper, self-standing works (Chemla and Most 2022), which allowed the preservation and continuation of mathematical practice ("deuteronomic texts," as Netz (1998) puts it).

6 Socio-historical and Material Aspects of Ptolemy's Mathematical Practice

The case study offered above shows a shift of focus from the construction of tables toward the use of tables. At the same time, a huge change occurs at the epistemological level, after the second century, as an outcome of what practitioners of Ptolemaic mathematics were experiencing. The proofs that one was supposed to master to become a true philosopher were perceived as being too difficult. Therefore, practitioners chose to focus on the use of tables and to omit theorems and proofs. To be sure, the ideals of Ptolemy's philosophy of mathematical practice were betrayed, already in the Alexandrian context. Pappus and Theon provide evidence of the difficulty of the subject and the problems encountered by students in Alexandria. The social history of astronomy in the fourth century in Alexandria is recounted in Theon's *Commentary on the Almagest*, Book 1:

> What has been left without a commentary seems difficult. As for those people about whom Ptolemy, at the beginning of the work [i.e. the *Almagest*], says are willing to prove the whole topic geometrically, *they end up managing most of the topics in handy tables through simple methods*. On the contrary, we have not only put our best efforts into going through each topic through geometrical proofs as far as possible, but have also left nothing at all of difficult teaching, and we are not among those people who hold this opinion on the subject. (My translation; Rome 1931–1945: III, 318, lines 8–14, my emphasis)

Furthermore, Theon's *Little Commentary on the Handy Tables*:

> My Epiphanius, we have accurately provided a more rational method for the computation of the stars in the Handy Tables in another treatise in five books. However, because *the majority of those who followed our classes in order to learn the subject not only are unable to follow sufficiently the multiplications and divisions of the numbers*, *but also are completely ignorant of geometrical proofs*, we have done our best to systematically comment on them [the tables] by providing plain methods so that the exposition of the subject would be clearer to them [i.e. those who do not understand geometry]. (My translation; Tihon 1978, 199, lines 1–10; my emphasis)

Not surprisingly, students, forthcoming practitioners of Ptolemaic astronomy, recurred easily to tables and learned how to use them, without having any interest in the geometrical proofs underlying their construction. The Hellenistic world after the second century saw the emergence of the practice of tables without being taught in geometrical proofs become pervasive in scholarly realms. Such a practice was commonly perceived as "astronomy" and "science" in the population who had access to education. This conception of mathematical practice in astronomical sciences obviously became the most successful and spread in the subsequent centuries in several cultural contexts (Bardi 2021).

This social aspect of mathematical practice would attract criticism from all generations of scholars after Ptolemy. For instance, Ptolemaic works reached the Arabic world through translations from Syriac and Middle Persian. The elements one finds in Ptolemy and his commentators in terms of theory, practice, and differences in the nature of the two branches of the science of the stars were transmitted almost unvaried. For example, in ninth-century Baghdad, Abū Maʿšar composed *The Great Introduction to Astrology*, where he claims that true mathematicians/philosophers must have a full command of the *Almagest* and not just be able to use the astrological part and that, moreover, they must be able to construct tables and not just learn how to use them (Burnett 2002). Another later and emblematic example, in a European context this time, is Regiomontanus (Johannes Müller) in the fifteenth century, in his lecture at the University of Padua, where he lamented that the scholars of his time claimed to be astrologers just because they knew how to use the tables without having mastered the mathematics behind them (Byrne 2006; Rutkin 2019, 369–373; Omodeo 2021).

The shift from proofs to procedures as well as the increased attention being paid to the practical dimension of tables are also reflected in a semantic shift in the Greek word to designate the table, *canon*, a shift attested to in commentaries on the tables.

Passages which define the concept of astronomical table show that, while for Pappus and Theon the word *canon* means both the object and the knowledge contained in the table, in the first commentary modeled on Theon's, the word *canon* means just the object, not the mathematical knowledge provided by and systematized into the table. More precisely, in the fourth century, Theon of Alexandria wrote:

> we call "table" the whole knowledge of this kind of work, which is arranged into the table, even if it happens to occupy more than one page; [we call] "line" what is labeled with this name in the current fashion; [we call] "column" what is placed in vertical, which happens to have only degrees or only sixties – we call them again both "column." (My translation; Tihon 1978, 201, lines 4–9)

In the seventh century, the Byzantine scholar Stephen of Alexandria, rephrasing Theon's definition, built a new *Commentary on the Handy Tables*:

> we call "table" this kind of work, even if it happens to occupy more than one page; [we call] "line" what is labeled with this name in the current fashion, that is read from the left to the right; [we read] "column" what is read from above to the bottom. (My translation; Lempire 2016, p. 84, lines 15–18)

The semantic shift occurs in parallel with a material-historical factor, which is understandable when the historical role assumed by tables is seen in the light of the transmission process of Ptolemaic works between the second and fourth centuries. Briefly, the history of Ptolemy's mathematical practice took shape alongside one of the most prominent changes in the transmission of knowledge, namely, the shift from scroll to codex (Cavallo 1975; Roberts and Skeat 1983). This transition was not direct, similar to what one experiences nowadays in the transition from the physical book to the digital format. Initially, the *Almagest* was thought to be displayed in scrolls, but Theon redacted his commentaries with the codex in mind. In fact, in his commentary on the tables, Theon claims that tables might be contained by several pages, and scrolls do not have pages. This transition from scroll to codex is also reflected in the change in the format of the tables in the *Handy Tables* compared to that of the *Almagest*, which goes from 45 to 30 lines in each table, as well as in the change in the numbers of year cycles, from 18 to 25 (Acerbi 2020, 597). Ptolemy's commentators adapted Ptolemaic mathematical content to the new material support, thus irreversibly influencing the minds of new practitioners of Ptolemaic astronomy.

7 Conclusion

The historical and philosophical features of Ptolemy's mathematical practice are the outcome of exceptionally favorable historical, cultural, and material circumstances, which emerged in Alexandria in Egypt under Roman rule from the second to the fourth centuries CE, and thanks to patrons of arts and sciences and institutions which preserved culture, such as the Museion and the Library. In this context, astronomical knowledge stemming from Babylonian, Egyptian, and Greek traditions could be

preserved and employed to develop a new mathematical astronomy, whose major products are to be found in Ptolemy's *Almagest*, the *Handy Tables*, and Pappus's and Theon's commentaries thereon.

The astronomical tables became the main object of the culture of mathematical practice which took the *Almagest* as its main reference. In fact, all the examined texts refer to the tables in one way or another, but the way in which they referred to tables changed significantly. On this account, examinations of mathematical practice should consider the *Almagest* and the *Commentaries* as a *corpus* of texts as well as a collective effort of Alexandrian mathematicians to produce and use the best set of tables in their institutional and scholarly contexts.

To be sure, Ptolemy's tabular format has been successful over the tradition of tables in Alexandria. Non-Ptolemaic tables are generally not based on the principle of regular circular motion and are aimed at tracking apparent motions of celestial bodies, while the novelty of Ptolemy consists in providing tables which correspond to the mathematical principles of the geometrical models of the motions (Jones 1997a, b, 1998, 2000; Goldstein 2007). However, the success of the *Handy Tables* and, consequently, that of Theon's *Little Commentary on the Handy Tables* lies not just in the tabular format and its user-friendliness; it also derives from the stylistic resources within the language of Greek mathematics. The transmission of Theon's *Little Commentary* is huge in comparison with other commentaries, because his text employed a prose that was most accessible to its readers and could completely omit the theory. In a way, it is a fully practical text.

From the second to the fourth century, the relevance of the tables for Ptolemy's mathematical practice is present in the *Almagest* and increased in the commentaries thereon. This historical phenomenon is explainable by two interrelated factors: first, the difficulty of Ptolemy's mathematical theory (and the laziness of the majority of students and practitioners); second, the material requirement to have tables and to merely learn how to use them (practitioners of astrology did not need any knowledge of mathematical theory and astrological activity was an easy source of income). To further explore material aspects, the enhanced focus on tables in the history of Ptolemaic mathematical practice is parallel to the passage from scroll to codex, which has shaped the format and organization of the mathematical content of the tables. Therefore, the history of Ptolemy's mathematical practice is inextricably linked to the material features of scrolls and manuscripts, which are indeed "stratified social objects" (Ronconi 2018).

The changes that occurred in the practice of tables form part of a social history which is contradictory in comparison to Ptolemy's ideals of philosophy of mathematical practice. For Ptolemy, the true mathematical practice consists of a combination of a full command of mathematical astronomy and astrology, and it is aimed at reaching a contemplative status of divine realms (in a Platonic sense). In both branches of the Ptolemaic science of the heavens, the role of tables is indeed crucial, for the true philosopher must be able to master the theorems and proofs to build tables in addition to having the knowledge to interpret the horoscopes that one can obtain from them. However, as noted, the practical dimension of tables became prevalent, and most practitioners aimed at learning how to use them.

In sum, the astronomical tables, as an object of study of mathematical practice, are the place in which ideal, social, and material layers of Ptolemaic mathematics meet. It is no accident, then, that history would go on to see the table become the most influential object for the reception of Ptolemaic astronomy in later centuries, in the Islamicate, Byzantine, and European contexts of the Middle Ages, as well as in Early Modernity.

The commentaries on the *Almagest* and the *Handy Tables* are part of the culture of "deuteronomic texts" (Netz 1998), which preserved mathematics and its canons in Late Antiquity but also changed mathematical practice. In the case of Ptolemy and his commentators, the change is evident in the role that tables assumed from the *Almagest* through to Theon's commentary on the *Handy Tables*: initially thought of as objects inserted into a discursive mathematical prose, and as goals of the mathematical prose of the *Almagest*, they became the very first object of the mathematical practice, thus deemed worthy of commentaries thereon and sufficient knowledge to be mastered in order to become a professional astronomer.

8 Cross-References

- ▶ Shaping, Revisited
- ▶ The Metaphysics of Platonism
- ▶ The Social Epistemology of Mathematical Proof
- ▶ What Happens, from a Historical Point of View, When We Read a Mathematical Text?

References

Acerbi F (2020) Interazioni fra testo, tavole e diagrammi nei manoscritti matematici e astronomici greci. La conoscenza scientifica nell'Alto Medioevo, Centro Italiano di Studi sull'Alto Medioevo, Settimane di Studio del CISAM 67:585–621

Acerbi F (2021) The logical syntax of Greek mathematics. Springer, Cham

Bardi A (2021) Hybrid knowledge and the historiography of science: rethinking the history of astronomy between second-century CE Alexandria, Ninth-Century Baghdad, and Fourteenth-Century Constantinople. Transversal: Int J Historiogr Sci 2021(11):1–13

Berggren JL, Jones A (2000) Ptolemy's geography. An annotated translation of the theoretical chapters. Princeton University Press, Princeton

Blum R (1991) Kallimachos. The Alexandrian Library and the origins of bibliography. Translated from the German by Hans H. Wellisch. University of Wisconsin Press, Madison

Boll F (1894) Studien über Claudius Ptolemäus. Ein Beitrag zur Geschichte der griechischen Philosophie und Astrologie. Jahrbücher für classische Philologie, Suppl 21:49–224

Brown P (1971) The world of late antiquity. Thames and Hudson, London

Burnett C (2002) The certitude of astrology: The scientific methodology of al-Qabīṣī and abū Maʿshar. Early Sci Med 7(3):198–213

Byrne JS (2006) A humanist history of mathematics? Regiomontanus's Padua Oration in context. J Hist Ideas 67(1):41–61

Cavallo G (ed) (1975) Libri, editori e pubblico nel mondo antico. Laterza, Roma-Bari

Chabás J, Goldstein BR (2012) A survey of European astronomical tables in the late middle ages. Brill, Leiden/Boston
Chemla K, Most GW (eds) (2022) Mathematical commentaries in the ancient world. A global perspective. Cambridge University Press, Cambridge
Cribiore R (2001) Gymnastics of the mind: Greek education in Hellenistic and Roman Egypt. Princeton University Press, Princeton
Cuomo S (2000) Pappus of Alexandrian and the mathematics of late antiquity. Cambridge University Press, Cambridge
Feke J (2018) Ptolemy's philosophy: mathematics as a way of life. Princeton University Press, Princeton
Fraser PM (1972) Ptolemaic Alexandria. Clarendon Press, Oxford
Goldstein BR (1967) The Arabic version of Ptolemy's Planetary hypotheses. American Philosophical Society, Philadelphia
Goldstein BR (2007) What's new in Ptolemy's Almagest? Nuncius 22(2):261–285
Goldsworthy A (2016) Pax Romana. War, peace, and conquest in the Roman World. Yale University Press, New Haven
Graßhoff G (1990) The history of Ptolemy's star catalogue. Springer, New York
Hamm EA (2011) Ptolemy's planetary theory. An English Translation of Book One, Part A of the Planetary hypotheses with introduction and commentary. PhD dissertation, University of Toronto
Jones A (ed) (1986) Pappus of Alexandria. Book 7 of the collection. Sources in the history of mathematics and physical sciences, vol 8. Springer, New York
Jones A (1997a) Studies in the Astronomy of the Roman period. Centaurus 39:1–36
Jones A (1997b) Studies in the Astronomy of the Roman period. Centaurus 39:211–229
Jones A (1998) Studies in the Astronomy of the Roman period. Centaurus 40:1–41
Jones A (1999) Uses and users of astronomical commentaries in antiquity. In: Most GW (ed) Commentaries – Kommentare. Vandenhoeck & Ruprecht, Göttingen, pp 147–172
Jones A (2000) Studies in the Astronomy of the Roman period. Centaurus 42:77–88
Jones A (2020) The ancient Ptolemy. In: Juste et al (eds) Ptolemy's science of the stars in the middle ages. Brepols, Turnhout, pp 13–34
Juste D, van Dalen B, Hasse DN, Burnett C (eds) (2020) Ptolemy's science of the stars in the middle ages. Brepols, Turnhout
Kennedy ES (1956) A survey of Islamic Astronomical tables. Trans Am Philos Soc, New Series 46(2):123–177
King DA, Samsó J, Goldstein B (2001) Astronomical handbooks and tables from the Islamic World (750–1900): an interim report. Suhayl 2:9–105
Lempire J (2016) Le commentaire astronomique aux *tables faciles* de Ptolémée attribué à Stéphanos d'Alexandrie
Mark Smith A (1996) Ptolemy's theory of visual perception: an English translation of the "Optics" with introduction and commentary. Trans Am Philos Soc, New Series 86(2) p vii-xi+1-300
Mogenet J, Tihon A (1985) Le "Grand Commentaire" aux Tables Faciles de Théon d'Alexandrie aux Tables Faciles de Ptolémée. Livre I. Biblioteca Apostolica Vaticana, Città del Vaticano
Morrow GR (1970) Proclus. A commentary to the first book of Euclid's elements (trans: Morrow GR). Princeton University Press, Princeton
Netz R (1998) Deuteronomic texts: late antiquity and the history of mathematics. Revue d'histoire des mathématiques 4(2):261–288
Netz R (1999) The shaping of deduction in Greek mathematics. A study in cognitive history. Cambridge University Press, Cambridge
Netz R (2009) Ludic proof. Greek mathematics and the Alexandrian aesthetic. Cambridge University Press, Cambridge
Netz R (2020) Scale, space and canon in ancient literary culture. Cambridge University Press, Cambridge
Neugebauer O (1957) The exact sciences in antiquity, 2nd edn. Dover, New York

North J (1986) Horoscopes and history. The Warburg Institute, University of London, London
Omodeo PD (2021) Johannes Regiomontanus and Erasmus Reinhold. Shifting perspectives on the history of Astronomy. In: Brentjes S, Fidora A (eds) Premodern translation: comparative approaches to cross-cultural transformations. Brepols, Turnhout, pp 165–186
Page DL (1981) Further Greek epigrams. Cambridge University Press, Cambridge
Pedersen O (2011) A survey of the Almagest. With annotation and new commentary by Alexander Jones. Springer, New York/Dordrecht/Heidelberg/London
Renn J (2020) The evolution of knowledge. Princeton University Press, Princeton
Robbins F (1940) Ptolemy. Tetrabiblos. Harvard University Press, Cambridge, MA
Roberts CH, Skeat TC (1983) The birth of the Codex. The British Academy, London Rome A (1931–1943) Commentaires de Pappus et de Théon d'Alexandrie sur l'Almageste, 3 vols. Biblioteca Apostolica Vaticana, Città del Vaticano
Rome Adolphe (1931–1945) Commentaires de Pappus et de Théon d'Alexandrie sur l'Almageste, I-III vol. Città del Vaticano: Biblioteca Apostolica Vaticana
Ronconi F (2018) Manuscripts as stratified social objects. Scandinavian J Byzantine Modern Greek Stud 4:19–40
Rutkin D (2019) Sapientia Astrologica: Astrology, magic and natural knowledge, ca 1250–1800. I. Medieval structures (1250–1500): conceptual, institutional, socio-political, theologico-religious and cultural. Springer, Cham
Sidoli N (2004) Ptolemy's mathematical approach: applied mathematics in the second century. PhD dissertation, University of Toronto
Sidoli N (2014) Mathematical tables in Ptolemy's Almagest. Hist Math 41:13–37
Swerdlow NM (2004) Ptolemy's Harmonics and the "Tones of the Universe" in the Canobic inscription. In: Burnett C, Hogendijk JP, Plofker K, Yano M (eds) Studies in the history of the exact sciences in honour of David Pingree. Brill, Leiden, pp 137–180
Taub L (1993) Ptolemy's universe. The natural philosophical and ethical foundations of Ptolemy's Astronomy. Open Court Publishing Company, Chicago/La-Salle
Tihon A (1978) Le "Petit Commentaire" de Théon d'Alexandrie aux Tables Faciles de Ptolémée. Biblioteca Apostolica Vaticana, Città del Vaticano
Tihon A (1991) Le "Grand Commentaire" de Théon d'Alexandrie aux Tables Faciles de Ptolémée. Livres II et III. Biblioteca Apostolica Vaticana, Città del Vaticano
Tihon A (1999) Le "Grand Commentaire" de Théon d'Alexandrie aux Tables Faciles de Ptolémée. Livres IV. Biblioteca Apostolica Vaticana, Città del Vaticano
Tihon A, Mercier R (2011) Πτολεμαίου Πρόχειροι Κανόνες. Les Tables Faciles de Ptolémée, 2 vol. Peeters, Louvain
Toomer GJ (1975) Ptolemy. In: Gillispie CC (ed) Dictionary of scientific biography, vol XI. Scribner, New York, pp 186–206
Toomer GJ (1976) Theon of Alexandria. In: Gillispie CC (ed) Dictionary of scientific biography, vol XIII. Scribner, New York, pp 321–325
Toomer GJ (1984) Ptolemy's Almagest. Duckworth, London
Van Brummelen G (1993) Mathematical tables in Ptolemy's Almagest. PhD dissertation, Simon Fraser University
Watts E (2012) Education: speaking, thinking, and socializing. In: Johnson SF (ed) The Oxford handbook of late antiquity. Oxford University Press, Oxford, pp 467–486

Menelaus' Spherics in Greek and Arabic Mathematics

Athanase Papadopoulos

Contents

1 Introduction 1746
2 Two Predecessors: Theodosius and Autolycus 1748
3 Menelaus' *Spherics*: Preliminary Remarks 1758
4 The Propositions 1760
5 Pappus 1773
6 Translations, Rectifications and Difficulties 1775
7 On the Impact on Arabic Mathematics 1778
8 On the Impact on Astronomy 1780
9 As a Word of Conclusion 1782
Appendix I: A Proof of Proposition 27 by Spherical Trigonometry 1784
Appendix II: A Proof of Proposition 71 by spherical Trigonometry 1785
References 1785

Abstract

We present the history of Menelaus' *Spherics* (second c. AD), explaining its importance and quoting in detail some of the significant propositions. We include this work in the general context of Greek and Arabic mathematics, highlighting the far-reaching connections with modern mathematical ideas.

Keywords

Menelaus of Alexandria · Spherics · Spherical geometry · Spherical trigonometry · Greek mathematics · Arabic mathematic · Theodosius of Tripoli · Autolycus of Pitane · Pappus of Alexandria

A. Papadopoulos (✉)
Institut de Recherche Mathématique Avancée (IRMA) and Centre de Recherche et d'Expérimentation sur l'Acte Artistique (ITI CREAA), Université de Strasbourg and CNRS, Strasbourg Cedex, France
e-mail: athanase.papadopoulos@math.unistra.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_109

1 Introduction

Menelaus' *Spherics* (second c. AD) constitute one of the most remarkable achievements of Ancient Greek mathematics, along with Euclid's *Elements*, Apollonius *Conics*, Diophantus' *Arithmetica*, Archimedes' treatises on geometry and statics, Theodosius' *Spherics*, Pappus' *Collection*, to name only some of the most important Greek mathematical works. Among these treatises, Menelaus' *Spherics* is certainly the least understood by historians and the least familiar to mathematicians of the modern period. As a matter of fact, we know that Fermat was nourished by the work of Diophantus, that Christiaan Huygens was deeply immersed in the writings of Archimedes and Apollonius, that Euler was an assiduous reader of Diophantus and Pappus and that he was well acquainted with Theodosius' *Spherics*, not to mention of course Euclid's *Elements*, a work which all mathematicians of the past thoroughly studied. But very few of the modern mathematicians mention Menelaus' *Spherics*,[1] a fact that might seem surprising, especially since a relatively large number of medieval Arab mathematicians were well acquainted with this work.

I can see two reasons for which Menelaus' *Spherics* remains unknown. First, there are problems inherent to the available versions: no Greek manuscript has survived, and the ones we possess, in Arabic or translated from the Arabic, present intricate differences in form and wording. Another reason is that this work is difficult to access from the mathematical point of view. Indeed, the existing manuscripts do not contain proofs of several of the difficult propositions. I will say more on this below.

Menelaus' *Spherics* opened up a new research field. Saying this, we are not talking about the subject of the geometry of the sphere which, as a general topic, already existed (we mentioned Theodosius' *Spherics*, but the sphere was studied before Theodosius, by Autolycus, Euclid, Archimedes, and others), but of the field known today as the "intrinsic geometry of surfaces," which can be described, precisely like Menelaus investigated the geometry of the sphere, as the study of the properties of curvilinear triangles on a surface which is not the Euclidean plane. Indeed, the vast majority of Menelaus' propositions are concerned with such

[1] An exception is Marin Mersenne, who was fond of old mathematical texts and who was aware of this work. Indeed, Mersenne published in (Mersenne 1626) a Latin edition of the third book of the *Spherics*, together with works of Autolycus, Archimedes, Apollonius, Theodosius, Diophantus and Pappus. But one century later, Anders Lexell, who was Euler's closest collaborator, and who was a specialist of spherical geometry, writes in his memoir (Lexell 1784): "From the time where the *Spherical elements* of Theodosius were handed in written form, it is very difficult to find, in the treatises of the geometers, other elements concerning further refinements of the theory of the figures drawn on the surface of the sphere, than those that are usually exposed in the *Elements of Spherical Trigonometry* and that are aimed at solving spherical triangles." It is not completely clear to what work with the title *Elements of Spherical Trigonometry* Lexell refers (this may be Euler's *Principes de la trigonométrie sphérique tirés de la méthode des plus petits et des plus grands* (Euler 1753)), but it is obvious that he was not aware of Menelaus' work, written two centuries after that of Theodosius, nor of the works of Naṣīr al-Dīn al-Ṭūsī and the other Arab mathematicians of the eleventh to thirteenth century who worked extensively on spherical trigonometry.

properties. Menelaus' major achievement is that he highlighted for the first time the notion of spherical triangle, and he developed the bases for its study.

The Arabs included this work into their body of knowledge and they understood its novelty. Let us quote right away Aḥmad ibn Sa'd al-Harawī (tenth century) from the very beginning of his edition of Menelaus' treatise (English translation in (Rashed & Papadopoulos, 2017, p. 500)):

> The way that Menelaus followed in that book is marvelous; nobody preceded him and nothing had been known to his predecessors on that subject. Did he foresee that this kind of geometrical science is unique in itself, that it has laws suitable to it and lemmas that lead to the goal which are different from those that one is willing to demonstrate by using straight lines and the determination of planes by their intersections?

The sentence about "straight lines and the determination of planes by their intersections" refers to the fact that Menelaus avoided as much as possible the use of Euclidean solid geometry which is the ambient geometry of the sphere; the expression "straight lines and planes" refers to Euclidean lines and planes. Let us also quote Menelaus himself, from the oldest fragment known of the *Spherics*, which we also edited and translated into English in (Rashed & Papadopoulos, 2017, p. 504):

> I have invented a kind of proof which is excellent and admirable. Several things relative to surface <properties> on the spherical surfaces appeared to me; I don't think they were offered to anybody else. And I ordered the lemmas and the proofs in an order so that it will be easy for those who are fond of science to rise up and attain universal and honorable sciences. In this way, with the help of particular propositions that occur to them from the experiences they acquire from that art, they will become skillful and outstanding in the universal propositions, given the multitude of universal propositions that can occur to them, and that become for them proofs and indications for the inaccessible and the difficult directions in the principles of Spherics, their role and what ensues from them.

The field that Menelaus opened in his *Spherics* can be especially appreciated in the context of non-Euclidean geometry, a domain of mathematics that became a particularly active research field in the last quarter of the twentieth century.

In the pages that follow, I will review this work in its context, discussing at the same time the mathematics it contains, the history of the texts that we possess, and the impact of this work on the Arab mathematicians of the period ninth to thirteenth century, who are certainly, among the mathematicians who read this work, those who have most understood its importance.

The content of the rest of this chapter is the following:

In Sect. 2, I briefly review works on the geometry of the sphere by two predecessors of Menelaus, namely, Theodosius of Tripoli and Autolycus of Pitane, pointing out relations between these works and Euclidean geometry, making connections with the work of Menelaus, with a later work by al-Ḥasan Ibn al-Haytham and hinting to some relations with the development of modern analytic spherical geometry.

Section 3 contains some preliminary general facts about Menelaus' *Spherics*, including its division and its main characteristics.

Section 4 is the heart of this chapter. I discuss in it some significant propositions of Menelaus' *Spherics* and at the same time I give an overview of the major themes that are addressed in this work. I believe that before commenting on the history of this book, one has to understand the subject it treats, and for this, a knowledge of at least some of the important statements and methods of proof is necessary.

I have divided this book, as in my edition with Rashed of al-Harawī's rectification (Rashed and Papadopoulos 2017), into seven major themes. The reader interested in a more complete discussion of the mathematical content of this work is referred to Part II (pp. 125–395) of our edition.

It may be noticed that some of the statements of Menelaus' propositions are analogues of propositions in Euclid's *Elements* (e.g., the triangle inequality, equality and inequality criteria for triangles, etc.), but one must be aware of the fact that this is only an apparent similarity in the sense that the proofs of the spherical propositions are very different from those of their Euclidean analogues. This is due to the fact that the properties of the underlying surfaces are very different: On the one hand, we have the Euclidean plane, in which two lines may intersect or not, and where any two lines starting at the same point diverge, and on the other hand, we have the sphere, in which any two lines intersect and where two lines starting at a point start by diverging and then converge to the same point. Furthermore, in the Euclidean plane, the angle sum in a triangle is constant and equal to two right angles, whereas on the sphere, the angle sum in a triangle is non-constant and is always greater than two right angles. There are many other substantial differences between the two geometries.

In Sect. 5, I discuss the influence of Menelaus' *Spherics* on Pappus. This is a rare occasion where the influence of this work on a later author from Greek Antiquity is visible.

Sections 6 and 7 constitute an important part of this chapter; they concern the influence of Menelaus' *Spherics* on Arabic mathematics of the period that starts at the ninth century and that ends at the fifteenth, where a considerable amount of work was done in translating this treatise and in trying to complete the proofs of the difficult propositions that are contained in it. This effort led to a development of important mathematical theories by the Arabs, especially in spherical geometry.

In Sect. 8, I discuss the connection of the *Spherics* with astronomy, a relation which, in my opinion, has been exaggerated by modern authors.

In the following, when I quote mathematical propositions, I have sometimes paraphrased the original text to make it more easily readable by a modern mathematician. The translations close to the original together with the original text of Menelaus' *Spherics* are contained in the volume (Rashed and Papadopoulos 2017).

2 Two Predecessors: Theodosius and Autolycus

Before talking about Menelaus' *Spherics*, I will make a short review of the works of two forerunners, Theodosius of Tripoli (first c. AD) and Autolycus of Pitane (fourth century BC). I will start with Theodosius, the immediate predecessor of Menelaus,

mentioning his influence on the latter, and then, I will consider Autolycus, on whose work was based that of Theodosius.

Theodosius' *Spherics* was written in the first century BC, that is, two centuries before Menelaus. The two works, which carry the same title, are concerned with the properties of the sphere. Theodosius' treatise was obviously not foreign to Menelaus. In fact, Menelaus in a more than one sense has placed himself in the continuity of Theodosius: It suffices to recall that at the beginning of Book II al-Harawī's rectification (a version of this work to which we shall often refer), Menelaus declares that his goal is to improve and generalize some propositions of Theodosius (Rashed & Papadopoulos, 2017, p. 684). Furthermore, Menelaus, at some places of his *Spherics*, refers either explicitly or at least implicitly to Theodosius' work, and uses some of his propositions. We have commented at length on this fact in our edition (Rashed and Papadopoulos 2017) of al-Harawī's rectification, and we have also highlighted there the major differences between the two works.

The best modern-language edition of Theodosius' *Spherics* is still Ver Eecke's (Theodosius, Les Sphériques de Théodose de Tripoli 1927). It is very close to the text from which it is translated, it reproduces faithfully the mathematical content of the work, and it is accompanied by a solid mathematical commentary.[2] Let me start with some general remarks on this work.

Theodosius, in his *Spherics*, makes a distinction between a *theorem* and a *construction*. Such a distinction, which has already been made explicit and commented upon by Proclus in his *Commentary to the first book of Euclid's* Elements, is highlighted by Heiberg and by Ver Eecke.[3] The latter, after the number of each proposition, writes whether it is a theorem or a construction. An example of a construction in Theodosius' *Spherics* is Proposition 2 of Book I, whose statement is: *To find the center of a given sphere*. This is an analogue of Proposition I of Book III of Euclid's *Elements* whose statement is: *To find the center of a given circle*. An important remark that fits here is that in Menelaus' *Spherics*, which is the main topic of the present chapter, all the propositions are meant to be constructive.

Theodosius' *Spherics* is divided into three books. None of the statements of the propositions in the first two books is close to any statement of a proposition of Menelaus' *Spherics*. In this respect, the two treatise are very different. If we want to summarize in very few words the difference between them, we can say that the first one is concerned with circles drawn on the surface of the sphere whereas the second one is concerned with triangles drawn on the surface of the sphere. This is especially true for the first two books of Theodosius.

Let us highlight some statements of Theodosius' *Spherics*.

Book I starts with definitions: that of the sphere (by which is meant the solid sphere), its surface, its center, and the notion of diameter. After the definition of the

[2] There are editions and commentaries done by authors with a poor knowledge in mathematics, and I prefer not to dwell upon this.

[3] Ver Eecke's edition is based in part on Heiberg's, see the comments in (Theodosius, Les Sphériques de Théodose de Tripoli 1927) p. LII of the Introduction.

latter, it is said that it is a rotation axis of the sphere, a sentence which Ver Eecke considers as a later addition ([36, p. 1, Footnote 3]). Indeed, the property that a diameter of the sphere is a rotation axis is not used in the treatise. It is used in Autolycus' work which we review below.

Theodosius defines the pole of a circle on the surface of the sphere as a point on this surface such that all the Euclidean segments starting at this point and ending at a point on the circle are equal. In particular, no attempt is made to define the pole intrinsically on the surface of the sphere (for instance, as a center with respect to lengths of arcs of great circles on that surface).

Theodosius then gives the definition of "two planes having the same inclination." His definition is the same as Euclid's. We recall that the definition of the angle made by two planes is contained in Euclid's Book XI (Definition 6), and that the definition of two pairs of planes having the same inclination is the content of the next definition of the same book. We also note that unlike in Euclid's *Elements*, besides this relative notion of inclination, there is no absolute notion of angle made by two planes in Theodosius' *Spherics*, neither in the definitions nor in statements or in the proofs. In his propositions and in his proofs, Theodosius only talks about *perpendicular* great circles and *oblique* great circles. The notion of angle between two lines is only used for the Euclidean lines in the solid sphere, and never for two arcs of circles on the surface of the sphere. These constitute fundamental differences between the *Spherics* of Theodosius and Menelaus. Indeed, Menelaus uses the dihedral angle between two planes to define the angle made by two circles on the sphere, and hence, to define the notion of angle at the vertex of a spherical triangle, and he applies this notion in an essential way in every proposition and in every proof, with very few exceptions.

After the definitions, the first book of Theodosius' *Spherics*, in Ver Eeckes' edition, consists of 23 propositions.[4] These propositions concern basic properties of the sphere and its intersections with planes. For instance, Proposition I.1 says that the intersection of the sphere with a plane is a circle. We already mentioned Proposition I.2, which is a construction problem: *To find the center of a given sphere*. A great circle is the intersection of the sphere with a plane passing through the origin. Proposition I.11 says that on the sphere, two great circles cut each other in two equal parts. Proposition I.12 is a converse and is less trivial: it says that if two circles on the sphere cut each other into halves, then these two circles are great circles. Proposition I.13 says that if, on a sphere, a great circle cuts another circle with right angles, then it cuts it into two equal parts and passes by its poles. Proposition I.16 says that the distance from a great circle to a pole is equal to the side of a regular inscribed quadrilateral (i.e., a square). Proposition I.17 is a converse. Proposition I.18 is a construction problem: *To construct a segment equal to the diameter of a given circle on the sphere*. Proposition I.21 is also a construction problem: *To find the pole of a given circle on the sphere*.

[4] The attribution of the last two propositions to Theodosius, according to Ver Eecke, is doubtful, see [36, Footnotes p. 29 and 30].

From the mathematical point of view, all this is Euclidean geometry; there is no spherical geometry at all.

Let us pass now to Book II.

This book, in Ver Eecke's edition, also consists of 23 propositions. Most of them deal with properties of *small* circles. The book starts with the definition of tangent circles: two circles are tangent to each other if the line that is the intersection of the two planes that contain them is tangent to both of them. Proposition II. 6 says that if a great circle is tangent to a small circle, then it is also tangent to another small circle equal to the first small circle. I have reproduced the figure of this proposition from an Arabic manuscript, Ms. 3464, Topkapi Saray, Ahmet III (Istanbul). This is the drawing on the top in Fig. 1. I would like to take this opportunity to make a comparison with the figure of the same proposition, as redrawn in the plate of the figures of Clavius' English edition of the *Spherics* (Clavius's commentary on the spheriks of Theodosius Tripolitæ 1721), which I reproduced here in Fig. 2. In this plate, this is No. 33.[5] The figure in the ancient Arabic manuscript shows a sense of abstraction which, in my opinion, is absent from the realistic figures in the more recent editions of Clavius, Ver Eecke, etc.

Let me mention two other propositions from the second book of Theodosius' *Spherics*, which are both construction problems. Proposition II.14 says the following:

Given a circle on the sphere which is not a great circle and given a point on it, to describe a great circle tangent to the given circle at the given point.

Proposition II.15 is more involved:

Given on the sphere a circle which is not a great circle and given a point of the sphere situated in the region between this circle and a parallel equal circle, to describe a great circle tangent to the first small circles and passing through the given point. These propositions have a taste of Apollonian constructions of circles tangent to each other in the plane. The figure that comes with the last construction, in the Istanbul 3464 Ms., is reproduced in Fig. 3 below. I have included it here to suggest the complication of the construction. This is also Fig. 42 in Clavius' version (Fig. 2 here).

The content of Book II also belongs to the realm of three-dimensional Euclidean geometry.

Book III of Theodosius' Spherics is concerned with important monotonicity properties that we discuss at length in (Rashed and Papadopoulos 2017). Such properties constitute the culmination of both Spherics, in terms of degree of difficulty. Let me explain by an example what I mean by a monotonicity property. I take as example Proposition III.9 in Theodosius. The corresponding figure, from the manuscript mentioned above, is reproduced here in Fig. 4. The 3 sub-figures correspond to three cases in the proof, and we do no not need to worry about this now. In each case, the action takes place in a hemisphere, bounded by the exterior great circle

[5] In Clavius' edition, the total number of propositions contained in each book is not the same as in Ver Eecke's edition.

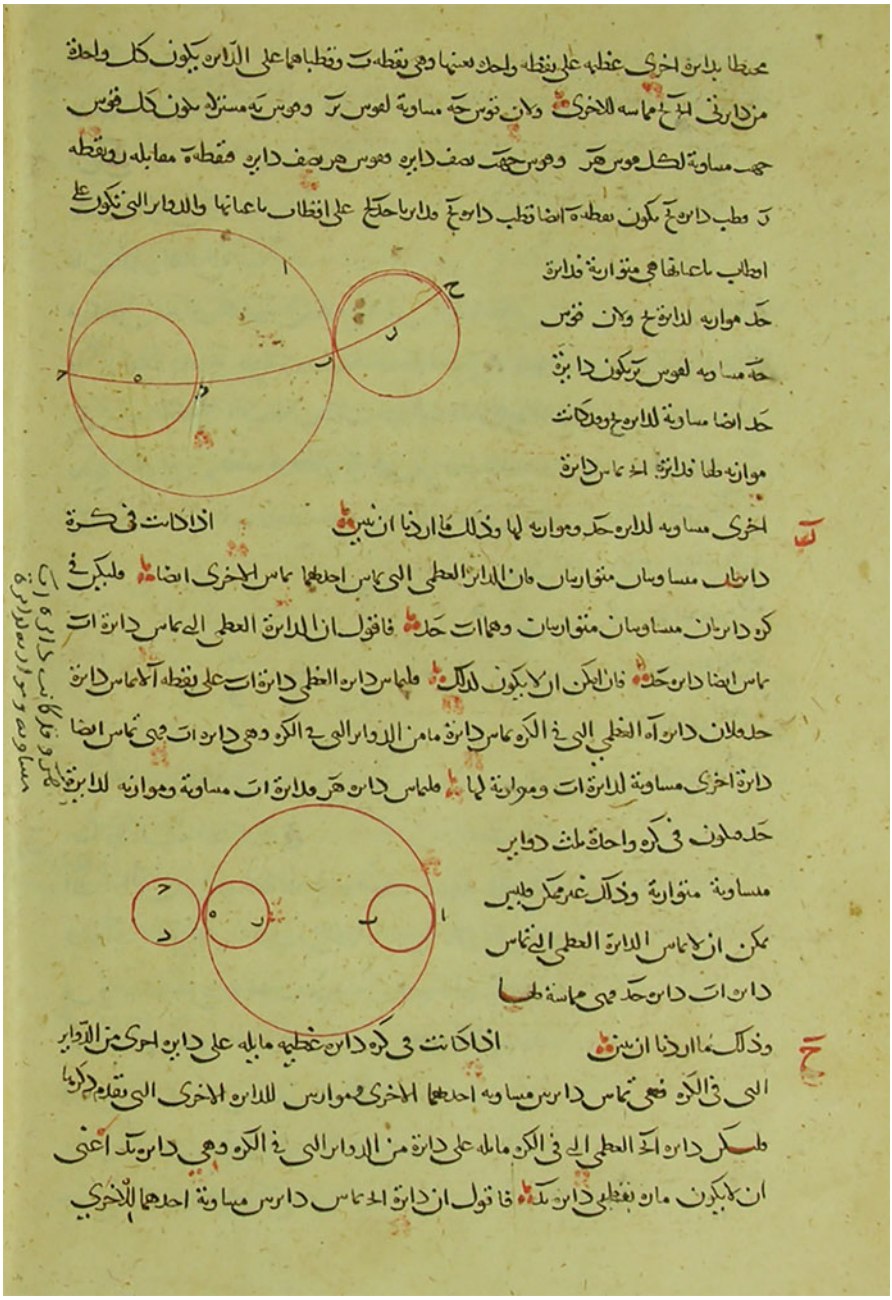


Fig. 1 From Theodosius' *Spherics*. On the top, the figure for Proposition II.6. Ms. 3464, fol. 30, Topkapi Saray, Ahmet III

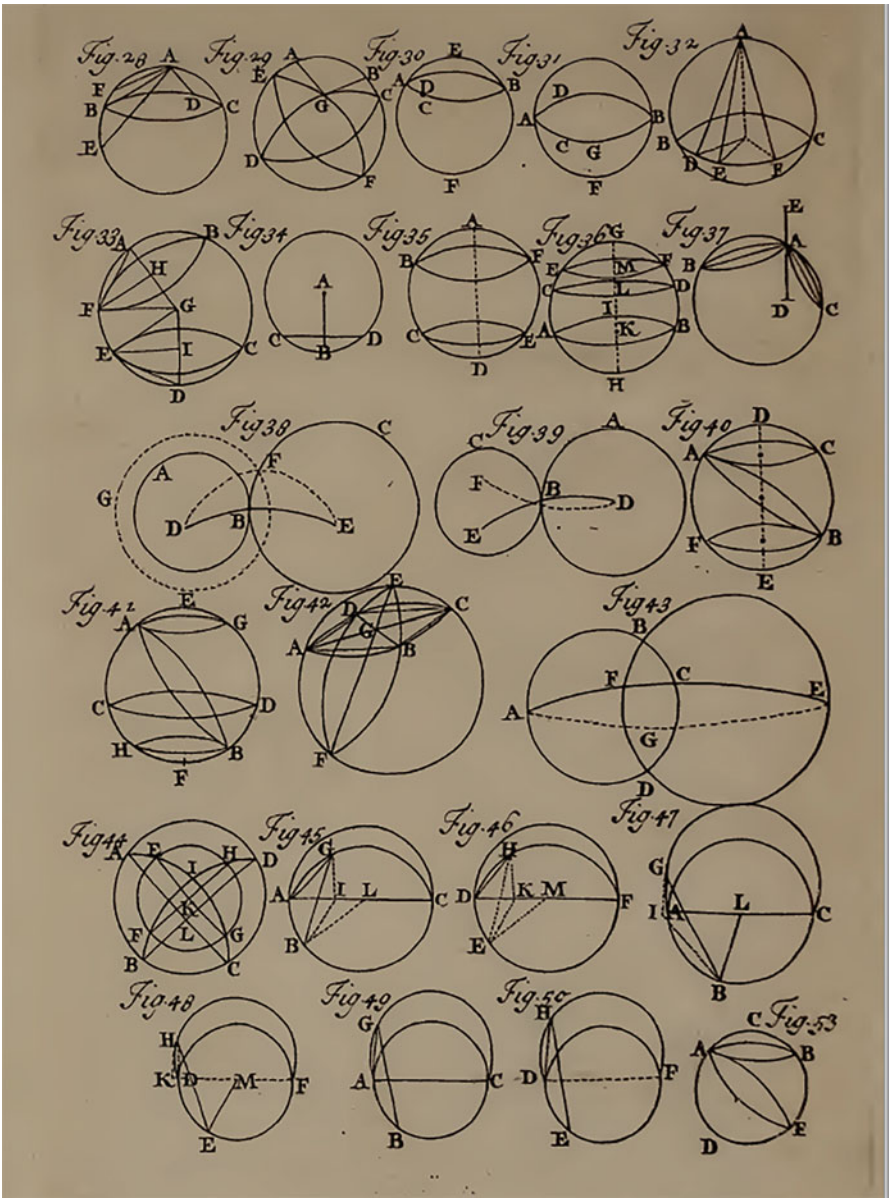


Fig. 2 A plate of figures for Book II of Theodosius' *Spherics*, Clavius' edition (Clavius's commentary on the Spherics of Theodosius Tripolitæ 1721)

represented, and the proposition concerns two arcs of great circles intersecting each other with an acute angle, one which we shall call the equator, and another one, the oblique arc (referring naturally to Fig. 4). In each of the three sub-figures, the point

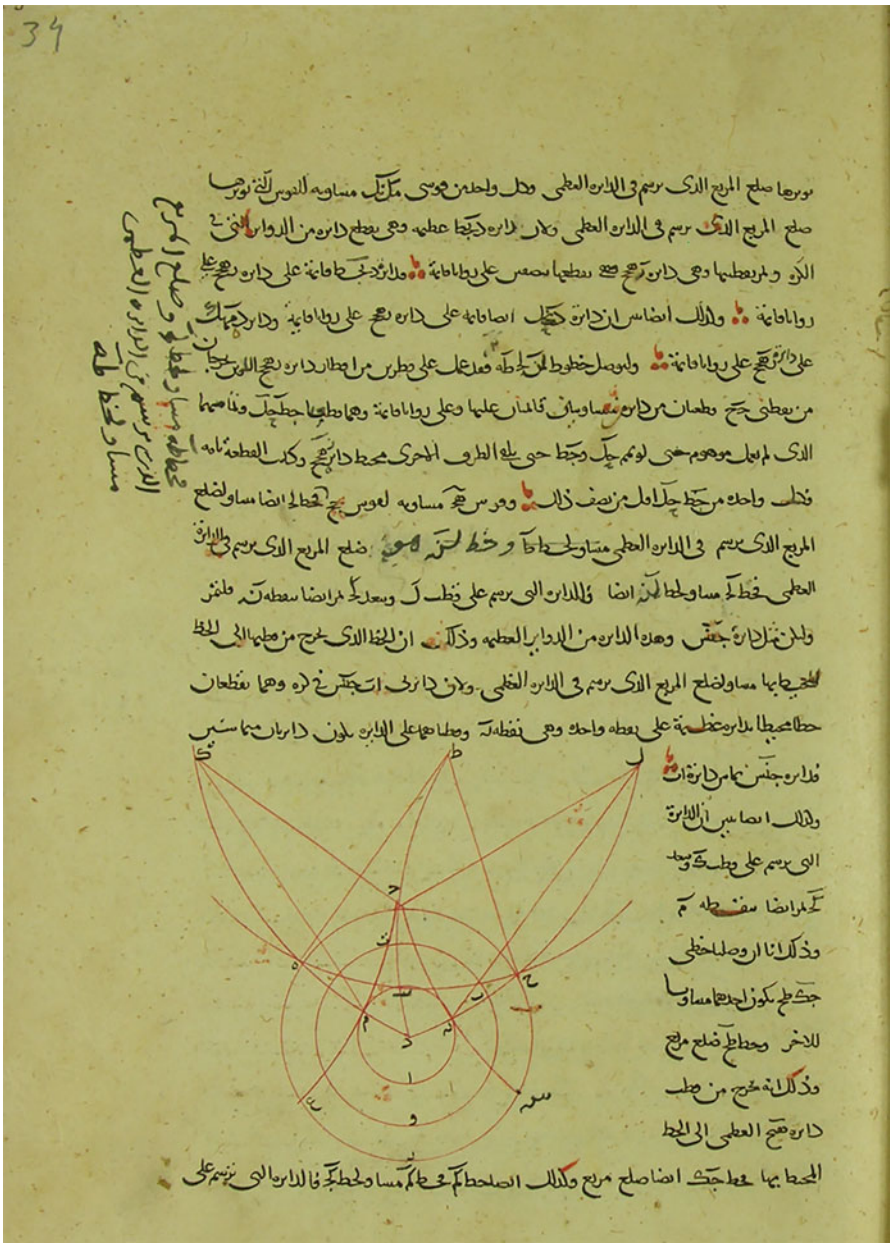


Fig. 3 From Theodosius' *Spherics*, Proposition II.15. Ms. 3464, fol. 30, Topkapi Saray, Ahmet III

on the upper left is a pole of the equator. On the oblique arc we take three points and we consider the two consecutive segments they bound. (There are several segments represented in the figure, and they are all used in the proof, but the statement

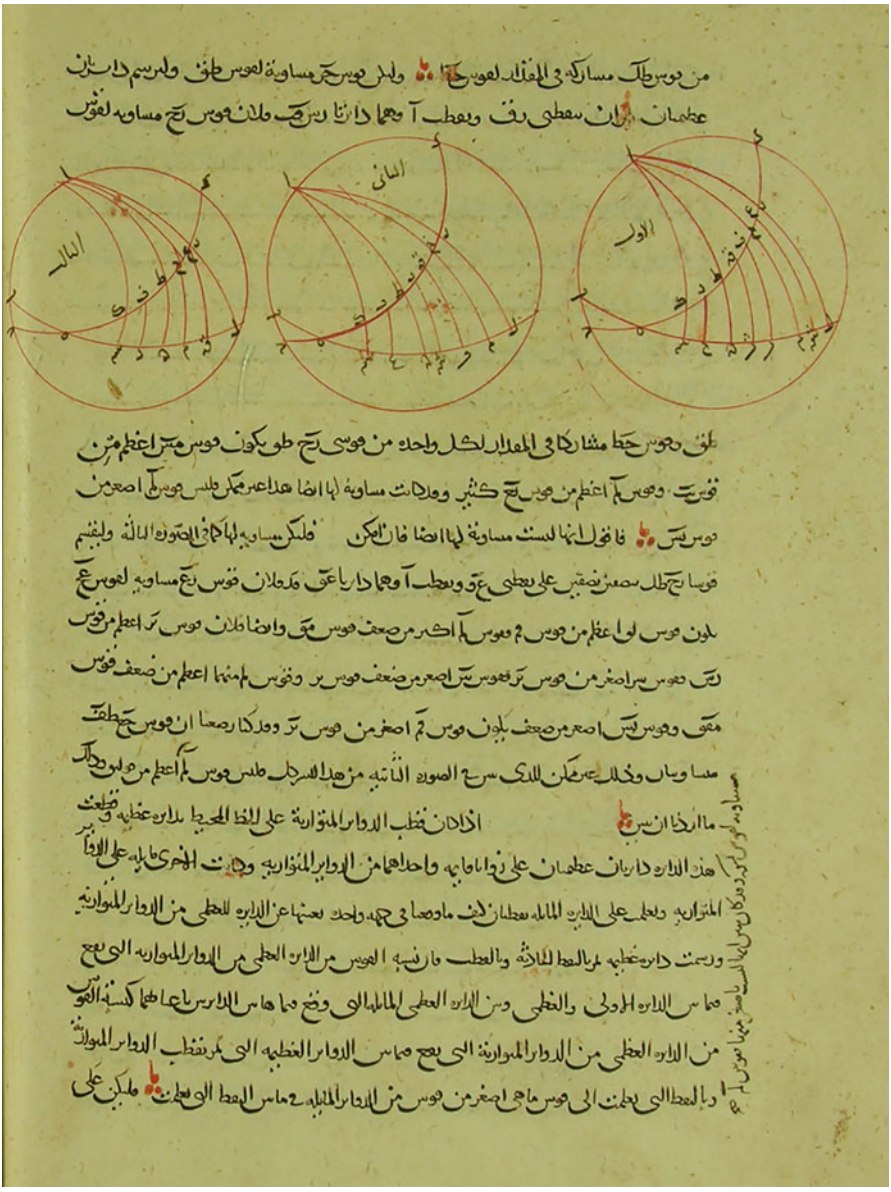


Fig. 4 From Theodosius' *Spherics*, Proposition III.10. Ms. 3464, fol. 50, Topkapi Saray, Ahmet III

involves only two of them.) We project these two segments perpendicularly (that is, using lines emanating from the pole), onto the equator. The proposition says that the ratio of the segment which is closer to the acute angle to its projection is greater than the ratio of the other segment to its projection. The word "monotonicity" that we used to describe the theme of this proposition refers to such an inequality. More

precisely, this word refers to the fact that if we start, instead of two consecutive segments on the oblique line, with a sequence of more than two consecutive segments on this line, then the sequence of ratios of one segment to its projection are all different and they vary monotonically: the more a segment is close to the acute angle of the triangle, the more the corresponding ratio is greater. We note that in the analogous Euclidean situation where we project perpendicularly segments on a Euclidean straight line onto another straight line, the conclusion of the statement is that the ratios are all equal.

Despite the fact that his methods were criticized by Menelaus who obtained more general results and who proved them using intrinsic methods, Theodosius remains, to the best of our knowledge, the first mathematician who highlighted such monotonicity properties which are proper to spherical geometry. We mentioned above that from the purely mathematical point of view, in Books I and II of Theodosius' *Spherics* all the statements are Euclidean. On the contrary, this proposition and other similar ones in Book III brings us to the heart of spherical geometry.

Between the Ancient Greek period and the modern one, monotonicity properties in Theodosius' Book II have been featured by Ibn al-Haytham, the preeminent tenth-century Arab mathematician, see (Rashed 2006–2014), in particular pp. 49–105 of the French edition. Roshdi Rashed, the author of this edition of Ibn al-Haytham, considers (rightly) that the developments by Ibn al-Haytham that were based on this proposition of Theodosius and similar ones constitute an early chapter in infinitesimal geometry. Note by the way the related multiplicity of lines in Fig. 4, which suggest a passage to the limit.

Let me mention here that several modern treatises on hyperbolic geometry start with propositions similar to Theodosius' III.9 that we just mentioned, and that such propositions are used in the foundations of the infinitesimal theory of non-Euclidean theory. For instance, in the doctoral dissertation of Louis Gérard, *Sur la géométrie noneuclidienne* (Gérard 1892), defended in 1892 and written under the direction of Poincaré, a monotonicity property similar to III.9 is used for establishing a functional equation which gives directly the trigonometric formulae of non-Euclidean geometry. In fact, the development in Gérard's thesis is done in the context of Lobachevsky geometry, but the same arguments, with inequalities in the reverse direction, give the trigonometric formulae for spherical geometry. I mention this here in order to suggest the power of these inequalities. Later authors have also used the same approach in the analytic development of non-Euclidean geometry, see, e.g., Barbarin's book (Barbarin 1902), one of the early French treatises on non-Euclidean geometry, in which the author makes at the same time a comparison between the spherical and the hyperbolic cases, starting precisely with these monotonicity properties.

From Proposition III.9 and its figure, it is also clear that spherical triangles without the name are present in Theodosius' Book III. Ver Eecke, on p. x-xi of the Introduction to his edition of Theodosius' *Spherics*, rightly states that the latter was surely aware of some of the properties of spherical triangles, but that he did not include this object in his work because he did not succeed in identifying their important properties.

Book III of Theodosius' *Spherics* contains 14 propositions, some of which are much more involved than the one we mentioned. Menelaus, in his *Spherics* gave

improved and stronger versions of some of these propositions. We shall review this in the next section, but let me mention here, in this connection, the beginning of Book II of Menelaus' *Spherics*, in al-Harawī's edition. We read the following on the critique of Menelaus toward Theodosius, regarding the method used to prove these propositions (Rashed & Papadopoulos, 2017, p. 684):

> Aḥmad ibn Abī Sa'd al-Harawī said that Menelaus overcame the difficulties of this science, something which was not within the reach of somebody else. Despite the fact that he mastered it, and albeit the majesty of what he was undertaking, he bypassed numerous lemmas which are needed by someone who examines this book and does not have the rank of Menelaus.
>
> We can see him criticizing Theodosius in his book on the Spherics; he sees that the method that he followed is not satisfactory, since it comprises complications and the drawing of numerous <straight> lines, and that Theodosius did not take into account, by this method, the properties of the figures which take place in the sphere, that is, the states of the angles generated by the intersection of circles.
>
> [I swear] on my life that everything that Theodosius proved in that book, Menelaus proved it in an easier way, and he intended that the proof is by the direct method.[6]

Theodosius' work was influenced by that of Autolycus, in particular, his two treatises *On the moving sphere* and *On risings and settings* written about 300 hundred years before.

Autolycus lived in the fourth century BC, about the same time as Euclid. His treatise *On the moving sphere* contains 12 propositions and is considered as the oldest complete Greek treatise that reached us. The expression "moving sphere" in the title denotes the Celestial sphere which, according to the theory that was in use at that time, rotates uniformly around an axis that crosses the Earth which sits at the center of the universe and which passes through the poles.

From the mathematical point of view, the twelve proposition of Autolycus are easy. I will just limit myself to three propositions from Autolycus' work, my aim being just to give an idea of some of the background of the two *Spherics*, especially the one of Theodosius. The reference I use is (Autolycus de Pitane 1979).

Proposition 1 says the following:

> *If a sphere rotates with a uniform velocity around an axis, all the points situated on the surface of the sphere and that are not on the axis describe parallel circles that have the same poles as the sphere and that are perpendicular to the axis.*

Proposition 6 says:

> *If on the (moving) sphere the fixed great circle (that is, the horizon) is oblique to the axis, it is tangent to two equal parallel circles: the one, situated on the side of the visible pole, will always be visible, and the other one, situated on the side of the invisible pole, will always be invisible.*

[6] "By the direct method" means without using a proof by contradiction.

One may compare this proposition to Theodosius' Proposition II.8 which we quoted above.

Proposition 12 of Autolycus says:

> *If on the (moving) sphere a fixed circle always cuts into two equal parts an arbitrary moving circle of the sphere and if none of the two is neither perpendicular to the axis nor passes by the pole of the sphere, then the two circles are great circles.*

The reader may notice the relation between the last proposition and Proposition I.12 of Theodosius' *Spherics* which we quoted above.

Some propositions of Autolycus' *Moving sphere* are used in Theodosius' astronomical treatises, *On habitations* and *On the days and nights* which we mention in Sect. 8.

We pass now to Menelaus' work.

3 Menelaus' *Spherics*: Preliminary Remarks

Talking about Menelaus' *Spherics*, one must necessarily mention the Arab commentators of this work, since this work reached us only through a certain number of Arabic versions, all of them containing more or less substantial differences among themselves.

We start by talking about the division of this work.

There is no general agreement on the division of Menelaus' *Spherics*. Naṣīr al-Dīn al-Ṭūsī, in the preface of his edition which we quote in (Rashed & Papadopoulos, 2015, p. 13), writes the following:

> This treatise consists of three books in some copies, and of two in some others. Concerning the three books, in most cases, the first one contains thirty-nine propositions, the last one twenty-five propositions, and the middle one, in many copies, contains twenty-four propositions, and in the copy of Ibn 'Irāq, twenty-one propositions. And in some of them, the first one contains sixty-one propositions, the second one nineteen propositions, and the third one twelve propositions. Regarding the copies with two books, the first one contains sixty-one propositions and the last one thirty propositions. And there are differences in some propositions, because some made two propositions a single one and conversely. But in general, the total number of propositions in the treatise is between eighty-five and ninety-one. I referred in the margin to some of the books and the number of propositions they contain, in red and in black, and sometimes I did it in the text.

Ibn Abī Jarrāda, in a manuscript we quote in (Rashed & Papadopoulos, 2017, p. 17) (see also Sect. 6 below), mentions a division into four books.

We have discussed at length the divisions of the various versions of the *Spherics* that survive in Chapter I of our edition of al-Harawī's rectification (Rashed and Papadopoulos 2017), which is divided into two books. At the beginning of Book II, a sentence explains this last division (Rashed & Papadopoulos, 2017, p. 684):

> As we have shown the lemmas which we needed (i.e., in the first 61 propositions), let us return now to what Theodosius wanted to show; we prove it using a universal assertion, without admitting in it the impossible.[7] We shall then show his error and rectify what is spoiled.

We conclude from this that Menelaus suggests that his goal in Book II is to improve some propositions of Theodosius. A rapid inspection of the two treatises shows that the propositions in question are those of Book III of Theodosius' *Spherics*. Indeed, Book II of Menelaus' Spherics contains a certain number of propositions that are concerned with the comparison of projections of segments of a great circle onto another such segment, in the spirit of Proposition III.9 of Theodosius' *Spherics* which we recalled above.

Although al-Harawī's rectification of the *Spherics* is divided into two books, the numbering of the proposition is continuous, from 1 to 91.[8] In what follows, it will be practical for us to use this numbering.

Now about the content.

The most important object of study in Menelanus' *Spherics* is the spherical triangle.[9] This is a figure made by three arcs of great circles called its *sides*, each two sides intersecting at a point called a *vertex* of the triangle. A side of a triangle is always assumed to be less than a semicircle. In this way, a side is the unique shortest arc between its vertices. This justifies the fact that in what follows, we shall sometimes call *segment* an arc of great circle which is smaller than a semi-circle, and *angle* an angle made by two segments at an intersection point.

Menelaus avoided as much as possible proofs that use solid Euclidean geometry. This is the point of view of intrinsic geometry of surfaces which we mentioned in the introduction. Furthermore, Menelaus always sought for new proofs. In the few cases where a proof similar to that of an analogous proposition in Euclid's *Elements* is possible in the spherical setting, he either does not give any proof at all (like in Proposition 6 which says that if we draw from the two vertices of the base of a triangle two arcs of great circles that meet within the triangle, then the sum of these arcs is smaller than the sum of the two legs of the triangle), or he gives a different proof (like in Proposition 8 which says that if two triangles have two pairs of equal sides such that the angle contained by one pair belonging to one of these triangles is greater than the angle contained by the corresponding pair in the other triangle, then the base of the triangle that has a greater angle is greater than the base of the other triangle).

It is conceivable that Menelaus did not give a proof of the famous *Sector Figure* (Proposition 66) because he did not have a proof which does not use the ambient

[7] This means by a direct method, without using proofs by contradiction.

[8] Regarding the number of propositions, one should note that there are statements for which Menelaus gives two different proofs, and each proof is counted as a new proposition (this is the case, for example, of Propositions 8 and 9). Thus, in some sense, there are less than 91 propositions.

[9] Two Arabic expressions مُثَلَّث (threefold) and شَكل ذو ثَلَاثَة أضلَاع ("trilateral figure") are used by the Arabic translators of Menelaus for the Greek term *τρίπλευρον* ("trilateral"). Let us recall that Euclid, in the *Elements* (Definition I.3), defines trilateral figures as those contained by three straight lines.

Euclidean geometry. It is also possible that Ptolemy, who uses extensively this proposition in his *Almagest*, gave a proof of it (see (Halma, 1813, pp. 50–55)) because of the absence of proof in Menelaus' treatise. We shall mention several times this proposition in what follows.

Furthermore, a large number of arguments of Menelaus rely heavily on polarity properties on the sphere that do not hold in the Euclidean place.

Thus, in most of the cases, the proofs of Menelaus' propositions of spherical geometry cannot be used for their Euclidean analogues. We shall give explicit examples in the next section.

4 The Propositions

There is more than one way of grouping the propositions of Menelaus' *Spherics* into topics: such a grouping depends on the importance that we attach to each particular theme and on the relations we can make between the various topics. I have collected the propositions of Menelaus' *Spherics* into seven groups, corresponding to seven themes, which I review now.

Group 1: Construction of Angles; Propositions 1, 41–44
I start with Proposition 1: *Given an arbitrary angle, and given an arbitrary point on a given line, we can construct an angle equal to the given angle having its vertex at the given point and a side on the given line.*

This proposition is of paramount importance, and it is abundantly used in the proofs of the rest of the propositions of the *Spherics*. Its proof, unlike the vast majority of the other propositions, uses the solid geometry of the ambient space. More precisely, it relies on constructions from Theodosius' *Spherics* and a criterion from Euclid's *Elements* for the equality of two angles seen from the center of a circle.

A remark is necessary here. We mentioned that most of Menelaus' proofs are intrinsic, that is, they use notions defined on the surface of the sphere (lines, angles, circles, etc.) and not from the ambient space. However, this cannot be true for *all* propositions. The reason is that Menelaus did not start with axioms for spherical geometry, since he thought of the sphere as a surface embedded in Euclidean 3-space, therefore the axioms of solid Euclidean geometry are sufficient for the study of such an object. Thus, some of the propositions, and in particular the first one, necessarily rely on Euclidean geometry. Establishing an autonomous set of axioms for spherical geometry is a modern idea.[10]

Let us now pass to another proposition from the first group, namely, Proposition 41. The general idea is to construct, in a spherical triangle, a segment (that is, an arc of great circle) from a vertex to its opposite side, making with this side an angle equal

[10] An early formulation of a set of axioms for spherical geometry is in (Halsted 1885).

to one of the angles of the triangle at the extremity of this side. The construction is not possible in the general case, and I will explain why, but let me first state the proposition under Menelaus' hypotheses:

> *Given a triangle ABC such that AB + BC is smaller than a semi-circle and such that the angle A is obtuse, and given a point D on BC, to construct a line from D that cuts the side AB with an angle* $\widehat{DIC}$ *equal to the angle A of this is triangle.*

The proof is contained in (Rashed & Papadopoulos, 2017, pp. 628–630). I have reproduced here Menelaus' figure (Fig. 5), just to suggest the fact that the construction is involved. This proposition is used in later propositions of the *Spherics*, and in particular in the last propositions.

Let me make a few comments on Proposition 41.

The Euclidien analogue of this proposition does not need any particular hypothesis: one takes the parallel to AB passing through D; it makes with AC an angle equal to the angle at A of the triangle ABC. But in the spherical case, parallel lines do not exist, and the problem is much more complicated. Without specific assumptions, the desired line does not always exist: See Fig. 6 for the particular case where each of AB and BC is a quarter of a circle. In this case, a great circle from D cannot make with AC a right angle, because D is not a pole of AC (since the pole is at B). In the case where we allow $AB + BC$ to be greater than a semi-cercle, the situation is not better: Take for example the case where the angle A is acute and C right, with AB and BC both greater than a quarter of a circle and let D be a pole of AC (Fig. 7). Then every line from D to the side AC makes a right angle with it, therefore this angle cannot be equal to the angle at A.

In Proposition 42, the same construction is done under the hypothesis that the angles at B and C are both acute. This proposition is used later (e.g., in the proof of

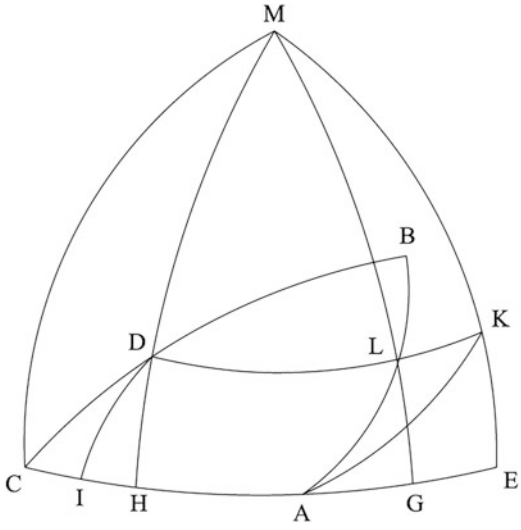


Fig. 5 Figure for Proposition 41

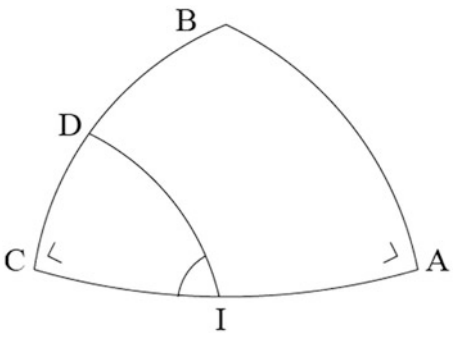


Fig. 6 Proposition 41 fails with the assumption that *AB* and *BC* are equal to a quarter of a circle

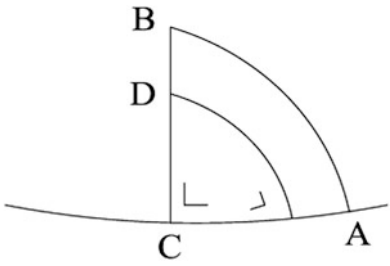


Fig. 7 Proposition 41 fails when *AB* and *AC* are greater than a quarter of a circle

Proposition 45). I have reproduced Menelaus' figure in Fig. 8 here. I have also reproduced (Fig. 9) the page of an Arabic manuscript containing the three propositions, 42, 43, 44 on the same subject. Notice the concision of the writing (the three propositions fit in one page).

Group 2: Comparison Between Triangles; Propositions 4, 8, 9, 13–19 and 20–23 This group consists of congruence and inequality criteria for spherical triangles. More precisely, there are two sorts of theorems in this group: (1) congruence theorems, i.e., theorems saying that if we have two triangles such that some elements (side lengths or angles) of one are equal to the corresponding elements of the second, then the two triangles are congruent; (2) non-congruence theorems, i.e., theorems saying that if we have two triangles such that some elements (side lengths or angles) of one are greater than the corresponding elements of the second, then some other element of the first is greater (or smaller) than the corresponding element of the second.

An example of a congruence theorem is Proposition 19, which has no analogue in Euclidean geometry. It says that if two spherical triangles have equal angles, then they are congruent.

We note incidentally that Menelaus never makes statements such as "the triangles are congruent," or "the triangles are equal," or equivalent expressions. Instead, he always gives precise statements, saying what side in the first triangle is equal to what

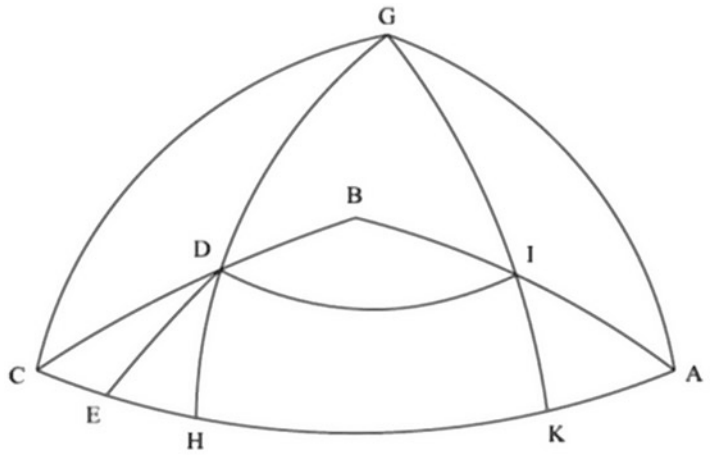


Fig. 8 Proposition 42

side in the second triangle and what angle in the first one is equal to what angle in the second one.

Examples of non-congruence theorems are Propositions 8 and 9, which are in fact a single theorem concerning two triangles having two pairwise equal pairs of sides. They say that if in one of these triangles, the angle contained by two equal sides (the legs) is greater than the angle in the second triangle contained by a pair of legs equal to the pairs of legs in the first triangle, then the first triangle has a greater base than the second. This proposition also holds in Euclidean geometry (Euclid's Proposition I.24).

Group 3: Geometry of Triangles; Propositions 2, 3, 5–7, 10–12 and 24–26
These propositions concern the geometry of individual triangles. For instance, Propositions 2 and 3 concern isosceles triangles. They say that if a triangle has two equal sides (the legs), then these legs subtend with the base equal angles, and conversely, if a triangle has two equal angles at the base, then the two legs subtending them are equal. These are analogues of Propositions I.5 et I.6 of Euclid's Elements. The proofs though are very different in the spherical case; they are based on the duality between lines and their poles, and they also use some propositions in solid geometry from Theodosius' *Spherics*. Proposition 5 is the triangle inequality for spherical triangles. Let us dwell on Proposition 11, which is used in a crucial way at several places in the *Spherics*:

Let ABC be a spherical triangle. Consider the exterior angle $\widehat{BCD}$*. Then we have the following three equivalences:*

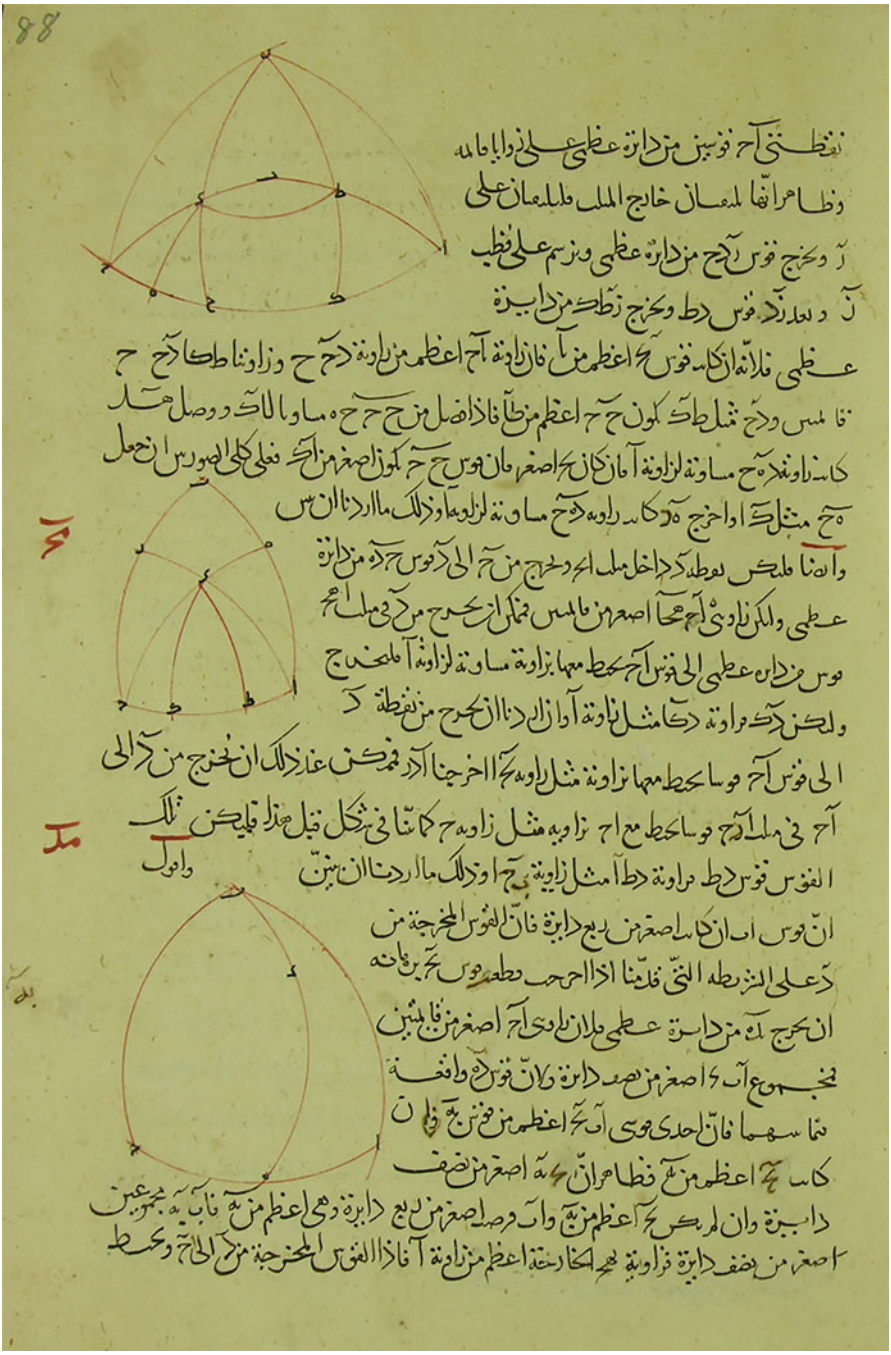


Fig. 9 Propositions 42, 43 and 44. Ms. 3464, fol. 88, Topkapi Saray, Ahmet III

1. $\widehat{BCD} = \widehat{A}$ *if and only if* $AB + BC$ *is a semi-circle;*
2. $\widehat{BCD} > \widehat{A}$ *if and only if* $AB + BC$ *is smaller than a semi-cercle;*
3. $\widehat{BCD} < \widehat{A}$ *if and only if* $AB + BC$ *is greater than a semi-cercle.*

Let us prove the first statement. We produce the sides AB and AC until we obtain a lune. One of the vertices of this lune is A, and we may set the point D to be second vertex (Fig. 10). If $\widehat{BCD} = \widehat{A}$, then $\widehat{BCD} = \widehat{D}$. In the triangle CBD, since $\widehat{C} = \widehat{D}$, we have $BC = BD$ (this is the congruence criterion of Proposition 3). Since $AB + BD$ is a semi-cercle, $AB + BC$ is also a semi-cercle.

The converse is obtained by reversing the arguments and using the congruence criterion of Proposition 2. The proofs of (2) and (3) are of the same type, based on the notion of lune and using Proposition 7 (inequality criteria for triangles).

Let me remark that if one proves (1) and (2), then (3) follows immediately by a *reductio ad absurdum*, but Menelaus does not allow this kind of reasoning.[11] This is why a proof of (3) is needed.

Proposition 12 of the *Spherics* is of paramount importance in spherical geometry; it says that the angle sum in any spherical triangle is greater than two right angles. It is a direct consequence of Proposition 11. We reproduce this proof, because it is also a noteworthy illustration of the intrinsic methods of Menelaus.

To prove Proposition 12, let ABC be a spherical triangle and let us consider the exterior angle $\widehat{BCD}$. It suffices to prove that $\widehat{BCD} < \widehat{A} + \widehat{B}$.

We suppose without loss of generality that $\widehat{BCD} > \widehat{A}$ (otherwise the conclusion we seek is obviously true). Then, we construct an angle $\widehat{DCE} = \widehat{A} < \widehat{DCB}$ with E on the extension of the segment AB, as in Fig. 11. Now Proposition 11 applied to the

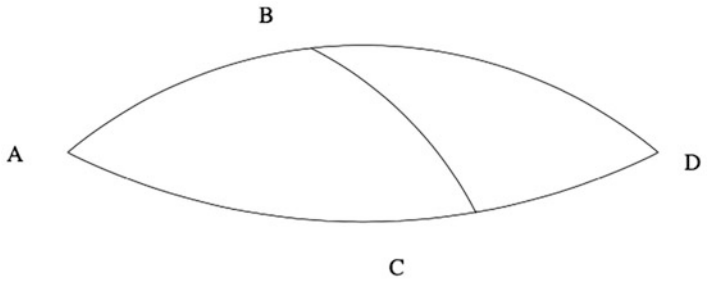


Fig. 10 Proposition 11

[11] Regarding this fact, let us quote Ibn Abī Jarrāda: "I said that Menelaus presented numerous lemmas for the proof of this proposition; but al-Harawī had shown some of them by a *reductio ad absurdum*; and we said that this does not belong to Menelaus' method. I showed all what was presented in the course of the proof without a *reductio ad absurdum*." (see the exact reference in (Rashed and Papadopoulos 2017) p. 41).

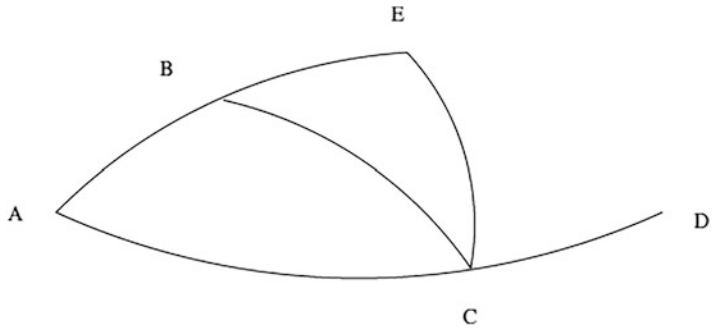


Fig. 11 Proposition 12

exterior angle $\widehat{DCE}$ of the triangle AEC implies that $AE + EC$ is a semicircle. Thus, $BE + EC$ is smaller than a semicircle. Applying again Proposition 11, this time to the triangle BCE, gives $\widehat{CBA} > \widehat{BCE}$. Adding $\widehat{A}$ to both sides gives $\widehat{BCD} < \widehat{A} + \widehat{B}$, which is the desired result.

Group 4: Division of Triangles; Propositions 27–40

This group consists of a set of propositions that concern inequalities obtained by dividing a spherical triangle by a segment (that is, an arc of great circle) connecting a vertex to the opposite side. The arc may be a side bisector, or an angle bisector, and there are also other kinds of division of a triangle, for instance, by an arc of great circle joining the midpoints of two sides. I will report more precisely on some of these propositions.

I start with a proposition which I find particularly interesting, namely, Proposition 27. It says the following:

> *In a spherical triangle ABC, let D and E be the midpoints of AB and BC, and DE an arc of great circle joining them. Then DE > AC/2.*

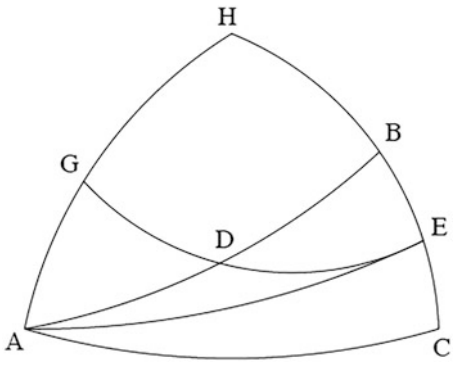
The importance of this kind of proposition in modern metric geometry cannot be overemphasized. It is equivalent to a concavity property for distances in spherical geometry, and more generally, in a space of positive curvature. Herbert Busemann took the conclusion of Proposition 27 as a definition of a metric of positive curvature, see (Busemann 1948; Busemann 1955; Busemann 1961).[12]

The proof of this proposition is a beautiful example of the intrinsic methods of Menelaus:

We produce ED until a point G such that $DG = DE$ and we produce AG and CB until they meet at a point H (Fig. 12).

[12] Busemann was not aware of the fact that the property was stated explicitly in a theorem of Menelaus. I highlighted this fact in (Papadopoulos, 2019).

Fig. 12 Proposition 27



Since $DG = DE$, $AD = DB$ and $\widehat{EDB} = \widehat{ADG}$, the two triangles EDB and ADG are congruent (This is one of the congruence theorems, Proposition 10 of the *Spherics*). Thus, we have $EB = AG$ and $\widehat{EBA} = \widehat{DAG}$.

From Proposition 11 on exterior angles which we recalled above, applied the triangle HBA and its exterior angle $\widehat{B}$, $AH + HB$ is a semi-circle. Thus, $AH + HE$ is greater than a semi-circle. Now we join A and E by a segment of great circle. Applying again Proposition 11 to the triangle HEA, we get $\widehat{CEA} < \widehat{EAG}$.

Now the two triangles CEA and EAG have a common edge EA, and we have $CE = AG$, and $\widehat{CEA} < \widehat{EAG}$. Therefore, $AC < EG$ (this is Proposition 8 of the *Spherics*). Since $DE = \frac{1}{2}GE$, then $DE > \frac{1}{2}AC$ and the proposition is proved.

Let us note that to perform the construction in the proof of Proposition 27, the sides of the triangle ABC must be small enough (for instance, we do not want the point H to fall inside the triangle ABC, which could happen if the segment AH is very long). We also note that in the definition of positive curvature one can assume the triangles small (the definition is local).

It is also important to know that there is an analogue of this proposition in the setting of hyperbolic geometry (the other classical non-Euclidean geometry), with the inequality in the conclusion reversed.

A more precise form of Proposition 27 can be proved using spherical trigonometry and a formula of Euler giving the area of a spherical triangle. I have included this proof in the first appendix to this chapter.

Among the other propositions in the same group, Propositions 28 and 29 concern again a triangle ABC with a segment DE joining the midpoints of AB and CB, but here, instead of comparing the segment DE to the side AC, one compares angles. More precisely, the propositions give the inequality $\widehat{BDE} < \widehat{BAC}$ under various conditions on the angles of the triangle ABC. Note that in the Euclidean case, we have the equality $\widehat{BDE} = \widehat{BAC}$.

A recent paper provides analogues of propositions 28 and 29 of Menelaus' *Spherics* in the setting of hyperbolic geometry, see (Senapati 2019). I am mentioning this to point out that such statements are still actual.

Group 5: Monotonicity; Propositions 45–62, 64, 65, 81–91
This group contains 32 propositions that express monotonicity properties. We already explained, in Sect. 2, what we mean by such properties and we talked about their importance. At the basis of such a property, we generally have a spherical triangle and projections of segments contained in one side onto another side. Such projections may use arcs of great circles making equal angles with the base side on which we project, like in Proposition III.9 we quoted from Theodosius *Spherics*, but there are other kinds of projections: using arcs of great circles perpendicular to the base, using small circles parallel to one side, etc. All these monotonicity results provide inequalities that show fundamental differences between spherical geometry and the two other classical geometries (Euclidean and hyperbolic). Furthermore, like for Proposition 41 which we recalled above, various kinds of hypotheses on the shapes of the triangles (angles or side lengths) are necessary in the statements of these propositions.

There is a progression in the results of Group 5 in terms of the sophistication of the statements and difficulty of the proofs. Some of these propositions are improvements of propositions of Theodosius on the same topic. I will mention two of these propositions, limiting myself to the statements.

Proposition 82 says the following:

Let ABC be a triangle with a right angle at A and acute angle at B, and with BC at most equal to a quarter of a circle. Consider on BC two arcs BD and EG and take the perpendiculars DH, EI, GK to the base AC (Fig. 13). *Then,*

if $BD = EG$ *then* $AH > IK$ *and* $AB - DH < EI - GK$;

if $AB - DH = EI - GK$, *then* $BD > EG$;

if $BD + (AB - DH) = EG + (EI - GK)$, *then* $BD < EG$;

if $BD - (AB - DH) = EG - (EI - GK)$, *then* $BD < EG$.

Furthermore, in every case, we have

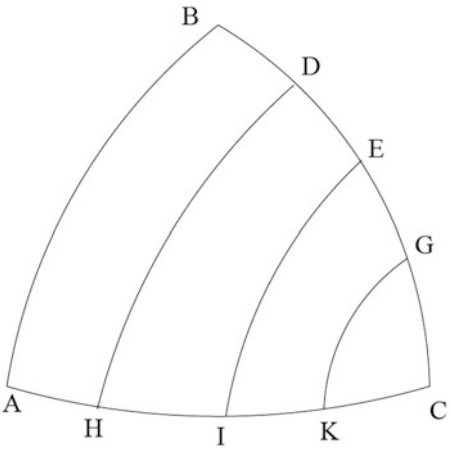


Fig. 13 Proposition 82

$$\frac{\text{crd } 2BD}{\text{crd } 2EG} > \frac{\text{crd } 2(AB - DH)}{\text{crd } 2(EI - GK)}.$$

Proposition 89, in the same group, says the following:

Let AB and BC be two arcs of great circles. Consider two points D and E on AB and take the perpendiculars DC, EH on BC (Fig. 14). *Then,* $\frac{\text{crd}2CH}{\text{crd}2DE}$ *is equal to the ratio of the product of the diameter of the sphere by the diameter of the (small) circle parallel to BC and tangent to AB to the product of the diameters of the two circles parallel to BC and passing by D and E.*

In Fig. 15, we have reproduced the page containing the last three propositions (89 to 91) from Ms. 3464; the reader can notice again the conciseness of the writing. Let me mention in this regard that I know of no written proofs of these three propositions, or of any of the last eleven propositions of the treatise. Furthermore, all the texts that reached us contain some ambiguities in the statements of these propositions. We have discussed the problems raised by these propositions in our edition (Rashed and Papadopoulos 2017).

Group 6: The Sector Figure and Its Applications; Propositions 66–70 and 72–80

Proposition 66 is the so-called Sector Figure. establishes a relation between six sides of a figure called "complete quadrilateral." This is the figure obtained by producing two pairs of opposite edges of a convex spherical quadrilateral until they pairwise meet at a point.[13] A picture of a complete quadrilateral is given in Fig. 16, where we start with the quadrilateral *ADEG* and we produce *AG* and *DE* until they meet at the point *B*, and *AD* and *GE* until they meet at the point *C*. Notice that each convex spherical quadrilateral gives rise to four complete quadrilaterals, depending on the choice of two adjacent sides. Using the notation of Fig. 16, the proposition says that we have:

$$\frac{\text{crd } 2AG}{\text{crd } 2BG} = \frac{\text{crd } 2AC}{\sin \text{ crd } 2CD} \cdot \frac{\text{crd } 2DE}{\text{crd } 2EB}.$$

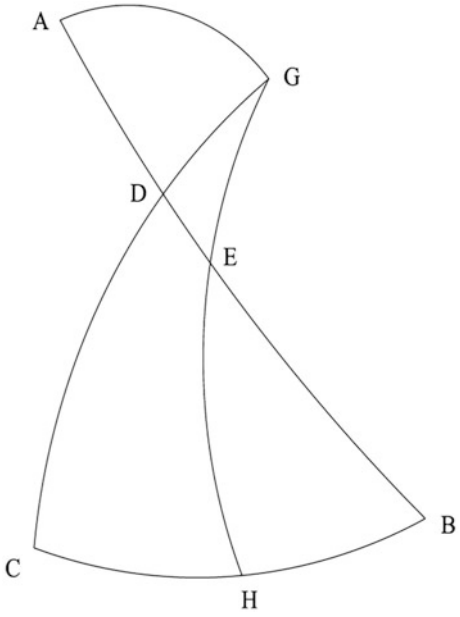
Expressed in the language of sines—which Menelaus did not have, but which was used by the Arab commentators of the *Spherics*—the relation is:

$$\frac{\sin\ 2AG}{\sin\ 2BG} = \frac{\sin\ 2AC}{\sin\ 2CD} \cdot \frac{\sin\ 2DE}{\sin\ 2EB}.$$

In some sense, this spherical proposition is close to the spherical Sine Rule for triangles; indeed the Sine Rule can be deduced easily from the Sector Figure, and conversely, there is a very short proof of the Sector Figure using the spherical Sine

[13] Convexity is understood here with respect to the intrinsic geometry of the sphere: A figure is said to be convex if (1) it is contained in an open hemisphere; (2) for any two points in this figure, the arc of great circle joining them is contained in the figure.

Fig. 14 Proposition 89



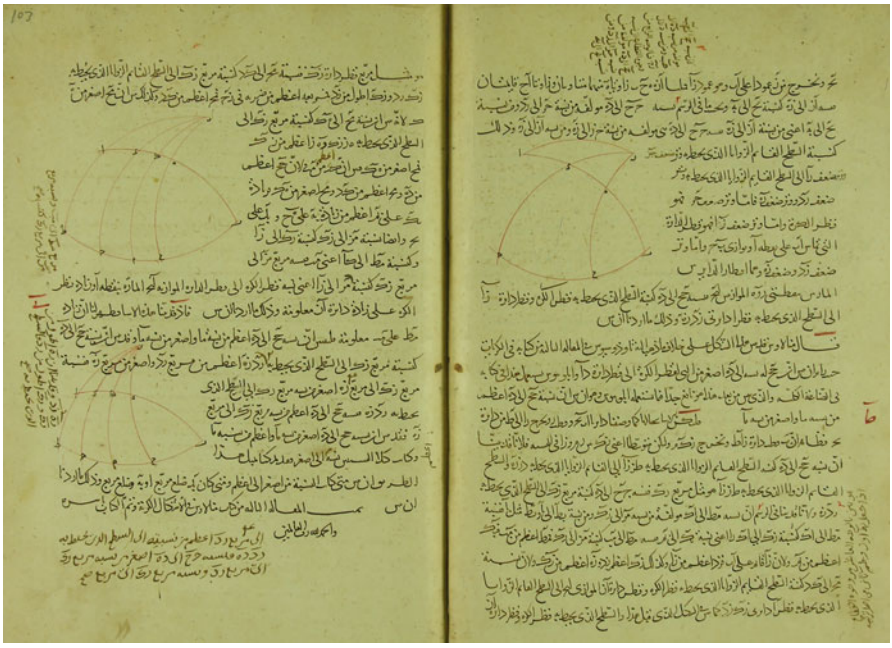


Fig. 15 Propositions 89 to 91 of the *Spherics*, Ms. 3464, fol. 99, Topkapi Saray, Ahmet III

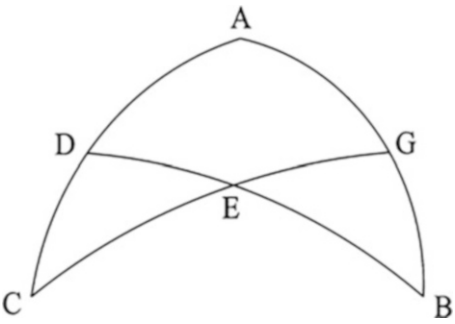


Fig. 16 The Sector Figure (Proposition 66)

Rule. One should note however that the discovery of the spherical Sine Rule occurred only in the eleventh century. In the meanwhile, the Sector Figure has been the only trigonometrical tool[14] used in astronomy. In particular, the proposition was heavily used by Ptolemy in the *Almagest* as the main ingredient for his astronomical calculations. Furthermore, it is the only result of the *Spherics* which is quoted under the name "Menelaus' Theorem."[15]

Proposition 66 occupies a particular place in the *Spherics*. It is situated among a series of propositions whose proofs are completely intrinsic to the surface of the sphere, but the proof given to this proposition in al-Harawī's redaction and in all the other Arabic manuscripts that survive is, like the one given by Ptolemy in his Almagest, purely Euclidean. (In fact, the proofs given by the Arabs are either Ptolemy's or a small variation it.) At the same time, such a proof was considered as faulty by the various Arab commentators and editors of the *Spherics* because it is not in the spirit of Menelaus' methods. Whether Menelaus had a proof of this proposition that does not rely on solid Euclidean geometry remains a mystery.

The other propositions in this group are proved using the Sector Figure. They involve divisions of spherical triangles by arcs of great circles joining vertices to the sides that subtends them. These results concern angle and external angle bisection (Propositions 73 and 79), angle division into three equal parts (Propositions 74 and 75), division by altitudes (Propositions 78, 80), and various other kinds of divisions. These propositions may all be proved using the spherical Sine Rule, and the Sector Figure Theorem is used in the absence of this theorem. We have discussed in detail all these propositions in our commentary in (Rashed and Papadopoulos 2017). In particular, Proposition 80 needs a careful discussion, and we comment on it here. In al-Harawī's version, the proposition says the following:

[14] We are using here the the word "trigonometry" for a formula that gives relations between the elements of a figure which is not necessarily a triangle.

[15] There are other names for this proposition. Ptolemy, in the *Almagest*, used the expression "secant figure," which is also the name that was generally adopted by the Arabs الشَكل القَطّاع in their commentaries on Menelaus' and Ptolemy's works and in the treatises they wrote on spherical geometry.

The altitudes that are produced from the angles of a triangle to their opposite sides intersect in one point.

The statement as it stands is not correct and needs a hypothesis. Let me explain this. Consider the example of a spherical triangle whose three angles are right. In this case, any arc of great circle drawn from a vertex to the opposite side that subtends it is an altitude. Obviously, three arbitrary such arcs drawn from the three vertices to the opposite sides do not necessarily intersect in a common point. Thus, the theorem fails in this case. There are much more general cases where the theorem fails.

The statement becomes correct if we add to it a further hypothesis. One possibility is the following:

Suppose that in a given triangle two altitudes intersect in a point within the triangle. Then there exists a third altitude passing through this point.

With such a hypothesis, the proof in al-Harawī's edition of Menelaus, of the fact that the three altitudes intersect in one point, becomes correct. This proof involves several applications of the Sector Figure Theorem.

Group 7: Proposition 71

I have dedicated a special section to Proposition 71 because I consider it as very special, both mathematically and historically, and I will explain why.

First, the statement is particularly interesting. It expresses a non-trivial property of spherical triangles, and I have not seen anywhere else, in the old and modern literature, a statement that bears any resemblance to it. Secondly, the proof of this proposition has given rise to an abundant literature by the Arab mathematicians who have revised and commented Menelaus' text. It seems that it was precisely in trying to prove this proposition that these mathematicians discovered the Sine Rule and the other formulae of spherical trigonometry, see the comments in (Rashed and Papadopoulos 2017) and Sect. 7 below. Moreover, the proof uses an important property that was obviously known to Menelaus, namely, the invariance of the spherical cross ratio, although he used it without proof. The invariance of the cross ratio is one of the main tools of projective geometry.

I have given a detailed proof of this proposition and reviewed several points in the history of the important mathematical activity that took place around it, between the ninth and the thirteenth centuries, in my two papers with Rashed (Rashed and Papadopoulos 2014; Rashed and Papadopoulos 2015).

Let me recall the statement of this proposition:

Consider two triangles ABE and LMO with right angles at B and M, and where the angles A and L are acute and equal. Then,

$$\frac{\text{crd}\,(AE+AB)}{\text{crd}\,(AE-AB)}=\frac{\text{crd}\,(LO+LM)}{\text{crd}\,(LO-LM)}.$$

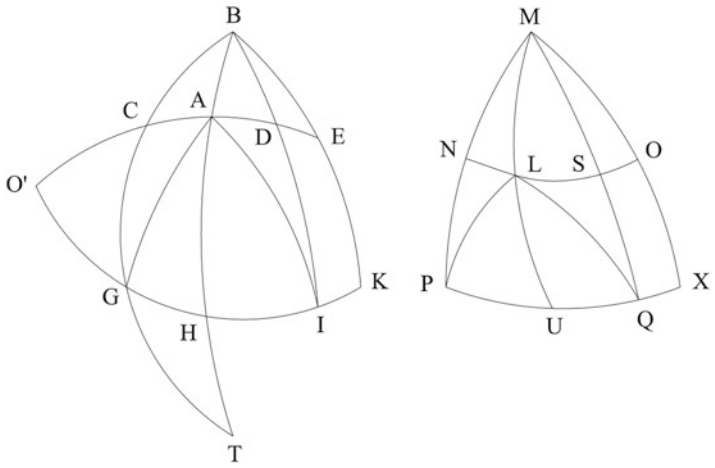


Fig. 17 Proposition 71

I have reproduced the figure that accompanies this proposition in Fig. 17 here.

As a matter of fact, inspired by this kind of result in spherical geometry, several theorems in hyperbolic geometry are proved in the recent paper (Papadopoulos and Su 2017). In particular, a more precise analogue of Proposition 71, in the setting of hyperbolic geometry, is given there (Theorem 2.1), based on hyperbolic trigonometry. Going backwards, a more precise form of the same proposition in the setting of spherical geometry can be given using the same outline, and we have reproduced it in the first appendix to the present chapter. Let me also mention here that Menelaus' theorem as well as the invariance of the cross ratio which is used in the proof of Proposition 71 have been adapted and used in the case of hyperbolic geometry, see (Papadopoulos and Su 2017; Papadopoulos and Yamada 2013). I am saying all this to give a hint of the freshness of the ideas behind this and other propositions of the *Spherics*.

5 Pappus

Pappus of Alexandria is among the rare mathematicians from Greek Antiquity who have embraced Menelaus' intrinsic methods. Book VI of his *Mathematical collection* contains a sequence of theorems on spherical geometry, some of which are proven using the intrinsic geometry of surfaces, in the pure style of Menelaus. As a matter of fact, the first proposition in this book concerns the triangle inequality, for a spherical triangle. Pappus' statement is the following:

> *If, on a spherical surface, three arcs of great circles which are all smaller than the semicircle mutually cut each other, then two of these arcs, taken in an any way, are greater than the remaining arc.*

The reader may notice that there is no mention of spherical triangle in this statement, but right after the proof of this proposition, Pappus notes: "Menelaus, in his *Spherics*, calls such a figure a trilateral (τρίπλευρον)." Pappus uses this word in the subsequent propositions. Another remark that fits here is that whereas Menelaus' proof of the triangle inequality is purely intrinsic, Pappus' proof uses the ambient Euclidean solid geometry (it is based on Proposition XI.20 of the *Elements*, which is a triangle inequality that concerns the three dihedral angles that form a solid angle.) The reason for which Pappus' proof must necessarily use Euclidean geometry is that this is the first proposition he gives on spherical geometry. As we already said, regarding Menelaus, spherical geometry, for both authors, based on the axioms of Euclidean geometry, and the first proposition, in any case, must use the Euclidean setting.

But the proofs of the next propositions of Pappus are intrinsic to the sphere. The statements are not always equivalent to those of Menelaus, and some of Pappus' propositions complement some propositions of his predecessor.

Let us now mention a few propositions of Pappus.

Proposition 2 says the following (Ver Eecke, 1982, p. 371):

> *If two arcs of great circles are drawn internally, above one side of a trilateral, then they are smaller than the remaining two sides of the trilateral.*

In other words, the proposition says that if we take an interior point in a triangle and if we join by two arcs of great circles that point to the vertices of the base, then the sum of these two arcs is smaller than the sum of the two legs of the triangle. This proposition is equivalent to Proposition 6 of the Spherics whose statement is (Rashed & Papadopoulos, 2017, p. 518):

> *If we produce from two extremities of a side among the sides of a triangle two arcs of great circles that meet within the triangle, then they are smaller than the remaining sides.*

Then, follow 9 propositions which are proved using Menelaus' methods. Some of the statements are close to propositions in Menelaus' Spherics, with slightly different hypotheses regarding the bounds on the lengths of the sides, and some other statements are contained in Theodosius *Spherics* (with proofs using Menelaus' intrinsic methods). As an example of Pappus' intrinsic methods, let us consider his proposition VI.4, which is equivalent to Proposition 38 of Menelaus (with a slightly different hypothesis concerning the lengths of the sides). The latter belongs to Group 4 of the Spherics, concerning the division of a triangle.

Menelaus' statement is the following (Rashed & Papadopoulos, 2017, p. 618):

> *If the sum of two unequal sides in a triangle is smaller than a semicircle, if we separate from the two extremities of the remaining side two equal segments, and if we produce from the two resulting points two arcs till the angle that the two sides contain, then they contain with the two sides two unequal angles, the greater at the smaller and the smaller at the greater, and the two produced arcs are smaller than the two sides of the triangle.*

Pappus' statement is the following (Ver Eecke, 1982, p. 373):

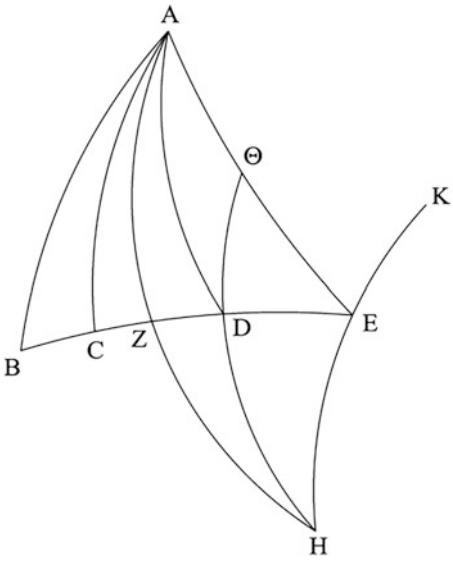


Fig. 18 The figure for Pappus' Proposition VI.4

Let four arcs of great circles AB, AC, AD, AE cut an arc of great circle BE, let the arc BC be equal to the arc DE, and let each of the arcs AB, AC, AD, AE be smaller than a quadrant. Then, the sum of the arcs BA, AE is greater than the sum of the arcs CA, AD.

We reproduce Pappus' proof in order to show the method (see Fig. 18).

Let us cut the arc *CD* into two equal parts at the point *Z*. Let us draw through the points *A*, *Z* the arc of great circle *AZH*, and let us set the arc *ZH* to be equal to the arc *AZ*. Let us draw through the points *H*, *E* the great circle *HEK* and through the points *H*, *D* the great circle *HD*Θ. Since the arc *HZ* is equal to the arc *ZA* and the arc *DZ* is equal to the arc *ZC*, it follows that the arc *DH* is also equal to the arc CA. For the same reason, the arc *EH* is also equal to the arc *BA*. But since the two arcs *AD*, *DH* are established internally on one of the sides *HA* of the trilateral *HEA*, it follows that the arcs *AD*, *DH* are smaller than the arcs *AE*, *EH*. Thus, the arcs *AE*, *EH* are greater than the arcs *AD*, *DH*. But the arc *EH* is equal to the arc *AB* and the arc *HD* is equal to the arc *AC*. Therefore, the sum of the arcs *BA*, *AE* is greater than the sum of the arcs *CA*, *AD*. This is what we had to prove.

I know of no other mathematician than Pappus before the modern times (that is, before the nineteenth century) who integrated in his writings Menelaus' methods of intrinsic geometry of the sphere.

6 Translations, Rectifications and Difficulties

Everything we know about the history of the translations and the various versions of Menelaus' *Spherics* is reported on in the first part of the edition (Rashed and Papadopoulos 2017), pp. 3–123 and in the two articles (Rashed and Papadopoulos

2014; Rashed and Papadopoulos 2015). I will summarize here some important points on that history.

Several Arabic versions of the *Spherics* were written starting in the ninth century. Al-Harawī's version, which relies on a version by al-Māhānī and which is a rectification[16] rather than a translation, is published for the first time, together with an English translation, in the volume (Rashed and Papadopoulos 2017). Besides al-Harawī's version, this volume contains the first critical edition of a fragment containing the 36 first propositions of a translation done in the same century, which we think is the oldest translation that survives. We know that another translation was also done in the same century, by the Arab mathematician and Hellenist Abū 'Uthmān al-Dimashqī. A Latin translation by Gerard of Cremona (1114–1175) and a Hebrew one by Jacob ben Makhir Ibn Tibbon (ca. 1273) were done from Arabic texts. According to the latter, there existed another Arabic translation by Ishāq ibn Ḥunayn, or by his father Ḥunayn ibn Ishāq. Besides these translations, there are also rectifications by Abū Naṣr Manṣūr Ibn 'Alī Ibn 'Irāq (born around the year 950), Naṣīr al-Dīn al-Ṭūsī (thirteenth century), Ibn Abī Jarrāda (thirteenth century), Muḥyī al-Dīn al-Maghribī (thirteenth century), al-Yazdī (fifteenth century), and certainly others. A German translation of the rectification by Irāq was published in 1936 by M. Krause, (Krause 1936). These versions differ on several points, with additions of lemmas and intermediate propositions, etc., but the overall mathematical content is the same. There are also writings and developments based on Menelaus' *Spherics* by Ibn Hūd, Thābet ibn Qurra, al-Bīrūnī, al Qiftī, Ibn al-Haytham, and others. We have discussed all this in the volume (Rashed and Papadopoulos 2017) where we have also made comparisons between significant passages of the existing Arabic versions that flourished from the ninth century on and continued till the thirteenth, and the Latin and Hebrew versions.

We have quoted, in the introduction of the present chapter, an excerpt of the Preamble of al-Harawī's rectification, in which the latter expresses his wonder and admiration for Menelaus' geometric reasoning, which he considers as completely new. In the same Preamble, al-Harawī mentions the difficulties inherent to the various translations, due to the fact that the geometers of his time were not capable of understanding several propositions and proofs of this treatise. He writes (Rashed & Papadopoulos, 2017, p. 24):

> When, from that book, the geometers received what they were not used to, and to nothing similar from all other kinds of geometries they knew, they found it difficult. The bad

[16] the Arabs used the words Iṣlāḥ إِصلَاح, which means "reform," or "revision" (an example is the treatise by Abū Naṣr Manṣūr ibn 'Alī ibn 'Irāq, *Iṣlāḥ Kitāb Mānālāwus fī al-ashkāl al-kuriyya*, Leiden, Or. 930), and Taḥrīr تَحرير, which means "redaction" (an example is Nasīr al-Dīn al-Ṭūsī's Taḥrīr Kitāb Mānālāwus fī al-ashkāl al-kuriyya [Hyderabad, 1359 H.]).

> translation and the fact that it was far from being understandable added to their apprehension, so they discarded it and they did not work on it. Nevertheless, no one among them denied the eminence of that book, but they acknowledged that what he showed in this way is easier and more valuable that anything which anybody else invested in.
>
> It was said that a group of geometers had the intention to rectify it, and since they were unable to do it, they asked the help of al-Māhānī, who rectified the first book and some propositions of the second, and he stopped at a proposition of which they said that its aim was difficult, and whose explanation was also hard.
>
> [. . .] I meditated on what al-Māhānī rectified and I saw during this short period of time that there were defects, and I carefully reviewed what needed to be corrected in the expression, in the meaning, and in the proof. And I persisted until the end of the book, until the proposition which the geometers imagined that al-Māhānī left aside and was not able to rectify. I also found a rectification that was far from being correct due to some modern geometer who claimed he rectified part of it, and left aside the rest. In the thing he said that he rectified there are some corrupt things, from which it was clear to anybody who examined them that he did not understand the purpose of this man [i.e., Menelaus].

According to the testimony of Naṣīr al-Dīn al-Ṭūsī, the situation of the existing versions of the *Spherics* in the thirteenth century was not better, except for the one by Ibn 'Irāq. He writes:[17]

> When I reached the book of Menelaus on the spherical figures, I found it in many different copies, with undetermined problems, together with disappointing rectifications, like the rectifications of al-Māhānī, of Abī al-Faḍl Aḥmad ibn abī Sa'ad al-Harawī and others, some of which are incomplete and others wrong. I was puzzled for what concerns the proof of some problems, until I obtained the rectification of Abū Naṣr Ibn 'Irāq.

It is likely that Ibn 'Irāq's is a rectification of the translation due to the Hellenist and mathematician Abū 'Uthmān al-Dimashqī. We refer again the reader to our discussion in (Rashed and Papadopoulos 2017).

In addition to his redaction of the *Spherics*, Ibn 'Irāq wrote a short treatise which concerns Proposition 71 which we discussed above (in our classification on Sect. 3, we included it in Group 8), titled *The Rectification of a Proposition of the Book of Menelaus on the Spherics*. A translation of this treatise is provided in the article (Rashed and Papadopoulos 2014).

Ibn Abī Jarrāda, a contemporary of Tūsī, considered that the most popular rectification that existed at his time was still that of al-Māhānī/al-Harawī, which, despite the fact that, according to him, it has many short-comings. Let us quote him (see (Rashed & Papadopoulos, 2017, p. 18)):

> I found that book (Menelaus' *Spherics*) so altered as none of the other geometry books, for reasons that were specific to it, among which are the difficulties in the proofs, for someone who was supposed to clarify them, because of the necessity of these conditions. And among these difficulties is the fact that Menelaus over-summarized it, as he mentioned in the Preamble. And also, among these difficulties, the multiplicity of the problems hidden in the course of the proofs, and of the abandoned ones to which he refers. Among these, that he

[17] Reference in (Rashed & Papadopoulos, 2017, p. 16)

> meant by the ratio of the arcs the ratio of the double of their chords in the entire third book and in part of the fourth, and he meant by the ratio of the arcs their own ratio in other parts of the fourth book, thinking, because of his extreme power in that field, that this would not be hidden to anyone who will undertake the clarification of his book. That was his basis, and it became difficult to anyone who came after him to recognize this. Part of <the difficulties> is also the fact that he refers in the proofs of the propositions of the fourth book to what he has proved in one of his books, titled *The Measurable Figures*, a book which is unknown, or else it is foreign and did not reach us.
>
> A group <of people> proposed to rectify this book; all of them had failed the aim, they stopped before reaching their goal and they did not comply with Menelaus' method. Among them are those who were not able to penetrate it. There are also those who rectified one part of the book, in the given order, and were incapable of going through it, like al-Māhānī, who reached a proposition which he was incapable of solving, and thus he did not complete it. Among them are also those who pretended they have rectified all the book, deluding and disappointing anyone who examines it. They relied on the fact that when one encounters a difficulty in understanding it, he attributes this lack either to himself or to the corruption of the transcription.
>
> Even though [al-Harawī's] rectification is the most famous one among those which the book was subject to, the best that he brings is where he limits himself to repeat Menelaus' expression. But as soon as he proposes himself to prove a neglected <problem> or to clarify a difficulty, or something similar, then he leaves the conditions of Menelaus, like in the first lemma by which he proved the proposition where al-Māhānī stopped, proving it with common sections whereas what he brought does not suffice for the completion of that proposition. Or as he did in the thirty-seventh proposition of the first book: he resorted to a *reductio ad absurdum* for what is hidden in the proof of Menelaus. Regarding the fourth book, which is the third in his rectification, he did not understand anything of it except the ninth proposition and he attributed everything which is in that book concerning the ratios of arcs to the ratios of chords of the double. Everything that I mentioned should be clear to anyone who examines the rectification of al-Harawī.

This translation and rectification work on Menelaus' *Spherics* generated an important mathematical activity that changed the course of Arabic mathematics. This is the subject of the next section.

7 On the Impact on Arabic Mathematics

Among Menelaus' propositions that had a major impact on Arabic mathematics is the Sector Figure (Proposition 66). Several treatises were written on the applications of this proposition. One such oldest known complete treatise is of Thābit ibn Qurra (d. 901), titled *A Treatise on the Sector Figure*, is reviewed and commented on in detail by H. Bellosta in her article (Bellosta 2009). We quote from the Introduction to this treatise (Bellosta, 2009, p. 365):

> I know that people never discussed any geometrical proposition used in astronomy as much as they did for this proposition. They did not overrate its reputation and the reason of their interest for it; this is what we know from the number of times it was used and the extreme need that we have for it in the spherical science. This is certainly the main direction on which

> many works in astronomy are based. Even though others than Ptolemy obtained this proposition and talked about it, nobody, to our knowledge, improved its statement or < gave> any arrangement or revision that matches its proof.

Other developments concerning the Sector Figure were reported on by Abū al-'Abbās al-Nayrīzī and Abū Ja'far Muḥammad ibn al-Ḥusayn al-Khāzin, in their commentaries of the Almagest. There is a large literature on the impact of this proposition on Arabic mathematics. The reader may refer to the work of M.-Th. Debarnot (Debarnot 1996; Debarnot 2002), and the latter's edition of the tenth century work of al-Bīrūnī (Debarnot and Al-Bīrūnī 1985). We have commented on this proposition and its applications in the commentary contained in our edition of the *Spherics*; see (Rashed & Papadopoulos, 2017, pp. 292–337).

In the thieteenth century, Naṣīr al-Dīn al-Ṭūsī, wrote a treatise on the Sector Figure (Naṣīr al-Dīn al-Ṭūsī 1891). in which he showed how a complete trigonometric system of spherical geometry can be derived from this proposition. Al-Ṭūsī's treatise was translated into French and edited by Alexandre Carathéodory[18] in 1891 (Naṣīr al-Dīn al-Ṭūsī 1891). See the detailed review in (Papadopoulos 2018). It is only in the eighteenth century that Leonhard Euler published a treatise containing such a complete system, obviously, without being aware of al-Ṭūsī's work, see (Euler 1753). Naṣīr al-Dīn uses in his work the notion of polar triangle. This notion and the spherical Sine Rule were already introduced in the tenth century by Ibn 'Irāq and Abū al-Wafā' al-Būzjānī (940–998). We have recalled, in (Rashed & Papadopoulos, 2017, pp. 795–801), the story of these discoveries and we have pointed out the confusion that reigned in the eighteenth and nineteenth century literature concerning the history of spherical trigonometry, in particular concerning the discovery of the polar triangle.

Naṣīr al-Dīn's treatise is a remarkable historical document that gives a very good idea of the state of spherical geometry, as developed by the Arabs between the ninth and the thirteenth centuries. At the end of Book IV, Naṣīr al-Dīn says that the Ancients have used the quadrilateral figure with confidence, without treating the large number of cases, and he mentions as examples Menelaus' *Spherics* and Ptolemy's *Almagest*. But the moderns, he says, in order to avoid engaging themselves in the examination of the multitude of ratios and the lengthy considerations that the use of compounded ratios requires, imagined and studied other figures, which replace the quadrilateral, which are as much useful and which do not lead to a multitude of cases. The presentation of these methods is the subject of Book V of his treatise.

We already mentioned another theorem of the *Spherics* that had a large impact on Arabic mathematics, especially on the development of spherical geometry with works by al-Būzjānī, al-Khujandī, Ibn Labbān, Ibn 'Irāq, and al-Bīrūnī. This is Proposition 71. As we already recalled, Menelaus did not give a proof of this

[18] Alexandre Carathéodory was a great-uncle of the famous mathematician Constantin Carathéodory.

proposition, but only an outline of a proof. Writing a complete proof became a big challenge that was taken up by several mathematicians, in Baghdad and elsewhere, between the ninth and the thirteenth centuries. We already mentioned that the proof of this proposition uses an invariance property of the spherical cross ratio, and to prove this invariance, the mathematicians of the tenth century used the spherical Sine Rule which they had established. It is conceivable that the development of this rule was motivated by the research in spherical geometry due to the effort to prove this proposition of Menelaus. In fact, the search for a proof of this proposition acted as a motivation for reshaping the whole field of spherical trigonometry. Likewise, it is possible that the discovery by Ibn ʻIrāq of the notion of polar triangle was also motivated by his work on the proof of Menelaus' Proposition 71. In his *Rectification of a Proposition of the Book of Menelaus on the Spherics*, which is edited and translated in the paper (Rashed and Papadopoulos 2014), Ibn ʻIrāq presents a set of propositions on spherical geometry that are useful in the proof of this proposition, but that are also of independent interest. At the same time, he gives a proof of the proposition under assumptions that are weaker than the ones that Menelaus made. While he worked on the proofs, Ibn ʻIrāq declared that not only Proposition 71, but the whole treatise of Menelaus' *Spherics* needs to be corrected. He reported on the attempts by al-Māhānī and al-Harawī on correcting Menelaus' treatise and he discussed several points in their proofs. We have discussed these points in our article (Rashed and Papadopoulos 2014).

Understanding the proof of Proposition 71 became an important matter, not only for mathematicians, but also for astronomers and other scientists. The last section of our chapter concerns astronomy.

8 On the Impact on Astronomy

In his introduction to his rectification, al-Harawī writes (Rashed & Papadopoulos, 2017, p. 501): "Ptolemy relies on this book, especially in the second book of the *Almagest* concerning angles and triangles produced by intersections of great circles." In the same introduction (Rashed & Papadopoulos, 2017, p. 504), he writes: "This book has an immense value in astronomy and the properties of the figures that occur when the horizons intersect the ecliptic, the orbs of the meridian, the equator and the circles called the ephemerides that the sun draws at the beginning of the zodiac." Concerning Menelaus' Proposition 66 (the Sector Figure), al-Harawī; writes (Rashed & Papadopoulos, 2017, p. 684):

> And he worked out in that book the figure which Ptolemy calls the "sector", and he constructed upon it several figures. But Ptolemy uses, from that book, numerous propositions in the second book of his treatise the *Almagest*, without attributing them to anybody and without proving anything. Indeed, all what he uses, for what concerns angles, arising from the intersection of the ecliptic and the horizons, and other things, can be proved by this book. The lemmas we need are among those which Ptolemy introduced, on the intersection of two lines and on the composition of the ratios which are made up thereof. And we find that

> Menelaus moved in that book to the Sector Figure in a way which does not fit his method, since he did not set for this passage any lemma or any discourse and he also did not make it the starting of a book. So either the lemmas of that figure were known by them, widespread, or they disappeared from that book.

The astronomer Ibn al-Haytham, who was familiar with the Sine Rule and who used it, also used the Sector Figure Theorem in his astronomical works.[19]

With all this, we must say that the astronomical side of Menelaus' *Spherics* has been so much exaggerated to the point that some authors have considered that the work (or at least part of it) was written for its use in astronomy. The confusion started with authors like Bjørnbo and Heath who (mistakenly) declared that a substantial part of the *Spherics* is astronomical, or that it has been written for the purpose of astronomy.[20] These statements have been repeated tirelessly and superficially by several modern authors. The first historians to put forward these ideas were probably convinced that among the Greeks, the only sphere worth studying was the celestial sphere, an idea that we, mathematicians, find hard to believe, knowing that the sphere and its intersections by planes is as natural a field of mathematics to study, as are the cones and their intersections by planes, which are the subject of study of Apollonius' *Conics*, and, in fact, as the ordinary plane with its straight lines. It is probable that Heath was influenced in his opinion by A. A. Bjørnbo (Bjørnbo 1902). In (Rashed & Papadopoulos, 2017, p. 3) we have given a list of authors from before the 1950s who repeated the statement that Menelaus' *Spherics* (or at least a substantial part of it) is a work on astronomy. The names of many modern authors may be added to this list, we prefer not to dwell on this.

Another statement of Heath which led to a misinterpretation is that he declared that Book III of Menelaus' *Spherics* belongs to "trigonometry,"[21] a statement which is also absurd. Heath writes in (Heath, 1921, p. 265): "It will have been noticed that, while Book I of Menelaus gives the geometry of the spherical triangle, neither Book I nor Book II contains any trigonometry. This is reserved for Book III." In fact, spherical trigonometry has been developed only several centuries later. The only reasonable meaning that can be given to Heath's statement is that Book III (like Books I and II) constitute a study of spherical triangles, but if such a study

[19] See e.g. Proposition 5 of the treatise of Ibn al-Haytham titled *The Configuration of the Motion of Each of the Seven Wondering Stars*, p. 294 sqq., in: Roshdi Rashed, *Les mathématiques infinitésimales du IXe au XIe siècle*, vol. V: *Ibn al-Haytham. Astronomie, géométrie sphérique et trigonométrie*, London, 2006.

[20] Heath's statement, concerning Menelaus' *Spherics*, in (Heath, 1921, p. 265), is bewildering: "Book II has practically no interest for us. The object of it is to establish certain propositions, of astronomical interest only, which are nothing more than generalizations or extensions of propositions in Theodosius *Sphaerica*, Book III."

[21] In Heath's treatise (Heath 1921), the study of Menelaus' book is part of the chapter called *Trigonometry*.

is termed as trigonometry, then several books of Euclid's *Elements* should also be termed as *trigonometry*. In fact, Heath says (p. 260) that "Greek trigonometry in its highest development [is] in the *Spherica* of Menelaus," which is correct, but it is only a relative statement. It is correct if we mean by this "trigonometry" in its etymological sense, the study of triangles, but then we have to include also Books I and II of the work.

On the other hand, taking about Greek astronomy, it is fair to recall that Theodosius wrote (at least) two astronomical treatises in which the celestial phenomena are studied from a geometric point of view, namely, *On habitations* and *On the days and nights*. The treatise *On habitations* contains 12 propositions which are all concerned with how the stars in the celestial sphere are seen from different places of the Earth. A brief paraphrase of the propositions is contained in Delambre's *Histoire de l'astronomie ancienne*, (Delambre 1817), vol. 1, pp. 235–7. I have reproduced in Fig. 19 a page containing Proposition 9 of the Habitations, which concerns a comparison between the location of the stars that are visible and the duration of the visibility of these stars, at two places that are not situated on the same meridian. For a recent edition of the Arabic translation of the *On habitations*, due to Qusṭa ibn Lūqā appeared in 2011, together with the Latin translation (done from the Arabic) Gerard of Cremona and an English translation, see (Kunitzsch and Lorch 2011).

This being said, I would like to stress, first, on the fact, that Heath remains one of the most valuable specialists of Greek mathematics in modern history, and secondly, that, by an ironic twist of history, it is true that Ptolemy, in his great astronomical treatise, relied heavily on a proposition of Menelaus' *Spherics*, which is not an astronomical work, rather than on the astronomically oriented *Habitations* or *Days and nights* of Theodosius, or on the *Moving sphere* or the *Risings and settings* of Autolycus, the reason being of course that this proposition of Menelaus could be used in an effective way in astronomical calculations.

9 As a Word of Conclusion

In the preceding pages, I have talked about the history of Menelaus' *Spherics*, discussing its relation with ancient and modern research on geometry.

In the next two appendices, I have included modern proofs of two propositions of Menelaus, using spherical trigonometric formulae. Such modern proofs are straightforward and they sometimes give more precise statements than the synthetic arguments. However, there is an enormous chasm between these automatic proofs and the depth of Menelaus' geometric proofs, in which the etymology of the word *Theorem*[22] as a contemplation finds it full force. It is in such instances which after all are rare in the history of mathematics that one can recognize the seal of a real mathematician.

[22] From the Greek, θεωρέω, "I watch," the same origin than the word *theatre*.

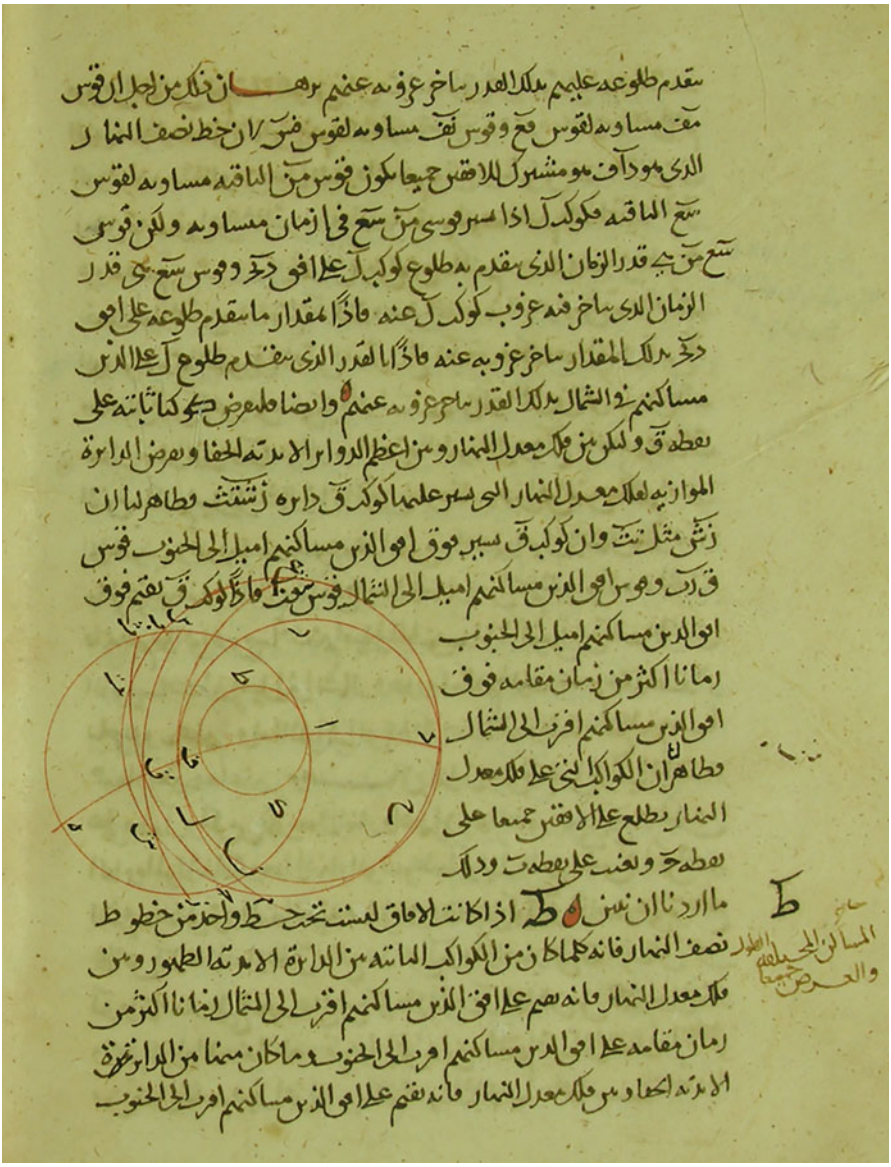


Fig. 19 From Theodosius' *Habitations*; at the bottom of the page, the statement of Proposition 9 and the corresponding figure. Version of Qusṭa ibn Lūqā, Ms. 3464, fol. 121, Topkapi Saray, Ahmet III

Acknowledgments This work obviously owes a lot to Roshdi Rashed with whom I collaborated for the edition (Rashed and Papadopoulos 2017), and I take this opportunity to express my admiration for all the work he has done for the history of Greek and Arabic mathematics.

Appendix I: A Proof of Proposition 27 by Spherical Trigonometry

We start with a formula of Euler from his memoir (Euler 1797) which says that the area of a spherical triangle ABC in terms of the lengths of sides a, b, c is equal to:

$$\frac{1+\cos b+\cos a+\cos c}{4\cos\frac{b}{2}\cos\frac{a}{2}\cos\frac{c}{2}}.$$

Consider now a triangle ABC, let D and E be the midpoints of AB and BC, and DE the arc of great circle joining them.

The cosine formulae in the triangles ABC and ADE give:

$$\cos B=\frac{\cos a\cos c-\cos b}{\sin a\sin c}$$

and

$$\cos B=\frac{\cos\frac{a}{2}\cos\frac{c}{2}-\cos DE}{\sin\frac{a}{2}\sin\frac{c}{2}}.$$

Equating the two values of $\cos B$, we obtain

$$\cos DE=\frac{(\cos b-\cos a\cos c)\left(\sin\frac{a}{2}\sin\frac{c}{2}\right)+\cos\frac{a}{2}\cos\frac{c}{2}\sin a\sin c}{\sin a\sin c}.$$

Replacing in the right hand side $\cos a$ and $\cos c$ by their values in terms of $\sin\frac{a}{2}$, $\sin\frac{c}{2}$, $\cos\frac{a}{2}$ and $\cos\frac{c}{2}$, and after some cancellations, we get:

$$\begin{aligned}\cos DE&=\frac{\cos b+2\cos^2\frac{a}{2}+2\cos^2\frac{c}{2}-1}{4\cos\frac{a}{2}\cos\frac{c}{2}}\\&=\frac{\cos b+(1+\cos a)+(1+\cos c)-1}{4\cos\frac{a}{2}\cos\frac{c}{2}}\\&=\frac{1+\cos b+\cos a+\cos c}{4\cos\frac{a}{2}\cos\frac{c}{2}}\\&=\frac{1+\cos b+\cos a+\cos c}{4\cos\frac{b}{2}\cos\frac{a}{2}\cos\frac{c}{2}}.\cos\frac{b}{2}\end{aligned}$$

Using Euler's formula we get

$$\cos DE=\cos\frac{1}{2}\mathrm{Area}(ABC)\cos\frac{b}{2},$$

which implies $DE<AC/2$.

Appendix II: A Proof of Proposition 71 by spherical Trigonometry

The following formula gives a more precise version of Menelaus' proposition 71:

Theorem Given a spherical triangle ABC with right angle at A and acute angle at C, and such that AC is smaller than a quarter of a circle, then we have:

$$\frac{\sin(BC + AC)}{\sin(BC - AC)} = \frac{1 + \cos C}{1 - \cos C}.$$

In particular, the ratio of sines in the conclusion depends only on the angle C, which implies the result of Proposition 71.

The proof is given in the setting of hyperbolic geometry in the paper (Papadopoulos and Su 2017). That proof is based on some non-Euclidean trigonometric transformations. It is easily adapted the spherical case, replacing the hyperbolic functions sinh and cosh by the functions sin and cos. We use the following lemma, which follows from the spherical cosine and Sine Rules.

Lemma We have

$$\tan BC = \cos C \cdot \tan AC.$$

Then we expand $\sin(BC + AC)$ and $\sin(BC - AC)$:

$$\frac{\sin(BC + AC)}{\sin(BC - AC)} = \frac{\sin BC \cos AC + \sin AC \cos BC}{\sin BC \cos AC - \sin AC \cos BC}$$
$$= \frac{\dfrac{\sin AC}{\cos AC} + \dfrac{\sin BC}{\cos BC}}{\dfrac{\sin BC}{\cos BC} - \dfrac{\sin AC}{\cos AC}}.$$

Using the lemma, the last term becomes

$$\frac{1 + \cos C}{1 - \cos C},$$

which is what we wanted to prove.

References

Autolycus de Pitane A (1979) In: Aujac G, de Brunet JP, Nadal R (eds) La sphère en mouvement, Levers et couchers héliaques, Testimonia. Les Belles Lettres, Paris

Barbarin P (1902) La géométrie non euclidienne. C. Naud, Paris

Bellosta H (2009) Le traité de Thābit ibn Qurra sur la figure secteur In: Rashed R (ed) Thābit ibn Qurra: Science and Philosophy in Ninth-Century Baghdad, Scientia Graeco-Arabica 4, Berlin, New York, pp 335–390

Bjørnbo AA (1902) Studien über Menelaos' Sphärik: Beiträge zur Geschichte der Sphärik und Trigonometrie der Griechen, Leipzig

Busemann H (1948) Spaces with non–positive curvature. Acta Math 80:259–311

Busemann H (1955) The geometry of geodesics. Pure Appl. Math. 6. Academic Press Inc., New York

Busemann H (1961) The synthetic approach to Finsler spaces in the large. In: Geometria del calcolo delle variazioni. C.I.M.E. Summer Schools, vol. 23. Springer, Berlin, Heidelberg

Clavius's commentary on the Spherics of Theodosius Tripolitæ: or, Spherical elements, necessary in all parts of mathematics, wherein the nature of the sphere is considered, English transl. E. Stone, London, Senex, Taylor and Sisson, 1721

Debarnot MTH (1996) Trigonometry. In: Rashed R (ed) Encyclopedia of the history of Arabic science, 3 vols, vol. II, London and New York, pp 160–203

Debarnot M-T (2002) Trigonometria. In: Storia della scienza, vol. III: La civiltà islamica. Roma, Enciclopedia Italiana, pp 432–447

Debarnot M-T (ed.) (1985) Al-Bīrūnī, Kitāb Maqālīd 'ilm al-hay'a: La trigonométrie sphérique chez les Arabes de l'Est à la fin du Xe siècle, ed. Institut Français du Proche-Orient, Damascus, Vol. 114, 314 p

Delambre J-B (1817) Histoire de l'astronomie ancienne. Courcier, Paris

Euler L (1753) Principes de la trigonométrie sphérique tirés de la méthode des plus grands et des plus petits, Mémoires de l'académie des sciences de Berlin, Vol. 9, 27, pp 223–257. Opera Omnia, Series 1, 277–308. English translation by A. Papadopoulos and A. Zhukova, In: *Spherical geometry in the eighteenth century: Euler, Lagrange and Lambert*, Springer, 2024, to appear

Euler L (1797) Variae speculationes super area triangulorum sphaericorum. *Nova Acta Academiae Scientarum Imperialis Petropolitinae* 10(1797):47–62. In: *Opera Omnia*, Ser. 1, Vol 29, pp 253–266. English translation by R. Caddeo, In: *Spherical geometry in the eighteenth century: Euler, Lagrange and Lambert*, Springer, 2024, to appear

Gérard L (1892) Sur la géométrie non euclidienne, Thèse de doctorat présentée à la Faculté des Sciences de Paris. Gauthier-Villars, Paris

Halma N (1813) Composition mathématique de Claude Ptolémée, traduite par Nicolas Halma et suivie des notes de Jean Baptiste Delambre, 2 vols. Henri Grand, Paris. Reprint, 1988, Librairie Blanchard, Paris

Halsted GB (1885) The elements of geometry. J. Wiley and Sons, New York

Heath TL (1921) A History of Greek Mathematics, vol. II: From Aristarchus to Diophantus, Oxford. repr. Dover, 1981

Krause M (1936) Die Sphärik von Menelaos aus Alexandrien in der Verbesserung von Abū Naṣr Manṣūr b. 'Alī b. 'Irāq, Abhandlungen der Gesellschaft der Wissenschaften zu Göttingen, phil.-hist, vol 3. Klasse, Berlin

Kunitzsch P, Lorch R (2011) Arabic and Medieval Latin Translations. In: Kunitzsch P, Lorch R (eds) Bayerische Akademie der Wissenschaften, phil.-hist. Sitzungsberichte, Klasse. Heft i; C.H. Beck, Munich, 2011

Lexell L (1784) Solutio problematis geometrici ex doctrina spaericorum, Acta academiae scientarum imperialis Petropolitinae 5(1): 112–126, English translation by R. Caddeo in: Spherical geometry in the Eighteenth century: Lexell, Schubert and Fuss (R. Caddeo, A. Papadopoulos and A. Zhukova ed.), Springer, 2024, to appear

Mersenne M (1626) In: Mersenni FM (ed) Universae geometriae mixtaeque mathematicae synopsis, et bini refractionum demonstratarum tractatus. Studio & opera

Naṣīr al-Dīn al-Ṭūsī (1891) Naṣīr al-Dīn al-Ṭūsī, Traité du quadrilatère, attribué à Nassiruddin-el-Toussy, d'après un manuscrit tiré de la bibliothèque, ancien grand vizir, traduit par Alexandre Pacha Carathéodory, ancien ministre des affaires étrangères. Constantinople, Typographie et Lithographie Osmanié

Papadopoulos A (2018) Naṣīr al-Dīn al-Ṭūsī's treatise on the quadrilateral: The art of being exhaustive. Gaṇita Bhārātī, (Indian Mathematics) the Bulletin of the Indian Society for History of Mathematics 40(1):153–180

Papadopoulos A (2019) Three theorems of Menelaus. Am Math Mon 126(7):610–619
Papadopoulos A, Su W (2017) On hyperbolic analogues of some classical theorems in spherical geometry. In: Fujiwara K, Ohshika K, Kojima S (eds) Hyperbolic geometry and geometric group theory. Advanced Studies of Pure Mathematics, No. 73, Mathematical Society of Japan, Tokyo, pp 225–253
Papadopoulos A, Yamada S (2013) The Funk and Hilbert geometries for spaces of constant curvature. Monatshefte für Mathematik 172(1):97–120
Rashed R (2006–2014) Ibn al-Haytham, Les Mathématiques infinitésimales du IXe au XIe siècle, vol. V. Ibn al-Haytham: Astronomie, géométrie sphérique et trigonométrie, London. Al-Furqān Islamic Heritage Foundation, New Spherical Geometry and Astronomy, A History of Arabic sciences and mathematics, vol. 4. London, New York
Rashed R, Papadopoulos A (2014) On Menelaus' *Spherics* III.5 in Arabic mathematics, I: Ibn 'Irāq. Arabic Science and Philosophy 24:1–68
Rashed R, Papadopoulos A (2015) On Menelaus' *Spherics* III.5 in Arabic mathematics, II: Naṣīr al-Dīn al-Ṭūsī and Ibn Abī Jarrāda. Arabic Science and Philosophy 25:1–32
Rashed R, Papadopoulos A (2017) Menelaus' Spherics: early translation and al-Māhānī/al-Harawī's version. De Gruyter, Series: Scientia Graeco-Arabica 21:890
Senapati H (2019) Monotonicity in spherical and hyperbolic triangles. In: Eighteen Essays in Non-Euclidean Geometry, European Mathematical Society
Theodosius, Les Sphériques de Théodose de Tripoli (1927) Œuvres traduites pour la première fois du grec en français avec une introduction et des notes par P. Ver Eecke. Desclée de Brouwer, Bruges. repr. Paris, Librairie Albert Blanchard, 1959
Ver Eecke P (1982) La Collection mathématique de Pappus d'Alexandrie. Librairie Albert Blanchard, Paris

Copernicus and Axiomatics

Alberto Bardi

Contents

1 Introduction 1790
2 Axiomatics Before Axiomatics 1791
3 Questions About Copernicus's Axioms 1793
3.1 Terminology 1794
3.2 Logic 1797
4 Copernicus's Axiomatic Approach 1798
5 An Interpretive Framework 1799
6 Conclusion 1801
7 Cross-References 1803
References 1803

Abstract

The debate about the foundations of mathematical sciences traces back to Greek antiquity, with Euclid and the foundations of geometry. Through the flux of history, the debate has appeared in several shapes, places, and cultural contexts. Remarkably, it is a locus where logic, philosophy, and mathematics meet. In mathematical astronomy, Nicolaus Copernicus's axiomatic approach toward a heliocentric theory of the universe has prompted questions about foundations among historians who have studied Copernican axioms in their terminological and logical aspects but never examined them as a question of mathematical practice. Copernicus provides seven unproved assumptions in the introduction of the brief treatise entitled *Nicolaus Copernicus's draft on the models of celestial motions established by himself*, better known as *Commentariolus* (ca. 1515), published circa 30 years before the final composition of his heliocentric theory (*On the revolutions of the heavenly spheres*, 1543). The assumptions deal with the renowned Copernican hypothesis of considering the Earth in motion and the Sun, not affected by motion,

A. Bardi (✉)
Department of the History of Science, Tsinghua University, Beijing, China
e-mail: brd@tsinghua.edu.cn

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_110

near the center of the universe. Although Copernicus decides to omit the proofs for the sake of brevity, the deductions in the *Commentariolus* are supposed to be drawn from the initial seven assumptions. Questions on the nature (are they postulates or axioms?) and the logic (is there an internal rigor?) of those assumptions have yet to be fully explored. By examining Copernicus's seven assumptions as a question of mathematical practice, it is possible to hold historical, philosophical, and logical aspects of Copernican axiomatics together and understand them as part of Copernicus's intuition and creativity.

Keywords

Axioms · Axiomatics · Axiomatization · Axiomatic method · Copernicus · Commentariolus · Creativity · Discovery · Euclid · Feferman · Heliocentrism · History of axiomatics · Mathematical astronomy · Pluralism · Postulates · Proclus · Schlimm · Use of axioms

1 Introduction

The first written draft of the heliocentric theory by Nicolaus Copernicus dates back to circa 1515 and is attested in an unpublished manuscript, which was known – already among Copernicus's peers and acknowledged then among later generations of scholars and historians – as *Commentariolus* (brief commentary). The full title sounds *Draft on the hypothesis on celestial motions established by Nicolaus Copernicus* (*Nicolai Copernici de hypothesibus motuum caelestium a se constitutis commentariolus*). This work was likely composed between 1510 and 1515. It would later attract the attention of historians of astronomy, especially in the twentieth century (Swerdlow 1973; Goddu 2010, 243–274, 291; Folkert et al. 2019, 7–111). In history, among the scholars who had in their hands a copy of the *Commentariolus*, some renowned astronomers and mathematicians are to be mentioned: Joachim Rheticus, Duncan Liddel, and Tycho Brahe (Dobrzycki and Szczucki 1989; Omodeo 2016, 11; Folkert et al. 2019, 8–14). This aspect is particularly telling of the relevance that the *Commentariolus* assumed in the sixteenth century, let alone for the wide circulation and discussion of Copernican theory in Europe. Particularly noteworthy is the so-called Aberdeen edition, because it contains the pages of *Commentariolus* and *De revolutionibus orbium coelestium* (*On the revolutions of the heavenly spheres*, Copernicus's major work, see Copernicus 1978) in parallel, according to apparent affinity between the topics. This shows that the generation immediately following Copernicus considered the two works to be complementary and not mutually exclusive, though they do not exactly match: in fact, Copernicus refined, polished, and improved his theory between the composition of *Commentariolus* and the publication of *De revolutionibus* in 1543 (Dobrzycki 1973; Folkert et al. 2019: 8–9).

The *Commentariolus* has left historians puzzled on several issues. On the one hand, the aspects regarding the details of the motion of the Earth and the mathematics implicit in the text as well as the question on the derivation of the heliocentric

hypothesis (i.e., whether Copernicus arrived at the models independently or was possibly inspired by geometric models developed by Arabic astronomers) are still debated to this day (Swerdlow 1973; Blåsjö 2014; Ragep 2016; Swerdlow 2017; Blåsjö 2018). On the other hand, the seven propositions that introduce the heliocentric hypothesis in the *Commentariolus* have been the object of much scholarly discussion (Swerdlow 1973; Rosen 1976; Goddu 2010; Vesel 2014; Lerner and Segonds 2015; Folkert et al. 2019, 24–28). On the latter issue, the basic questions are as follows: What kind of propositions are they? Postulates, axioms, or something else? How should they be conceived and what kind of meaning does Copernicus attach to *axioms*? Do they have an internal rigor? Was Copernicus looking for logical rigor?

Copernicus does indeed adopt an axiomatic approach to his mathematical–astronomical theory, but the nature of his principles has confounded historians of science and generated a debate which continues to this day. Questions about Copernicus's axioms have hitherto been confined to historians of science, and studies of mathematical practice have never been employed to tackle such a problem.

This chapter examines Copernicus's axiomatic approach as a question of mathematical practice. First, it illustrates the main features of pre-modern axiomatics, evaluates the interpretations on *Commentariolus*'s principles advanced so far by historians, and eventually offers a new interpretive framework in which historical, philosophical, and logical details are comprehensively considered.

2 Axiomatics Before Axiomatics

One may ask: What does Copernicus have to do with axiomatics? Copernicus and his time seem far too distant from us today to suggest that there is any link to questions concerning "axiomatics" (in the modern sense of the term). In fact, the kind of questions that recent literature on mathematical practice has been tackling regarding axioms and their use is present already in the history of the reflections, discussions, and debates on the foundations of mathematical sciences (Schlimm 2013). These debates have their roots in Greek antiquity. For instance, there are traces of debates on foundational issues of mathematics in ancient and Hellenistic times (Acerbi 2010; Acerbi 2013; De Risi 2016). For centuries, the most successful reference work for Western mathematical sciences has been Euclid's *Elements*, and this work has also been taken as emblematic of axiomatic rigor. On the one hand, the theorems taken from the *Elements* were employed to build mathematical theories; on the other hand, the hypothetico-deductive structure of *Elements* was deemed the prototype of reason and certainty. On these grounds, it soon became an important object of study not only for mathematicians, but also for philosophers and logicians. The *Elements*' deductive system is based on principles, from which all subsequent deductions are to be drawn. A set of principles opens Euclid's *Elements*, Book One: definitions of the subject addressed, postulates, and axioms (common notions). The distinction between postulates and axioms was a topic of debates and controversies already in ancient times. However, the ancient conception of axiomatization was very different from our own. For centuries, mathematicians, logicians, and

philosophers did not look upon the definition of axioms as logically necessary, but rather saw them in light of the need to justify a "weak" science, which would not need to justify its principles if it was already consistent and, roughly speaking, logically coherent. This explains why the foundations of arithmetic had been (largely though not entirely) neglected until the nineteenth century, while the foundations of geometry were discussed since Antiquity, making Euclid's *Elements* the canon for Western mathematics (De Risi 2016).

Euclid's prevalence notwithstanding, the questions raised about principles were quite heterogeneous. Notably, influential scholars of the Greek tradition, such as Archimedes, Aristarchus, Apollonius, and others, did not assume principles in the same manner as Euclid did in the *Elements* and, in applied mathematics, such as optics, Euclid himself did not employ the same axiomatics as in the *Elements* (Capecchi 2018: 6–13). However, there was no homogeneity on the nature of principles even at a more general level. Notably, Aristotle's definition of postulates does not coincide with the Euclidean use, and Proclus, reprising Geminus (first century BCE), differs from other Greek philosophical schools (Bobzien 1996; Acerbi 2010; Acerbi 2013; Bobzien 2019).

In sum, the history of axiomatics has never been monolithically Euclidean. However, the focus on axiomatization of arithmetic in the nineteenth and twentieth centuries, after the works of Dedekind, Klein, Hilbert, Peano, Frege, and so on and the reflections on axiomatics by logicians like Carnap, Gödel, and Tarski and his followers, has led historians, philosophers, and logicians to study traces of "axiomatics before axiomatics" in mathematical sciences (optics, astronomy, mechanics), logic, and philosophy of antiquity till early modernity (Henkin et al. 1959; Hintikka et al. 1981; Novaes 2020; Cantù and Luciano 2021). Some influential works are worthy of note in this regard. Arpad Szabó was pioneering in his studies of the interplay of philosophical and mathematical influences in the development of Greek axiomatics and claimed that the Eleatic school influenced the rise of the deductive method in early Greek mathematics (Szabó 1964). Contrary to Szabó, Wilbur R. Knorr argued that the deductive procedures evident in fifth-century Greek works are intrinsic to mathematics and that the influence of the Eleatic dialectic on the development of axiomatics in arithmetic and geometry could only have occurred later, in the environment of the fourth-century Academy, in which the interest in mathematics merged with a renewal of the logical and philosophical views espoused by the Eleatics (Knorr 1981). Later on, Patrick Suppes shed light on axiomatics in Greek and Hellenistic science, pointing out their similarities with contemporary philosophical and scientific endeavors:

> the admiration many of us have for the rigor and relentlessness of the axiomatic method in Greek geometry has given us a misleading view of the role of this method in the broader framework of ancient Greek mathematical sciences. By stressing the limitations of the axiomatic method or, more explicitly, by stressing the limitations of the role played by the axiomatic method in Greek mathematical science, I do not mean in any way to denigrate what is conceptually one of the most important and far-reaching aspects of Greek mathematical thinking. I do want to emphasize the point that the use of mathematics in the mathematical sciences and in foundational sciences, like astronomy, compare rather closely with the contemporary situation. (Suppes 1993, 25)

As the above claims suggest, the possibility of external influences (from philosophy or logic) on an axiomatic method in mathematical sciences is a widely explored topic, and questions on this issue are also raised by the use of axioms in *Commentariolus*. Whereas this text has indeed pushed historians to search for possible philosophical influences on Copernicus's axiomatic approach, the possibility of this approach as intrinsic to Copernicus's mathematical practice has yet to be considered.

3 Questions About Copernicus's Axioms

Philadelphia, PA, 1973. At a symposium marking the 500th anniversary of Copernicus's birth, the eminent historian of science Noel Swerdlow presents a paper on the *Commentariolus* along with a translation into English and a commentary thereon. He states that Copernicus incorrectly called the seven sentences in the first section of the *Commentariolus* "axioms" and that those assumptions are illogically structured – the question on axioms in the *Commentariolus* became a topic of interest immediately following Swerdlow's translation of the Latin text (Swerdlow 1973). I will therefore report Swerdlow's translation since it is the one from which the discussions on axioms proceeded.

Swerdlow's critique refers to the first passages of *Commentariolus*, where Copernicus claims that Claudius Ptolemy and his predecessors were not coherent with the axiom of uniform circular motion, for they introduced geometrical devices which ultimately contradicted that principle and made cosmology excessively complicated. Ultimately, his predecessors failed to conceive of the universe as an integral whole (Goddu 2009). In the following, the words of Copernicus, writing in first-person singular, account for this state of the art:

> Therefore, when I noticed these [difficulties], I often pondered whether perhaps a more reasonable model composed of circles could be found from which every apparent irregularity would follow while everything in itself moved uniformly, just as the principle of perfect motion requires. After I had attacked this exceedingly difficult and nearly insoluble problem, it at last occurred to me how it could be done with fewer and far more suitable devices than had formerly been put forth if some postulates, called axioms, are granted to us, which follow in this order:
>
> First Postulate There is no one center of all the celestial spheres (*orbium*) or spheres (*sphaerarum*).
>
> Second Postulate The center of the earth is not the center of the universe, but only the center towards which heavy things move and the center of the lunar sphere.
>
> Third Postulate All spheres surround the sun as though it were in the middle of all of them, and therefore the center of the universe is near the sun.
>
> Fourth Postulate The ratio of the distance between the sun and earth to the height of the sphere of the fixed stars is so much smaller than the ratio of the semidiameter of the earth to the distance of the sun that the distance between the sun and earth is imperceptible compared to the great height of the sphere of the fixed stars.
>
> Fifth Postulate Whatever motion appears in the sphere of the fixed stars belongs not to it but to the earth. Thus the entire earth along with the nearby elements rotates with a daily motion on its fixed poles while the sphere of the fixed stars remains immovable and the outermost heaven.

> Sixth Postulate Whatever motions appear to us to belong to the sun are not due to [motion] of the sun but [to the motion] of the earth and our sphere with which we revolve around the sun just as any other planet. And thus the earth is carried by more than one motion.
>
> Seventh Postulate The retrograde and direct motion that appears in the planets belongs not to them but to the [motion] of the earth. Thus, the motion of the earth by itself accounts for a considerable number of apparently irregular motions in the heavens. (Translation Swerdlow 1973, 435–436)

In response to Swerdlow's criticism, Edward Rosen attempted to do justice to Copernicus's axiomatics (Rosen 1976). The problem would also be taken up again later (Goddu 2010; Vesel 2014; Lerner and Segonds 2015; Folkert et al. 2019; Goddu 2019, 162–163; more on this below). In sum, the questions on Copernicus's axiomatic approach touch on two key areas: *terminology* and *logic*.

3.1 Terminology

Let us recall Copernicus's words in *Commentariolus*:

> [. . .] it at last occurred to me how it could be done with fewer and far more suitable devices than had formerly been put forth if *some postulates, called axioms*, are granted to us, which follow in this order. (Swerdlow 1973, 435, emphasis added)

The original Latin text reads *petitiones quas axiomata vocant*, where *petitio* conveys a variety of meanings: requirement, statement, assumption, axiom, and postulate.

According to Swerdlow (1973, 437):

> The seven postulates, incorrectly called axioms since they are hardly self-evident, take the place of the general description of the universe in the opening chapters of the *Almagest*, the *Epitome*, and later, *De revolutionibus*. [. . .] There is no reason to doubt that he also believes these postulates to be true.

Swerdlow was undoubtedly referring to a precisely connotated meaning of axioms and postulates, where axiom is exclusively connected to self-evidence. This distinction is not modern; it traces back to the debates on Euclid's *Elements*, more precisely to the Greek philosopher Geminus (first century BCE), as attested in Proclus's *Commentary on the First Book of Euclid's Elements* (Acerbi 2010). According to Geminus-Proclus, the postulates of Euclid's *Elements* can be divided into two groups, reflecting their different nature: constructions are required in postulates 1–3, while postulates 4–5 state properties of particular geometric objects. As for the axioms (or common notions), they were generally conceived as assumptions conveying self-evident truths, hence requiring no proofs. Yet, not every author of mathematics assumed principles in the same manner as Euclid did. As anticipated above, Archimedes used axioms as true statements to describe the physical world, for example, Archimedes's *On the Sphere and the Cylinder* opens with axioms (Netz 2004, 34–36); Aristotle thought that "an

axiom is a primary proposition which must be possessed by whoever is to gain any knowledge"; "the axioms are the primary propositions from which a proof proceeds" (Aristotle 2002, I, 2: 72a17; I, 10: 76b14–15); the Stoics considered axioms as kinds of assertibles but did not employ the category of self-evidence in their logic, much to Proclus's disappointment (Ierodiakonou 2006).

Edward Rosen, author of translations of several Copernican treatises, *Commentariolus* included, accused Swerdlow of anachronism (Rosen 1976): Copernicus was not taking part in any Hilbertian program. Rosen's reply was beneficial to future generations of Copernicus scholars. André Goddu (Goddu 2010) undertook detailed research on the making of Copernican cosmology and his possible sources, detected his logical and philosophical backgrounds, and surmised a plausible method of reasoning in a Socratic-dialectic process, used by Copernicus in setting out his axioms. Goddu claimed that axioms are to be intended in the sense of common notions, assumptions to be taken for granted, where no self-evidence is needed:

> Copernicus began *Commentariolus* controversially with "petitiones" that he also called "axioms." If we set aside the personal attacks by some commentators, the experts agree, even if inadvertently, that he did not mean the word "axiom" in the sense of self-evident principles but rather in the sense of assumptions or common notions. As Copernicus himself made abundantly clear, the rest follows only if the seven postulates or assumptions are granted him. It is evident that he arrived at these seven propositions by working his way back to them as the ones necessary and sufficient from which to derive the remaining propositions. (Goddu 2010, 243)

Matjaž Vesel proposed a different interpretation. The axioms are set out and Copernicus draws his conclusions directly from them, and his source should be Proclus's commentary on Timaeus:

> I believe Copernicus' manner of exposition and the nature of *petitiones quas axiomata vocant* find their explanation in Proclus' *Commentary on Plato's Timaeus* II, 3 [. . .]. Proclus explains that Plato is not an empiricist: Plato will not start with experiences and then draw conclusions from them. Plato's *methodos* is hypothetical, or, rather, Plato uses the method of the hypothesis. He sets out fundamental *axiômata* and *hypotheseis* and draws conclusions from them. Proclus presents first a list of five axiomata, and then follows another list of seven *axiômata*. Describing Plato's "hypothetical method" Proclus does not refer to Plato's own description of hypothetical method but explicitly refers to the *methodos* used by geometers. They first postulate, define and name their key principles before proceeding to their demonstrations based on them. And he cites an example from Euclid. On the basis of the fundamental principles or hypotheses Plato's Timaeus then proceeds, in Proclus' reading of the text, to a number of "demonstrations" (*apodeixeis*) based on them and required in order to solve the problems [. . .]. Copernicus' method in the Commentariolus is highly reminiscent of Proclus: he first establishes seven *petitiones quas axiomata vocant* and then promises to provide mathematical *demonstrationes* in a larger book. (Vesel 2014, 269–270)

Michel-Pierre Lerner and Alain Segonds agree in part with Swerdlow on the first objection, for they claim that Copernicus's axioms are not self-evident (Lerner and Segonds 2015, 233). To some extent, they agree with Goddu when he writes that "the postulates are not axiomatic in the Euclidean sense, there is not a logically

deductive relationship between the postulates, and the results described later in the text derive from the postulate of a moving Earth" (Goddu 2019, 163).

All things considered, two things are noteworthy here. First, there is no occurrence of the equivalence *petitio* as a synonym of *axioma* in the Euclidean tradition (De Risi 2016). It is likely that Copernicus knew that axioms were known and evident to everybody, which could explain why he names his principles such. Second, the word *axioma* in Latin was extremely rare around 1515, when Copernicus composed the *Commentariolus*. Indeed, the only occurrence is in Giorgio Valla's *De expetendis et fugiendis rebus* (Valla 1501: Book 10, ch. 110), an encyclopedic work of sciences and arts. Valla was a humanist who was keen on transliterating Greek words into the Latin alphabet (De Risi 2016, 643; Goddu 2010, 229–236). In the passage in question, Valla is rephrasing the distinction between postulates and axioms, drawing from Proclus:

> But a postulate prescribes that we construct or provide some simple or easily grasped object for the exhibition of a character, while an axiom asserts some inherent attribute that is known at once to one's auditor. (Proclus 1970, 142)

It is likely that Copernicus read this work by Valla and thus wished to communicate to his readers that some humanists call the mathematical principles "axioms" and that it is not a case of distinguishing the Euclidean meaning of postulates and axioms because he is using axioms as a generic term for a principle to be taken for granted. In addition, as a trainee of mathematics and logic in Italy and Poland, Copernicus was certainly aware of the Euclidean meaning.

All this confirms that Swerdlow's interpretation has aprioristically assumed that axioms must be intrinsically connotated with self-evidence. As mentioned, Copernicus does not adhere to any Hilbertian project, nor is he dealing with geometry in the fashion of Peano's program on the foundations of geometry (Rosen 1976). Therefore, there is good reason to leave aside the option of "self-evidence" and instead look for an interpretation in the literature that Copernicus had at his disposal.

What has gone unnoticed so far is that Copernicus could refer to the Greek tradition of mathematical sciences, which was keen on using the axiomatic approach. For instance, several Greek works of mathematical sciences often adopted seven unproved assumptions to open treatises, such as Euclid's *Optics* (Burton 1945) and Archimedes's *On the Equilibrium of Planes* (Heath 2002, 189–190). It is likely that Copernicus arranged his treatises on seven principles in order to refer to that tradition, which could well have served as a model for him. Archimedes's *On the Equilibrium of Planes* was circulating in Italy thanks to the Latin translation of William of Moerbecke and that work opens with the word *petimus* (the verb linked to *petitio*, meaning "we require," "we state some principles"), which is followed by exactly seven principles (Clagett 1976, 116). Moreover, *Elementa Jordani* (*Elements of mechanics* by Jordan of Nemore) provides seven initial unproved assumptions (Clagett and Moody 1952, 154–155). Further research will clarify whether Copernicus could have had access to these sources, but it is likely that he could have assimilated, if not read, this literature in excerpts or classes at the universities he attended in Poland and Italy.

3.2 Logic

The second point of criticism aimed at Copernicus's axiomatic approach deals with the internal logic of his assumptions. According to Swerdlow:

> He [i.e. Copernicus] has, however, made more assumptions than were necessary. Since his adoption of heliocentric theory followed from his study of possible rearrangements of planetary models, only postulates 3 and 6 are temporally prior to his planetary theory. In fact only postulates 3 and 6 are logically prior while postulates 2, 4, 5, and 7 are consequences of 3 and 6, and postulate 1 stands by itself. Thus, if 3 and 6 are true, then 2, 4, 5, and 7 can be proved. Postulates 3 and 6 cannot be proved, and the evidence for their truth is the sense of the heliocentric theory itself, that is, the fixing of the order and distances of the planets, and the explanation of the second anomaly, gratuitously given here as the seventh postulate. It is also worth noting that the postulates, with the exception of the first, have no connection with the objections to Ptolemy's representation of the first anomaly stated earlier nor with Copernicus's own model for the first anomaly, but are concerned only with the heliocentric theory. Since Copernicus has raised no objection to Ptolemy's representation of the second anomaly, the introduction of these postulates at this point appears unmotivated. Perhaps this flaw and the logical error of stating postulates 2, 4, 5, and 7 as postulates rather than deductions from postulates 3 and 6 are intelligible if one considers that the *Commentariolus* may well have been written in haste with no revision. (Swerdlow 1973, 437)

It is true that some axioms can be derived from others, but the whole problem cannot be comprehended if one does not advance another question, which in the logic of historical inquiry precedes Swerdlow's question: Is Copernicus seeking internal rigor in his assumptions? And what kind of rigor is required in the sixteenth century for a brief text of mathematical astronomy? Is he striving to give a list with as few assumptions as possible?

There is indeed a logic in Copernicus's axioms, although they are – strictly logically speaking – redundant, and this logic was brilliantly exposed by Edward Rosen (1976). The passage is reported here in full, and axioms are mentioned by numbers between parentheses.

> (1) undermines Eudoxus' principle of homocentricity by denying that everything in the universe is centered on the earth: "There is no one center." Consequently there are multiple centers, of which the earth is one. (2) puts the earth at the center, not of the universe, but of the moon's motions. Another center is the sun, placed by (3) near the center of the universe. The distance between these two centers, earth and sun, is said by (4) to be imperceptible in comparison with their distance from the stars. (5) The stars do not move, their apparent daily motion being due to the earth's axial rotation. (6) The sun does not move, its apparent motion being due to the earth's orbital revolution. (7) The apparent retrogression of the planets is likewise only an appearance due to the motion of the terrestrial observer. By common consent, (7) is Copernicus' single most valuable contribution to technical astronomy. (Rosen 1976, 47)

Later, the examination of Copernicus's logical background will allow André Goddu to convincingly describe the Copernican heliocentrism as an attempt toward envisioning the universe as an integral whole, thus in need of a "unique,

non-arbitrary, and commensurable structure benefitting the exquisite craftsmanship of its architect" (Goddu 2009, 332). Copernicus was indeed trained in *mereology*, the branch of medieval logic which deals with the relations between the whole and the parts (Goddu 2009).

Moreover, as anticipated above, the need for axioms in Copernicus's times was not seen as a logical necessity; indeed, it could even be seen to signal a lack of rigor, since a science should be justified without principles. The main axiom of astronomy was that celestial bodies move along uniform and circular motions, and Copernicus did not need to mention it in his list. In sum, Copernicus and his contemporaries would not deem the "redundancy" of axioms to be a mistake or a problem in the context of *Commentariolus* and the literature to which it refers (Ptolemy and ancient astronomers).

By considering all terminological and logical concerns in relation to Copernicus's assumptions, it is evident that Swerdlow (1973) was moving his objections from applying a part of the developments of modern axiomatics to a sixteenth-century context. It is therefore expedient to explore what contemporary studies on axiomatics might have to say about Copernicus's mathematical practice. Moreover, it is worth examining whether one should look for an external source or to study the axioms in *Commentariolus* as an integral part of Copernicus's creativity, a key feature in his development of the mathematical theory for heliocentrism.

4 Copernicus's Axiomatic Approach

Historians have focused on the terminology and the logic that the seven assumptions bring to the fore, but nobody has yet stressed that there is another axiom, which precedes the seven axioms and which was the common assumption for all scholars dealing with astronomy since antiquity. The additional axiom runs as follows: all celestial bodies move according to uniform and circular motions (Taub 1993; Feke 2018). In this instance, reflections on axiomatics in mathematical practice provide new tools to build an interpretive framework which unifies the shared fundamental axiom along with the seven axioms and their historical and philosophical aspects.

Dirk Schlimm (2013) has analyzed the use of axioms in contemporary mathematical practice and argued that "axioms can play many roles in mathematics and that viewing them as self-evident truths does not do justice to the ways in which mathematicians employ axioms" (Schlimm 2013, 73). As the ancient world was not monolithically Euclidean in assuming principles in mathematical sciences, so the contemporary mathematical, philosophical, and logical disciplines are similarly not attached to the notion of self-evidence and perfect rigor when it comes to issues of axiomatics. As for questions of redundancy and internal logic among axioms, Schlimm's study offers a significant remark: "Popular criteria for axioms, e.g., that they should be as few, simple, and self-evident as possible, are highly idealized desiderata and by no means necessary conditions for systems of axioms" (Schlimm 2013, 39).

Most importantly, before the time of Hilbertian and formalistic interpretations of Copernicus's axioms, professional mathematicians were also raised on the axiomatic approach. Notably, the renowned collective of mathematicians under the name of Nicolas Bourbaki stated that the essential aim of the axiomatic method "is exactly that which logical formalism by itself can not supply, namely the profound intelligibility of mathematics" (Bourbaki 1950, 223). The same Bourbaki argued that formalization is "but one aspect of this [axiomatic] method, indeed the least interesting one" (Bourbaki 1950, 223).

All this speaks in favor of considering Copernicus's method as axiomatic even from a contemporary perspective. On this account, nothing should prevent historians from regarding Copernicus's *petitiones* as proper axioms.

As noted previously, historians of astronomy have inquired into Copernicus's possible sources in his use of axioms. By dealing with this puzzle in mathematical practice, another question arises: Does Copernicus need a source for establishing axioms? On the one hand, the historian can answer that Copernicus is likely seeking to contribute to the Greek tradition of mathematical sciences, with which he could have come into contact via Latin translations of Archimedean works, in which axioms are used in a somewhat generic sense and as true statements about the universe and nature. Thus, Copernicus is striving to give a list of seven assumptions to engage with that tradition, conforming to a model. On the other hand, Copernicus knows his mathematics: his intuition might well form part of his practice, that is, from the experience he assimilated with ancient astronomical mathematics and the observations he conducted. Copernicus's axiomatic approach might be the outcome of his mathematical practice, not necessarily the derivation from philosophical notions or other scientific sources he could have encountered during his career. This hypothesis is further supported by an idea that several professional mathematicians have shared on the intelligibility of mathematics and its intrinsic creativity. For instance, André Weil has claimed that:

> The views of Greek philosophers about the infinite may be of great interest as such; but are we really to believe that they had great influence on the work of Greek mathematicians? Some universities have established chairs for "the history and philosophy of mathematics": it is hard for me to imagine what those two subjects can have in common. (Weil 1980, 230)

5 An Interpretive Framework

It is clear that Copernicus employs a non-Euclidean approach in axiomatics, driven by the needs of reforming Ptolemy and his predecessors toward a new vision of the universe as an integral whole. Interestingly, Copernicus would refine his heliocentric theory after completing *Commentariolus*, outlining his definitive theory only in 1543 (in *De revolutionibus orbium coelestium – On the revolutions of the heavenly spheres*). The latter work does not list the seven axioms but presents their content along with theorems and proofs. On this account, it is likely that the seven axioms fed Copernicus's creativity and were engines for new discoveries during the composition of the final heliocentric theory. Further research will surely shed light on this issue.

The details of Copernicus's axiomatic approach could be described by adopting the guidelines set out by Dirk Schlimm in his study on the use of axioms in contemporary mathematical practices (Schlimm 2013). Such an attempt toward a new interpretation is also needed to counteract a Copernican literature which has been almost Euclidean or has tended to see Copernican axioms in a single dimension, that is, as dependent on an external source, and not in terms of their intrinsic mathematical and practical dimension – here one could well recall Bourbaki's and Weil's words on mathematical intuition and creativity. According to Schlimm, there are four main points in a comprehensive study on the use of axioms (Schlimm 2013, 39–40; points quoted in the following).

(a) "Axiomatics is an epistemic and methodological tool that can be employed in various ways. In other words, there is no single role that axioms play in mathematical practice; rather, the same set of axioms can be employed in different roles."

This is in accordance with the non-monolithically Euclidean axiomatics since Greek Antiquity, which is useful to recall in order to understand why Copernicus considered "postulate" as being equivalent to "axiom."

(b) "Popular criteria for axioms, e.g., that they should be as few, simple, and self-evident as possible, are highly idealized desiderata and by no means necessary conditions for systems of axioms."

This point has been extensively developed in treating the logic of Copernican axioms.

(c) "The practical usefulness of axioms goes well beyond the context of justification and the aim of clarifying and providing foundations for mathematical theories; they are also engines for discovery in mathematics."

This point can be seen in the fact that Copernicus continued to develop his theory after redacting *Commentariolus*. It took him around 30 years to publish his final heliocentric theory.

(d) "The dimensions of *presentation*, *role*, and *function* allow us to characterize different uses of axioms and to compare them along these three axes, with the purpose of clarifying discussions of axiomatics in mathematical practice."

Schlimm provides this framework for one interpretation among many, but this is in keeping with the historical case presented above. While it is evident how points a, b, and c have been already touched on in Copernicus's case, point d needs to be further articulated. It is necessary to recall that the Copernican axioms are seven plus the common fundamental axiom of ancient astronomy. On this account, the *function* of the axioms is to set the foundations for the geometry describing a heliocentric universe and to develop it further; the Copernican axioms *present* the main points of a new

theory, marking a strong break from the astronomers of the past. The axiom of uniform and circular motion and the seven others have different *roles*. To describe this feature, it is worth introducing the distinction between *structural* and *foundational* axioms as suggested by Solomon Feferman (Feferman 1999) and adapt it to a context of sixteenth-century mathematical astronomy and logic. As for the structural axioms, Feferman claims:

> When the working mathematician speaks of axioms, he or she usually means those for some particular part of mathematics such as groups, rings, vector spaces, topological spaces, Hilbert spaces, etc. These axioms have nothing to do with self-evident propositions, nor are they arbitrary starting points. They are simply definitions of kinds of structures that have been recognized to recur in various mathematical situations. (Feferman 1999, 100)

This is an apt definition for the seven Copernican assumptions, for in Copernicus's case there are seven not self-evident propositions to define a heliocentric structure of the universe, and he conceives them as non-arbitrary assumptions about physical hypotheses to envision the universe as an integral whole (Goddu 2009). As such, it makes sense to define them as *structural* in their context.

Concerning foundational axioms, Feferman states that:

> in contrast to the working mathematician's structural axioms, when the logician speaks of axioms, he or she means, first of all, laws of valid reasoning that are supposed to apply to *all* parts of mathematics, and, secondly, axioms for such fundamental concepts as number, set, and function that underlie all mathematical concepts; I call the latter *foundational axioms*. (Feferman 1999, 100)

Commentariolus contains the fundamental axiom of ancient astronomy, which underlies all the other assumptions, namely, all motions of celestial bodies must be uniform and circular. This could well serve as a foundational axiom in Copernicus's composition, and it is also coherent with his vision of a whole and integral universe (Goddu 2009). On these grounds, adapting Feferman's distinction to *Commentariolus* would build a comprehensive interpretive framework for Copernicus's axiomatic approach.

A Schlimm-Feferman framework integrates aspects of mathematical practice with the historical questions noted above on terminology and logic. In such a framework, the foundational axiom of uniform and circular motion can be deemed the one which triggers Copernicus's creativity. As Copernicus would continue to further elaborate his theory leading up to the publication of his major work, circa 30 years later, the seven axioms can be considered engines of discovery and aids to mathematical creativity.

6 Conclusion

Copernicus's axiomatic approach toward a heliocentric theory of the universe, as exposed in his *Commentariolus*, has triggered a range of questions among historians, who have studied his axioms in their terminological and logical aspects. Terminologically, given the rarity of the term *axioma* in Latin literature before the first half of the

fifteenth century, Copernicus likely encountered the term in some Latin humanist work (probably Giorgio Valla's *De expetendis*) and thus became aware that somebody was referring to *petitiones* as *axiomata*. As a result, he added the term *axioma* with the intention of conveying that some humanist authors may call mathematical principles *axiomata*. Humanism was indeed the cultural climate in which Copernicus was living during the years of his Italian apprenticeship. He did not need to enter the post-Euclidean debate on the difference between postulate and axiom because he was not writing a treatise on geometry but a draft on astronomy, that is, mathematics applied to a branch of natural philosophy: he considered his models as physically real, not instrumental.

It is notable that there was no controversy concerning Copernicus's axioms prior to 1973. What Swerdlow provided was a kind of formalistic criticism that morphed into a debunking of Copernicus's logic, which proved to be anachronistic: such criticism was based on a strict interpretation of axiomatics applied to history, which, in turn, does not reflect the current scenario of axiomatics in logic and mathematics. For instance, the rigor and logic for which Swerdlow was advocating with regard to axioms did not correspond to what professional mathematicians such as Nicolas Bourbaki and André Weil were thinking on axiomatics and the potential of the axiomatic method in the second half of the twentieth century.

Moreover, this case study has shed light on the similarities between axiomatics in early modern and contemporary times, a sort of bridge with the axiomatics of the present and the past, a task which was already undertaken some time ago by Patrick Suppes. Indeed, Suppes claimed that:

> There was certainly a sense of methodology deeply embedded in Euclid, Archimedes, and Ptolemy, but it was not a sense of methodology that was completely explicit or totally worked out, just as Aristotle's own general principles are never exemplified in any detailed and complicated scientific examples of an extended sort. The gap between philosophical analysis, canons of axiomatic method, and actual working practice was about the same order of magnitude that it is today. What is surprising, I think, from a philosophical standpoint is that the gap seems, if anything, to have widened rather than narrowed over the past 2000 years. (Suppes 1993, 40)

The gap between historical–philosophical, logical, and mathematical (practical) plans is what is evident in the history of the scholarship on Copernicus's *Commentariolus*. Studies on mathematical practice have proven beneficial to better comprehend the Copernican axiomatic approach and provide some order to such plans. In sum, to speak about Copernicus's assumptions as axioms and to consider the assumptions (*petitiones*) as synonymous with axioms (*axiomata*) is to give a correct translation of the Copernican text even from a contemporary and not mono-dimensional (self-evident) perspective on axiomatics, thereby avoiding the pitfalls of anachronism.

Unlike past considerations of Copernican axioms, this study has pointed out that the seven unproved assumptions set out by Copernicus are linked to the axiom of uniform circular motion. To better comprehend this feature in *Commentariolus*, an interpretive framework based on Schlimm and Feferman sheds new light on the different features of Copernican axioms and enables an understanding of them through the lens of pluralism. The utility of the Schlimm-Feferman framework is at least twofold: first, it holds

together the plurality of the uses of axioms (by considering the use of the term in the literature before Copernicus) and the difference in the kinds of axioms (foundational and structural); second, it points to the potentiality of Copernicus's axioms as tools for mathematical creativity and engines of discovery in light of Copernicus's final elaboration of the heliocentric theory.

Ultimately, Copernicus's mathematical practice adds another nuance to the possibility of considering the importance and applicability of mathematics and its history: as Wittgenstein famously observed, "mathematics is a colorful conglomerate of techniques of proof." [Full original quotation: "Ich möchte sagen: Die Mathematik ist ein BUNTES Gemisch von Beweistechniken.-Und darauf beruht ihre mannigfache Anwendbarkeit und ihre Wichtigkeit."] (Wittgenstein 1956, p. 84, author's translation). Indeed, the Copernican axioms set out in *Commentariolus* would form the basis on which to draw proofs at a later stage, in *De revolutionibus* (1543), but the history of mathematical creativity toward heliocentrism – between the compositional stages of *Commentariolus* and *De revolutionibus* – remains to be written. At any rate, if the *Commentariolus* per se does not explicitly contain techniques of proof, historians dealing with the Copernican axioms have certainly given birth to a colorful conglomerate of interpretations. Taking a closer look at axiomatics, drawing on the work of mathematicians, logicians, and philosophers of mathematical practice, has offered a new interpretation and has hopefully brought some order and clarity to such a historical conglomerate.

7 Cross-References

- ▶ Euclidean and Non-Euclidean Geometry in the History and Philosophy of Mathematical Practice
- ▶ Historiography of Mathematics from the Mathematician's Point of View
- ▶ Implicitly Defining Mathematical Terms
- ▶ Introduction to the Handbook of the History and Philosophy of Mathematical Practice
- ▶ Logic in the History and Philosophy of Mathematical Practice
- ▶ Non-deductive Justification in Mathematics
- ▶ One Mathematic(s) or Many? Foundations of Mathematics in Twentieth-Century Mathematical Practice
- ▶ The Values of Mathematical Proofs
- ▶ What Mathematicians Do: Mathematics as Process and Creative Rationality

References

Acerbi F (2010) Two approaches to foundations in Greek mathematics: Apollonius and Geminus. Sci Context 23(2):151–186

Acerbi F (2013) Aristotle and Euclid's postulates. Class Q 63(2):680–685

Aristotle (2002) Posterior analytics. Translated with a commentary by Jonathan Barnes. Oxford University Press, Oxford. [1975, 1993, reprinted 2002]

Blåsjö V (2014) A critique of the arguments for Maragha influence on Copernicus. J Hist Astron 45(2):183–195
Blåsjö V (2018) A rebuttal of recent arguments for Maragha influence on Copernicus. Studia Historiae Scientiarum 17:479–497
Bobzien S (1996) Stoic syllogistic. Oxf Stud Anc Philos 14:133–192
Bobzien S (2019) Stoic sequent logic and proof theory. Hist Philos Logic 40:234–265
Bourbaki N (1950) The architecture of mathematics. Am Math Mon 57:221–232
Burton HE (1945) The optics of Euclid. J Opt Soc Am A 35(5):357–372
Cantù P, Luciano E (2021) Giuseppe Peano and his school: Axiomatics, symbolism and rigor. Philosophia Scientiæ Travaux d'histoire et de philosophie des sciences 25(1):3–14
Capecchi D (2018) The path to post-Galilean epistemology. Reinterpreting the birth of modern science. Springer, Cham
Clagett M (1976) Archimedes in the middle ages. Volume two. The translations from Greek by William of Moerbecke. The American Philosophical Society, Philadelphia
Clagett M, Moody EA (eds) (1952) The medieval science of weights (*Scientia de Ponderibus*) treatises ascribed to Euclid, Archimedes, Thabit ibn Qurra, Jordanus de Nemore and Blasius of Parma. The University of Wisconsin Press, Madison
Copernicus N (1978) Complete works. Volume 2. On the revolutions, edited by Jerzy Dobrzycki, translated by Edward Rosen with commentary. Polish Scientific Publishers, Warsaw
De Risi V (2016) The development of Euclidean axiomatics. The systems of principles and the foundations of mathematics in editions of the elements in the early modern age. Arch Hist Exact Sci 70:591–676
Dobrzycki J (1973) The Aberdeen copy of Copernicus's Commentariolus. J Hist Astron 4(2): 124–127
Dobrzycki J, Szczucki L (1989) On the transmission of Copernicus's Commentariolus in the sixteenth century. J Hist Astron 20(1):25–28
Feferman S (1999) Does mathematics need new axioms? Am Math Mon 106:99–111
Feke J (2018) Ptolemy's philosophy: mathematics as a way of life. Princeton University Press, Princeton
Folkert M, Kirschner S, Kühne A (eds) (2019) Nicolaus Copernicus Gesamtausgabe, Band IV, Opera minora: Die kleinen mathematisch-naturwissenschaftlichen Schriften. Editionen, Kommentare und deutsche Übersetzungen. With assistance from Uwe Lück and translations by Fritz Krafft. Walter de Gruyter Oldenbourg, Berlin
Goddu A (2009) Copernicus's Mereological vision of the universe. Early Sci Med 14:316–339
Goddu A (2010) Copernicus and the Aristotelian tradition. Brill, Leiden/Boston
Goddu A (2019) The (likely) last edition of Copernicus's Libri revolutionum. Variants [Online] 14. http://journals.openedition.org/variants/908, https://doi.org/10.4000/variants.908
Heath TL (2002) The works of Archimedes. Dover, New York
Henkin L, Suppes P, Tarski A (eds) (1959) The axiomatic method with special reference to geometry and physics. In: Proceedings of an International Symposium held at the University of California, Berkeley, December 26, 1957 January 4, 1958. North-Holland Publishing Company, Amsterdam
Hintikka J, Gruender D, Agazzi E (eds) (1981) Theory change, ancient axiomatics, and Galileo's methodology. In: Proceedings of the 1978 Pisa conference on the history and philosophy of science, vol 1. D. Reidel Publishing Company, Dordrecht/Holland
Ierodiakonou K (2006) Stoic logic. In: Gill MP, Pellegrin P (eds) A companion to ancient philosophy. Blackwell, Malden, pp 505–529
Knorr WR (1981) On the early history of axiomatics: the interaction of mathematics and philosophy in Greek antiquity. In: Hintikka J, Gruender D, Agazzi E (eds) Theory change, ancient axiomatics, and Galileo's methodology. Proceedings of the 1978 Pisa conference on the history and philosophy of science, vol 1. D. Reidel Publishing Company, Dordrecht/Holland, pp 145–186
Lerner MP, Segonds AP (eds) (2015) Nicolas Copernic, De revolutionibus orbium coelestium, Des revolutions des orbes célestes. Vol. 1: Introduction by Michel-Pierre Lerner and Alain-Philippe

Segonds with the collaboration of Concetta Luna, Isabelle Pantin, and Denis Savoie. Les Belles Lettres, Paris
Netz R (2004) The works of Archimedes. Translated into English, together with Eutocius' commentaries, with commentary, and critical edition of the diagrams. Volume I. The two books on the sphere and the cylinder. Cambridge University Press, Cambridge
Novaes CD (2020) The dialogical roots of deduction. Cambridge University Press, Cambridge
Omodeo PD (2016) Science and medicine in the humanistic networks of the northern European renaissance. In: Omodeo PD (ed) Duncan Liddel (1561–1613) Networks of polymathy and the Northern European Renaissance. Brill, Leiden/Boston
Proclus (1970) A commentary to the first book of Euclid's elements (trans: Morrow GR). Princeton University Press, Princeton
Ragep FJ (2016) Ibn al-Shāṭir and Copernicus: the Uppsala notes revisited. J Hist Astron 47(4): 395–415
Rosen E (1976) Copernicus' Axioms. Centaurus 20:44–49
Schlimm D (2013) Axioms in mathematical practice. Philos Math 3(21):37–92
Suppes P (1993) Limitations of the axiomatic method in ancient Greek mathematical science. In: Suppes P (ed) Models and methods in the philosophy of science. Springer, Dordrecht, pp 25–40
Swerdlow N (1973) The derivation and first draft of Copernicus's planetary theory. Proc Am Philos Soc 117:423–512
Swerdlow N (2017) Copernicus's derivation of the heliocentric theory from Regiomontanus's eccentric models of the second inequality of the superior and inferior planets. J Hist Astron 48(1):33–61
Szabó A (1964) Transformation of mathematics into deductive science and the beginnings of its foundation on definitions and axioms. Scripta Mathematica 27:113–139
Taub L (1993) Ptolemy's universe. The natural philosophical and ethical foundations of Ptolemy's astronomy. Open Court Publishing Company, Chicago/La-Salle
Valla G (1501) De expetendis et fugiendis rebus opus. Aldus Manutius, Venice
Vesel M (2014) Copernicus: Platonist astronomer-philosopher. Peter Lang GmbH Internationaler Verlag der Wissenschaften, Frankfurt am Main
Weil A (1980) History of mathematics: why and how. In: Letho O (ed) Proceedings of the international congress of mathematicians, Helsinki 1978. Academia Scientiarum Fennica, Helsinki, pp 227–236
Wittgenstein L (1956) Bemerkungen üner die Grundlagen der Mathematik – Remarks on the foundations of mathematics. In: von Wright GH, Rhees R, Anscombe GEM (eds) trans. Anscombe GEM. The M.I.T. Press, Massachusetts Institute of Technology, Cambridge, MA/London

Visual Representations of Euclidean Geometry: Diagrammatic Reasoning in Oliver Byrne's Work

Andrea Pedeferri

Abstract

Around the half of the nineteenth century the Irish mathematician and civil engineer Oliver Byrne published a peculiar version of *Euclid's Elements* where he displays Euclid's proofs in form of pictures with little to no explanatory text (or the usual letter symbols) associated. That is, Byrne uses graphical representations (encoded using standardized colors and shapes), corresponding with diagrammatic parts of geometrical entities, which inhere in propositions. In this way, Byrne proves by *picturing* Euclid's propositions. A recent, renewed interest in the importance of mathematical practice had led to a reconsideration of the role and scope of diagrams and diagrammatical reasoning in mathematics. In this context, Byrne's peculiar use of diagrammatical representations can be seen as both closer to the original Greek and Euclidean ideals of rigor and justification in mathematics, as well as an effective example of how visualizations in mathematics may serve as efficient cognitive-representational tools.

Keywords

Geometry · Euclid · Diagrams · Diagrammatic reasoning · Geometrical visualizations

Towards the middle of the nineteenth century, Irish mathematician and civil engineer, Oliver Byrne, published a peculiar version of *Euclid's Elements*: *The First Six Books of the Elements of Euclid* in which coloured diagrams and symbols are used instead of letters for the greater ease of learners. The work's prolix title affords some idea as to its quirks. This eccentricity aside, what renders Byrne's edition both unusual and appealing is that he displays Euclid's proofs in pictorial form, providing

A. Pedeferri (✉)
Union College, Schenectady, NY, USA
e-mail: pedefera@union.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_25

little to no associated explanatory text, nor any expected algebraic symbols. That is, Byrne uses graphical representations (encoded using standardized colors and shapes), corresponding with diagrammatic parts of geometrical entities, which inhere in propositions. In this way, Byrne proves by *picturing* Euclid's propositions. The proofs are carried out through the graphical disclosure of inferential process by way of diagrams. The result is visually striking. However, as Byrne points out, this protocol

> [. . .] has a greater aim than mere illustration; we do not introduce colours for the purpose of entertainment, or to amuse by certain combinations of tint and form, but to assist the mind in its researches after truth, to increase the facilities of instruction, and to diffuse permanent knowledge. (Byrne 1847)

In addition to the *pedagogical* import of Byrne's new presentation of geometrical proofs and constructions, the equipotency of his methodology with respect to traditional approaches to geometry resulted in the work's dismissal as a mere curiosity (as can be found in works, for example, by McLean, Cajori, and DeMorgan). However, his most celebrated work, this version of the *Elements*, would eventually gain notoriety, albeit owing largely to its esthetic appeal; its novel color combinations predating the likes of Piet Mondrian.

A recent, renewed interest in the importance of mathematical practice had led to a reconsideration of the role and scope of diagrams and diagrammatical reasoning in mathematics. In this context, Byrne's peculiar use of diagrammatical representations can be seen as both closer to the original Greek and Euclidean ideals of rigor and justification in mathematics, as well as an effective example of how visualizations in mathematics may serve as efficient cognitive-representational tools.

From this perspective, Byrne's work acquires a significance far exceeding its pedagogical value. To this end, Byrne provides an interesting case study to the history of diagrammatic mathematical representations, as well as the study of diagrammatic reasoning in mathematics, interest in which is currently enjoying a revival. Byrne's work is moreover refreshing insofar as it constitutes a departure from traditional accounts of Euclidean geometry, as proffered on the mathematic scene in Victorian Britain.

A mathematician and civil engineer, Oliver Byrne was born in 1810 in Ireland. He lived a rather adventurous life between England and the USA, dying in hardship in 1880. Although he was not part of the British academic environment (at least not consistently), his work enjoyed reasonable levels of circulation, as evidenced in several reviews and commentaries, which can be found in contemporary journals and books. Although Byrne's writings are highly peculiar, even *eccentric*, they form part of – and should be read within – the context of developments and debates that permeated the field of mathematics in Britain at the time. While Byrne did not play a significant role in the development of mathematics in the nineteenth century, it is nonetheless interesting to note how his works sit with some of the most important debates on the Victorian mathematical scene. In Byrne's books and contributions especially, we find – more or less directly – echoes of discussions and controversies

regarding how to teach mathematics, as well as the role of *Euclid's Elements* in secondary and postsecondary education. Byrne's interest in the practical applications of mathematics and geometry (particularly in the field of civil engineering) greatly influenced his work, together with a marked pedagogical intent. His approach was modeled on creating usable mathematical practices. Still, whatever underwrote his interest in discovering new techniques for mastering geometry and logarithms, we can detect a push towards more flexible representations and uses of inferential processes and proofs. This is especially evident when we examine his edition of the *Elements* and a number of other works, in particular *The Young Geometrician* or *Practical Geometry without Compasses*. While the former constitutes an attempt to "make visible" the inferential structure of proofs, the latter is a peculiar but highly practical approach to geometry, whose methodology is characterized by visual, "mobile" representations of geometric constructions.

Byrne's adult life coincides with the Victorian era, a period of British political and economic preeminence. The same, however, cannot be said about the state of British mathematics at the time, which could boast none of the successes enjoyed by the bulk of British society. The weak state of British mathematics was not, in fact, peculiar to the Victorian period. In fact, during that time (especially in its second half) we witness a sort of mathematical renaissance across the British Empire. Undoubtedly, however, following the heights achieved by Isaac Newton, British mathematics became stagnant, at least in comparison to many parts of the Continent.

In particular, when we take a closer look at the discipline, we see how *Euclid's Elements* constituted the textbook of choice at all levels of British education, a proclivity that continued until after the mid-nineteenth century. This near-total reliance on Euclid was, however, no longer commonplace on the Continent. In France, for example, even by the seventeenth century several influential scholars and educators began to remove *Euclid's Elements* from courses of study, otherwise invoking only heavily redacted versions of it. This move away from Euclid resulted from the influence of La Logique de Port-Royal, which advocated major revisions in the approach to geometry, pushing towards more clarity and comprehension. This resulted in scholars distancing themselves more and more from Euclid, who was now considered finicky and cumbersome. Meanwhile in Britain, Euclidean geometry was still regarded as the paradigm for the logical structure of mathematics. It was only after the 1860s that Euclid's preeminence began to wane in Britain, prompting a vigorous debate from anti-Euclidean academics and teachers.

Byrne wrote his version of the *Elements* in 1847, in a climate wherein Euclid prevailed over the teaching of geometry in both schools and universities. However, Byrne's book was radically different from long-established editions of Euclid, which were traditionally deployed in educational settings. A comparison with the more traditional and influential accounts of geometry serves to highlight Byrne's innovative approach, which aimed to appeal directly to our geometrical intuitions, in particular through bypassing algebraic representations. In Byrne's time, one of the most popular versions of *Euclid's Elements* was that edition by the Scottish mathematician John Playfair. Playfair's views were sympathetic to those of the Scottish Enlightenment group called the Common Sense philosophers, who lobbied for

Euclidean geometry as the exemplar of true knowledge. Playfair's *Elements* were published in 1795 and were largely based on the foregoing popular edition of Euclid, edited by Robert Simson in 1756, which was regarded as the handbook of elementary geometry in Great Britain. Both Simson's and Playfair's editions aimed to present a more accurate and rigorous version of Euclid's work, which would exemplify how best to do geometry. Geometry was still regarded as the essential component of not only mathematics but, more generally, of how to properly train pupils in the art of reasoning. In his book, Simson writes that his review of the *Elements* was aimed at restoring "the principal Books of the *Elements* to their original accuracy" (Simson 1756). To this end, his main concern was that of fixing the errors that were carried over by previous editions. The presentation of Euclid's work followed the traditional presentation of propositions through lettered diagrams accompanied by explanatory text.

To this end, Playfair's Euclid constitutes a notable improvement. For, it not only represents a modernized reprint of previous editions, but moreover introduces elements of analysis, and so approaches the discipline with greater fidelity to the real-time developments in mathematics beyond British borders. While Playfair's approach retained the traditional diagrammatic representations (with lettered diagrams of Euclid's propositions), as Amy Ackerberg-Hastings points out, he

> intended to rejuvenate the content of the textbook. He believed that the influence of Simson, and of Colin Maclaurin, had led the next generation of British mathematicians to emphasize so-called 'synthetic' methods to such an extent that those mathematicians foolishly ignored the discoveries made by Continental mathematicians in the eighteenth century. (Ackerberg-Hastings 2002)

Playfair was therefore well aware of the divide between British and Continental (especially French) mathematics. Indeed, he charged Britons with focusing too much on the *synthetic* aspect of proofs (that is, the need to work on sets of concrete, finished proofs), by contrast with the Continental focus on abstractions and algebraic generalizations (what he terms the *analytic* method). British mathematicians were very much against Continental analytic methods since they would apparently lead to "the illogical habit of generalizing too soon, of suspending conviction upon some remote analogy, instead of founding it upon irrefragable evidence" (Ackerberg-Hastings 2002). Across his works, Playfair attempts to merge the two approaches. To the traditional Euclidean use of diagrams he adds algebraic symbolism and techniques, thereby appealing to the Continental proclivity for algebra as a valid way of representing the concept of rigor and as a model for pure knowledge. The synthetic-centered approach was structured to provide a model for the "justificatory" aspect of mathematics, while a more analytic-centered approach caters to mathematics' "discovery" aspect, something that British mathematics lacked. However, Playfair's approach did not result in a renewed interest in mathematics' "discovery" side, nor did it change the mainstream British mentality towards geometry. Playfair's own versions of Euclid's proofs were indeed presented following the synthetic method, that is, the traditional model. In general, there was some pushback against

the idea of the "analytic" methods, with arguments centering around those ideas given expression by Thomas Young above.

Therefore, albeit with some interesting differences, the use and understanding of *Euclid's Elements* still fell on the traditional side, which centered upon reading him as the paradigm of rigor and reasoning in the abstract. The geometrical nature of Euclid's work was essentially seen as a model for highlighting the nature of correct inferential processes. This permitted a fairly sharp distinction between "science and [the] art of geometry" (Hydin Price 1981), with the study of the *diagrammatic* nature of Euclidian proofs being largely overlooked. Owing to (and partially evidencing) this, it is interesting to note how, at the time, there were growing criticisms regarding the fact that students in both secondary schools and college were unable to work *practically* with geometrical constructions:

> A person may be a Cambridge Wrangler, and yet unable to make a simple graphical construction with accuracy. The ordinary schoolboy's knowledge of practical geometry is generally worthless or nil (Remark from a review in *Nature*, 1903, p. 147). (Hydin Price, 1981).

Thus, as an improvement upon Euclid, Byrne's work combined both the rigor of the original, albeit combined with a novel, graphical representation of visual constructions, which did not merely constitute a visual corollary but an essential element of the reasoning process. Unfortunately, this was not well communicated by the author and so the public were largely unreceptive to the work. Byrne's *Euclid* has therefore been relegated to near-invisibility. Indeed, even recently renewed interest in his work has focused almost entirely on the *aesthetic* appeal of his book, to the exclusion of his pioneering recasting of the original.

A famous saying that is invoked to describe familiar obstacles in our quest for understanding is that "The essential is invisible to the eye" (the Little Prince may have been the first to make this explicit). The fact that there are certain features of reality (broadly understood as both natural and artificial), which seem to evade the possibility of clear, linguistic attribution or description, is not new. From physical features like the existence of "invisible" forces in nature, which regulate the cosmos, to abstract concepts, which are essential for defining our rationality, such as those of inferential processes, the "invisibility of the essential" is a phenomenon pertaining to many different contexts. This has always been the case, especially in mathematics, given its paradigmatically abstract nature. Indeed, ontological questions regarding what numbers *are* remains to this day an open question. Understandably, then, symbolic languages were initially adopted in disciplines like mathematics, for, besides the necessity of simplicity and rigor, it has always been important to *visually* grasp a domain's essential features, thereby making the invisible visible.

To this end, geometry has always been exemplary. In this domain, the challenge of both representing geometrical shapes and proof processes, which operate on geometrical figures, have typically been met through the use of diagrams. However, especially on the contemporary understanding of mathematics, diagrams have been largely regarded as merely visual aids, that is, as helpful, pedagogical (and sometimes heuristic) complements. Accordingly, they have been habitually confined to the

moment of mathematical discovery, while being considered nugatory with regard to the justificatory process. One of the reasons why diagrams and diagrammatic reasoning in general have been so disregarded (especially in the last century) is a consequence of the influence that the movements (schools) on the foundations of mathematics had on our understanding of mathematics itself. Their core idea that rigor could only be achieved through logical means is brilliantly summarized by Gödel:

> The development of mathematics toward greater precision has led, as is well known, to the formalization of large tracts of it, so that one can prove any theorem using nothing but a few mechanical rules. (Gödel 1931)

In 1931, Gödel summarized in but a few sentences around a century's worth of revolutionary mathematical developments, which had moreover led to the birth of modern logic. The arithmetization of analysis, the birth non-Euclidean geometries, and the elaboration of new axiomatic and formal systems were all examples of achievements making the period one of the most exciting and progressive in the history of mathematics. This optimism was grounded in the idea that logical consistency should be at the core of mathematics. Consequently, the reputed "certainty" of geometry, as based on "evident truths," had to relinquish its former role as the model for the discipline, thereby precluding any "serious" place for diagrams and diagrammatic reasoning in proof processes. For these pedagogical tools were considered to be insufficiently rigorous, owing to which they were consigned to entry-level mathematical practice. (By and large, mathematical practice did not boast any relevant status of its own – whether foundationally or epistemologically.)

The eventual collapse of this approach did not correspond with an immediate sea change of attitudes within the philosophy of mathematics. It nonetheless opened up a space for a reconsideration of the value of mathematical practice in the philosophical analysis of mathematics. As articulated by Mancosu, different lines of thought in the philosophy of mathematics eventually converged to support the "belief that attention to mathematical practice is a necessary condition for a renewal of the philosophy of mathematics" (Mancosu 2010). Accordingly, since diagrammatic reasoning and the use of diagrams were generally regarded as mere aspects of mathematical practice, the newfound interest in the value and importance of these parts of mathematics moreover brought renewed attention to the nature and scope of diagrammatic reasoning itself. Indeed, while the status of *Euclid's Elements* as a paradigm of rigor and rigorous argumentation has been questioned by foundationalists, what is *unquestionable* is that the *Elements* have always been exemplary with regard to the utility of diagrams, as well as to what we have been calling "diagrammatic reasoning."

From the nineteenth century onwards, several challenges have been leveled against *Euclid's Elements*. In particular, attention has historically been directed towards the fact that, in Euclidean propositions, conclusions are drawn directly from diagrams by appealing to intuition rather than through rigorous inferential rules. Moreover, it was difficult to accept the widespread presence of unwarranted generalizations from particular diagrams. These reasons, among others, brought about new and rigorous axiomatizations in order to render geometry independent

of figures. As Pasch claims, "the theorem is only truly demonstrated if the proof is completely independent of the figure" (Pasch 1882).

More recently, however, a new understanding of the importance of mathematical practice has led to reconsiderations regarding appeals to geometrical intuitions in the form of different kinds of visualization. For instance, Mancosu has pointed out that, in order to have a *genuine* philosophical inquiry into mathematics, it is important to address several issues, in particular the role that visual images and diagrams play in mathematical discourse. Thanks to scholarly moves in this direction, Euclid and his *Elements* have been "relegitimized." Indeed, through a series of works, the most influential of which is Manders' *The Euclidean Diagram*, "the role of diagrams in the inferential standards of Euclidean practice" has been rehabilitated (Manders 2010). Within this paradigm shift, it is interesting to note how the peculiarities of Byrne's presentation of Euclid can be read using new theoretical tools and models proposed by many scholars who are sympathetic to this revised position regarding mathematical practice in the philosophy of mathematics.

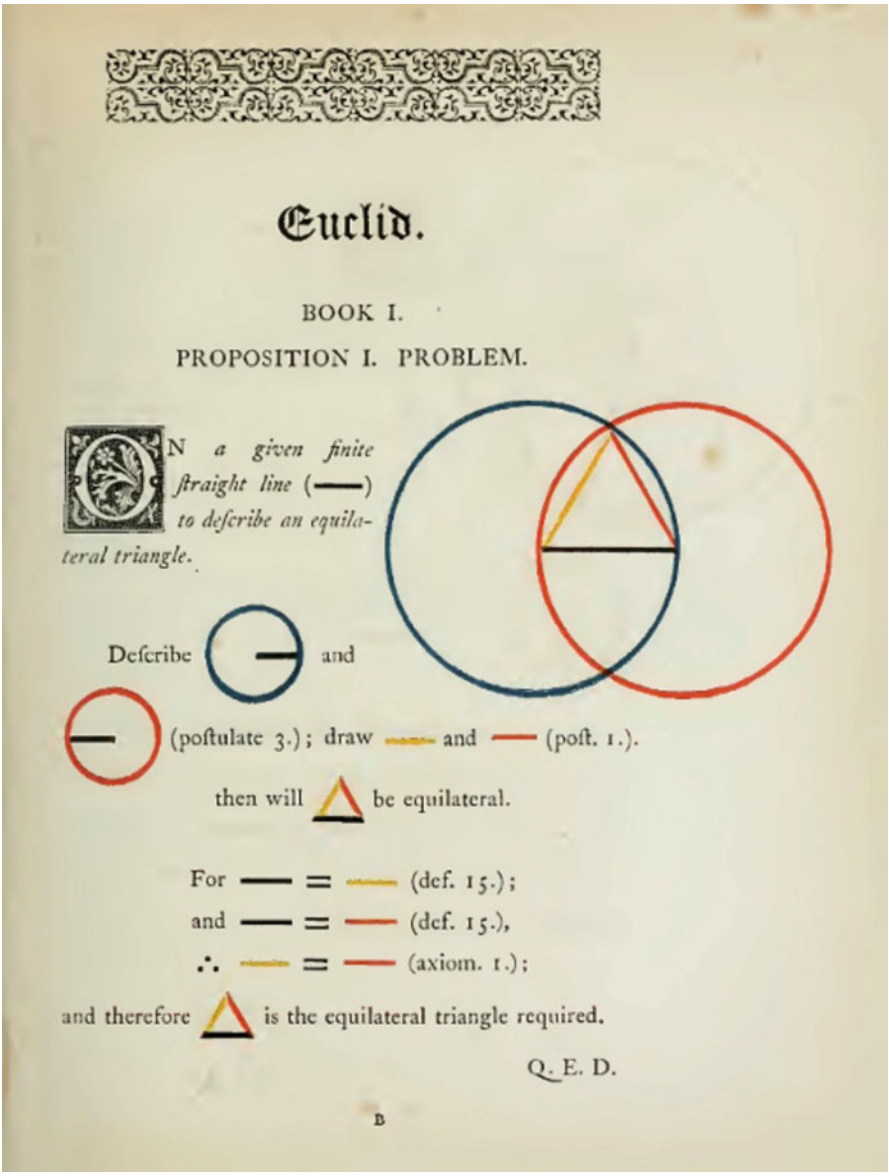
Byrne's use of diagrammatical representations is a peculiar feat within the history of mathematics and of the mathematical practice of his time. However, his innovative use of graphical representations – which appeal directly to our geometrical intuitions, thereby bypassing any algebraic representation – can be taken as an early example of how visualizations in mathematics can be regarded as efficient cognitive-representational tools. As Giardino puts it, these "external and material representations are used as inferential tools, that is, as tools for thinking about what they represent" (Giardino 2018a).

Byrne replaces the traditional, analytical presentation of geometry with color-coded visual illustrations, thereby doing away with traditional letter indices (an example being Euclid's Book 1 Proposition 1, see Fig. 1 below). Byrne's characteristic color proofs represent – directly and intuitively – factual relations between the relevant geometrical entities/constructions and the logical inferences of the proofs themselves. By providing unmediated, visual access to the material references of proofs, as well as their logical structure, Byrne's color proofs facilitate the study of the geometrical objects to which they refer, thereby enhancing both theoretical and applied reasoning about their mathematical contents.

In the Introduction to the *Elements*, Byrne writes:

> This work has a greater aim than mere illustration; we do not introduce colors for the purpose of entertainment or to amuse *by certain combinations of tint and form*, but to assist the mind in its researches after truth, to increase the facilities of instruction, and to diffuse permanent knowledge. (Byrne 1847)

When we compare this to Giardino's definition of diagrams, Byrne's own conception becomes clearer, for on neither account are diagrams mere proxies for mathematical entities. Diagrams have "to be intended [...] as *cognitive tools* that are meant to spatially display information in order to improve memory and promote inference, and not necessarily to depict mathematical objects" (Giardino 2017). Similarly, Byrne writes:



Euclid.

BOOK I.

PROPOSITION I. PROBLEM.

ON *a given finite ſtraight line* (——) *to deſcribe an equilateral triangle.*

Deſcribe ◯ and ◯ (poſtulate 3.); draw —— and —— (poſt. 1.).

then will △ be equilateral.

For —— = —— (def. 15.);

and —— = —— (def. 15.),

∴ —— = —— (axiom. 1.);

and therefore △ is the equilateral triangle required.

Q. E. D.

B

Fig. 1 Proposition I

the formalities and paraphernalia of rigour are so ostentatiously put forward, as almost to hide the reality. Endless and perplexing repetitions, which do not confer greater exactitude on the reasoning, render the demonstrations involved and obscure, and conceal from the view of the student the consecution of evidence. (Byrne 1847)

Byrne's reference to the concealment of the consecution of evidence highlights the gulf that existed between mathematic practice and foundationalist ideals regarding rigorous demonstration. Byrne's words are in fact against the foundationalist idea that a demonstration, or proof, should be a seamless, inferential process, which can be followed step-by-step and inspected, albeit not as something that can be "seen" (Giardino 2017).

Byrne's explanatory introduction of his work shows an apparent naiveté in not providing a rigorous justification for the use of his special colored diagrams. Indeed, he focuses almost entirely on his work's pedagogical perks. However, his diagrams reveal themselves simultaneously as visual representations of the logical inferential procedures of proofs and the actual representations of the *objects* of those proofs. This makes them the perfect example of the dual nature of diagrams as cognitive-representational tools, that is, as mathematical objects that are "intended both as representations and as instruments to think about what they represent" (Giardino 2017).

Byrne is clear that the visual aspect of diagrams is not only helpful but is intimately connected with the practice of mathematics, indeed, with the *moment* of mathematical justification.

His choice of shapes and colors furthermore reveals a subtle conception of how to make the invisible visible. Take, for example, his definition of a line as an object with length without breadth. It is apparent that, in trying to draw a line, we immediately incur the inconsistency of having an object without breadth. Byrne's "solution" is to use colors to reinforce our geometrical intuition about linearity:

> A mathematical line, which is length without breadth, cannot possess colour, yet the junction of two colours on the same plane gives a good idea of what is meant by a mathematical line; [...] Let two colours meet and cover a portion of the paper; where they meet is not red, nor is blue; therefore the junction occupies no portion of the plane, and therefore it cannot have breadth, but only length: from which we can readily form an idea of what is meant by a mathematical line. (Byrne 1847)

Accordingly, diagrams are flexible objects (or, in Giardino's terminology, "kinesthetic") insofar as they may be used and manipulated to trigger a process of discovery, on the one hand, while providing evidence for justification, on the other. The dual role of diagrams and diagrammatic reasoning is what Byrne had in mind when he stated that traditional, propositional descriptions of inferences ended up "hiding reality." Diagrams, by contrast, actively *disclose* reality, permitting more immediate access, owing to which they are better vehicles for revealing demonstrative inferences:

> Geometrical diagrams are not signs, but the materials of geometrical science, the object of which is to show the relative quantities of their parts by a process of reasoning called Demonstration. (Byrne 1847)

Byrnes' experimentation with a new visual representation of Euclid's propositions aims to reconnect Euclid's work with its native sensibility for evident and direct

rigor. Indeed, Byrne does so without taking recourse to one of the distinguishing features of Greek mathematics, viz. the addition of letters to geometrical shapes, thereby forming what are normally called "lettered diagrams." Lettered diagrams combine the instinctive, visual appeal of figures with the analytic rigor afforded by algebraic notation. This hybrid has been instrumental in supporting a key conception of Greek mathematics with regards to the notion of validity, which is far removed from the notion of validity we use today. Nowadays, we link validity to the logical notion of *deductive* validity, that is, the idea of providing conclusive evidence for the truth preservation of an inference, thanks to which a valid inference can be used as a model that may thereafter be generalized. In Greek mathematics, the emphasis was more on the possibility of repeating processes that could deliver the desired results. Lettered diagrams were a key element of those processes, for they allowed the creation and use of geometrical constructions that were, by their nature, reproduceable ad infinitum. Thus, the standard of validity was intimately connected with the idea of practical reproducibility. Byrne's prescinding from the use of letters does not, however, undermine this. On the contrary, it forces the reader to do bona fide diagrammatic reasoning, that is, using the diagrams not only as justificatory buttresses but as structural tools that enable reasoning *within* the diagram.

Proposition 1 is a paradigmatic example of Byrne's peculiar use of diagrams (see Fig. 1). It is well known that Proposition 1, which concerns the construction of regular triangles, has been the target of criticism about logical gaps and inconsistencies, which subverts the supposed rigorousness associated with Euclid's propositions. The most common objection concerns the existence of point C (in lettered diagrams), that is, the logical justification of the existence of the point where two circles intersect. This has been taken as evidence that Euclid's system is "hopelessly flawed," as Daniel MacBeth puts it. However, on MacBeth's view, this is not actually the case. This is because the role of diagrammatic representations becomes something very different from a system of axioms and more like a system whereby the diagram is used to reason in it, owing to which we can "see" that the inferences are correct:

> First, the *Elements* is not best thought of as an axiomatic system but is more like a system of natural deduction; its Common Notions, Postulates, and Definitions function not as premises from which to reason but instead as rules or principles according to which to reason. Secondly, demonstrations in Euclid do not involve reasoning about instances of geometrical figures, particular lines, triangles, and so on; the demonstration is instead general throughout. The chain of reasoning, finally, is not merely diagram-based, its moves, at least some of them, licensed or justified by manifest features of the diagram. It is instead diagrammatic; one reasons *in* the diagram in Euclid, or so it will be argued. (Macbeth 2010)

This conception of diagrams and diagrammatic reasoning is present in Byrne's presentation of Proposition 1, in which we can also clearly see the "justificational adequacy" (Manders 2010, p. 66) of diagram-based reasoning. This is because Byrne's presentation exemplifies what Manders calls the diagram disciplines, that is, *the* "proper exercise of skill in producing (still imperfect) diagrams required by the practice" (Manders 2010). According to Manders, traditional geometrical proofs

have a dual nature represented by two parts of the proof: a verbal part (the discursive text), and a graphical part (the diagram) (Giardino 2017). In Byrne, the absence of most of the discursive text, replaced by his colored figures, causes the two parts to collapse, which provides a direct and strong visual example of how the reader is "obliged" to be engaged in diagrammatic reasoning, that is, reasoning *in* the diagrams and not just *with* them. Manders says, in this regard, that:

> we take participants to be responsible for attaining uniformity of response in reading off features in diagrams. But the production and reading of diagrams cannot humanly be controlled so as to obtain a uniform response to diagrams directly on questions such as equality of lines. If a practice licenses a type of response to artifacts, then in order to attain uniformity of response, it must provide means by which participants can resolve their differences in responses to those physical artifacts they might reasonably produce. In traditional geometry, the division of responsibility between diagram and text, in which the text tracks equality information, is a way to meet this challenge. (Manders 2010)

Byrne's diagrams do the work of putting together those two aspects by allowing participants "to respond in a stable and stably shared fashion" (ibid.) when exposed to such artifacts. This enables readers (participants) to comply with the diagrammatic discipline.

A further example of this is Byrne's presentation of Proposition 10 of Book 3 (Fig. 2). The proposition states that "A circle does not cut a circle in more than two points." The proof is a reduction ad absurdum and it requires the construction of impossible figures. Byrne elegantly represents the "imperfect diagrams" allowing a straightforward engagement with diagrammatic reasoning.

Byrne's diagrams are able to capture another feature of Euclid's diagram-based presentation: the compositional nature of diagrams. Marco Panza (2012) argues that Euclidean arguments are diagram based since the objects of Euclidean geometry have a special relationship with their diagrammatical representations. They are compositional objects made of elementary objects such as "concrete lines representing segments and circles that are drawn through a single construction step" each of which is "one and divisible: these elementary objects can both compose complex diagrams and be decompose in parts" (Panza 2012). Panza refers to proposition 1.11 and 1.12 as a clear example of this interplay between the concrete geometrical objects and their diagrammatical representations. Those two propositions show how to construct a perpendicular to a line through a given point. They differ from each other in that in proposition 1.11 the given point lies on the line while in proposition 1.12 does not. If we look at Byrne's presentations of the two propositions (Figs. 3 and 4), it is clear how the idea of the compositional nature of diagrams finds a perfect representation. Byrne's diagrams allow the reader to both clearly follow the process of construction of the elementary geometrical objects and the process of proving how those objects satisfy the statement of the proposition.

Over the past decade, a distinction introduced by Manders has become very influential – if not standard – in the discussion about use of diagrams in geometry. The distinction is that of *exact* and *co-exact attributions*. Respectively, these are features that we get from visual representations, on the one hand, and text, on the

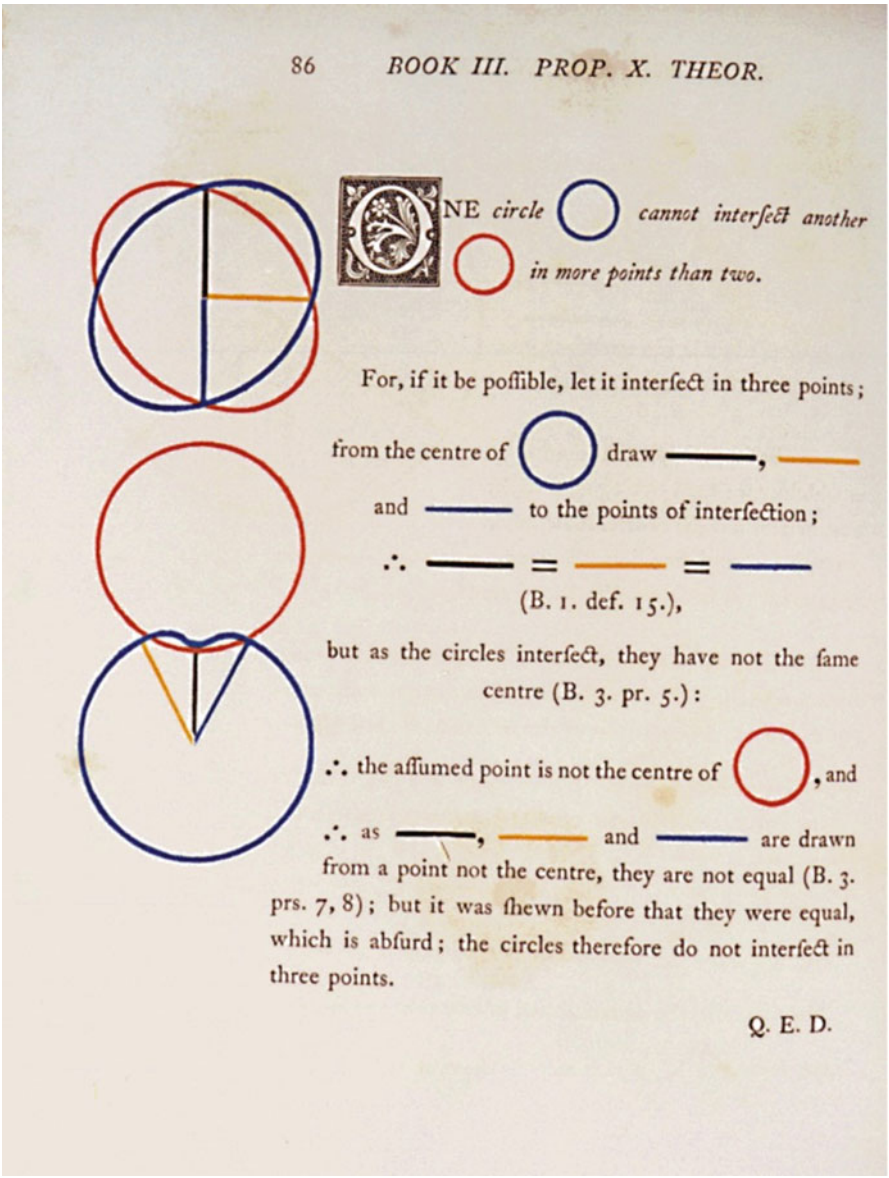


Fig. 2 Proposition X

other. *Exact* attributions are features such as metric information, usually given in the text, while *co-exact* attributions are topological details pertaining directly to the diagram. To import this distinction into the traditional discourse of Euclidean proofs, diagrams are used to infer co-exact information. Thus, in every act of geometrical

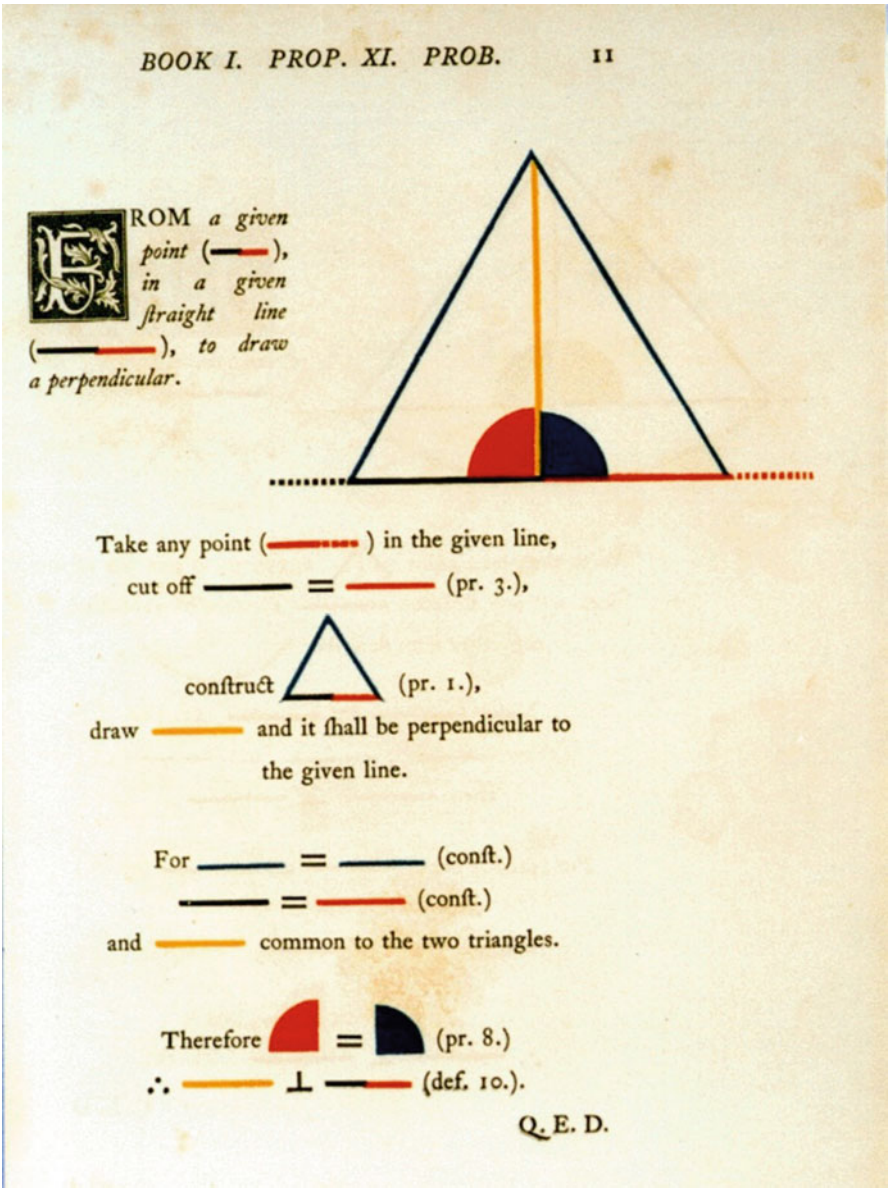


Fig. 3 Proposition XI

manipulation, co-exact aspects remain stable while exact aspects are amenable to change.

The above is especially clear in Byrne's diagrams, particularly on account of the absence of letters and (by and large) text. As Manders notes, Euclidian diagrams

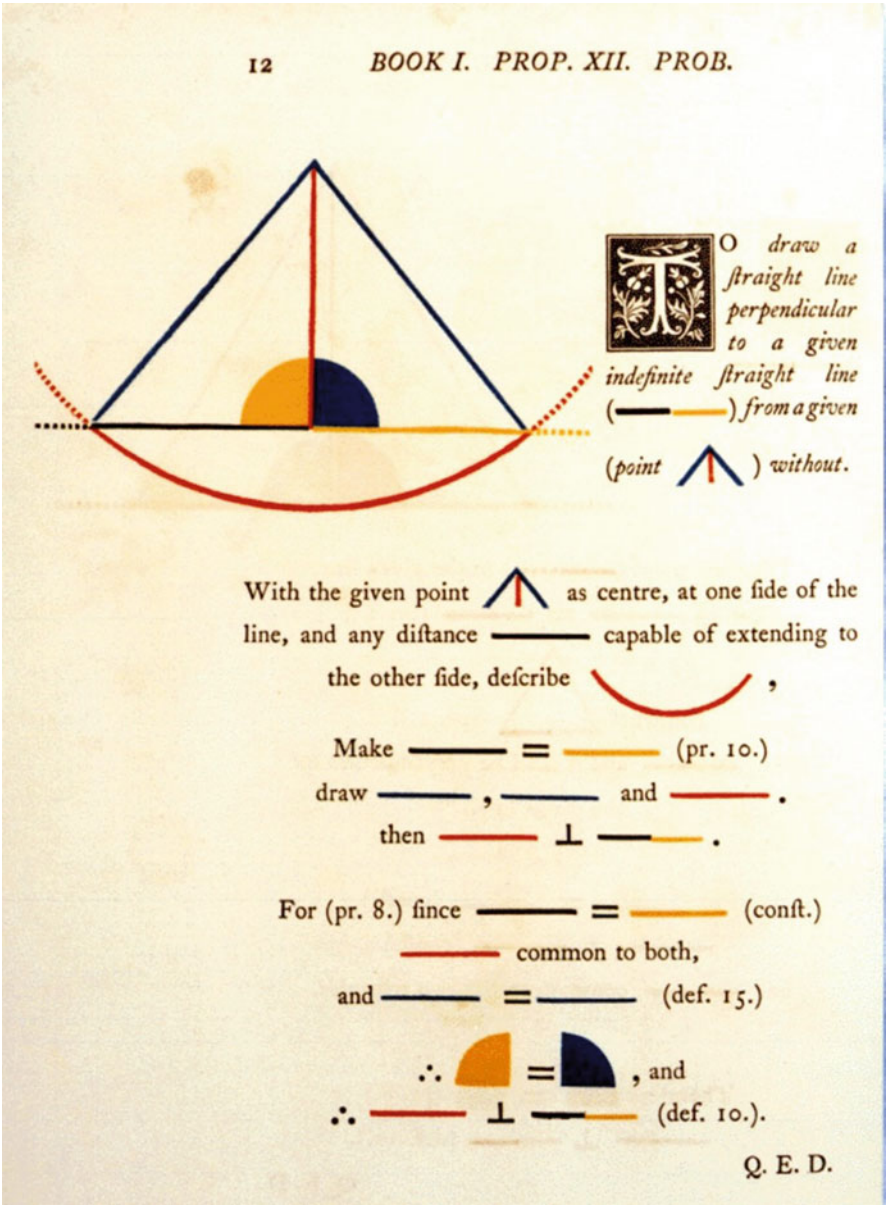


Fig. 4 Proposition XII

were designed to provide inferences about co-exact features. Thus, when we look at Byrne's Proposition 1, the problem of the existence of the point of intersection between two circles is something that incorporates a co-exact attribution, which we are able to infer without issues, namely by "seeing" how the two circular

constructions have been built. In this sense, Byrne's quasi-absolute reliance on visual representations provides the perfect model for avoiding "diagram fallacies," which, as Manders points out:

> Typical alleged 'fallacies of diagram use' rest on taking it for granted that an – to the eye apparently realized but false – exact condition would be read off from a diagram; but the practice never allows an exact condition to be read off from the diagram. Typical 'gaps in Euclid' involve reading off some explicit co-exact feature from a diagram; and this is permissible [...] (Manders 2010).

The very nature of Byrne's special graphical representations of the proposition (and of all the other Euclidean propositions), whereby both diagrams *and* text are symbolized by color-coded visual representations, actually force participants to focus on co-exact conditions. It further compels the reader to reason *in* the diagram, given the diagrammatic nature of the proof's steps. By following Byrne's version of Euclid's proofs, we can actually "see" the stability of Euclid's geometry and its proofs, as provided by the stability of co-exact conditions.

Therefore, for Byrne, what is largely a pedagogical expedient becomes a very efficient way (more so than traditional lettered diagrams) of representing the main features of the Greek archetype, thereby remaining faithful to the original purpose and use of diagrams and diagrammatic reasoning.

In a very brief book review of Byrne's book *The Young Geometrician*, published on the British journal *Atheneum* on July 1866, it is said, probably quite ironically, that "there is always something ingenious about Mr. Byrne's notions." This was the feeling at that time also for his version of *Euclid's Elements*. However, the recent literature on diagrams and diagrammatic reasoning shows how the "ingenious" color proofs used by Byrne cannot just be relegated to the domain of mathematical oddities. In fact, they possess the necessary features of representational cognitive tools and, accordingly, they efficiently allow theoretical and practical reasoning in the diagrams.

References

Ackerberg-Hastings A (2002) Analysis and synthesis in John Playfair's "elements of geometry". Br J Hist Sci 35(1):43–72

Avigad J, Dean E, Mumma J (2009) A formal system for Euclid's elements. Rev Symb Log 2 (4):700–768

Byrne O (1847) The first six books of the elements of Euclid in which coloured diagrams and symbols are used instead of letters for the greater ease of learners. William Pickering, London

Byrne O (1865) The young geometrician. Chapman and Hall, London

Dal Magro T, García-Pérez M (2019) On Euclidean diagrams and geometrical knowledge. Theoria 34(2):255–276

Flood R, Rice A, Wilson R (eds) (2011) Mathematics in Victorian Britain. Oxford University Press, Oxford

Giaquinto M (2007) Visual thinking in mathematics. Oxford University Press, Oxford

Giardino V (2017) Diagrammatic reasoning in mathematics. In: Magnani M, Betrolotti T (eds) Hanbook of model-based science. Springer, New York, pp 499–521

Giardino V (2018a) Tools for thoughts: the case of mathematics. Endeavour 42:172–179
Giardino V (2018b) Manipulative imagination: how to move things around in mathematics. Theoria 32(2):345–360
Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatsh Math Phys 38:173–198
Hanna G, Sidoli N (2007) Visualization and proof: a brief survey of philosophical perspectives. ZDM 39:73–78
Hawes S, Kolpas S (2015) Oliver Byrne: the Matisse of mathematics. Convergence. https://www.maa.org/press/periodicals/convergence/oliver-byrne-the-matisse-of-mathematics
Hydin Price M (1981) The reform of English mathematica education in the nineteenth and early twentieth centuries. PhD thesis, University of Leicester
Macbeth D (2010) Diagrammatic reasoning in Euclid's elements. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 235–267
Mancosu P (2010) The philosophy of mathematical practice. Oxford University Press, Oxford
Manders K (2010) The Euclidean diagram. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 80–133
Panza M (2012) The twofold role of diagrams in Euclid's plane geometry. Synthese 186(1):55–102.
Pasch M (1882) Vorlesungen uber neuere Geometrie. Teubner, Leipzig
Playfair J (1795) Elements of geometry, containing the first six books of Euclid, with two books on the geometry of solids. Bell and Bradfute, Edinburgh
Saito K, Sidoli N (2012) Diagrams and arguments in ancient Greek mathematics: lessons drawn from comparisons of the manuscript diagrams with those in modern critical editions. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 135–162
Simson R (1756) The elements of Euclid, viz. the first six books, together with the eleventh and twelfth. Robert and Andrew Foulis, Glasgow

From Friezes to Quasicrystals: A History of Symmetry Groups

Franka Miriam Brückler and Vladimir Stilinović

Contents

1 Introduction 1824
2 Point Groups 1829
3 Space Groups in Three Dimensions 1844
4 Space Groups in Lower and Higher Dimensions 1849
5 Color Symmetry Groups 1854
6 Conclusion 1858
7 Cross-References 1859
References 1859

Abstract

Even if, since many thousands of years, humans were fascinated by symmetry, which is reflected in many preserved ornaments on buildings, paintings, pottery, etc., the mathematical study of symmetrical patterns is just a few hundred years old. In fact, until the nineteenth century, when it received an impulse from crystallography, no systematical approaches were developed. In this chapter, we follow the history of (Euclidean) symmetry from its first appearances and connections with art, over geometric classifications by nineteenth-century crystallographers, the coalescence of this crystallographic-geometrical approach with theory of groups in the second part of the nineteenth century, until various twentieth-century generalizations to symmetry groups in higher dimensions, magnetic, and colored groups.

F. M. Brückler (✉)
Department of Mathematics, Faculty of Science, University of Zagreb, Zagreb, Croatia
e-mail: bruckler@math.hr

V. Stilinović
Department of Chemistry, Faculty of Science, University of Zagreb, Zagreb, Croatia
e-mail: vladimir.stilinovic@chem.pmf.hr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_132

Keywords

Euclidean symmetry · Mathematical crystallography · History of group theory · Point groups · Space groups · Color groups · Magnetic groups

1 Introduction

Once upon a time, the oldest known temple was built (Fig. 1). It features several symmetrical aspects in alignments and shapes found in its remains (Haklay and Gopher 2020). Highly symmetrical ornaments have been preserved from Paleolithic and Neolithic times, for example, friezes (Fig. 2) exhibiting all seven symmetry types ("frieze groups") can be found already in the Magdalenian period (ca. 12,000–10,000 BCE) (Jablan 2002). These are just some of the many pieces of evidence that since the earliest of times humans have been fascinated by symmetry.

The word *symmetry* itself is of ancient Greek origin: *συμμετρια*, originally referring to harmony of proportions (Weyl 1952). Ancient Greek art is well known for its symmetrical features, and the Greek sculptor Polykleitos (fifth century BCE) wrote on proportions, in particular the "ideal" ones for a human body. Through history, symmetry became to be associated with beauty and regularity, as numerous famous examples in art and architecture history confirm. In some cases, one might

Fig. 1 The archeological site of Göbekli Tepe (ca. 9500–8000 BCE), in modern Türkiye

conjecture that, at least on a subconscious level, attempts were made to systematically explore and design different symmetry *arrangements*, and not only different symmetrical figures themselves. The most studied of such examples are from the medieval Islamic period. The Islamic art is famous for its decorative patterns (Fig. 3). Even if the question if all the 17 possible "wallpaper groups," i.e., possible symmetry arrangements in two-dimensional periodic tilings, are represented in the palace of Alhambra in Granada, modern Spain, is not settled (Grünbaum 2006; Grünbaum et al. 1986; Rønning 2009), the Islamic art provides an abundance of examples of various types of symmetry in the plane, which are nowadays described mathematically by two-dimensional point groups, frieze groups, and wallpaper groups (Jablan 2002; Laigo et al. 2014; Weyl 1952). Indications of attempts of classification by symmetry types can also be found in ethnographic history, e.g., in traditional Hungarian needlework (Hargittai and Lengyel 1985).

As for many other scientific fields and branches, the roots of the systematic research of symmetry arrangements can be traced back to the period of Renaissance. A particularly well-known example is Johannes Kepler's attempt to explain the outer symmetry of snow crystals by inner symmetry (Kepler 1611). Kepler's

Fig. 2 Examples of friezes on pottery fragments found in the same city (Osijek, Croatia) spanning ca. 3500 years of history: (**a**) bronze age, (**b**) ancient Roman, (**c**) medieval

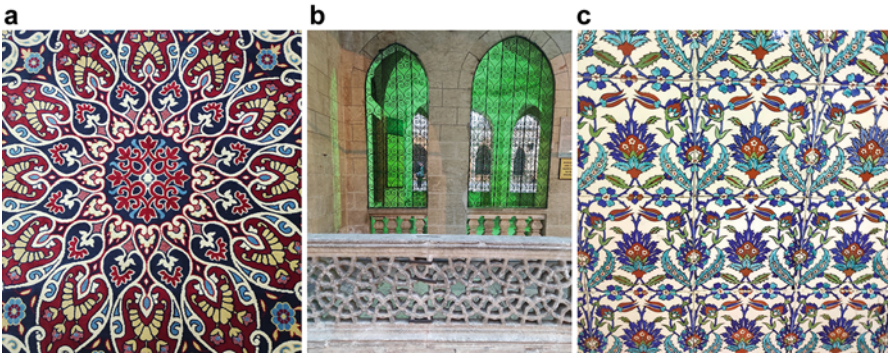


Fig. 3 Examples of highly symmetric Islamic art from Dergah Camı in Şanliurfa, Türkiye: (**a**) a rosette on a carpet, (**b**) a frieze in the front yard, (**c**) a tessellation on a pillar

booklet is also often considered to mark the beginning of mathematical crystallography, which became the main motor for the development of symmetry group theory (Shafranovskii 1975).

There were two main directions developed in crystallography in the seventeenth and eighteenth centuries. One was the study of the external shape of the crystals (Fig. 4), including the measurements of the geometric features of naturally occurring crystals, leading to the *law of constant angles* (the First law of crystallography) discovered in 1669 by Nils Steensen (Nicolaus Steno, 1638–1686) and rediscovered in 1772 by Jean-Baptiste Louis Romé de l'Isle (1736–1790) and Arnould Carangeot (1742–1806). This eventually led to the analysis of symmetries of (three-dimensional) geometric figures. The other were the speculations on the internal structure of the crystals, generally by assuming periodic arrays of constituents (usually, in line with Kepler, presumed to be spheres). The two directions coalesced in the work of René Just Haüy (1743–1822), who proposed that internal structure of crystals comprised a periodic assembly of polyhedra, and derived the various shapes of the crystals, as well as the values of the observed interfacial angles (Burke 1966; Kubbinga 2012; Lalena 2006; Lima-de-Faria 1990; Molčanov and Stilinović 2014; Schuh 2007).

Even if the classifications of types of symmetry arrangements in one to three dimensions required only elementary geometry, they were not fully explored before the nineteenth century. In particular, the emergence of group theory simplified the formulation of the results, as well as the enumeration of different symmetry arrangements. There were several lines of mathematical thought that culminated in the creation of theory of groups, the most important one being the search for solutions in radicals for algebraic equations of degree 5 and higher. The algebraic solutions for linear equations, as well as quadratic ones, were in essence known to ancient

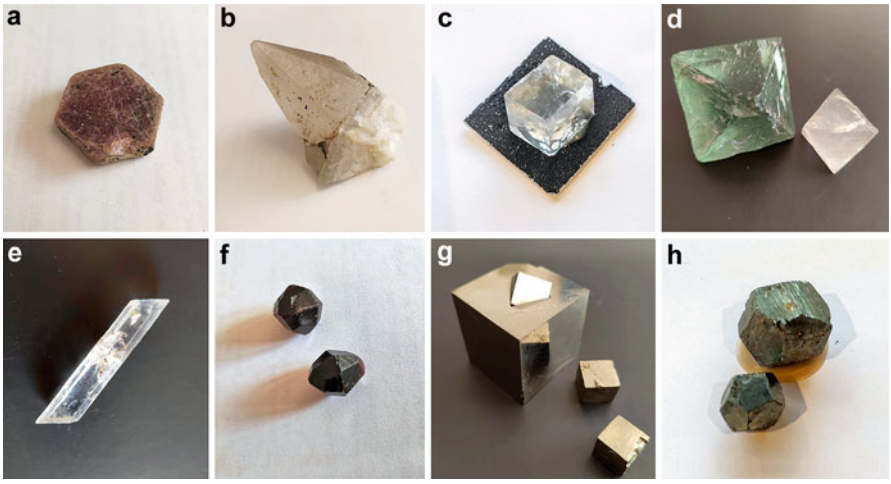


Fig. 4 Examples of crystals from various crystal systems: (**a**) corundum – hexagonal, (**b**, **c**) calcite – trigonal, (**d**) fluorite – cubic, (**e**) gypsum – monoclinic, (**f**) garnet – cubic, (**g**, **h**) pyrite – cubic

Babylonians and Egyptians, but for the first time fully described in Al-Khwarizmi's book *Al-Kitāb al-mukhtaṣar fī ḥisāb al-jabr wal-muqābala* (ca. 825 AD; from its title, the term *algebra* was derived). Solutions in radicals for algebraic equations of degree 3 and 4 were found by the sixteenth-century Italian mathematicians Scipione del Ferro, Niccolò Fontana Tartaglia, Girolamo Cardano, and Lodovico Ferrari. In the second half of the eighteenth century, French mathematicians Joseph-Louis Lagrange and Alexandre-Théophile Vandermonde were researching this topic and tried to find the solutions for equations of higher degrees by analyzing *why* solutions in radicals exist for degrees up to 4. They noted that for degrees 1 to 4, the coefficients in the equations are expressible as symmetric functions of the solutions, i.e., invariant with respect to permutations of solutions. Also, rational functions of the coefficients are symmetric, while the formulas for solutions in radicals are not, which suggests that the root functions are the ones "breaking" the symmetry. This approach via permutations was adopted by other mathematicians, who began to investigate properties of permutations themselves. In particular, the Italian mathematician Paolo Ruffini, around 1800, discovered several properties of sets now known as permutation groups. Ruffini was the first mathematician to claim that fifth-degree equations are in general not solvable in radicals; however, his proofs were incomplete, and the first full proof was given in 1824 by the Norwegian mathematician Niels Henrik Abel. His French contemporary, Évariste Galois, in turn investigated a more general question: Given an algebraic equation of order 5 or higher, how can one determine if that particular equation is solvable in radicals or not? His research marks the establishment of the theory of groups (of permutations). However, even if he developed his results in the period 1828–1832, they only became known after 1846, when Joseph Liouville posthumously published Galois'. Two years before, 1844, Augustin-Louis Cauchy, who was almost surely familiar with at least a part of Galois' work, published a text on *systémes des substitutions conjuguées*, a term for permutation groups which was used for several decades later. Still, it was not until the 1860s and the works of Camille Jordan, Arthur Cayley, and others that the mathematics community fully recognized the importance and power of the theory of groups. Jordan's papers used the term *groups*, which was adopted by subsequent mathematicians, and Cayley was the first who defined and studied abstract groups (Brückler 2017; Kleiner 1986; Van der Waerden 1985).

Meanwhile, other scientists successfully classified some of the symmetry arrangements, and after Jordan's and Cayley's popularization of the theory of groups, these were re-formulated, re-proven, and extended in group theory terms – symmetry group theory emerged. In this chapter, the main moments and results from this development are presented. The aim is to present a history of discovery of various types of symmetry groups as well as the history of their classifications in various dimensions. This survey is suitable not only for specialists and, in most part, is not formal, but a familiarity with basic group theory and Euclidean geometry, on the level presented, e.g., in Chapter 2 of Artin (1991) and the first three chapters of Ryan (1986), is of advantage. However, the most important definitions shall be stated and suitable explanations given as needed.

It has to be emphasized here that in this survey only *Euclidean* symmetry groups, i.e., only symmetries which are isometries (distance-preserving mappings) with

respect to the Euclidean metric, are considered. Thus, for the purposes of the present text, the following definition is used:

Definition 1 A *symmetry group* is a subgroup of the isometry group $E(n)$ of a Euclidean space E of dimension n. A symmetry group is the symmetry group of a set of points F in E if it contains all isometries of the ambient Euclidean space under which F is invariant (these are called symmetries of F).

Since most of the early history of symmetry group theory was motivated by crystallography, and since atoms and ions in crystals cannot be separated by infinitesimal distances, most of this text is restricted to *discrete* symmetry groups (Artin 1991; Brückler 2019; Burckhardt 1988), i.e., symmetry groups with the property that there exists a positive number ε such that:

- The translation vectors of all non-trivial translations in the group (if any) are longer than ε.
- The rotation angles of all non-trivial (both proper and improper) rotations in the group are greater than ε.

For the purposes of this survey, an orthogonal transformation is an isometry possessing a fixed point; for a more formal approach, see e.g., Artin (1991). From general Euclidean geometry, it is well known that every isometry, i.e., symmetry, can be described as a composition of an orthogonal transformation with a translation. This was proven for movements of rigid bodies (i.e., isometries preserving chirality) by Leonhard Euler in 1758. Somewhat later, in 1840, Olinde Rodrigues improved this result by showing that the mentioned translation can be taken to be in the direction of the rotation axis (Van der Waerden 1985, p. 120).

Consequently, symmetry groups can be classified into two basic types:

Definition 2 A *point group* is a symmetry group with the property that all its elements possess at least one common fixed point (i.e., it is a symmetry group containing only orthogonal transformations). A *space group* is a symmetry group containing non-trivial translations.

To be noted here is that since bounded objects cannot possess non-trivial translational symmetries (i.e., translational symmetries which are not the identity transformation), symmetry groups of bounded sets of points are always point groups. Also, point groups in n-dimensional Euclidean spaces are always subgroups of the orthogonal group $O(n)$, which is the subgroup of $E(n)$ consisting of all orthogonal transformations. On the other hand, if a geometric pattern also possesses non-trivial translational symmetry, then it cannot be bounded. The most prominent examples of periodical patterns in nature are (idealized) crystal structures. This is the motivation for defining *crystallographic groups* as a type of space groups which are both discrete and cocompact (Vince 1997). Alternatively, n-dimensional crystallographic groups can be defined as groups containing a normal, torsion-free (i.e., the only

group element of infinite order is the identity), maximal Abelian subgroup of rank n and a finite index (Cid and Schulz 2001). However, for the purposes of this survey, it is sufficient to say that a discrete space group is a crystallographic group if it is the symmetry group of an object (e.g., idealized crystal structure) in n-dimensional Euclidean space that contains translational symmetries in n linearly independent directions, i.e., of a periodic object in n-dimensional Euclidean space. Besides crystallographic groups, in the course of this exposition (discrete) space groups, which describe objects possessing translational symmetry in less than n linearly independent directions (subperiodic objects), as well as some generalizations of the notion (known as magnetic and color groups), shall also be discussed.

2 Point Groups

The first systematic determination of possible symmetries with a fixed point, in modern terms described by point groups, has been attributed to **Leonardo da Vinci** (1452–1519): "Leonardo da Vinci engaged in systematically determining the possible symmetries of a central building and how to attach chapels and niches without destroying the symmetry of the nucleus. In abstract modern terminology, his result is essentially our table of the possible finite groups of rotations (proper and improper) in two dimensions" (Weyl 1952, p.66). According to Weyl (1952), Leonardo has discovered that bounded plane figures either possess only rotational symmetries (with a common center) or both rotational symmetries around a center and line (mirror) reflections about axes passing through the common center of rotations (see Fig. 5). Thus, he noted that all symmetries of a bounded plane figure have (at least) one common fixed point. Bounded plane figures are usually known as *rosettes*; a circle would technically fit the definition of a rosette, but it is usually exempted from

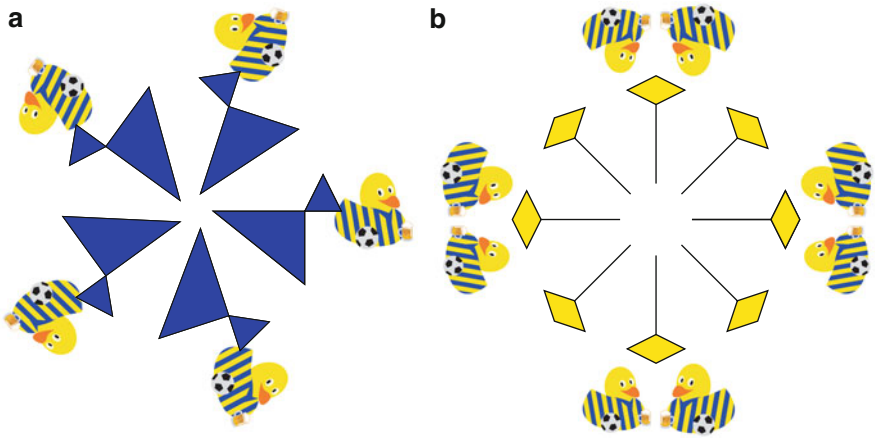


Fig. 5 Rosettes with symmetry groups C_5 (**a**) and D_4 (**b**)

it, since the symmetry group of a circle is the full orthogonal group $O(2)$ (Anglin and Lambek 1995), consisting of all rotations of the plane about a fixed point (the center of the circle) and reflections about any axis through that fixed point. More precisely, it is usually assumed that a rosette is a bounded two-dimensional figure with a discrete symmetry group (Sibley 2015). Thus, in modern terms, the discovery of the following theorem is attributed to Leonardo:

Theorem 1 (Leonardo's Theorem) The symmetry group of a rosette is either C_n or D_n, with n being a positive integer.

In Theorem 1, the symbol C_n is used to denote the *cyclic group* of order n, generated by a rotation with a rotation angle of $360°/n$ (a so-called n-fold rotation). The cyclic group C_1 is also known as the *trivial group*, and it is the symmetry group of all objects that are considered asymmetric in everyday language (mathematically speaking, it only possesses the trivial symmetry, which can be described as a rotation for $360° = 360°/1$ about an arbitrary axis). The symbol D_n is used to denote *dihedral groups* of order $2n$, generated by an n-fold rotation R and a reflection R' about a line through the center of R (in the three-dimensional case, R' is a twofold rotation about an axis intersecting the axis of R). Thus, D_1 is the symmetry group of rosettes possessing only one mirror symmetry, while in three dimensions, D_1 is the same as C_2. Here, as in the rest of the chapter, the *Schoenflies notation* for symmetry groups is used; other common notations are the Hermann-Mauguin notation and Conway's orbifold notation (Conway et al. 2008; Hahn 2005).

In crystallography, it is common to classify crystals into crystal classes. All crystals having the same point group belong to the same crystal class, which can be generalized into:

Definition 3 A *symmetry class* is the collection of all (bounded) objects having the same symmetry (point) group.

In other words, symmetry classes are equivalence classes with respect to "having the same point symmetry group" as an equivalence relation. Classes are usually denoted by the same symbol as the corresponding point groups. Due to Leonardo's systematic explorations of geometric shapes with different symmetries (Reynolds 2008), it might be more appropriate to say that he discovered that the possible symmetry classes of rosettes are C_n and D_n.

Leonardo's contribution apparently remained the only systematic attempt of classifying symmetry types for several centuries. In the beginning of the nineteenth century, even mathematicians did not consider (geometric) symmetry further than mirror-symmetry (Scholz 1989). In the eighteenth century, the definition of symmetry still resembled the ancient Greek one; for example, in Diderot's and d'Alemberts *Encyclopédie* one can read: "SYMMETRY (Architect.) is the relation, the proportion and the regularity of parts that are necessary to compose a nice whole [...]" (Kubbinga 2012, p. 7). The first mathematical definition of mirror symmetry was given by the French mathematician **Adrien-Marie Legendre** (1752–1833), who in

his *Géométrie* (1794) considered only this type of symmetry: "I will call two polyhedra symmetric polyhedra which, having a common base, are constructed similarly such that one is above the plane of this base and the other is below it, with the condition that the summits of the corresponding solid angles are located at equal distances from the plane of the base, on the same line perpendicular to this plane"[1] (Burckhardt 1988, p. 14; Hon 2005, p. 288).

It was crystallography which directed the attention to other types of symmetry of polyhedra and its systematic exploration. The regular (macroscopic) shapes of crystals (Fig. 4) have attracted attention since the antiquity, however the importance of crystal shapes – and thus a need to study their symmetry – has not been widely recognized until the late eighteenth century. In particularly, the idea of symmetry of crystals, although vaguely present, required almost a century to fully "crystallize."

The first to suggest that the shapes of the crystals might be an important feature (and a basis for a classification) of minerals was the Swiss physician and naturalist Moritz Anton Kappeler (or Mauritius Antonius Cappeler, 1685–1769) in his 1723 treatise *Prodromus crystallographiae de crystallis improprie* sic *dictis commentarium* (which also provides the first appearance of the word *crystallographia* = crystallography) (Lima-de-Faria 1990). However, he was only observing the crystal shape in most general terms (e.g., rounded, pyramidal, prismatic, polyhedral, dendritic, ...). Half a century later, Abraham Gottlob Werner (1749–1817) described crystal shapes as derived from six "primary" forms (icosahedron, dodecahedron, prism, pyramid, table, and lens), later adding also cube as the seventh. He also had a vague idea that certain forms were related (by symmetry) – e.g., cube and "perfect double tetragonal pyramid" (i.e., the regular octahedron) – as certain minerals may crystallize in both forms as well as a sequence of intermediary forms. This idea was further developed by Jean-Baptiste Louis Romé de l'Isle (1736–1790) who, basing his hypothesis on much closer observation of crystal forms as well as measurements of interfacial angles, defined the regular tetrahedron, cube, regular octahedron, rhombohedron, rhombic octahedron, and trigonal dodecahedron as the basic forms. From these, he derived further ca. 450 forms, which can be achieved by truncation of vertices or edges of the primitive (or subsequently derived secondary) forms, as possible shapes of crystals. When performing these truncations, he generally preserved some or all of the symmetry of the original form (Fig. 6), showing he had intuited the symmetry of different forms, although he never expressed it explicitly (Burke 1966; Capeller 1723; Romé de l'Isle 1783; Werner 1774).

Based on his description of the inner structure of crystals, R. J. Haüy in 1801 noted that not all five Platonic solids are possible as forms of crystals, but just the regular tetrahedron, hexahedron (cube), and octahedron. He also noted that rotational symmetry of order 5 is not possible in crystals (Burckhardt 1988, p. 18–19). In 1815, he

[1] *J'appelerai polyèdres symétriques deux polyèdres qui, ayant une base commune, sont construits semblablement, l'un au-dessus du plan de celte base, l'autre au-dessous avec celte condition que les sommets des angles solides homologues soient situes a egales distances du plan de la base, sur une meme droite perpendiculaire à ce plan.*

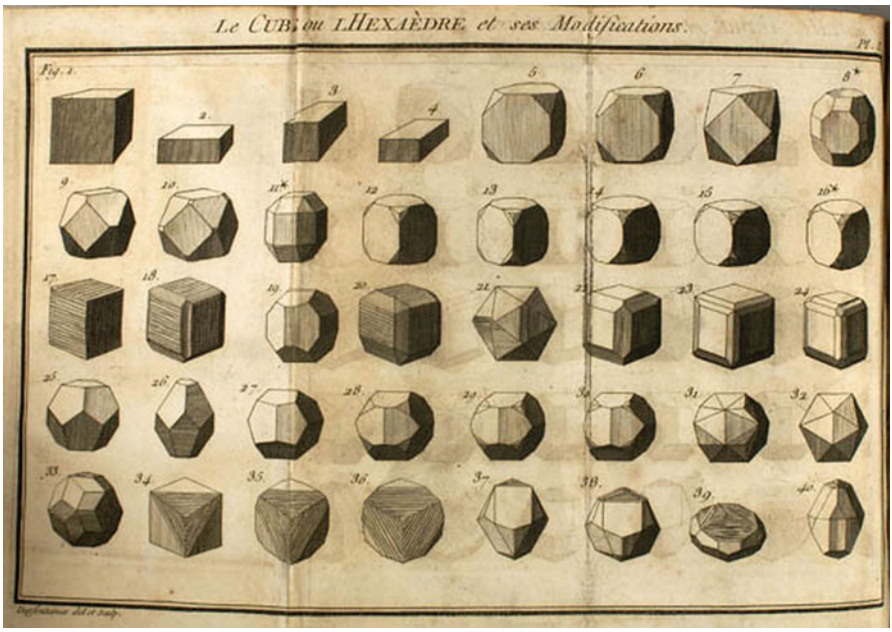


Fig. 6 Table from Romé de l'Isle's *Cristallographie* (1783) depicting various truncations of a cube leading to a series of different forms. Picture courtesy of Virtual museum of the history of mineralogy (https://mineralogy.eu/index.html)

published a series of memoirs concerning what he called the "law of symmetry" (*loi de symmétrie*), which postulated the idea of rotational symmetry in crystals, however without developing it in much detail. Also, he generally considered only a single (vertical) axis of rotation, which made it difficult to explain the observed crystal forms with additional (horizontal) axes of rotation (Haüy 1815). Haüy is also credited with the discovery (1784) of the crystallographic *law of rational indices* (the Second law of crystallography). In mathematical terms, this law says that if a face of the crystal is chosen to lie in the plane with equation $x + y + z = 1$ (thus implicitly defining the coordinate system), all intercepts of all other face planes of the same crystal always have rational coordinates. At Haüy's time, this law was conjectured and supported by observations of macroscopic crystals, but it has later been proven from the periodicity of the inner structure of crystals (Rogers 1912).

The actual classification of (not only) crystal symmetry begins with the so-called German school of crystallography: "The Germans more particularly assessed the outward geometrical appearance of the crystal, just as it presents itself to the observer. From 1815 to 1830 various attempts aiming at an exhaustive classification of the crystalline forms using this key concept had already been reported." (Kubbinga 2012, p. 13). The German mineralogist **Christian Samuel Weiss** (1780–1856), who visited Haüy in 1807/8, did not accept Haüy's molecular model of crystal structure, but argued that the macroscopic form of the crystal is determined

by directional forces leading to anisotropic crystal growth rates and therefore a polyhedral form – the form as well as the forces describable by means of a system of non-coplanar axes (Burke 1966). Building on the (at that time still not proven) fact that crystals cannot exhibit axes of rotational symmetry other than 1, 2, 3, 4, and 6, in 1815, he discussed their possible positions relative to one another and concluded that there are 4 types of possible configurations of axes (7, including sub-types): regular or isometric – modern cubic; four-tier or monodimetric – modern tetragonal; two-tier or trimetric – modern triclinic, monoclinic, and orthorhombic; six-tier or monotrimetric – modern hexagonal and trigonal (Burckhardt 1988; Burke 1966). In 1822, Friedrich Mohs (1773–1839), apparently unaware of Weiss's contribution, arrived also at seven types of configurations, avoiding some errors committed by Weiss, thus defining the modern seven *crystal systems* (Burke 1966; Scholz 1989). The symmetries of the seven configurations of the axes (being also the highest possible symmetries of the crystals belonging to corresponding crystal systems) later became known as *holoedries* (Burckhardt 1988, p. 31–32). After the periodicity of crystal structure was confirmed by X-ray diffraction in 1912, this was found to be equivalent to the modern group-theoretical definition of holoedries (Souvignier 2003; Streitwolf 1967):

Definition 4 A *lattice* in Euclidean space is the set of all points with integer coordinates with respect to a chosen – in general, skew – coordinate system. The *holoedry* is the group of all symmetries of the lattice which fix the origin (also known as the point group of the lattice). A *crystal system* is the set of all crystal structures with the same *holoedry*.

The names of the 7 crystal systems and the corresponding holoedries are given in Table 1.

In the 1820s, it was realized that there "are limits as to the geometry of crystals, and that a rigorously strict deduction of the distinct possibilities in terms of symmetry was a feasible enterprise" (Kubbinga 2012, p. 13). The German physicist, geographer, and crystallographer **Moritz Ludwig Frankenheim** (1801–1869) in 1826 departed from the law of rational indices for crystals and its consequence, the impossibility of rotational symmetries other than those of orders 1, 2, 3, 4, or 6. Using Weiss's classification of crystal axial systems into four types/seven subtypes,

Table 1 The 7 crystal systems and the corresponding holoedries

Crystal system	Holoedry
Triclinic	C_i (the symmetry group of a general skew four-sided prism)
Monoclinic	C_{2h} (the symmetry group of a skew rectangular prism)
Orthorhombic	D_{2h} (the symmetry group of a general cuboid)
Trigonal	D_{3d} (the symmetry group of a trigonal trapezohedron)
Hexagonal	D_{6h} (the symmetry group of a regular hexagonal upright prism)
Tetragonal	D_{4h} (the symmetry group of a square upright prism which is not a cube)
Cubic	O_h (the group of all symmetries of a regular octahedron or cube)

and analyzing which arrangements of lesser symmetry are possible for each, he arrived at 32 crystal classes. In fact, he essentially determined the 32 crystallographic point groups as subgroups of the seven holoedries (Burckhardt 1988, p. 34–39; Frankenheim 1826). However, Frankenheim's contribution remained generally unknown at the time, was rediscovered only ca. 150 years later, and therefore had no influence on the further developments of symmetry theory in the nineteenth century. For this reason, the first proof of the existence of exactly 32 crystal classes (crystallographic point groups) is usually attributed to the German physician and mineralogist **Johann Friedrich Christian Hessel** (1796–1872). Hessel's work also remained unnoticed for ca. 60 years, until L. Sohncke drew attention to it (Burckhardt 1988; Coxeter 1948; Cromwell 1997; Scholz 1989).

Hessel, who was familiar with Weiss's and Haüy's work, described the systematic derivation of possible symmetries of (macroscopic) crystals, and more general solids, in his more than 300-page-long essay *Krystall*, published in the 1830 edition of *Gehler's physikalisches Wörterbuch* (Scholz 1989). He described (macroscopic) crystal shapes as polyhedral and commented on the insufficiency of mathematical symmetry theory at that time: "A scientific knowledge of the forms of the crystals that satisfies all requirements presupposes that one is familiar with the general principles which the theory of shapes must follow when investigating the various possible shapes. However, since these principles are not yet clearly expressed in works on pure mathematics, and since the general theory of shapes in the extent it is now available is still a young science, it has to first be developed here, before continuing to its application to crystal shapes"[2] (Hessel 1830, p. 1034).

Thus, Hessel develops a systematical study of symmetries of two- and three-dimensional (bounded) figures. Besides mirror symmetry, he also describes p-fold rotational symmetry, explicitly allowing rotational symmetry of infinite order ($p = \infty$), see (Hessel 1830, p. 1045). First, he studies mirror and rotational symmetries of (bounded) plane figures. By studying the possibilities of rotating the figure itself or its mirror image into positions that coincide with the original, Hessel arguments that the second is only possible if the figure itself is mirror-symmetric and continues with the proof that every plane figure is either *1fach pgliedrig* (exhibits only p-fold rotational symmetry), or it is *2fach pgliedrig* (also exhibits mirror symmetry). In other words, in this part (Hessel 1830, pp. 1040–1043), Hessel has proven Leonardo's theorem (including circular symmetry), and this is apparently the earliest published proof of it. Hessel showed even more: For a chosen line in the plane of the figure (intersecting the p-fold rotational axis in the midpoint), in the first case one obtains p equivalent positions of that line (each two neighboring ones separated by an angle of $360°/p$), and

[2] *Eine wissenschaftliche Kenntniss von den Formen der Krystalle, die allen Anforderungen genügt, setzt voraus, dass man mit den allgemeinen Principien vertraut sey, welche die Gestaltenlehre bei Untersuchungen der verschiedenen möglichen Gestalten überhaupt zu befolgen hat. Da diese Principien jedoch in Werken über reine Mathematik noch nicht so bestimmt aus gesprochen sind, indem die allgemeine Gestaltenlehre in der Ausdehnung, wie sie jetzt vorliegt, eine noch jugendliche Wissenschaft ist, so muß sie hier erst entwickelt werden, um so dann die Anwendung derselben auf die Krystallgestalt folgen zu lassen.*

in the second case $2p$ equivalent positions line (each two neighboring ones separated by an angle of $360°/(2p)$). In group-theoretic terms, at that time unknown to Hessel (as noted in the introduction, Galois developed the foundations of group theory in the period 1828–1832, but it was not publicly known at that time, and besides, these foundations only considered groups of permutations), he has proven that the cyclic groups C_p are of order p, and that the dihedral groups D_p are of order $2p$.

Hessel then continues to study symmetries of solids. Here, he notes the possibility of a polyhedron possessing more than one rotational axis of symmetry and introduces names for rotational axes of various orders, including a systematic investigation of possibilities of rotations into mirror-symmetric positions. In other words, he considered both proper and improper rotations, where *improper rotations* in three-dimensional space are defined as the composition of a rotation with a mirror symmetry (reflection on a plane). In particular, mirror symmetries are examples of improper rotations of order 1. Hessel distinguishes classes of solids depending on if they possess a principal axis of symmetry (*Hauptaxe*) or not: A principal axis of symmetry, according to Hessel's definition, is a rotational axis that has no symmetrically equivalent positions for the same solid. An example of a solid with a principal axis would be a regular hexagonal prism, while a cube is an example of a solid without it. By systematically investigating the possibilities, Hessel enumerates (describes) symmetry types – de facto point groups – of solids with a p-fold principal axis. Again, he explicitly mentions the case of $p = \infty$ and the possibility of conical and cylindrical symmetry.

He then continues with the more abstract classification of the axial systems (possible configurations of rotational axes as symmetry axes of solids) and describes the forms (*Gestalten*) which correspond to each of the axial systems with a principal axis. To achieve this, he arrives at many geometric results on possible polyhedral shapes, or in crystallographic terms, Hessel developed another important notion in crystallography, the crystal forms. The description of this theory is not directly connected to our topic, and the interested readers are referred either to the original text (Hessel 1830) or the more modern exposition (Nelson 2011). Then Hessel continues to *hauptaxenlose Strahlensysteme*, axial systems without a principal axis. Developing his systematic theory, he geometrically proves several results about the possible orders and angles between rotational axes in such systems, e.g., that the highest possible order of a rotational axis in an axial system without a principal axis cannot be higher than 5 (Hessel 1830, p. 1116). Even if Hessel's approach is – from a modern perspective – cumbersome, and involves no group theory, he successfully and geometrically proved (the geometric equivalent of) the following generalization of Leonardo's theorem to three dimensions:

Theorem 2 (Hessel's Theorem) The only possible point groups of bounded sets of points in three-dimensional Euclidean space are:

1. With the principal axis:
   - C_n (contain only symmetries, which are rotations about one axis)
   - D_n *(only symmetries are rotations about one principal axis of order n and about n horizontal twofold axes)*

- *C_{nh} (like C_n with additionally one horizontal plane of reflection)*
- *C_{nv} (like C_n with additionally n planes of reflection containing the principal axis)*
- *D_{nh} (like C_{nv} with an additional horizontal plane of reflection)*
- *D_{nd} (like D_n with additional planes of reflection containing the principal axis and bisecting the angles between neighboring horizontal twofold axes)*[3]
- *S_{2n} (like C_n, but the principal axis is an improper one; not to be confused with the symmetric group of all permutations of a set with 2n elements, which has the same notation)*
- *Symmetry groups of a circular cylinder ($D_{\infty h}$) and circular cone ($C_{\infty v}$)*

2. Without the principal axis:
   - *T, O, I (groups consisting only of rotational symmetries of a regular tetrahedron, octahedron, or icosahedron)*
   - *T_d, O_h, I_h (groups of all symmetries of a regular tetrahedron, octahedron, or icosahedron)*
   - *T_h (like O_h, but with fourfold axes replaced by twofold ones)*
   - *Symmetry group of a sphere (K_h)*

   These point groups are illustrated in Fig. 7 by examples of solids in the corresponding symmetry classes.

Equivalently and more concisely, Hessel's theorem states that all finite subgroups of O (Artin 1991) are subgroups of one of the following four types of groups: D_{nh}, T_d, O_h, I_h (Sibley 2015, p. 290). The modern, group-theoretical proof of Hessel's theorem can be found, e.g., in Grove (1997).

Apparently unaware of Hessel, Auguste Bravais (more famous for his classification of plane and space lattice, see next section) also obtained the classification of polyhedral symmetries (i.e., Hessel's theorem) in 1849. In his two 1849 papers (Bravais 1849a, b), he studied symmetries of polyhedra, departing from Legendre's definition of (mirror) symmetry, adding his own definitions of central and rotational symmetry, and proving many classical geometrical results on symmetry, for example, that every rotational symmetry of a polyhedron is of integer order, i.e., that its angle is commensurable with 360° (Bravais 1849b, p. 141–142). In group theoretical terms, Bravais' results describe the properties of elements of point groups and their compositions.

In the second part of his essay (Hessel 1830), Hessel returned specifically to the study of crystals to arrive at the result he is most famous for: There are only 32 crystallographic point groups (crystal classes) in three-dimensional space. He proved this by combining his previous results on general symmetries of polyhedra and axial systems with arguments about directions of forces acting on crystal planes and the known macroscopic forms of crystals of his time. In this way, he arrived at

[3] Some of the groups listed above are not only isomorphic, but the same. For example, $C_{1h} = C_{1v}$ and $D_{1h} = C_{2v}$.

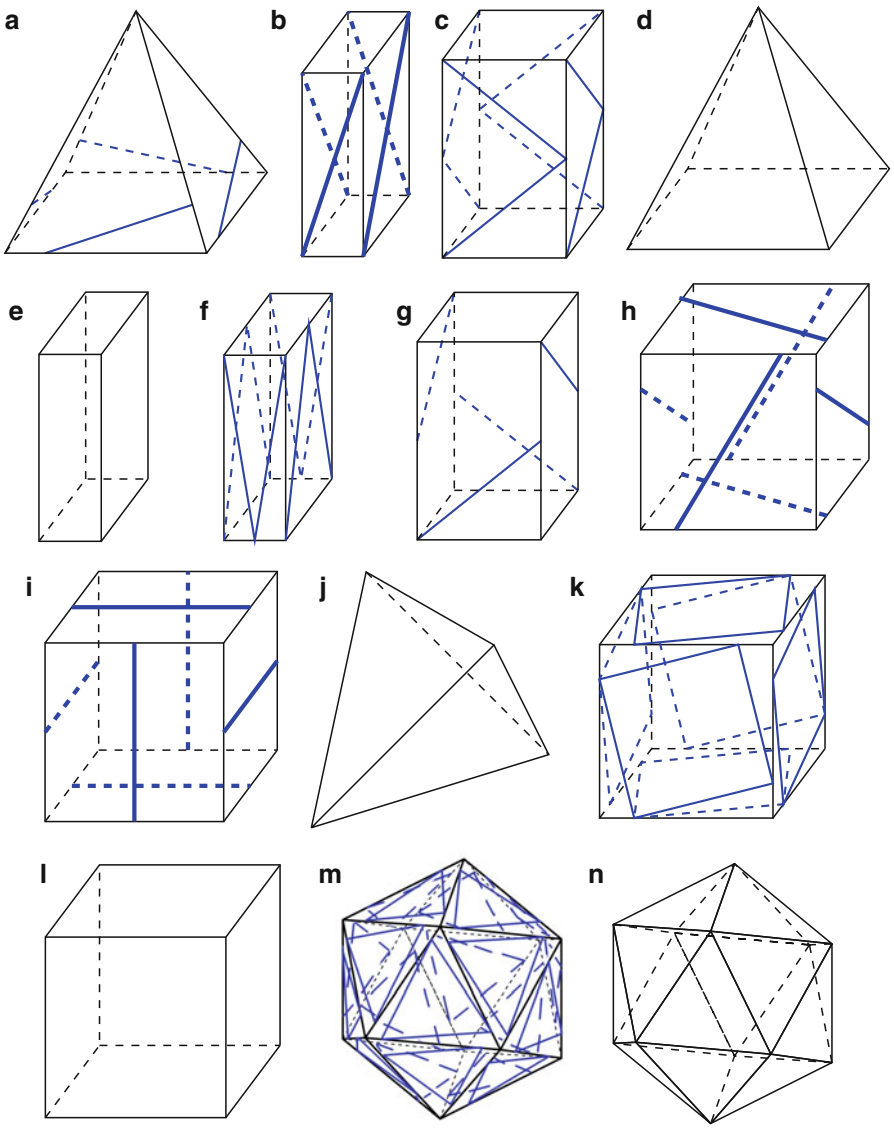


Fig. 7 Examples of solids with and without decoration demonstrating 14 basic types of three-dimensional point groups: (**a**) C_n (with $n = 4$), (**b**) D_n (with $n = 2$), (**c**) C_{nh} (with $n = 4$), (**d**) C_{nv} (with $n = 4$), (**e**) D_{nh} (with $n = 2$), (**f**) D_{nd} (with $n = 2$), (**g**) S_{2n} (with $n = 2$), (**h**) T, (**i**) T_h, (**j**) T_d, (**k**) O, (**l**) O_h, (**m**) I, n) I_h

the conclusion that not all the classes (point groups) from his theorem are possible for crystals and finally determined the possible crystal classes (Burckhardt 1988; Hessel 1830). Barlow refers to Hessel as the first to discover the 32 crystal classes – obviously Frankenheim's results were unknown to Barlow – highlighting the fact

that Hessel "treats crystal symmetry purely as a matter of similarity of position with respect to one another of lines and planes which obey Haüy's law of rational indices" (Barlow 1901, p. 2). In other words, Hessel's arguments lack the formal mathematical background, since his correct geometrical arguments depend on the inner structure of crystals, which implies the impossibility of rotational axes of orders other than 1, 2, 3, 4, and 6 (and the law of rational indices).

Definition 5 A point group (in two- or three-dimensional Euclidean space) is called a *crystallographic point group* if it does not contain (proper nor improper) rotations of orders other than 1, 2, 3, 4, and 6.

For example, of all the possible two-dimensional point groups from Leonardo's theorem, only 10 (C_1, C_2, C_3, C_4, C_6, D_1, D_2, D_3, D_4, D_6) are crystallographic. Independently of Hessel, Frankenheim, and Bravais, the Finnish-born Russian mineralogist **Axel Gadolin** (1828–1893) published his study, which included his proof of the existence of exactly 32 crystal classes, in 1869 in Russian (and 1871 in French). Like Hessel and Frankenheim, he based his argument for the impossibility of symmetries of orders 5 or higher than 6 on the law of rational indices and studied the possible combination of rotational axes and other symmetry elements to arrive at the result (Barlow 1901; Burckhardt 1988; Saurel 1911; Whitlock 1934):

Theorem 3 (Frankenheim-Hessel-Bravais-Gadolin) There exist exactly 32 crystallographic point groups (i.e., crystal classes, as illustrated in Fig. 8):

- 11 consisting only of rotations: C_1, C_2, C_3, C_4, C_6, D_2, D_3, D_4, D_6, T, O;
- 21 also including central or mirror symmetries: C_{1h} (often denoted C_s), C_{2h}, C_{3h}, C_{4h}, C_{6h}, C_{2v}, C_{3v}, C_{4v}, C_{6v}, D_{2h}, D_{3h}, D_{4h}, D_{6h}, D_{2d}, D_{3d}, S_2 (usually denoted by C_i), S_4, S_6 (usually denoted by C_{3i}), T_d, T_h, O_h.

As already mentioned, parallelly to the described developments in crystallography, group theory emerged, initially as the study of groups of permutations related to solutions of algebraic equations. Its importance and generality first began to be noted in the second half of the nineteenth century. Felix Klein used it in 1872 for "bringing order to geometry" (Kleiner 1986, p. 198). During the 1870s Sophus Lie, Henri Poincaré (and Felix Klein) introduced connections of group theory with analysis, and also Arthur Cayley's 1854 introduction of abstract groups began to have influence in the mathematical community (Brückler 2017; Kleiner 1986). However, for the present chapter, the most important line of thought, which enabled the reformulation of the described results on symmetry classes in group-theoretical terms and provided foundations for further classifications of symmetry patterns, began with one of many contributions to group theory by the French mathematician **Camille Jordan** (1838–1922). He was the first to study "groups of movements" (or in modern terms, symmetry groups) in his two-part study titled *Mémoire sur les groupes de mouvements* (1868/69). Jordan considered movements of space and his work was inspired by Bravais' *Etudes cristallographiques*. He defined groups of movements as follows: *The group formed by the movements having this*

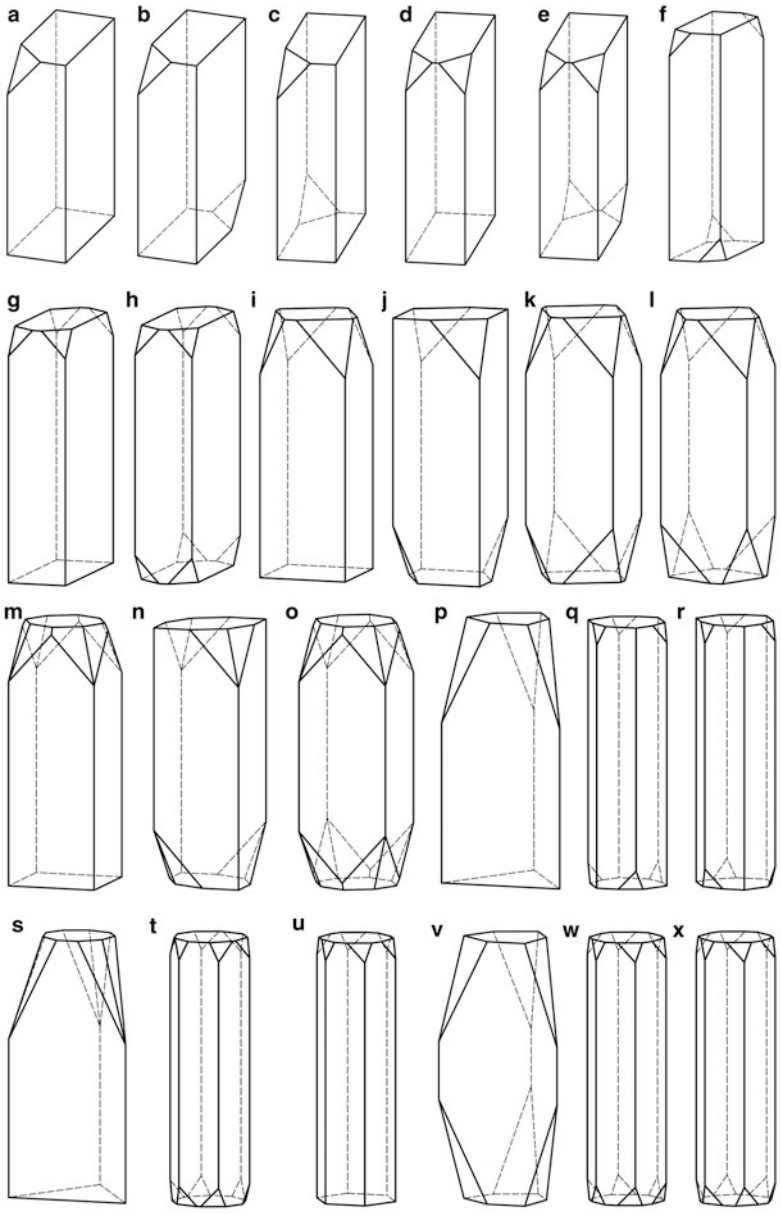


Fig. 8 Examples of polyhedra belonging to each of the 32 crystal classes: 1. Triclinic system (**a**) C_1, (**b**) C_i; 2. Monoclinic system: (**c**) C_2, (**d**) C_s, (**e**) C_{2h}; 3. Orthorhombic system: (**f**) D_2, (**g**) C_{2v}, (**h**) D_{2h}; 4. Tetragonal system: (**i**) C_4, (**j**) S_4, (**k**) C_{4h}, (**l**) D_4, (**m**) C_{4v}, (**n**) D_{2d}, (**o**) D_{4h}; 5. Trigonal system: (**p**) C_3, (**q**) C_{3i}, (**r**) D_3, (**s**) C_{3v}, (**t**) D_{3d}; 6. Hexagonal system: (**u**) C_6, (**v**) C_{3h}, (**w**) C_{6h}, (**x**) D_6, (**y**) C_{6v}, (**z**) D_{3h}, (**aa**) D_{6h}; 7. Cubic system: (**bb**) T, (**cc**) T_h, (**dd**) O, (**ee**) T_d, (**ff**) O_h. Pictures courtesy of Darko Tibljaš, generated by software Kristall2000 (Schilling 2019)

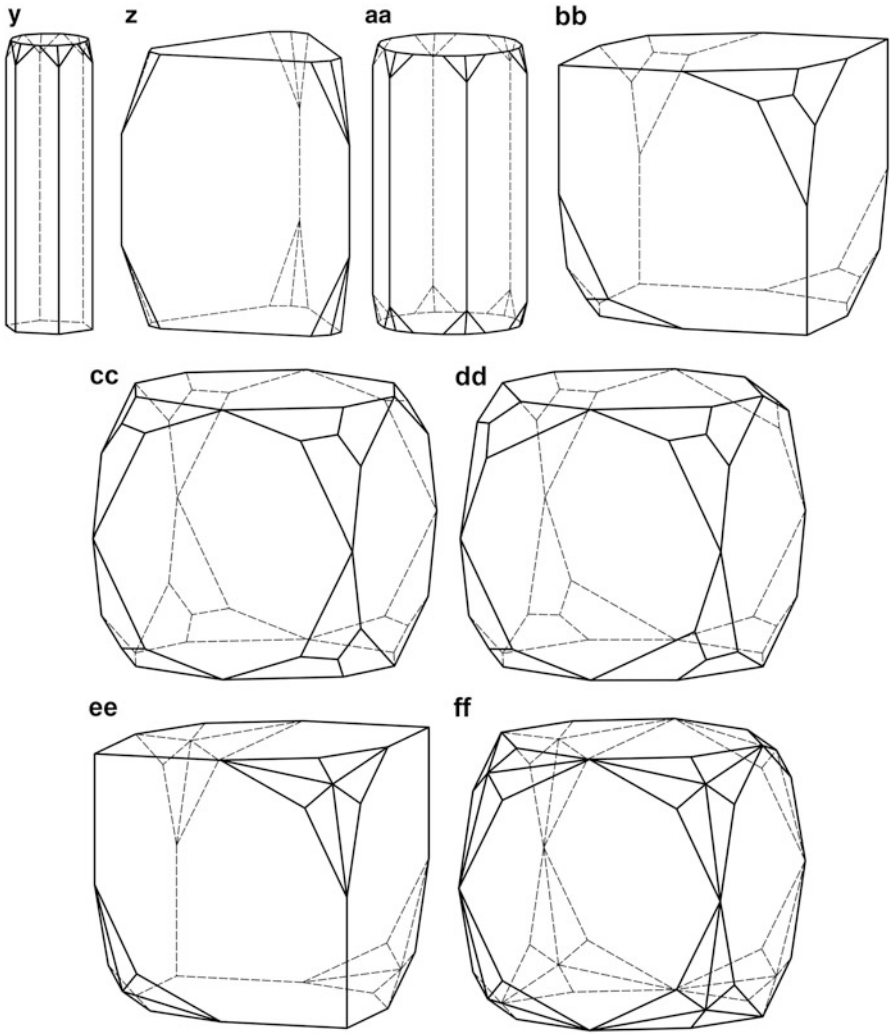


Fig. 8 (continued)

characteristic property that if M' and M" are any two movements forming part of this group, M'M" will also be part of it[4] (Jordan 1869, p. 168). Then he set out to enumerate all possible groups of movements. To achieve this, he first determined all (discrete *and* continuous) translation groups in three-dimensional space and in total obtains 10 types of topologically closed groups of translations. Then he found all

[4] *Le groupe forme par les mouvements cher de cette propriete earacteristique, quo si M' et M" sont deux mouvements quelconques faisant partie de ce groupe, M'M" en fera partie.*

possible groups of rotations in three-dimensional space: the discrete groups of type C_n, D_n, T, O, I, the two continuous rotation groups (all rotations about a fixed point or about a fixed axis); and finally the group consisting of all rotations about an axis *and* all "full turns" about axes perpendicular to the first one. In modern terms, Jordan determined all subgroups of the special orthogonal group SO (Artin 1991); these are often called *chiral point groups* (Jordan 1869; Kubbinga 2012; Van der Waerden 1985). As defined in the introduction, point groups are subgroups of orthogonal groups $O(n)$, and the name orthogonal group has apparently also been used first by Jordan a year later (Jordan 1870). Thus, it seems appropriate to say that Camille Jordan is *the* father of symmetry group theory.

By the beginning of the twentieth century, all of this became common knowledge in the crystallographic and mathematical community, as discernible, e.g., from the paper (Barlow 1901). Barlow begins his paper by emphasizing that no study of crystal symmetry can be complete without considering their internal structure and says: "it is somewhat remarkable that the important work of deducing the existence of 32 classes of crystal symmetry from first principles *has been based on crystal form and not on internal structure*" (Barlow 1901, p. 2). After criticizing the deficiencies of such proofs, in particular Gadolin's method, he explicitly bases his further considerations on Jordan's work. Barlow's starting point is the following "Fundamental definition. Homogeneity of structure consists in a likeness of ultimate parts or molecules of a body, both as to their nature and their relative arrangement, of the following kind: Corresponding to every mathematical point in the mass are found evenly distributed at finite intervals a number of points whose relation to the ultimate structure, regarded as of unlimited extent, is the same as that of the point selected" (Barlow 1901, p. 5). It is clear that Barlow's homogeneity is what is nowadays called periodicity. He then proves the existence of exactly 32 crystal classes (i.e., crystallographic point groups) combining arguments of Sohncke and Gadolin with mathematical formality of (geometric) arguments based on the periodic inner structure. Barlow does not explicitly use group theory techniques in his proofs, but his reference to Jordan and several formulations of partial results on "groups of movements" clearly show that he was well acquainted with the relation between symmetries (of crystals) and the corresponding groups.

Even if Barlow's criticism of Hessel's and Gadolin's arguments is well justified, it relates only to the reduction of all three-dimensional point groups to the ones relevant for crystallography. Hessel's theorem and its proof, even if formulated purely geometrically, is correct and only its original formulation (and the style of arguments, not their content itself) separates it from the modern group-theoretic one, so it is fully appropriate to say that Hessel was the first to prove it.

In the beginning of the twenty-first century, the English mathematician **John Horton Conway** (1937–2020) proved Hessel's theorem in an apparently different, but equivalent setting. Conway considers *spherical symmetry groups*, discrete symmetry groups of patterns on spheres and arrives at the same 14 types of (discrete) symmetry groups as Hessel (Conway et al. 2008). The equivalence of these two approaches is best explained by the following quote from (Belcastro and Yackel 2011, p. 151): "[. . .] if we had a tiling on a sphere made of soft clay, we could then flatten each tile (perhaps by smacking it with a pan) and make our previously tiled

sphere have a surface with flat sides [. . .]. If we then also straightened out the edges, the resulting object would be called a polyhedron. [. . .] Conversely, we could inflate a polyhedron until it is projected onto a sphere." For all readers interested in handicraft and/or art, it might be of interest to mention here that the Japanese temari-balls provide a beautiful illustration of Conway's spherical symmetry, i.e., three-dimensional point groups (Belcastro and Yackel 2011; Yackel 2012).

Before moving on to other dimensions than 2 and 3, reflection groups should be mentioned. These are point symmetry groups generated by reflections. While in dimensions up to 3, the usual notion of mirror symmetry is equivalent to reflection, it is customary to define reflections as a special type of isometries in arbitrary dimension:

Definition 6 A *hyperplane* in an n-dimensional Euclidean space E is any $(n-1)$-dimensional affine subspace of E. A *reflection* on a hyperplane Π is a non-trivial isometry of E for which the set of fixed points is Π. A *reflection group* is a discrete subgroup of the orthogonal group $O(n)$ generated by a set of reflections on hyperplanes in the n-dimensional Euclidean space.

Even if the motivation for considering reflection groups is most naturally described by kaleidoscopes (Fig. 9) and the corresponding symmetries of the

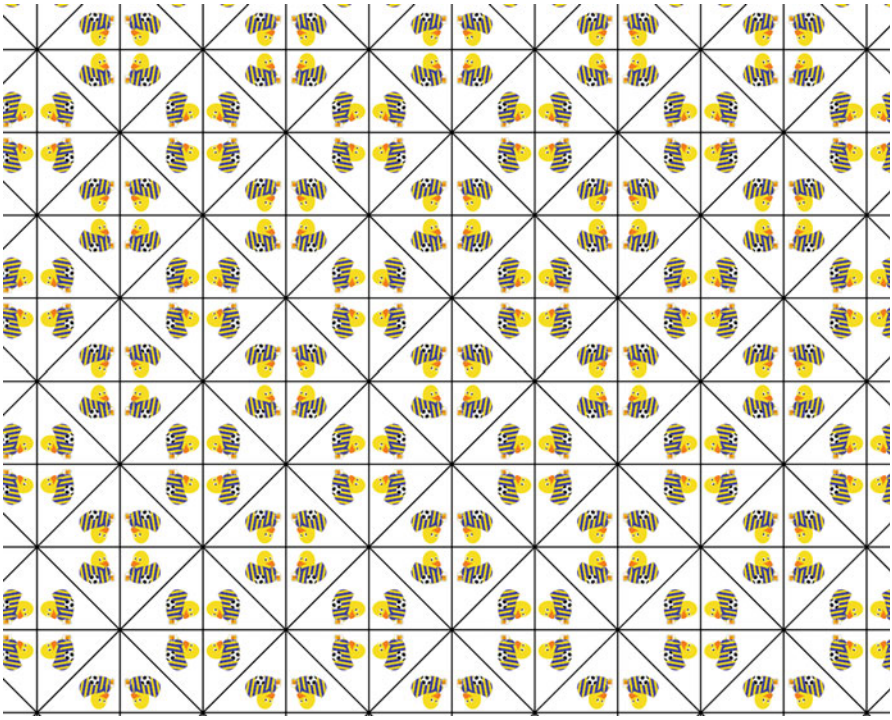


Fig. 9 Example of kaleidoscopic symmetry in the plane

plane, their separate study historically arose partly from the described line of investigation of polyhedral symmetries, partly from the research on Lie groups (in arbitrary dimensions) by W. Killing, E. Cartan, and H. Weyl. It was **Harold Scott MacDonald Coxeter** (1907–2003) who first classified reflection groups in arbitrary dimension, and introduced their abstraction, which is now known as Coxeter groups (Conway et al. 2008; Coxeter 1934, 1935; Goodman 2004).

Coxeter's studies were not the first to consider point symmetries in higher dimensions. He himself refers to an 1889 paper by the French mathematician **Édouard Jean-Baptiste Goursat** (1858–1936), who considered symmetry operations in four-dimensional space (Conway and Smith 2003; Coxeter 1934; Goursat 1889; Hurley 1951; Van der Waerden 1985). A few years before Coxeter's introduction of reflection groups, W. Threlfall and H. Seifert published the first enumeration of all four-dimensional point groups (Threlfall and Seifert 1931; Van der Waerden 1985). However, "Every group does appear in Seifert and Threlfall's list, but appears an uncertain number of times, since the equivalences between the groups that depend on parameters are not completely listed"(Conway and Smith 2003, p. 48). In 1949, C. Hermann provided the foundation for listing all crystallographic orthogonal transformations in arbitrary finite dimensions (Hermann 1949), and soon the first paper containing the enumeration of crystallographic point groups in dimension 4 followed (Hurley 1951). Here Hurley, departing from Goursat's results, deduced the 24 point-symmetry operations in four-dimensional space (MacKay and Pawley 1962). Using periodicity as the defining characteristic of crystal structure, i.e., determining crystallographic point groups in four-dimensional space as holoedries (see next section for crystallographic restriction in higher dimensions), Hurley found 222 crystallographic point groups in dimension 4. This was corrected in 1967, obtaining the final number of 227 crystal classes in four-dimensional space (Hurley et al. 1967). Conway and Smith (2003) finally described all point groups in four-dimensional space, i.e., all discrete subgroups of O (Bamberg et al. 2003). Starting from the general result that for an n-dimensional Euclidean space, every orthogonal transformation acting on it can be described as a composition of at most n mirror symmetries (reflections), they obtained a characterization of the elements in O (Bamberg et al. 2003) and classified all its subgroups to obtain the four-dimensional equivalent of Leonardo's and Hessel's theorem. In total, they obtained 71 types of discrete point groups in dimension 4 (Conway and Smith 2003, p. 44–47). In order to enumerate crystallographic (not only point) groups in higher dimensions, various algorithms were developed in the late twentieth and early twenty-first centuries, and the numbers of crystallographic point groups (geometric crystal classes) in dimensions 5 and 6 have been found to be 955 and 7103, respectively (Downward 2011; Plesken 1981; Plesken and Schulz 2000; Veysseyre and Veysseyre 2002; Veysseyre et al. 2002).

During the twentieth century, the approach to symmetry groups shifted even more toward abstract group-theoretic, and lists of crystallographic point groups are now understood simply as sets of representatives for the **Q**-classes of space groups: The point group P corresponding to a space group G can be defined as (isomorphic to) the quotient group G/T (T being the translational subgroup of G). Point groups can always be taken (represented) as finite subgroups of GL(n, **Z**), since every point

symmetry of an n-dimensional lattice is an invertible linear operator, and every invertible linear operator acting on an n-dimensional space can be, with respect to any chosen basis of the space, represented by an invertible square matrix of order n. If the basis is chosen to consist of n non-coplanar vectors connecting a fixed lattice point O (the origin) to n of its nearest neighbors, the matrix entries will be integers, i.e., the matrix will be an element of GL(n, **Z**). Two space groups are **Q**-equivalent if the corresponding point groups are conjugate in GL(n, **Q**). Besides the numbers of **Q**-classes (also called geometric crystal classes) the numbers of **Z**-classes (also known as arithmetical crystal classes) are known in dimensions up to 6 too (Brown et al. 1978; Plesken and Schulz 2000; Schwarzenberger 1984).

Finally, a short note on point groups in dimension 1. The only symmetry groups of bounded parts of a straight line are quite obviously C_1 and D_1, exemplified by the following two linear patterns:— - and –. The fact appears to be quite trivial, and there seems to be no reference as to who first explicitly noted it; implicitly, it was surely known to the mathematicians and crystallographers of the nineteenth century, and probably also much earlier.

3 Space Groups in Three Dimensions

Space groups describe the symmetries of periodic and subperiodic objects. In Fig. 10, the difference between periodicity and subperiodicity is illustrated by an example of a *wallpaper* (periodic object in the plane) and a *frieze* (subperiodic object in the plane). More formally:

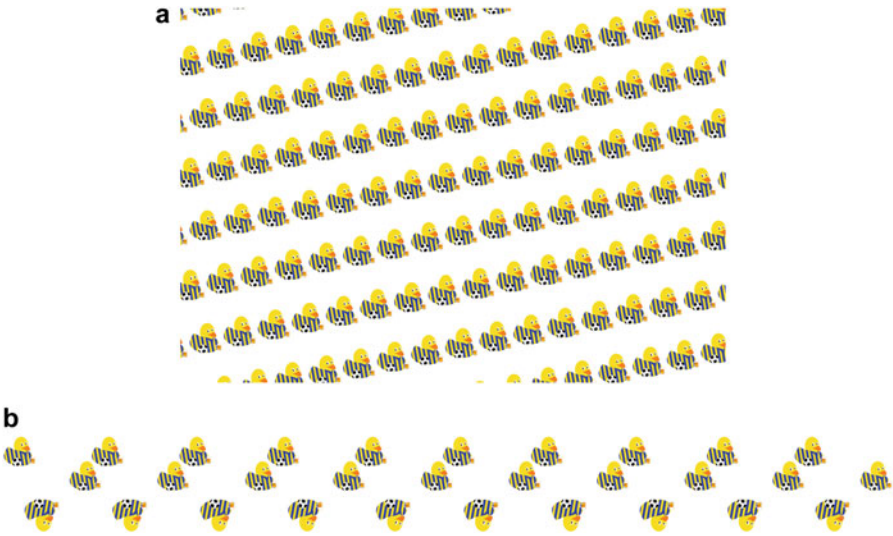


Fig. 10 Bounded clippings from a wallpaper (**a**) and a frieze (**b**) – a periodic and a subperiodic object in two-dimensional Euclidean space

Definition 7 A subset C of an n-dimensional Euclidean space is said to be *periodic* if there exists a basis B of the corresponding vector space such that translations for all basis vectors are symmetries of C, and it is said to be *subperiodic* if translations for at least 1 and at most n–1 basis vectors are its symmetries.

It is important to note that if C is periodic, then the basis B spans a parallelotope (this is the generalization of parallelograms and parallelepipeds to arbitrary finite dimension), which is usually called the *unit cell* of C. This unit cell is compact, thus the definition of a crystallographic space group as a discrete and cocompact space group is equivalent to saying it is the symmetry group of a periodic set. The fundamental models of periodic sets are lattices, since every periodic set can be described by superimposition (de facto, convolution) of the unit cell with the lattice (in a coordinate system with origin in one of the vertices of the unit cell and coordinate axes defined by B). In the following, the term "space group" is used when speaking of crystallographic space groups, i.e., symmetry groups of periodic objects. When speaking of space groups of subperiodic objects, the term "subperiodic space group" shall be used. These are discrete too, but not cocompact.

The early history of space groups is related to periodic plane tessellations (tilings). It is well-known, and often attributed to the Pythagoreans, that there exist only three regular tilings of the plane – with squares, isosceles triangles, and regular hexagons (Anglin and Lambek 1995). The oldest preserved proof of this fact appears to be Johannes Kepler's proof in his *Harmonice Mundi* (1619) (Burckhardt 1988; Brückler 2017; Grünbaum and Shephard 1987; Kepler 1619). This is a special case of the fundamental property of periodic objects in the background of the derivation of the space groups and known as the *crystallographic restriction*. The classical crystallographic restriction is a theorem which states that periodic objects in three-dimensional space cannot possess rotational axes of orders other than (1,) 2, 3, 4, or 6. Due to the laws of crystal growth, this property is also valid for macroscopic crystal shapes. As mentioned in the previous section, both Frankenheim and Hessel implicitly used it when they derived the existence of exactly 32 crystal classes, i.e., crystallographic point groups, even if explicitly their arguments were based on its consequence, the law of rational indices (Bamberg et al. 2003; Sibley 2015). The usual, simple geometric proof of the crystallographic restriction is generally attributed to W. Barlow (Barlow 1901, p. 17; Coxeter 1969, pp. 60–61), and it is equally valid for two-dimensional periodic objects. This, however, was not the first published mathematical proof of the crystallographic restriction: A. M. Schoenflies gave a proof already in 1891 (Schoenflies 1891), but it was lengthy and rather complicated.

The classification of lattices as models of periodic structures is the first step toward the determination of all space groups. The first to attempt a classification of space lattices was M. L. Frankenheim in 1835 and 1842. He arrived at 15 types of lattices, one too many – his error was that in the monoclinic system he counted one type of lattice twice. This error is not present anymore in Frankenheim's 1856 text, *Die Anordnung der Moleküle im Krystalle* (Burckhardt 1988, p. 48–51). Thus, the first full classification is attributed to the one who, in 1850, mathematically

rigorously achieved it both for dimensions 2 and 3: the French physicist **Auguste Bravais** (1811–1863). He was the first to systematically and purely mathematically investigated lattices. His famous 1850 essay *Mémoire sur les systèmes formés par des points distribués régulièrement sur un plan ou dans l'espace* contained a detailed formal mathematical proof of the existence of exactly 5 plane (Fig. 11) and 14 space lattice types, which are now known as *Bravais lattices* in his honor. Apparently, even if Bravais was familiar with results from the French and German schools of crystallography, he was not yet familiar with group theory, and his proofs use a combination of geometric and linear algebra arguments (Bravais 1850; Burckhardt 1988).

From the modern point of view, the definition of Bravais lattices is more abstract (Streitwolf 1967, p. 45):

Definition 8 Let T be the translational subgroup of a space group of a periodic object C, i.e., the subgroup consisting of all translational symmetries of C. If all translational vectors of translations from T are taken to have a common initial point O, then the set of all terminal points of all these vectors is the *Bravais lattice* of C.

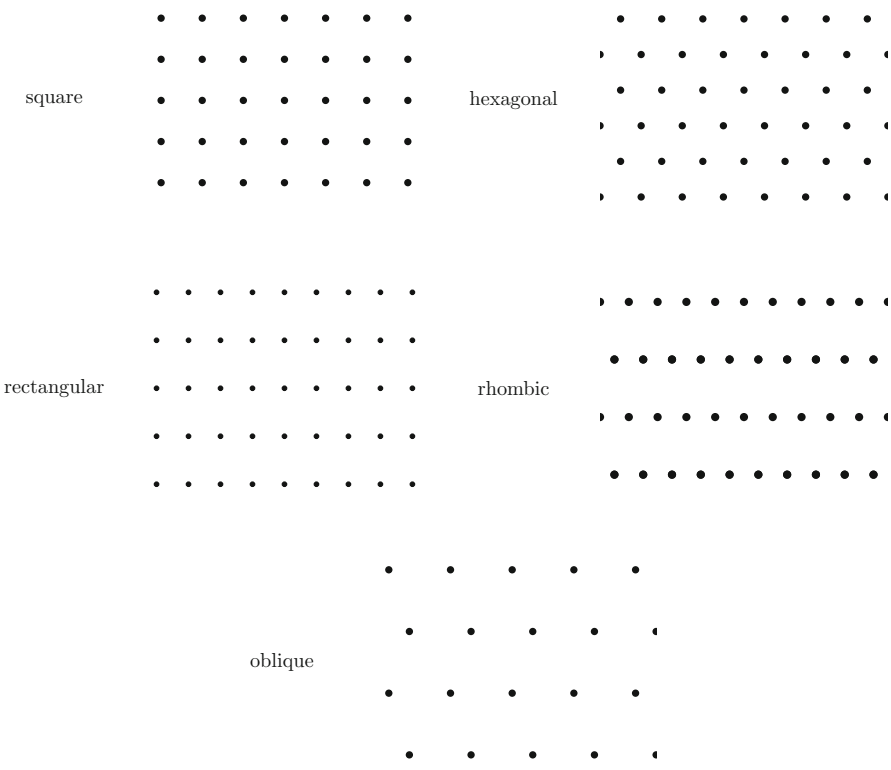


Fig. 11 Clippings of the 5 Bravais lattices in the plane

In even more modern approaches, the notion of Bravais lattice has become more abstract and gave rise to the definition of *Bravais groups*, see e.g., Souvignier (2003).

Bravais's writings influenced C. Jordan (Schwarzenberger 1984), who – as already described – introduced groups of movements, i.e., symmetry groups, in 1869. Jordan's work was very influential and soon became known in the crystallographic community. After determining all rotational (chiral) point groups in three-dimensional space, Jordan attempted to combine rotations and translations in order to determine all possible (chiral) groups of movements, but his enumeration was incomplete. Boisen and Gibbs (1976, p. 145) attribute the first group-theoretical classifications of crystal symmetry to Minnigerode (1884), Schoenflies (1891), and Weber (1896). However, an earlier usage of group theory in crystallography was due to German mathematician **Leonard Sohncke** (1842–1897), who first noted the incompleteness of Jordan's enumeration. In his major work *Entwickelung einer Theorie der Kristallstruktur* (1879), resulting from his studies from 1867 onward, Sohncke used group theory (as he calls it, *Jordan'sche Methode*) and obtained 66 chiral space groups in space, i.e., space groups in three-dimensional space containing only rigid motions: rotations, translations, and their compositions (Sohncke 1879). These are now known as *Sohncke groups*. He later found that he had counted one twice, so there are 65 of them (Burckhardt 1988; Van der Waerden 1985). Sohncke's classification was a herald of the soon-to-follow complete classification and enumeration of space groups in three dimensions, which was successfully completed independently by A. M. Schoenflies and E. S. Fyodorov.

The German mathematician **Arthur Moritz Schoenflies** (1853–1928), in his first works (1886/1887), re-proved many of the already mentioned results, including Sohncke's, but his proofs much more clearly emphasize the role of group theory. For Schoenflies, crystal classes are point groups (*Hilfsgruppen*), and he determined translations compatible with each of them. In the next years, he devoted himself to the determination of all *erweiterte Gruppen*, i.e., space groups which also include mirror-symmetries and were not considered by Jordan and Sohncke. By 1890, he determined the number of space groups in total to be, erroneously, 236. He also noted that each of these can be obtained by building up from one of the 32 *Hilfsgruppen*. Schoenflies's approach from these studies somewhat differed from the final one, as presented in his famous textbook *Kristallsysteme und Kristallstructur* (1891). He explicitly refered to Hessel, Bravais, and Sohncke in the preface, and after a quite complete introduction to the symmetry group theory of his time, he first gave the first fully group-theoretical deduction of all crystallographic point groups. Schoenflies divided them into groups of first kind (rotational groups) and groups of second kind (all the others, these he obtained from the groups of the first sort by systematically adding central and mirror symmetries and rotoinversions, i.e., compositions of rotations with central symmetry). He also proved that in each group of the second kind, the numbers of orientation-preserving and orientation-inverting elements are the same, and that the orientation-inverting elements can be obtained by composing just one of them with all the rotations in the group. All point groups are divided into seven crystal systems before, in the second part of his textbook, he systematically determined all the – now correctly – 230 three-dimensional space

groups, in each of the crystal systems: 2 in the triclinic system, 13 in the monoclinic system, 59 in the orthorhombic system, 68 in the tetragonal system, 25 in the trigonal system, 27 in the hexagonal system, and 26 in the cubic system. In the last chapter, Schoenflies finally proved the crystallographic restriction theorem (which was in the beginning just stated as a consequence of the law of rational indices) and the law of rational indices from the presupposed periodic structure of crystal, thus fully completing the arguments (Burckhardt 1967, 1988; Schoenflies 1891).

The Russian mineralogist and crystallographer **Yevgraf Stepanovich Fyodorov** (also often transliterated as Evgraf Fedorov, 1853–1919) was also familiar with Sohncke's classification of chiral space groups, and began his own researches in that direction already in 1881. During the 1880s, he developed his *Симметрия правилъных систем фигур* (*Simmetriya pravilnykh sistem figur*), which was published in 1890/1 with a follow-up containing corrections (Fyodorov 1890, 1891). His method arose from considerations of plane and space tessellations, and was developed independently of Schoenflies. However, from 1889 up to 1908, they communicated by letters. This correspondence was initiated by Schoenflies, who sent two of his publications to Fyodorov, but Fyodorov's 1890 classification of "regular systems of points" (periodic tilings) by symmetry was achieved independently and was more complete than Schoenflies's publication from the same year. Even if also based on group theory and departing from symmetry classes, his approach is different than Schoenflies's. Fyodorov classifies space groups as symmorphic, hemisymmorphic, and asymmorphic: "The systems themselves are classified into three groups: 1) The symmorphic systems are those whose elementary figures belong to the same kind of symmetry as the systems themselves. 2) The hemisymmorphic systems are to be understood as two analogous symmorphic systems, which together form a regular double (not symmorphic) system. 3) All the rest are asymmorphic systems. From these definitions follows that each figure in a symmorphic systems possesses a point in which all the symmetry elements of the system intersect, that in the analogous point of a figure in a hemisymmorphic system only elements of rigid symmetries (i.e. symmetry of the 1st kind) intersect, and that such points are missing in figures of asymmorphic sytems" (Fyodorov quote translated from its German translation: Burckhardt 1967, p. 235). By mid-October 1890, he arrived at the number of 228 "regular systems of points" (space groups), which he corrected in November 1890, after correspondence with Schoenflies, to 230 (Burckhardt 1967, 1971, 1988; Nespolo et al. 2018).

Thus, in 1891, the first two independent, fully group theoretical classifications of symmetry groups were published:

Theorem 4 (Fyodorov-Schoenflies) The number of crystallographic space groups in three-dimensional space is 230.

Some authors add the British crystallographer **William Barlow** (1845–1934) to the list of three originators of Theorem 4, since in 1894 he published his own (predominantly geometric) proof of it (Barlow 1894). In response to criticism, he responded that he had no access to Fydorov's and Schoenflies's papers (Burckhardt

1967; Lalena 2007; Whitlock 1934). By the beginning of the twentieth century, group theory became generally used in studies on symmetries of crystals, as is exemplified, e.g., by the textbook *Mathematical Crystallography and the Theory of Groups of Movements* (Hilton 1963), originally published in 1903, which more or less follows the exposition of (Schoenflies 1891). In fact, for a long time, it was the only book presenting Schoenflies's, Fydorov's, and Barlow's theory in English, since Schoenflies and Barlow published their results in German, and Fyodorov in Russian (Lalena 2007).

Finally, it should be noted that in literature, the number of space groups is often given as 219, and not as 230. Namely, from the point of view of abstract algebra, space groups describing two enantiomorphic periodic objects have the same group structure, and even their elements are the same (possibly differing only in the direction of a screw rotation). This approach, in which two space groups are equivalent if they are isomorphic, results in only 219 types of three-dimensional space groups, named *affine* (*classes of*) *space groups* (Nespolo et al. 2018). Of those 219, 11 are chiral (can appear in two enantiomorphic "copies"), and since in the physical world (including crystallography, chemistry and physics, and even art) the handedness is a relevant characteristic, it is common to state the number as 230 in most texts dealing with the history or application of space group symmetry.

4 Space Groups in Lower and Higher Dimensions

After such an intense period of studies of three-dimensional symmetry groups, it is no wonder that ideas of generalizations appeared. As noted in the section on point groups, toward the end of the nineteenth century, mathematicians – namely, Goursat – first considered symmetries in four-dimensional space. Among the famous 23 problems posed in 1900 by **David Hilbert** (1862–1943), one is directly related to generalization of symmetry group theory to other dimensions: In the first part of his 18th problem, Hilbert comments that Fyodorov and Schoenflies have shown that there are only finitely many symmetry groups with a "finite fundamental region" (i.e., crystallographic space groups) for the three-dimensional Euclidean space and asks if the same is true for every n-dimensional Euclidean space (Hilbert 1900). That the answer is "yes" was proven 10 years later by the German mathematician **Ludwig Bieberbach** (1886–1982). Bieberbach first proved that for a symmetry group, having a "finite domain" (which in modern terms would be a compact fundamental domain, i.e., a unit cell as described above) is equivalent to not containing "infinitesimal operations" (i.e., to being discrete). Further, he proved various other results, in particular that n-dimensional crystallographic groups contain translations in n linearly independent directions. Finally, he showed that for every finite dimension n, there exist only finitely many (up to isomorphism) different discrete subgroups of $E(n)$, the group of all isometries of the n-dimensional Euclidean space (Bieberbach 1910). This theorem is the reason that space groups for spaces of dimensions higher than 3 are sometimes referred to as *Bieberbach groups*; more commonly, this term is reserved only for torsion-free crystallographic groups (Cid and Shulz 2001). A year

later, another German mathematician, **Ferdinand Georg Frobenius** (1849–1917), noted that instead of proving the finiteness of the classes up to isomorphism, it would be more adequate to classify crystallographic groups according to "affine equivalence" and proved the corresponding variant of Bieberbach's theorem. Again, a year later, Bieberbach showed that for crystallographic groups, "affine equivalence" is the same as isomorphism and thus finalized the proof of the following theorem (Burckhardt 1988; Vince 1997; Zassenhaus 1948):

Theorem 5 (Bieberbach-Frobenius' Theorem) The number of classes of affine equivalence of crystallographic space groups in $E(n)$ is finite for every positive integer n.

In particular, Bieberbach's and Frobenius's results showed that crystallographic space groups G contain a normal, torsion-free, maximal Abelian subgroup T (the translational subgroup) of rank n and finite index, and that the corresponding point group $P = G/T$ is isomorphic to a finite subgroup of GL(n, $\mathbf{Z}$). On the basis of this insight, H. Zassenhaus in 1948 developed a fully algebraic group-theoretical algorithm for classification of space groups in arbitrary dimensions, more precisely, an algorithm for calculating the set of representatives of their isomorphism classes (Cid and Schulz 2001; Streitwolf 1967; Zassenhaus 1948).

Zassenhaus exemplified his algorithm by showing that there are 17 space groups in dimension 2, a result proven before by geometrical methods. Namely, in the 1920s, space groups with one- and two-dimensional periodicity also caught the interest of mathematicians. The two-dimensional analog of crystal structures, i.e., two-dimensional periodic objects in the plane, are known as wallpapers, and the corresponding two-dimensional space groups are known as *wallpaper groups*. Even if the depiction of (clippings of) wallpapers is common in the long history of art, their classification according to symmetry was only proven along with its three-dimensional counterpart. Their symmetries were first studied by Sohncke in 1874, who erroneously determined that the total number of wallpaper groups was 13 (Burckhardt 1988; Sohncke 1874). Fyodorov, along with his classification of three-dimensional space groups, published the correct classification of 17 wallpaper groups (1891), but this remained mostly unknown for some time. Thus, the first classification is usually attributed to **György (George) Pólya** (1887–1985), who published an independent proof in 1924 which soon became widely known (Burckhardt 1988; Pólya 1924; Speiser 1927).

Toward the end of his article, Pólya writes (Pólya 1924, p. 282): "A particularly easy problem is the following: to count all possible symmetry structures of a bordure (ribbon, frieze); there are seven of them."[5] These are now known as seven *frieze groups* (see Fig. 12), with frieze being the standard term for subperiodic patterns in the plane, i.e., objects in Euclidean space of dimension 2 possessing translational

[5] *Eine besonders einfache Aufgabe ist die folgende: es sind alle verschiedenen Strukturen der Symmetrie einer Bordüre (Band, Fries) aufzuzählen; es gibt deren sieben.*

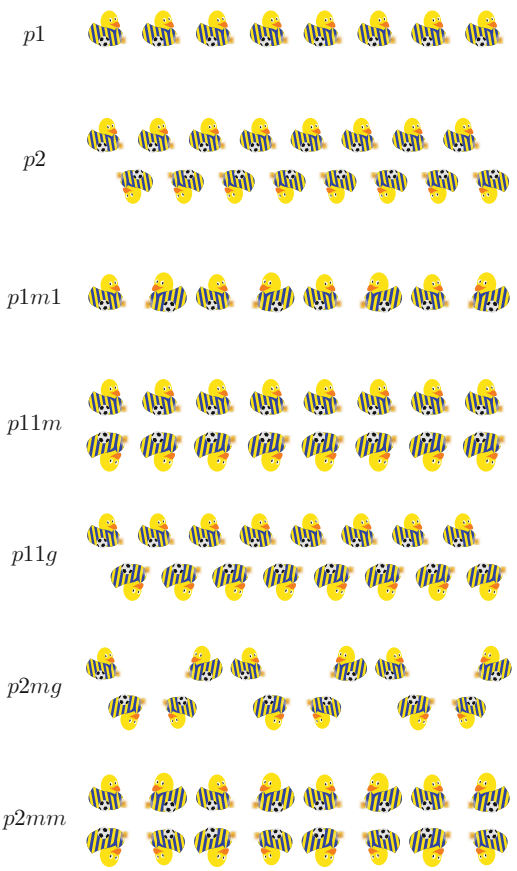


Fig. 12 Clippings from examples of seven frieze classes

symmetry in only one direction. A full proof of this fact was given by Swiss crystallographer **Paul Niggli** (1888–1953) in the same issue of the *Zeitschrift für Kristallographie* (Niggli 1924). According to (Burckhardt 1988, p. 141), in the period 1920–1923, A. M. Günzburg independently proved that there are seven frieze groups and 17 wallpaper groups, but he only published his results in 1929. Last, but not least, it is worth mentioning that Conway et al. (2008) arrive at the seven frieze groups as spherical symmetry groups by wrapping the friezes around a sphere.

To summarize the last two paragraphs:

Theorem 6 (Fyodorov-Pólya-Niggli) There exist 17 fully periodic and seven subperiodic space groups in dimension 2.

The three-dimensional equivalent of frieze groups, i.e., subperiodic space groups of objects in three-dimensional space with translational symmetry in only one direction, are known as *rod groups*. They were first, incompletely, studied in one

of the first influential textbooks on (finite) group theory, *Die Theorie der Gruppen von endlicher Ordnung, mit Anwendungen auf algebraische Zahlen und Gleichungen sowie auf die Kristallographie,* by the Swiss mathematician **Andreas Speiser** (1885–1970), which appeared in four editions. The sixth and seventh chapters in its second edition (Speiser 1927) are devoted to "Symmetries of ornaments" and "Crystal classes" and contain quite modern presentations of the mentioned results, and more. Besides giving the classification of frieze and wallpaper groups, Speiser considers friezes as reliefs, and by differentiating between mirror-symmetric copies with respect to the plane of the frieze, he arrived at an additional 24 space groups for "two-sided friezes," i.e., in total 31 frieze groups – 7 "one-sided" plus 24 "two-sided" (Burckhardt 1988; Speiser 1927). The full classification of rod groups was achieved in 1929, without reference to Speiser, but with reference to Niggli, by the German physicist and crystallographer **Carl Heinrich Hermann** (1898–1961): He showed that the number of rod groups, which he called chain groups (*Kettengruppen*), is 75 (Hermann 1929).

In general, space groups containing translations in only one direction (space groups of mono-periodic objects) are known as *line groups* (Damjanović and Milošević 2010). Even if, since antiquity, ornaments belonging to the class of friezes and all seven types of frieze symmetry were present in art (Jablan 2002), as described above, their mathematical classification did not happen until the twentieth century. Also, it is quite obvious that periodic "ornaments" on a line either have no other symmetry besides translation (e.g., — - - - - — - - - - — - - - -) or additionally possess mirror symmetries (e.g., – – – –), i.e., there are only two one-dimensional space (line) groups. As in the case of point groups in dimension 1, this seems to have been considered too trivial to merit a specific proof, and when mentioned in literature, it is stated without a specific reference, see e.g., Schwarzenberger (1984).

In three-dimensional space, one can also consider subperiodic space groups of biperiodic objects, i.e., of objects possessing translational symmetry in two (but not three) linearly independent directions. These are known as *layer groups* and were first classified in 1929 along with rod groups by C. H. Hermann, who called them net groups (*Netzgruppen*) and showed that there are 80 of them by systematically combining elements of symmetry. In the same year, the result was independently obtained by the Swiss crystallographer **Leonhard Weber** (1883–1968), who deduced them as special cases of Niggli's 1919 classification of space groups, which was based on the division of space groups into 73 *arithmetic crystal classes* (these are classes corresponding to the Fyodorov's 73 symmorphic space group types, usually defined as equivalence classes of ordered pairs of translational and point groups of periodic objects). Weber also noted that 17 of the 80 layer groups correspond to cases when the plane containing the pattern is itself a mirror-plane, i.e., there are 17 layer groups that correspond to "transparent" patterns, analogously to Speiser's "two-sided" friezes (Burckhardt 1988; Hermann 1929; Alexander and Hermann 1929; Heesch 1929; Nespolo et al. 2018; Pabst 1962; Schwarzenberger 1984; Weber 1929).

The study of space groups was also extended to higher dimensions. The above-mentioned Bieberbach's and Frobenius's results from the 1910s were the first step

toward a classification of n-dimensional crystallographic groups. A second important step is a suitable generalization of the crystallographic restriction, but that needed a few more decades. In general, for spaces of dimension n, the crystallographic restriction, i.e., the set of possible orders of rotational symmetry of an n-dimensional lattice, can be formulated as follows (Bamberg et al. 2003):

Definition 9 The *general crystallographic restriction* for dimension n (CR_n) is the set of all positive integers m for which an invertible matrix A of order with integer entries exists such that $A^m = I_n$ (the identity matrix of order n) and m is the smallest such integer.

The group-theoretical background of the definition is the following. For an element g in a group $(G,\cdot)$, the order of g, Ord(g), is defined as the smallest positive integer m such than $g^m = e$ (the identity element in the group), if such m exists (otherwise $\mathrm{Ord}(g) = \infty$ by definition). Thus:

$$CR_n = \{m \in \mathbf{N} : \exists A \in GL(n, \mathbf{Z})\ Ord(A) = m\}.$$

The classical crystallographic restriction can now be written in the form $CR_2 = CR_3 = \{1, 2, 3, 4, 6\}$. Also, quite obviously $CR_1 = \{1, 2\}$. The generalized crystallographic restriction was first studied in (Hermann 1949), even the first study of orders of invertible integer matrices was given in (Vaidyanathaswamy 1928). By extension of the number-theoretical method described by H. Hiller (Hiller 1985), who is one of several authors who independently discovered CR_n, one gets an efficient algorithm to compute $CR_n \setminus CR_{n-1}$, the additional possible orders when increasing the dimension for 1, for any n (Bamberg et al. 2003). Thus, it is known how to determine the values of CR_n for arbitrary n, and Table 2 contains the values for dimensions up to 11.

Even if such studies seem to be of purely abstract, mathematical interest, it has proven not to be so. The discovery of quasi-periodic crystals (usually referred to simply as *quasicrystals*) by Dan Schechtman and his colleagues in 1982, displaying many properties similar to crystals, but exhibiting "forbidden" axes of symmetry (e.g., of order 5) resulted in studies of possible models (Lifshitz 1996; Schechtmann et al. 1984). In the cases known to date, the exhibited rotational symmetry in

Table 2 The crystallographic restriction values for dimensions 1 to 11

N	CR_n
1	1, 2
2, 3	1, 2, 3, 4, 6
4, 5	1, 2, 3, 4, 5, 6, 8, 10, 12
6, 7	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 18, 20, 24, 30
8, 9	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 24, 28, 30, 36, 40, 42, 60
10, 11	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 21, 22, 24, 28, 30, 35, 36, 40, 42, 45, 48, 56, 60, 70, 72, 84, 90, 120

quasicrystals is either consistent with CR_3, or of one of the orders 5, 8, 10, or 12, consistent with CR_4 (Steurer 2018). A particularly often-used model is to describe quasiperiodic structures as projections of periodic ones from higher-dimensional spaces to dimension 3. Thus, for the description of currently known quasicrystals, space groups in spaces of higher dimensions than 6 do not seem necessary, but if, for example, a quasicrystal with rotational symmetry of order 11 would be discovered, the corresponding periodic model would (see Table 2) require a space of dimension at least 10 (Souvignier 2006; Stilinović and Brückler 2012).

In 1949, Hermann studied the possible elements of space groups for an arbitrary dimension (Hermann 1949), and by the turn of the centuries, a variety of results for dimensions up to 6 were proven (Janssen et al. 1999; Janssen et al. 2002). The final enumeration of four-dimensional space groups was achieved in the 1970s (with a small error) by H. Brown, R. Bulow, J. Neubüser, H. Wondratschek, and H. Zassenhaus. In that enumeration, one of the 4894 four-dimensional space groups was doubly counted, and this error was corrected in 2002. The error was related to the already mentioned distinction between chiral and achiral symmetry groups: Just as in three dimensions 11 of the 230 space groups appear as two enantiomorphic pairs each, and the total count is 219 if one does not distinguish enantiomorphic pairs, in dimension 4 there are 111 chiral groups, i.e., the total count is 4783 if one does not distinguish enantiomorphic pairs (Brown et al. 1978; Conway and Smith 2003; Neubüser et al. 2002).

The direct approach based on CR_n, however, seems unfeasible for higher dimensions since the number of space groups increases rapidly. To this effect, the original Zassenhaus algorithm was further refined, and then implemented by J. Opgenorth, W. Plesken, T. Schulz, and others as the CARAT (Crystallographic Algorithms and Tables) package in 2000 (Cid and Schulz 2001; Lehrstuhl B für Mathematik Aachen 2023; Opgenorth et al. 1998; Plesken and Schulz 2000). Using it, Plesken and Schulz arrived at 222,018 space groups in dimension 5, and 28,927,922 in dimension 6 (both not counting enantiomorphic pairs as different); the number for dimension 6 has later been corrected to 28,927,915 (Lehrstuhl B für Mathematik Aachen 2023). The distinction between enantiomorphic pairs was published by Souvignier in 2003, who arrived – correctly – at 79 chiral among the 222,018 five-dimensional space groups (making the total 222,097), and – erroneously due to old data from CARAT – at 7052 in dimension 6 (Souvignier 2003). The correct number of chiral space groups in dimension 6 is apparently still unknown at the present time, as are the numbers of space groups in higher dimensions.

5 Color Symmetry Groups

Besides the more abstract types of symmetries and symmetry groups in non-Euclidean spaces, which are outside the scope of this chapter, another type of Euclidean symmetry groups has been studied extensively during the last 100 years. Their study dates back to the papers written in 1929 and 1930 by German mathematician **Heinrich Heesch**

(1906–1985). His research began by investigating layer groups, inspired by Speiser's results, and he noted that the "two-sided" ("transparent") friezes and wallpapers can be interpreted as "normal" ones, but allowing coloring in two colors. Motivated by the idea of generalizations to higher-dimensional spaces, he writes (Heesch 1930a, p. 326): "This is the problem of extending the 17 Niggli's two-dimensional plane groups to the 80 three-dimensional ones, just as it is here about building up from the 230 three-dimensional R3 groups into four dimension."[6] Heesch begins with explaining that the step from 17 wallpaper to 80 layer groups can be achieved if we add a third coordinate to the plane of the ornament, but only use the sign (plus or minus) of that coordinate to discern between the mirror-symmetric copies, and comments that this can be viewed as being colored in two colors. He refers to Weber for the coloring representation, and to Speiser and Hermann for other ornamental representations of the additional coordinate. Furthermore, Heesch added a possible physical interpretation with positive and negative ions arranged in lattice points, thus essentially prophetically announcing future applications. Heesch successfully enumerated all three-dimensional "black-and-white" symmetry classes and arrived at now 122 "black-and-white" point groups in space, divided into the 7 usual crystal systems. These were rediscovered as part of Hurley's derivation of all crystallographic point groups in dimension 4 (Burckhardt 1988; Heesch 1930a; Hurley 1951; MacKay and Pawley 1963; Schwarzenberger 1984; Wills 2017; Zamorzaev 1988). In the follow-up paper (Heesch 1930b), Heesch also considered continuous and semidiscontinuous (meaning, continuous in some, discrete in other directions) symmetries, referring to the mentioned paper by C. Jordan and the paper (Shubnikov 1930), in which the future "father" of anti-symmetry, A. V. Shubnikov, first illustrated Speiser's 31 two-sided frieze groups. Belov later noted that the derivation of these groups becomes easier if instead of being monocolored and two-sided, the friezes are considered to be one-sided and two-colored (Belov 1956; Pabst 1962).

At Heesch's time, the idea did not seem to have possible applications, but – as often with mathematical theories – it has proven otherwise two decades later. It was the Soviet mathematician **Alexei Vasilievich Shubnikov** (1887–1970) who in 1951 brought the "black-and-white" symmetry groups into prominence for crystallography in his monograph *Симметрия и антисимметрия конечных фигур* (*Simmetriya i antisimmetria konechnykh figur*) (Bradley and Cracknell 1972; Shubnikov 1951). He developed the concept of *anti-symmetry* "as a tool to understand the spatial anisotropy of a crystal's physical properties, following in the direction laid by Curie" (Wills 2017, p. 152). According to Shubnikov, for a three-dimensional periodic structure an additional, two-valued (black/white, +/−, parallel/anti-parallel magnetic moment, . . .) coordinate is considered in every point, and the anti-symmetry operation is defined as changing the value of this coordinate in all points. He re-derived Heesch's result and showed that there are 32 "colorless" point groups (i.e., the usual crystallographic ones), 32 "grey" point groups (here every

[6] *Es ist dies das Problem der Erweiterung der 17 Nigglischen zweidimensionalen Ebenengruppen zu den 80 dreidimensionalen, so, wie es sich jetzt darum handelt, die 230 dreidimensionalen R3-Gruppen zur Gesamtheit der vierdimensionalen auszubauen.*

symmetry operation occurs both with and without antisymmetry, i.e., they are invariant under antisymmetry), and 58 "black and white" point groups (where some anti-symmetry operations are combined with some symmetry operations), in total 122 (Wills 2017; Zamorzaev 1988). Later, Shubnikov also extended the Hessel's list of point groups in dimension 3 by considering "limiting symmetry groups," i.e., two-colored point groups of objects possessing rotational axes of infinite order. The colorless version is often called *Curie limiting groups*, as it was Pierre Curie who first investigated them (Curie 1894; Shubnikov 1988a). There are seven Curie limiting groups (Fig. 13), which arise as limits of the point groups from Hessel's theorem: C_∞ (the symmetry group of a circular cone rotating in one direction), D_∞ (the symmetry group of a twisted circular cylinder), $C_{\infty h}$ (the symmetry group of a rotating cylinder), $C_{\infty v}$ (the symmetry group of a cone at rest), $D_{\infty h}$ (symmetry group of a cylinder at rest), $K_h = O$ (the symmetry group of a sphere), and $K = SO$ (symmetry group of a sphere without planes of reflection, i.e., such that all its diameters are twisted in the same direction) (Artin 1991; Curie 1894; Shubnikov 1988a). By adding the antisymmetry feature, Shubnikov arrived in 1988 to 14 *limiting antisymmetry point groups*, more precisely to seven gray and seven black-and-white ones. Together with the seven Curie groups, it makes 21 two-colored limiting point groups (Krishnamurty and Gopalakrishnamurty 1969; Palistrant 2005; Shubnikov 1988a, 1988b).

Shubnikov's (and Heesch's) ideas were extended to space groups by Soviet mathematician, **Alexander Mihailovich Zamorzaev** (1927–1997), in his PhD thesis in 1953, who found in total 1651 "generalized Fyodorov groups," divided into four classes – colorless, gray, black and white point groups with colorless Bravais lattices, and black and white point groups with black and white Bravais lattices (Burckhardt 1988; Litvin 2016; Wills 2017). They were also enumerated by Belov, Neronova, and

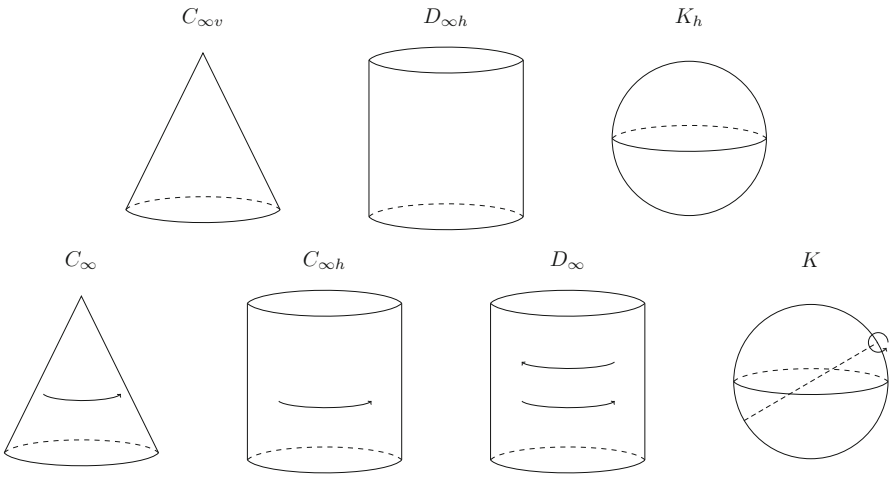


Fig. 13 Examples of solids belonging to symmetry classes of the seven Curie limiting groups (arrows indicate directions of rotations or twists)

Smirnova in 1957 and by Opechowski and Guccione in 1965. These groups became known as *Shubnikov groups* (so they were called, e.g., by Zamorzaev) or *magnetic space groups*, but also as *Heesch groups*, *Opechowski-Guccione groups, dichromatic, 2-color* or *black-and-white groups*. Shubnikov, on the other hand, called them *type III groups* (Ш is the Russian letter with pronunciation "sh"), and as Ш was often typeset as III in later works, the three subtypes became known as type I (colorless), type II (gray), and type III (black-and-white) groups (Bradley and Cracknell 1972; Lifshitz 2005; Wills 2017). The term magnetic groups was used by B. A. Tavger and V. M. Zaitsev, who described the magnetic structure of MnO, using antisymmetry as an operation that reverses atomic magnetic moments or electron spins (Wills 2017). Tavger and Zaitsev also tabulated the magnetic point groups, while the 1651 magnetic space groups (230 colorless, 230 gray, and 1191 black and white) were derived by a different method than Zamorzaev's by Belov, Neronova, and Smirnova in 1955, who called them Shubnikov groups (Bradley and Cracknell 1972; Burckhardt 1988; Tavger and Zaitsev 1956; Senechal 1988; Wills 2017; Zamorzaev 1988). The study of magnetic space groups has more recently been extended to higher dimensions too, and related to quasicrystals (Lifshitz 2005).

Another development, this time directly from art, simultaneously inspired the interest in color symmetry groups. It was the work of Dutch artist **Maurits Cornelis Escher** (1898–1972), whose many depictions of periodic tessellations of the plane started to be subject to mathematical analysis at the time. In particular, his usage of color symmetry attracted attention during the 1960s, and its value as teaching aids became recognized in the crystallographic community too (Senechal 1983, 1988). Soon the ideas were generalized to more than two colors, to multiple antisymmetry, to *P*-symmetry, to spin groups (with spins viewed as classical axial vectors which are free to rotate in all directions), and to arbitrary dimensions.

The first ideas on extensions of the described antisymmetry to a more general, multiple one, were considered by Shubnikov already in 1944 (but not mentioned in his 1951 monograph). They were further developed by Zamorzaev and others in the 1970s, and "the theory of simple and multiple antisymmetry is actually completed, the only thing left is to develop new methodical techniques for the purpose of its further incorporation into natural sciences" (Zamorzaev 1988, p. 558), such as S. V. Jablan's method of generating groups of simple and multiple symmetry types using his "antisymmetry characteristic" (Jablan 1986). Also, in the 1970s, the idea of spin groups was studied by Litvin, Opechowski, and others (Lifshitz 2005), while Zamorzaev and his group in about the same time generalized both antisymmetry and color symmetry to *P*-symmetry (Jablan 1994; Lungu 2002).

One of the first attempts of considering more than two colors appeared already in 1956, when N. V. Belov and T. N. Tarhova published a paper on "Groups of color symmetry." The definition was fully clarified, and the subject was theoretically founded in 1961 in a short paper called *Farbgruppen* by B. L. van der Waerden and J. J. Burckhardt (Jablan 2007; Schwarzenberger 1984; Van der Waerden and Burckhardt 1961; Zamorzaev 1988). Their definition is essentially the following:

Definition 10 Let S be a discrete set of points (or more general motifs), where each point $x \in S$ has an additional property ("color") $f(x)$ from a finite set F, i.e., a *coloring* $f: S \to F$ is given. We speak of a *colored set* S_F. The *color symmetry group* of S_F is the group of all ordered pairs (T, P), where T is an element of the symmetry group G of S and P a permutation of F, such that $fT = Pf$.

Van der Waerden's and Burckhardt's paper became the basis of much of the further research in the field (Senechal 1983). By now, the research of color symmetry groups is a rich field of study. Besides the fully periodic color groups in various dimensions, subperiodic have been studied as well, and applications in the study of quasicrystals have been discovered too (Lifshitz 1997; Litvin 2005; Palistrant 2000; Palistrant and Jablan 1994; Souvignier 2006).

6 Conclusion

Even if the oldest evidence of human interest in symmetry are related to art, and even if Euclidean geometry was already highly developed in Ancient Greece, the mathematical study of symmetrical patterns arose less from artistic interests or purely mathematical considerations: It was in most part inspired by the development of crystallography at the end of the eighteenth and the beginning of the nineteenth centuries. When the initially purely geometric approaches were integrated with the simultaneously, but independently, developed branch of group theory, the study of symmetry groups emerged. At first serving the purpose of classifying crystals according to symmetry, at the beginning of the twentieth century, it began a life of its own. Many aspects of this were described in the text above, others were omitted.

Of the omitted ones, two deserve a special mention. The scope of this exposition was limited to symmetry in Euclidean spaces, but symmetry group research was also, in the second half of the twentieth century, extended to non-Euclidean spaces (Wilkie 1966). Also, one of the most important applications of symmetry groups is in theoretical physics. Its beginnings date to Pierre Curie and his "symmetry principle," introduced in 1894 and related to the mentioned discovery of limiting groups. Curie says: "The elements of the symmetry of the causes are to be found in the effects they produce"[7] (Curie 1894, p. 394). A crucial moment regarding applications of symmetry group theory in physics was Emmy Noether's first theorem (1915/1918), which established the connection between symmetries of physical systems and the conservation laws. Readers interested in the history of this line of impact of symmetry group theory are kindly referred to (Kosmann-Schwarzbach 2018).

In the beginning, there were friezes (on temples and otherwise), which exemplify human age-old fascination by symmetry. Now, at the end of this time-travel through the history of Euclidean symmetry groups, there are quasicrystals, for which many facts – on symmetry and more – are yet to be discovered.

[7] *Les éléments de symétrie des causes doivent se retrouver dans les effets produits.*

7 Cross-References

► Leonhard Euler Between Mathematics and Natural Philosophy: An Introduction to Natural Science Anleitung zur Naturlehre

References

Alexander E, Hermann K (1929) XIII. Die 80 zweidimensionalen Raumgruppen. Z Krist 70: 328–345
Anglin WS, Lambek J (1995) The heritage of Thales. Springer, New York
Artin M (1991) Algebra. Prentice Hall, Upper Saddle River
Bamberg J, Cairns G, Kilminster D (2003) The crystallographic restriction, permutations, and Goldbach's conjecture. Am Math Monthly 110:202–209. https://doi.org/10.1080/00029890.2003.11919956
Barlow W (1894) Über die Geometrischen Eigenschaften homogener starrer Strukturen und ihre Anwendung auf Krystalle. Zeitschrift für Krystallographie und Minerologie 23:1–63
Barlow W (1901) I. Crystal symmetry. The actual basis of the thirty-two classes. Philos Mag 1:1–36
Belcastro SM, Yackel C (2011) Crafting by concepts. AK Peters, Natick
Belov NV (1956) On one-dimensional infinite crystallographic groups. Kristallografiya 1:474–476
Bieberbach L (1910) Über die Bewegungsgruppen des *n*-dimensionalen euclidischen Raumes mit einem endlichen Fundamentalbereich. Gött Nachr:75–84
Boisen MB, Gibbs GV (1976) A derivation of the 32 crystallographic point groups using elementary group theory. Am Min 61:145–165
Bradley CJ, Cracknell AP (1972) The mathematical theory of symmetry in solids. Clarendon Press, Oxford
Bravais A (1849a) Note sur les polyèdres symétriques de la géométrie. J de Mathematique 14: 137–140
Bravais A (1849b) Mémoire sur les polyèdres de forme symétrique. J de Mathematique 14:141–180
Bravais A (1850) Mémoire sur les systèmes formés par des points distribués régulièrement sur un plan ou dans l'espace. J École Polytechnique 19:1–128
Brown H, Bulow R, Neubüser J, Wondratschek H, Zassenhaus H (1978) Crystallographic groups of four-dimensional space. Wiley, New York
Brückler FM (2017) Geschichte der Mathematik kompakt – Das Wichtigste aus Arithmetik, Geometrie, Algebra, Zahlentheorie und Logik. Springer Spektrum, Berlin
Brückler FM (2019) Mathematische Grundlagen der Kristallographie. Springer Spektrum, Berlin
Burckhardt JJ (1967) Zur Geschichte der Entdeckung der 230 Raumgruppen. Arch Hist Exact Sci 4: 235–246
Burckhardt JJ (1971) Der Briefwechsel von E. S. von Fedorow und A. Schoenflies, 1889–1908. Archive Hist Exact Sci 7:91–141
Burckhardt JJ (1988) Die Symmetrie der Kristalle. Springer AG, Basel
Burke JG (1966) Origins of the science of crystals. University of California Press, Berkeley and Los Angeles
Capeller MA (1723) Prodromus crystallographiae de crystallis improprie sic dictis commentarium. Typis Henrici Rennwardi Wyssing, Lucerne
Cid C, Schulz T (2001) Computation of five- and six-dimensional Bieberbach groups. Exp Math 10: 109–115
Conway JH, Burgiel H, Goodman-Strauss C (2008) The symmetries of things. AK Peters, Wellesley
Conway JH, Smith DA (2003) On quaternions and octonions: their geometry, arithmetic, and symmetry. AK Peters, Natick
Coxeter HSM (1934) Discrete groups generated by reflections. Ann Math 35:588–621

Coxeter HSM (1935) The complete enumeration of finite groups of the form $r_i^2 = \left(r_i r_j\right)^{k_{ij}} = 1$. J London Math Soc 10:21–25
Coxeter HSM (1948) Regular polytopes. Methuen, London
Coxeter HSM (1969) Introduction to geometry. John Wiley & Sons, New York
Cromwell PR (1997) Polyhedra. Cambridge Univ Press, Cambridge
Curie MP (1894) Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique. J de Phys III:393–415
Damjanović M, Milošević I (2010) Line groups in physics. Springer-Verlag, Berlin
Downward M (2011) Crystallographic point groups in five dimensions. Z Kristallogr 226:32–39
Fyodorov YS (1890) Simmetriya pravilnykh sistem figur. Tipografiia A. Iakovsona, St. Petersburg
Fyodorov YS (1891) Simmetriya pravilnykh sistem figur. Zapiski St Petersburg Mineral Obsc 2nd series 28:1–146
Frankenheim ML (1826) Crystallonomische Aufsätze. Isis(Jena) 19:497–515. 542–565
Goodman R (2004) Alice through looking glass after looking glass: the mathematics of mirrors and kaleidoscopes. Am Math Mon 111:281–298
Goursat ME (1889) Sur les substitutions orthogonales et les divisions régulières de l'espace. Annales scientifiques de l'ÉNS 6:9–102
Grove LC (1997) Grups and characters. John Wiley & Sons Inc, New York
Grünbaum B (2006) What symmetry groups are present in the Alhambra? Notices AMS 53: 670–673
Grünbaum B, Grünbaum Z, Shephard GC (1986) Symmetry in Moorish and other ornaments. Comput Math with Appl 12B:641–653
Grünbaum B, Shephard GC (1987) Tilings and patterns. WH Freeman, New York
Hahn T (ed) (2005) International tables for crystallography, Volume A, Space-group symmetry, 5th edn. Springer, Dordrecht
Haklay G, Gopher A (2020) Geometry and architectural planning at Göbekli Tepe, Turkey. Camb Archaeol J 30:343–357. https://doi.org/10.1017/S0959774319000660
Hargittai I, Lengyel G (1985) The seventeen two-dimensional space-group symmetries in Hungarian needlework. J Chem Educ 62:35–36
Haüy M (1815) Première suite du mémoire sur la loi de symétrie. Mem Mus Hist Nat I:206–225
Heesch H (1929) Zur Strukturtheorie der Ebenen Symmetriegruppen. Z Kristall 71:95–102
Heesch H (1930a) XIX. Zur systematischen Strukturtheorie. III. Über die vierdimensionalen Gruppen des dreidimensionalen Raumes. Z Krist 73:325–345
Heesch H (1930b) XX. Zur systematischen Strukturtheorie. IV. Über die Symmetrien zweiter Art in Kontinuen und Semidiskontinuen. Z Krist 73:346–356
Hermann C (1929) XVII. Zur systematischen Strukturtheorie. III. Ketten- und Netzgruppen. Z Krist 69:250–270
Hermann C (1949) Kristallographie in Räumen beliebiger Dimensionszahl. I. Die Symmetrieoperationen. Acta Cryst 2:139–145
Hessel JFC (1830) Krystall. In: Brandes HW, Gmelin L, Horner J, Muncke GW, Pfaff CH (eds) Johann Samuel Traugott Gehler's Physikalisches Wörterbuch, Abtheilung 2: I bis K, vol Band 5. EB Schwickert, Leipzig, pp 1023–1340
Hilbert D (1900) Mathematische Probleme. Nachr K Ges Wiss Göttingen. Math-Phys Klasse (Göttinger Nachrichten) 3:253–297
Hiller H (1985) The crystallographic restriction in higher dimensions. Acta Cryst A41:541–544
Hilton H (1963) Mathematical crystallography and the theory of groups of movements (unabridged and corrected republication of the 1903 original). Dover Publ Inc, New York
Hon G (2005) Kant vs. Legendre on symmetry: Mirror images in philosophy and mathematics. Centaurus 47:283–297. https://doi.org/10.1111/j.1600-0498.2005.00027.x
Hurley AC (1951) Finite rotation groups and crystal classes in four dimensions. Proc Camb Phil Soc 47:650–661
Hurley AC, Neubuser J, Wondratschek H (1967) Crystal classes of four-dimensional space R4. Acta Cryst 22:605

Jablan SV (1986) A new method of generating plane groups of simple and multiple antisymmetry. Acta Cryst A42:209–212
Jablan SV (1994) Proportional colored symmetry. Acta Cryst A50:403–405
Jablan SV (2002) Symmetry, ornament and modularity. World Scientific Publishing, Singapore
Jablan SV (2007) *Farbgruppen* and Their Place in the History of Colored Symmetry. In: Proccedings of the conference "Forma y Simetría: Arte y Ciencia Congreso", Buenos Aires, 2007
Janssen T, Birman JL, Koptsik VA, Senechal M, Weigel D, Yamamoto A, Abrahams SC, Hahn T (1999) Report of a subcommittee on the nomenclature of *n*-dimensional crystallography. I. Symbols for point-group transformations, families, systems and geometric crystal classes. Acta Cryst A55:761–782. https://doi.org/10.1107/S0108767398018418
Janssen T, Birman JL, Dénoyer F, Koptsik VA, Verger-Gaugry JL, Weigel D, Yamamoto A, Abrahams SC, Kopsky V (2002) Report of a subcommittee on the nomenclature of *n*-dimensional crystallography. II. Symbols for arithmetic crystal classes, Bravais classes and space groups. Acta Cryst A58:605–621. https://doi.org/10.1107/S010876730201379X
Jordan C (1869) Memoire sur les groupes de mouvements. Annali di Matematica 2(167–215):322–345
Jordan C (1870) Traité des substitutions et des équations algébriques. Gauthier-Villars, Paris
Kepler J (1611) Strena seu de Niue sexangula. Godefridum Tampach, Frankfurt am Main
Kepler J (1619) Harmonices mundi. Sumptibus Godofredi Tampachii Bibl Francof, Linz
Kleiner I (1986) The evolution of group theory: a brief survey. Math Mag 59:195–215
Kosmann-Schwarzbach Y (2018) The Noether theorems. Invariance and conservation laws in the twentieth century. Springer, New York
Krishnamurty TSG, Gopalakrishnamurty O (1969) Magnetic symmetry and limiting groups. Acta Cryst A25:333–334
Kubbinga H (2012) Crystallography from Haüy to Laue: controversies on the molecular and atomistic nature of solids. Z Kristallogr 227:1–26
Laigo GR, Gebreyohannes HM, Al Khamisi FMH (2014) Symmetry groups of Islamic patterns at the Sultan Qaboos Grand Mosque. In: Greenfield G, Hart G, Sarhangi R (eds) Bridges Seoul conference proceedings 2014. Tessellations Publ, Phoenix (AZ), pp 183–190. Retrieved from https://archive.bridgesmathart.org/2014/
Lalena JN (2006) From quartz to quasicrystals: probing nature's geometric patterns in crystalline substances. Cryst Rev 12:125–180
Lehrstuhl B für Mathematik Aachen (2023) The CARAT Homepage. https://lbfm-rwth.github.io/carat/doc/index.html. Accessed 20 Mar 2023
Lifshitz R (1996) The symmetry of quasiperiodic crystals. Physica A 232:633–647
Lifshitz R (1997) Theory of color symmetry for periodic and quasiperiodic crystals. Rev Mod Phys 69:1181–1218
Lifshitz R (2005) Magnetic point groups and space groups. In: Bassani F, Liedl GL, Wyder P (eds) Encyclopedia of condensed matter physics. Elsevier, pp 219–226. https://doi.org/10.1016/B0-12-369401-9/00748-8
Lima-de-Faria J (ed) (1990) Historical atlas of crystallography. International Union of Crystallography and Kluwer Academic Publishers, Dordrecht, London and New York
Litvin DB (2005) Tables of properties of magnetic subperiodic groups. Acta Cryst A61:382–385
Litvin DB (2016) Magnetic subperiodic groups and magnetic space groups. In: Arroyo MI (ed) International tables for crystallography, volume a, space-group symmetry (2nd online edition). Retrieved from: https://it.iucr.org/Ac/
Lungu A (2002) Zamorzaev's *P*-symmetry and its further generalizations. In: Communications from the conference: International seminar on discrete geometry (dedicated to the 75th birthday of Professor AM Zamorzaev) 2002, Chisinau, p 45–53. Retrieved from https://www.researchgate.net/publication/341966106_ZAMORZAEV'S_P_-SYMMETRY_AND_ITS_FURTHER_GENERALIZATIONS_1. Accessed 25 Mar 2023
MacKay AL, Pawley GS (1962) Bravais lattices in four-dimensional space. Acta Cryst 16:11–19

Molčanov K, Stilinović V (2014) Chemical crystallography before X-ray diffraction. Angew Chem Int Ed 53:638–652
Nelson SA (2011) Crystal form, zones, crystal habit. In: Earth & environmental sciences 2110 – Mineralogy. https://www2.tulane.edu/~sanelson/eens211/forms_zones_habit.htm. Accessed 22 Mar 2023
Nespolo M, Aroyo MI, Souvignier B (2018) Crystallographic shelves: space-group hierarchy explained. J Appl Crys 51:1481–1491
Neubüser J, Souvignier B, Wondratschek H (2002) Corrections to *crystallographic groups of four-dimensional space* by Brown *et al.* 1978. [New York: Wiley and Sons]. Acta Cryst A58:301
Niggli P (1924) Die regelmäßige Punktverteilung längs einer Geraden in einer Ebene (Symmetrie von Bordürenmuster). Z Kristall 60:283–298
Opgenorth J, Plesken W, Schulz T (1998) Crystallographic algorithms and tables. Acta Cryst A54: 517–531. https://doi.org/10.1107/S010876739701547X
Pabst A (1962) The 179 two-sided, two-colored, band groups and their relations. Z Krist 117: 128–134
Palistrant AF (2000) On the rosette, tablet, and hypertablet groups of *P*-symmetries and their relation to the groups of multidimensional symmetries. Cryst Rep 45:887–893
Palistrant AF (2005) On a number of some categories of the limiting groups of multidimensional plane crystallographic point groups of symmetry. Cryst Reports 50:7–15
Palistrant AF, Jablan SV (1994) Limiting groups of multidimensional plane-point groups. FILOMAT 8:41–55
Plesken W (1981) Bravais groups in low dimensions. Match 10:97–119
Plesken W, Schulz T (2000) Counting crystallographic groups in low dimensions. Exp Math 9: 407–411
Pólya G (1924) Über die Analogie der Kristallsymmetrie in der Ebene. Z Kristall 60:278–282
Reynolds M (2008) The octagon in Leonardo's drawings. Nexus Netw J 10:51–76. https://doi.org/10.1007/s00004-007-0056-8
Rogers AF (1912) The validity of the law of rational indices, and the analogy between the fundamental laws of chemistry and crystallography. Proc Am Phil Soc 51:103–117
Romé de l'Isle J-BL (1783) Cristallographie. Chez Didot jeune, Paris
Rønning F (2009) Islamic patterns and symmetry groups. In: Ernest P (ed) Philosophy of mathematics education journal 24 – Special issue on mathematics and art. https://education.exeter.ac.uk/research/centres/stem/publications/pmej/pome24/index.htm
Ryan PY (1986) Euclidean and non-Euclidean geometry: an analytic approach. Cambridge Univ Press, Cambridge
Saurel P (1911) On the classification of crystals. Bull Amer Math Soc 17:389–397
Schechtmann D, Blech I, Gratias D, Cahn JW (1984) Mettalic phase with long-range orientational order and no translational symmetry. Phys Rev Lett 53:1951–1954
Schilling K (2019) Kristall2000 (ver. 09-19-21). Speyer
Schoenflies A (1891) Kristallsysteme und Kristallstructur. BG Teubner, Leipzig
Scholz E (1989) The rise of symmetry concepts in the atomistic and dynamistic schools of crystallography, 1815–1830. Revue d'histoire des sciences 42:109–122
Schuh CP (2007) Mineralogy & crystallography. On the history of these sciences from beginnings through 1919. Tucson (Arizona). Retrieved from https://archive.org/details/History_Mineralogy_2007/mode/2up
Schwarzenberger RL (1984) Colour symmetry. Bull Lond Math Soc 16:209–240
Senechal M (1983) Coloring symmetrical objects symmetrically. Math Mag 56:3–16
Senechal M (1988) Color symmetry. Comput Math Appl 16:545–553
Shafranovskii II (1975) 14.10. Kepler's crystallographic ideas and his tract "the six-cornered snowflake". Vistas Astron 18:861–876. https://doi.org/10.1016/0083-6656(75)90181-6
Shubnikov AV (1930) Über die Symmetrie des Semikontinuums. Z Krist 73:430–433
Shubnikov AV (1951) Simmetriya i antisimmetria konechnykh figure. AN SSSR, Moscow

Shubnikov AV (1988a) On the works of Pierre Curie on symmetry. Comput Math Applic 16: 357–364
Shubnikov AV (1988b) Antisymmetry of textures. Comput Math Appl 16:373–377
Sibley TQ (2015) Thinking geometrically. MAA, Washington DC
Sohncke L (1874) 1874. Die regelmässigen ebenen Punktsysteme von unbegrenzter Ausdehnung. J Reine Angew Math 77:47–101
Sohncke L (1879) Entwickelung einer Theorie der Krystallstruktur. B G Teubner, Leipzig
Souvignier B (2003) Enantiomorphism of crystallographic groups in higher dimensions with results in dimensions up to 6. Acta Cryst A59:210–220
Souvignier B (2006) The four-dimensional magnetic point and space groups. Z Kristallogr 221: 77–82
Speiser A (1927) Die Theorie der Gruppen von endlicher Ordnung, mit Anwendungen auf algebraische Zahlen und Gleichungen sowie auf die Kristallographie, 2. Auflage edn. Springer Verlag, Berlin
Steurer W (2018) Quasicrystals: what do we know? What do we want to know? What can we know? Acta Cryst A74:1–11
Stilinović V, Brückler FM (2012) Kvazikristali – otkriće, struktura i svojstva. Kem Ind 61:349–359
Streitwolf H-W (1967) Gruppentheorie in der Festkörperphysik. Akademische Verlagsgesellschaft Geest & Portig K-G, Leipzig
Tavger BA, Zaitsev VM (1956) Magnetic symmetry of crystals. JETP 30:564–568
Threlfall W, Seifert H (1931) Topologische Untersuchung der Diskontinuitätsbereiche endlicher Bewegungsgruppen des dreidimensionalen sphärischen Raumes. Math Annalen 104:1–70
Vaidyanathaswamy R (1928) Integer-roots of the unit matrix. J London Math Soc 3:121–124
Van der Waerden BL (1985) A history of algebra: from al-Khwarizmi to Emmy Noether. Springer-Verlag, Berlin
Van der Waerden BL, Burckhardt JJ (1961) Farbgruppen. Z Krist 115:231–234
Veysseyre R, Veysseyre H (2002) Crystallographic point groups of five-dimensional space. 1. Their elements and their subgroups. Acta Cryst A58:429–433
Veysseyre R, Weigel D, Phana D, Veysseyre H (2002) Crystallographic point groups of five-dimensional space. 2. Their geometrical symbols. Acta Cryst A58:32–39
Vince A (1997) Periodicity, quasiperiodicity, and Bieberbach's theorem on crystallographic groups. Am Math Monthly 104:27–35
Weber L (1929) XII. Die Symmetrie homogener ebener Punktsysteme. Z Krist 70–328
Werner AG (1774) Von den äusserlichen Kennzeichen der Fossilien. Siegfried Lebrecht Crusius, Leipzig
Weyl H (1952) Symmetry. Princeton University Press, Princeton
Wilkie HC (1966) On non-Euclidean crystallographic groups. Math Z 91:87–102
Wills AS (2017) A historical introduction to the symmetries of magnetic structures. Part 1. Early quantum theory, neutron powder diffraction and the coloured space groups. Powder Diffract 32: 148–155
Whitlock HP (1934) A century of crystallography. Am Mineral 19:93–100
Yackel C (2012) Teaching temari: geometrically embroidered spheres in the classroom. In: Bosch R, McKenna D, Sarhangi R (eds) Bridges Towson conference proceedings 2012. Tessellations Publ, Phoenix (AZ), pp 563–566. Retrieved from https://archive.bridgesmathart.org/2012/
Zamorzaev AM (1988) Generalized antisymmetry. In: Hargittai I, Vainshtein BK, Udalova VV, Rodin EY (eds) Crystal symmetries. Shubnikov centennial papers. Pergamon Press, Oxford, pp 555–562
Zassenhaus H (1948) Über einen Algorithmus zur Bestimmung der Raumgruppen. Comm Math Helvetici 21:117–141

The Notion of Space in Grothendieck

From Schemes to a Geometry of Forms

John Alexander Cruz Morales

Contents

1 Introduction 1866
1.1 A Brief Biographical Note 1867
2 A Walk Around Algebraic Geometry 1870
2.1 Schemes: Generalizing Algebraic Varieties 1870
2.2 Toposes: Spaces Without Points 1871
3 Searching for New Foundations: Tame Topology 1873
4 Towards a New Geometry of Forms 1875
4.1 Riemann's Heritage 1879
5 The Notion of Space: Mathematics, Physics, and Philisophy 1880
6 Conclusion 1882
References 1884

Abstract

In this essay we give a general picture about the evolution of Grothendieck's ideas regarding the notion of space. Starting with his fundamental work in algebraic geometry, where he introduces schemes and toposes as generalizations of classical notions of spaces, passing through tame topology and ending with the formulation of a geometry of forms, we show how the ideas of Grothendieck evolved from pure mathematical considerations to physical and philosophical questions about the nature and structure of space and its mathematical models.

Keywords

Schemes · Toposes · Tame spaces · Geometry of forms · Grothendieck

J. A. Cruz Morales (✉)
Universidad Nacional de Colombia, Bogotá, Colombia
e-mail: jacruzmo@unal.edu.co

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_5

1 Introduction

Alexander Grothendieck is one of most influential mathematicians of the last century and it is not difficult to argue that is one of the most influential through the history of mathematics. The reach and depth of Grothendieck's work have been extensively discussed in many places, see for instance (Cartier 2001; Lochak and Schneps 1997; Zalamea 2019) and the bibliography therein. Our goal is more modest since we do not pretend to cover fully Grothendieck's ideas but just focus on a specific, but central part of his work, namely the notion space (Since this text is thought for a general audience including mathematicians, philosophers, philosophers of mathematics, and, in general, anyone interested in the history and philosophy of the notion of space in the grothendieckean work, we will avoid the technicalities and will privilege a conceptual approach.).

It is very clear, even in a superficial reading of Grothendieck's work, that the search of an adequate definition of the notion of space is a leitmotif for Grothendieck, from the very beginning of his mathematical career when he was interested in nuclear spaces to the last mathematical reflections in the 1980s when he tried to develop a geometry of forms as a new foundations for topology and the physical space (It is known that during his retirement in Lasserre since 1991 until his death, Grothendieck wrote extensively about different topics and it would be plausible to think that one of these topics was the notion of space. However, this is a mere speculation, so for the purpose of this text, we do not take into account that period.), passing through his monumental work in algebraic geometry.

The question *What is a space?* has been one of the main questions the human being has asked since the ancient times. Many philosophers, natural philosophers (We will use the expression *natural philosopher* as it was understood in seventeenth and eighteenth century.), and mathematicians have tried to give an answer (There is no point to make an extensive list of the people who have been interested in the problem of the space but this list should include people like Plato, Aristotle, Archimedes, Leibniz, Newton, Riemann, and Einstein, just to mention some outstanding examples. Clearly Grothendieck follows this line.) and even though we have now a better understanding of the problem we are still far from a definitive answer. The question is hard since it involves not only mathematics but also physics and philosophy. This was cleverly noted by Grothendieck in the 1980s (Of course, Grothendieck was not the first one who notes this. However, we want to emphasize that Grothendieck realized the importance of philosophy and physics as complements of the mathematical approach to the notion of space, since traditionally the studies on Grothendieck's work have not paid attention to this aspect. On the contrary, it is usually believed that Grothendieck was not interested at all in the physical and philosophical aspects of the notion of space, and this is partially true, if we only focus on Grothendieck's ideas before the 1980s. Nevertheless, his mind changed a lot, regarding physical and philosophical issues, and this is one of the points we want to discuss in this essay.) and we will return to this point later in this text.

In fact, in (Grothendieck n.d.) Grothendieck writes about the notion of space:

> "*'La notion d' 'espace' est sans doute une des plus anciennes en mathématique. Elle est si fondamentale dans notre appréhension 'géométrique' du monde, qu'elle est restée plus ou moins tacite pendant plus de deux millénaires. C'est au cours du siècle écoulé seulement que cette notion a fini, progressivement, par se détacher de l'emprise tyrannique de la perception immédiate (d'un seul et même 'espace' qui nous entoure), et de sa théorisation traditionnelle ('euclidienne'), pour acquérir son autonomie et sa dynamique propres. De nos jours, elle fait partie des quelques notions les plus universellement et les plus couramment utilisées en mathématique, familière sans doute à tout mathématicien sans exception. Notion protiforme d'ailleurs s'il en fut, aux cents et mille visages, selon le type de structures qu'on incorpore à ces espaces, depuis les plus riches de toutes (telles les vénérables structures 'euclidiennes', ou les structures 'affines' et 'projectives', ou encore les structures 'algébriques' des 'variétés' de même nom, qui les généralisent et qui assouplissent) jusqu'aux plus dépouillées: celles où tout élément d'information 'quantitatif' quel qu'il soit semble disparu sans retour, et où ne subsistent plus que la quintessence qualitative de la notion de 'proximité' ou de celle de 'limite', et la version la plus élusive de l'intuition de la forme (dit 'topologique').*"

Following Deligne (Deligne 2020) we can think of a space as something for which *localization* makes sense. In this approach the key word is *localization*. Thus, the question *What is a space?* might be replaced by *What does localization mean?*. Along this essay we will try to show how Grothendieck approached this question. The basic idea of localization, the one Deligne had in mind, is exemplified by the following situation (This example is due to Deligne who wrote it to the author in (Deligne 2020).): Let X be a topological space. In order to define a continuous function $f: X \to \mathbb{R}$, it suffices to define it on open subsets $\mathcal{U}_i$ covering X and to check the agreement on two by two intersections. This way the notion of continuous function is defined by localization on X.

In the example above there are two important points that play an important role in the grothendieckean notion of space. The first one is the absence of points, replaced by another thing, in the example by open sets, and the second one is that those open sets cover the whole thing. It has been remarked in other places (see Cartier 2001) how problematic the notion of point could be in order to understand what a space is. One of the main features in the several generalizations Grothendieck proposes for the notion of space is that for Grothendieck the space is not made by points. Points are just marks that can be put on the space but not its fundamental elements. On the other hand, the idea of covering is essential in order to understand the relation between the local and the global aspects of the space and it also will be important in the discussion between the continuum and the discrete in Grothendieck's geometry of forms as we will see later.

1.1 A Brief Biographical Note

We will give a very short review of Grothendieck's life. More detailed accounts can be found in (Cartier 2001; Jackson 2004; Scharlau 2010; Zalamea 2019) (We include

this brief note in order to give the reader an idea of Grothendieck's particular life and shed light on how his personality as mathematician, and more general as human being, was formed.). Alexander Grothendieck was born in Berlin on March 28, 1928 to Alexander Schapiro and Hanka Grothendieck, both were active political militants. Due to the activities of his parents, Grothendieck lived nearby Hamburg (The years in Hamburg had a big impression in Grothendieck. We were informed by Winfried Scharlau (Scharlau 2020) that Grothendieck visited Hamburg around 2006. He knew this by some local people that informed him on the visit. In some sense, it seems that Grothendieck was recollecting his own steps.) under the tutelage of the family of a protestant priest until 1939 when he met his family (Grothendieck's parents took part in the Spanish Civil War enrolled in the anarchist militias.). However, his father was taken prisoner in the concentration camp Le Vernet and then deported to Auschwitz where he died and Alexander and his mother were sent to the camp Rieucros. This compromise of his parents with political issues was very influential in Grothendieck's point of view, both mathematical and nonmathematical.

From 1945 to 1948 Grothendieck studied mathematics at Montpellier. In Récoltes et Semailles (Grothendieck n.d.) he recollects the years in Montpellier and how in solitude developed a measure theory close to the one developed by Lebesgue many years ago. This episode marked his life as mathematician. As Grothendieck wrote (see Grothendieck n.d.) thanks to this he learnt the importance of solitude in the work of a mathematician. A solitude that accompanied him his whole life beyond mathematics.

In 1949 Grothendieck arrived to Paris to Cartan's seminar at the École normale supérieure. This was an important turn on Grothendieck's mathematical education, since he moved from the border (Montpellier) to the center of French mathematics (Paris). In Paris and Nancy, where he wrote his doctoral dissertation under Jean Dieudonné and Laurent Schwartz, Grothendieck could show his enormous mathematical talent. After this, he spent time as postdoctoral researcher in Sao Paulo and Kansas.

During his time in Kansas in 1955 Grothendieck's work took an important turn since he left analysis and started a new mathematical road in (algebraic) geometry. It is in this stay when he wrote the main ideas of his work in abelian categories (This work is known as *Tôhoku* because of the journal where it was published.) and also he started to work in his version of the Riemann–Roch theorem. Thus, his monumental reconstruction of algebraic geometry started and, even more, his quest of an adequate definition of the notion of space really started in an explicit way.

Since 1958 until 1970 Grothendieck occupied a position at IHES. With Dieudonné as his scribe, Grothendieck wrote the giant treatise *Elements of Algebraic Geometry* (EGA) and ran his celebrated *Seminar of Algebraic Geometry* (SGA). Grothendieck was positioned as the main expert in algebraic geometry in the world and won the Fields medal in 1966 for his fundamental work. During this time he introduced many important concepts (schemes and toposes among others) in order to carry out his program of providing a vast generalization of the notion of space.

Despite the success of his seminar and the fruitful years at the IHES, Grothendieck resigned his position in 1970 arguing ethical reasons because the founding the IHES

was receiving from the military agency (This is the universal accepted reason on Grothendieck's resignation to the IHES. However, some authors (see (Ruelle 2007), for instance) believe that there could be deeper reasons behind Grothendieck's demission.). After being involved in ecological activities in the group *Survivre et vivre* he moved back to Montpellier and took one position as professor. It is argued that Grothendieck left mathematics after his departing of IHES since he ceased any mathematical publication but this is far from being true. Certainly Grothendieck abandoned the establishment but this does not mean that he abandoned mathematics. In fact, many important Grothendieck's ideas were developed during the late 1970s and the beginning of the 1980s when he was in Montpellier.

In order to have an idea of Grothendieck's mathematical interests during the 1980s, it is good to see (Grothendieck 1997). This influential text (It is known the story on how Voevodsky, being a student in Moscow, started to learn French just to read Grothendieck's Esquisse. Certainly this reading was very important in Voevodsky's career.) contains the main ideas of Grothendieck after his reconstruction of the algebraic geometry. There we can find discussions about moduli spaces of Riemann surfaces and the Teichmüller tower, tame topology (another approach to a generalization of the notion of space), children drawings, and anabelian geometry. This text condensed many of the mathematical thoughts written in manuscripts that widely circulated between the mathematical community.

It is interesting to note, however, that some ideas concerning the geometry of forms that will be important for us, in the sense that they constitute an important grothendieckean approach to the notion of space and its relations with physics and philosophy, written in an relatively unknown manuscript (Grothendieck 1986), are not discussed at the Esquisse. These ideas were really important for Grothendieck as we can see from this extract of a letter sent to Tsuji on July fourth 1986:

> "*Excusez-moi d'avoir laissé passer quelques jours avant de répondre à votre lettre précédente - depuis près d'un mois je suis lancé sur un nouveau thème mathématique. Des nouveaux fondements de la topologie, dans un esprit très différent de celui de la topologie générale - j'ai envie d'appeler cette nouvelle topologie 'géométrie des formes' ou 'analysis situs', et il est bien possible que ce sera là le premier texte mathématique que je publierai, après mon départ de 1970 (et avant les divers textes prévus, que j'annonce dans l'introduction à Récoltes et Semailles).*"

Aside from the mathematical texts, Grothendieck wrote in the 1980s other general texts of great interest. His long text Récoltes et Semailles is a jewel in the whole sense of the word. There Grothendieck not only explored his past as mathematician but also discussed the intricate aspects of mathematical creativity and the ethics of a mathematician. In La clef des songes, Grothendieck analyzed his dreams and arrived to a personal discovering of God.

In 1991 Grothendieck decided to move to Lasserre and only kept contact with few people. During this time he wrote extensive texts about different topics, but it is still a mystery the exact content of the thousands of pages he wrote in his isolation. He passed away on November 13, 2014, at Saint-Girons, leaving behind him a big legacy for the future generations.

2 A Walk Around Algebraic Geometry

In this section we will discuss two generalizations of the notion of space that Grothendieck proposed as part of his program of reconstruction of algebraic geometry. Those generalizations are the notion of scheme and topos. These two concepts are central in Grothendieck's work, and particularly relevant if one wants to understand Grothendieck's idea of space. In (Grothendieck n.d.) Grothendieck examined his more important contributions in mathematics and regarding toposes and schemes he writes:

> "*Parmi ces thèmes, le plus vaste par sa portée me paraît être celui des topos, qui fournit l'idée d'une synthèse de la géométrie algébrique, de la topologie et de l'arithmétique. Le plus vaste par l'étendue des developpements auxquels il a donné lieu dès à present, est le thème des schémas. C'est lui qui fournit le cadre 'par excellence' de huit autres parmi ces thèmes envisagés, en même temps qu'il fournit la notion centrale pour un renouvellement de fond en comble de la géométrie algébrique, et du langage algébrico-géométrique.*"

It is clear from this quote the unifying role played by schemes and toposes inside Grothendieck's mathematical vision. Thus, these two concepts allow to study not only topological or geometric aspects of the space but also algebraic and arithmetic ones. Here, we would like to emphasize that from the very beginning Grothendieck's idea of space involve a particular entity of arithmetic and geometric nature. That is what Zalamea (2019) calls the space-number (In (Cruz Morales Submitted) we associate Zalamea's idea of the space-number to the idea of form in Grothendieck. We will go back to this point later in this text, since it is important in this discussion about the notion of space.) in analogy with the space-time in physics. We agree with Zalamea's interpretation, so we believe we need to understand Grothendieck's idea of space in terms of his idea of merging arithmetic and topology/geometry (This is a manifestation of the duality discrete/continuum. When we discuss Grothendieck's ideas of the space coming from his work in the 1980s, this dichotomy will be clearer. Now, it is interesting to notice that some mixed discrete/continuum structure for the space is present in Grothendieck's first works in algebraic geometry, at least in an implicit way, and it will be more relevant when Grothendieck tries to think in the physical space.). Therefore, schemes and toposes can be seen as the first instances of the space-number that Grothendieck had in mind.

2.1 Schemes: Generalizing Algebraic Varieties

In his ICM talk in 1958 (Grothendieck 1958) Grothendieck introduced his vast program of generalization of algebraic geometry that took him the next decade. During his years at the IHES he dedicated his life (Grothendieck's capacity of work is legendary. It is said that he could spend until 12 h of continuous work.) to develop these ideas. One of the central notion he introduced is the notion of scheme and for us this is important since it is his first generalization of the notion of space.

As it is discussed in (Deligne 1998; Zalamea 2019) Grothendieck's intrepid idea is to accept that for every commutative ring A (with unit) it is possible to define an affine scheme $X = \mathrm{Spec}(A)$. What is needed here is to define the structure sheaf $\widetilde{A}$. In this case, the structure sheaf is the sheaf of rings over X with a basis of open sets $\{x \in X : f(x) = [f]_x \neq 0\}$ with $f \in A$ and with fibers the local rings A_x with $x \in X$. Thus, considering a ringed space, i.e., a pair $(X, \mathcal{O}_X)$ with X a topological space and $\mathcal{O}_X$ a sheaf of rings over X, we say that it is an affine scheme if it is isomorphic to $\left(\mathrm{Spec}(A), \widetilde{A}\right)$. A general scheme is defined by gluing affine schemes.

Note that in the definition above the points of a scheme are not relevant. The important thing is that a scheme is an object where we can make *localization*, in the sense we discussed before. Thus, a scheme is a space of algebraic geometric nature that generalizes the notion of algebraic variety. Indeed, Grothendieck characterized the notion of scheme as a metamorphosis of the notion of algebraic variety (see Grothendieck n.d.). The use of the word metamorphosis is interesting, since it suggests a natural transformation (Naturalness is important for Grothendieck. In his view, mathematical concepts must be natural, in contrast to artificial and intricate constructions.) from the notion of algebraic variety to that of scheme.

About naturalness of the notion of scheme, Grothendieck wrote in (Grothendieck n.d.).

> "*La notion de schéma est la plus naturelle, la plus 'évident' imaginable, pour englober en une notion unique la série infinie de notions de 'variété' (algébrique) qu'on maniait précédemment.*"

And also in the same text adds

> "*L'idée même de schéma est d'une simplicité enfantine - si simple, si humble, que personne avant moi n'avait songe à se pencher si bas. Si 'bébête' même, pour tout dire, que pendant des années encore et en dépit de l'évidence, pour beaucoup de mes savants collègues, ça faisait vraiment 'pas sérieux'! Il m'a fallu d'ailleurs des mois de travail serré et solitaire, pour me convaincre dans mon coin que 'ça marchai' bel et bien.*"

It is clear that the notion of scheme is for Grothendieck the natural way in which the notion of algebraic variety should evolve, but also he noticed that the notion was always there, but it was needed to have the right eyes (his eyes in this case) to see it. Grothendieck was interested in capturing in one notion the multitude of varieties that arise when we work over prime characteristics. Again, the dichotomy discrete/continuum is present here. This is the fundamental aporia associated to the notion of space and Grothendieck started to struggling with it by introducing the idea of scheme. However, the notion of space is richer than that of algebraic variety and Grothendieck knew that.

2.2 Toposes: Spaces Without Points

At this point we have just discussed a restricted idea of space since we have considered only those spaces that are algebraic varieties and look at their

metamorphosis in the idea of scheme. However, for Grothendieck the space is something more general and, in this sense, the notion of scheme is not enough to capture the subtleties of space. Thus, it is necessary to go beyond algebraic geometry and entering in the realm of topology. The right notion to start doing that is the notion of topos.

In order to understand what a topos is we need to start with the idea of site. Here we will follow the description in (Zalamea 2019). Let us consider a category $\mathcal{C}$ and let us define a topology on it. For any object X of $\mathcal{C}$ it is possible to define a set of sieves $J(X)$ (A sieve can be identified with subobjects of the dual category.) satisfying some axioms, namely:

1. For any object $X \in \mathcal{C}$, any sieve $R \in J(X)$ and any morphism $f : Y \to X \in \mathcal{C}$, the sieve $R \times_X Y$ is in $J(Y)$.
2. For any sieves of X, R, and R', if $R \in J(X)$ and for any $Y \to R$, $R' \times_X Y \in J(Y)$, then $R' \in J(X)$.
3. For any object X of $\mathbb{C}$, $X \in J(X)$.

A category $\mathcal{C}$ endowed with a topology that satisfies 1, 2, and 3 is called a site. We focus on the idea of site since it allows us to illustrate Deligne's point that a space is something where localization makes sense. However, a site is just the first germ in the generalization of topological spaces that Grothendieck had in mind. Then, what is needed to be considered is a sheaf over a site.

Thus, a topos (In this text, we use the word topos in order to refer Grothendieck topos. There is a more general definition of topos (elementary topos) due to Lawvere but we do not consider it here.) is the category of all sheaves on a site. Therefore, a topos can be seen as a generalized space. Like schemes that can be seen as a metamorphosis of the idea of algebraic variety, a topos can be seen as a metamorphosis of the idea of topological space. Again, the word metamorphosis is quite accurate, since it allows to see how toposes arise as a natural deformation of the classical idea of topological space (We will see later that this metamorphosis does not stop in the notion of topos and will take another forms in terms of moderate topology and geometry of the forms.).

Grothendieck introduced the idea in order to deal with cohomology theories in algebraic geometry, and in particular, toposes play an important role in Grothendieck's approach to Weil conjectures. This way, toposes also arise as an important objects in problems of arithmetic nature. Again, it is possible to see that for Grothendieck the notion of space is something beyond purely geometric/topological considerations and it is central in whole mathematics. In (Grothendieck n.d.), Grothendieck points out:

> "*Celle de topos constitue une extension insoupçonnée, pour mieux dire, une métamorphose de la notion d'espace. Par là, elle porte la promesse d'un renouvellement semblable de la topologie, et au delà de celle-ci, de la géométrie. Dès à présent d'ailleurs, elle a joué un rôle crucial dans l'essor de la géométrie nouvelle (surtout à travers les thèmes cohomologiques ℓ-adique et cristallin qui en sont issus, et à travers eux, dans la démonstration des conjectures de Weil).*"

In the same text Grothendieck adds:

> "*C'est le thème du topos, et non celui des schémas, qui est ce 'lit', ou cette 'rivière profonde', où viennent s'épouser la géométrie et l'algèbre, la topologie et l'arithmétique, la logique mathématique et la théorie des catégories, le monde du continu et celui des structures 'discontinues' ou 'discrètes'. Si le thème des schémas est comme le coeur de la géométrie nouvelle, le thème du topos en est l'enveloppe, ou la demeure. Il est ce que j'ai con'çu de plus vaste, pour saisir avec finesse, par un même langage riche an résonances géométriques, une 'essence' commune à des situations des plus éloignées les unes des autres, provenant de telle région ou de telle autre du vaste univers des choses mathématiques*."

From the last quote it is very clear that for Grothendieck the notion of space, with its avatar in the notion of topos, is ubiquitous and central in mathematics (The notion of topos is, in certain sense, marginal in mathematics nowadays. However, recent works of Alain Connes and his collaborators (see (Connes and Consani 2014), as an example) have brought new life to the idea of topos and we believe this concept will play a central role in major mathematical problems like the Riemann hypothesis, for instance.) and adequate device to study in a unified way arithmetic and geometry, a place where discrete and continuous structures can live together in harmony.

So far, we have seen a generalization of the notion of space running in two steps. First one corresponding to the notion of scheme and the second one corresponding to the notion of topos. This two-steps generalization occurs inside the domain of algebraic geometry which was the main object of Grothendieck's reflections between 1958 and 1970. Thus, starting with the natural spaces in algebraic geometry, i.e., algebraic varieties, Grothendieck tried to find an archetypical structure for the space in the notion of topos. However, as we will see below the notion of topos does not constitute the final step in Grothendieck's thoughts related to the notion of space. There are other aspects to be considered and the idea of going beyond the notion of topos will dominate Grothendieck's work during the 1980s.

3 Searching for New Foundations: Tame Topology

Around the middle of the 1980s Grothendieck wrote an interesting text (Grothendieck 1997) detailing several research programs he was thinking during the 1970s, after his retirement from the mathematical milieu, which he considered would play an important role in the future mathematical research. One of these programs concerns to the development of the so-called tame topology, a new foundations for topology. In Grothendieck's words (see Grothendieck 1997):

> "*I would like to say a few words now about some topological considerations which have made me understand the necessity of new foundations for 'geometric' topology, in a direction quite different from the notion of topos, and actually independent of the needs of so-called 'abstract' algebraic geometry (over general base field and rings). The problem I started from, which already began to intrigue me some fifteen years ago, was that of defining a theory of 'dévissage' for stratified structures, in order to rebuild them, via a canonical process, out of 'building blocks' canonically deduced from the original structure*."

Here, in this quote, we see how Grothendieck wanted to propose a new setting that goes beyond the idea of topos and, in fact, beyond algebraic geometry. As we have seen, the origin and the development of the notion of topos is linked to the development of algebraic geometry in the setting Grothendieck envisioned during the 1960s, so the notion of space developed during this period. Nevertheless, Grothendieck realized that in order to capture the essence of the space is necessary to considered a more general picture. Thus, with his idea of tame topology, the notion of space in Grothendieck started to be independent of the algebraic-geometric origin of Grothendieck's first approach.

Grothendieck chose an axiomatic approach in order to understand what a tame space could be. In his words:

> "*My approach toward possible foundations for a tame topology has been an axiomatic one. Rather than declaring (which would indeed be a perfectly sensible thing to do) that the desired 'tame spaces' are no other than (say) Hironaka's semianalytic spaces, and then developing in this context the toolbox of constructions and notions which are familiar from topology, supplemented with those which had not been developed up to now, for that very reason, I preferred to work on extracting which exactly, among the geometrical properties of the semianalytic sets in a space* $\mathbb{R}^n$*, make it possible to use these as local 'models' for a notion of 'tame space' (here semianalytic), and what (hopefully!) makes this notion flexible enough to use it effectively as the fundamental notion for a 'tame topology' which would express with ease the topological intuition of shapes*."

Even though, as Grothendieck wrote, a tame space is not a semianalytic space, it is illustrative, in order to understand the idea Grothendieck had in mind, to discuss a bit the notion of semianalytic set, since Grothendieck did not provide an explicit definition of tame space in (Grothendieck 1997).

Let M be a real analytic manifold and consider a subset X of M, the X is called semianalytic if each $a \in M$ has a neighborhood $\mathcal{U}$ such that $X \cap \mathcal{U} \in S(\mathcal{O}(\mathcal{U}))$, where $\mathcal{O}(\mathcal{U})$ denotes the ring of real analytic functions on $\mathcal{U}$ and $S(\mathcal{O}(\mathcal{U}))$ denotes the smallest family of subsets of $\mathcal{U}$ containing all $\{f(x) > 0\}, f \in \mathcal{O}(\mathcal{U})$, which is stable under finite intersections, finite unions, and complement.

Without a precise definition, but using the notion of semianalytic space as local model, it is possible to see that the new tame spaces fit in the general definition of space we have assumed following Deligne. However, there is an important point in the origin of the idea of tame spaces that it did not appear in the previous notions of scheme and topos. When Grothendieck formulated the idea of tame space he was thinking in a right framework to deal with the notion of geometrical shape. Thus, the main motivation here is to develop a notion of space that allows us to study the geometric properties of shapes (In the next section we will see that with the same motivation of understanding the elusive idea of form (being geometrical shapes just one instance of this archetypical notion), Grothendieck was beyond tame spaces and proposed the so-called geometry of forms.). This point is very clear from this quote (see Grothendieck 1958):

> "*After some ten years, I would now say, with hindsight, that 'general topology' was developed (during the thirties and forties) by analysts and in order to meet the needs of*

analysis, not for topology per se, i.e. the study of the topological properties of the various geometrical shapes. That the foundations of topology are inadequate is manifest from the very beginning, in the form of 'false problems' (at least from the point of view of the topological intuition of shapes) such as the 'invariance of domains', even if the solution to this problem by Brouwer led him to introduce new geometrical ideas."

Unlike the notion of scheme and topos, both central in algebraic geometry and widely known for mathematicians nowadays, the notion of tame space, and therefore tame topology, is marginal in mathematics. There are some developments in logic (van den Dries 1998) in relation to O-minimal structures, but besides this the idea of tame space is not that relevant (We consider that it would be a very interesting research program to try to develop Grothendieck's ideas in (Grothendieck 1997) related to tame topology. We think that this could open new geometric perspectives and new understanding of topological phenomena as Grothendieck envisioned.). However, in our study of the notion of space in Grothendieck is, in fact, very important, since it provides the first attempt of Grothendieck to formulate a theory of space with the clear motivation of providing a setting to study shapes as foundational objects for geometry/topology, as we already mentioned. In this sense, we believe that it is not possible to understand how the idea of space evolved in Grothendieck's work without passing through the discussion about tame topology. On the other hand, we will see that in order to tackle the problem of understanding shapes, this is not the final answer that Grothendieck wanted to provide.

It is important to remark that the notion of tame space does not pretend to generalize the idea of topological space (like toposes do) but provides a new understanding for it. In that sense, a tame space is a new form that the notion of space can take, it is new type of space. Thus, we can see that the notion of space is really protean for Grothendieck.

4 Towards a New Geometry of Forms

We saw in the last section how part of the motivation for developing a tame topology was to construct a right notion of space in order to deal with the various geometrical shapes. In fact, the notion of form (In (Cruz Morales Submitted) we have argued that there is an important difference between form and shape and we have discussed how the notion of form is more general than that of shape. In fact, Grothendieck himself chose to use the word form rather than shape, as we can see in a letter to Ronald Brown on June 17th, 1986:

"I've been very strongly involved with mathematics lately, working out another approach to 'topology' and 'form' (different from topological spaces, from topoi and from moderate spaces as proposed in Esquisse d'un Programme), just getting the basic language straight.") was one of the main subjects of Grothendieck's reflections in the 1980s, and in some sense tame topology was not enough for Grothendieck in order to provide a general setting for the notion of space if we want to study the geometric properties of forms. In this direction, Grothendieck

proposed a new approach for the foundations of topology. This is what he called geometry of forms. In a letter to Yamashita on July 9th, 1986, Grothendieck wrote:

"*I have been very intensely busy for about a month now, with writing down some altogether different foundations of 'topology', starting with the 'geometrical objects' or 'figures', rather than with a set of 'points' and some kind of notion of 'limit' or (equivalently) 'neighbourhoods'. Like the language of topoi (and unlike the so-called 'moderate space' theory foreshadowed in the Esquisse, still waiting for someone to take hold of the work in store), it is a kind of topology 'without points' - a direct approach to 'shape'. I do hope the language I have started developing will be appropriate for dealing with finite spaces, which come off very poorly in 'general topology' (even when working with non-Hausdorff spaces).*"

Grothendieck wrote a manuscript (Grothendieck 1986) about these ideas. The manuscript remained relatively unknown (except for a few mathematicians) until the online publication of Grothendieck's manuscripts by Université de Montpellier. However, since these were handwritten by Grothendieck, it is not easy to decipher its content. To the best of our knowledge the first places where there is an attempt to discuss the content of the manuscript are (Cruz Morales 2017, 2018) based on some correspondence of Grothendieck where he wrote about it. For the purpose of this essay, we consider that the ideas on the manuscript that we can extract from the correspondence are very important since they allow us to really understand how the notion of space was evolving in Grothendieck's work. We want to argue that a full understanding of the problem of space in Grothendieck is impossible without studying the manuscript or at least its main ideas. However, a deep study of the text is still a work in progress, but here we hope to give a good idea of what Grothendieck was thinking (Just to provide an idea of the text, we want to give the title of the different chapters of the manuscript.

Chap 1. On a Topology of the (Topological) Forms. Chap 2. Topological Realizations of Networks. Chap 3. Networks via Decompositions. Chap 4. Analysis Situs (First Attempt). Chap 5. Algebra of Figures. Chap 6. Analysis Situs (Second Attempt). Chap 7. Analysis Situs (Third Attempt). Chap 8. Analysis Situs (Fourth Attempt). Chap 9. Notes.

It is interesting to see how Grothendieck uses the expression analysis situs that evokes Leibniz and Poincaré. In a letter to Yamashita on September 16th, 1986, Grothendieck mentioned Poincaré conjecture in relation to the analysis situs he invented. It is still matter of study to understand the relation between Poincaré and Grothendieck (and also between Grothendieck's ideas and Leibniz' ideas), but we find illustrative this extract of the letter:

> "*I have been impressed, of course, by your second list of 'newest events'. I didn't even know about the solution of Poincaré's conjecture - comes just at the right time for me, to be able to justify within the framework of the 'geometry of forms' or 'analysis situs' I am developing in terms of extant 'general topology', a certain definition of 'regular figure' (the combinatorial substitute for 'variety') I had in mind*".
>
> Grothendieck refers the proof announced by Colin Rourke and Eduardo Rêgo. The proof turned out to be incorrect. However, some years after this, Gregory Perelman found a right proof for the conjecture. So it is tempting to ask: What is the relation between Poincaré conjecture and Grothendieck's geometry of forms?).

In the extract of the letter to Yamashita we have quoted there is one important point that we want to remark. Grothendieck pointed out that his new approach to the notion of space is closer to the idea of topos than the idea of tame space, in the sense that these new spaces are not made by points, but also because here he is looking for a direct approach to shapes. However, looking at the discussion in (Grothendieck 1958) concerning tame topology we can see that the motivation of Grothendieck was to understand the forms from a geometric point of view. Is there any contradiction here? We think the answer is no.

One thing is to be motivated by some idea and another thing is that this idea can be carried out by certain construction we propose. Our conjecture is that Grothendieck realized that his tame topology was not the right framework for the goal of understanding forms, taking forms in a very broad sense, and for this reason he saw the necessity of producing a new theory and that is what he expressed in the letter to Yamashita. Actually, Grothendieck explored different approaches during the years in order to understand what exactly a space is and the idea of tame space can be thought as one of them. Before his retirement at Lasserre in June 24th, 1991, Grothendieck wrote a letter to an unknown recipient (The unknown recipient is identified as A.Y. This letter and its relation with the manuscript about the geometry of forms have been discussed in (Cruz Morales 2017).):

> "*Il est vrai qu'us cours de dix dernières années, j'ai réfléchi ici et là à diverses extensions de la notion d'espace, en gardant à l'espirit la remarque pénétrante de Riemann. J'en parle dans quelques lettres à des amis physiciens ou 'relativistes'. Il ne doit pas être très difficile p. ex. de développer une sorte de calcul différentiel sur des 'variétés' qui seraient des ensembles finis (mais à cardinal 'très grand'), ou plus généralement discretes, visualisés comme format une sorte de 'réseau' trés serré de points dans una variété* C^{∞} *(p. ex. una variété riemannienne) - une sorte de géometrié différentielle 'floue', oú toutes les notions numériques son défines seulement 'à* ε *prés', pour un ordre d'approximation* ε *donné*."

This quote is particularly relevant. Three important points appear here that are not present in the reflections about the notion of space that Grothendieck proposed in his work in algebraic geometry and in tame topology. First of all, Grothendieck asks for the intrinsic structure of the space and its discrete or continuous structure. Of course, when he discussed about toposes the relation between the discrete and the continuum world was present, but in the sense on how toposes can put arithmetic and geometric phenomena together, but here the difference is that the discussion on the dichotomy continuum/discrete appears in a more foundational and structural way. Secondly, he was not only thinking of the mathematical space but also in the physical space. It is widely believed that Grothendieck was not interested in physics at all, but it seems that in his last period and regarding the problem of space he was interested in physics and, in fact, he had contact with some physicists and discussed with them his ideas. Finally, he recognize explicitly that some of his thoughts about the space came from Riemann. We will back to the

relation with Riemann later, so let us focus on the first and the second points we have mentioned.

What was the kind of space that Grothendieck had in mind? We can find a first answer in the letter to Yamashita on June 9th, 1986. There Grothendieck argued:

> "*As Riemann pointed out I believe, the mathematical continuum is a convenient fiction for dealing with physical phenomena, and the mathematics of infinity are just a way of approximating (by simplification through 'idealisation') an understanding of finite aggregates, whose structure seems too elusive or too hopelessly intricate for a more direct understanding (at least it has been so till now). Yet it may well be that we are approaching at present the point where the continuous models of the physical world fail to be adequate - but the physicists are so accustomed taking for granted the conceptual superstructure of 'continuous space' worked out by the Greeks and their successors up to Relativity and Quantum mechanics, and to confuse it with reality, that they may well never get aware of a widening gap between concepts and phenomena. This brings to my mind that not only the language of (so-called 'topological') shape is to be rethought from scratch, in order to be capable to account for 'shape' of finite spaces, but the same holds with vast sectors of geometry and analysis, if not all of it. Thus differential and integral calculus, differential equations and the like, differentiable or Riemannian varieties, the tensor calculus on it, etc - all this quite evidently (once you start thinking about it) should make and has to make sense within the framework of (say) finite aggregates as well. Just think, say, of* 10^{100} *or* 10^{1000} *points pretty densely located on the unit sphere or in the unit ball, a lot more than needed for exceeding by very far he accuracy of any conceivable physical measurement, so that no physicists nor anybody whosoever (except maybe God Himself) could ever possible distinguish between the ambient "continuous" space (granting it does have some kind of 'physical' existence, which I doubt) and this kind of network thought of as an 'approximation' (whereas in reality, the opposite seems to me more likely to be true: the continuum is the concept approximating the elusive finite - but very large - aggregate). This would mean rethinking too the notion of differentiability (say) and of differential, in wholly new terms, when 'passing to the limit' means now a finite process, namely passing to 'small' values of the parameter or parameters, where the rate of 'smallness' is precisely prescribed and should by all means not be exceeded, and where the 'margin of errors' admitted in the computation or definition of the 'differential' or 'derivative' is equally precisely prescribed, in keeping with the former prescription*."

Grothendieck was inclined for a discrete structure for the space and even he thought of the continuum as a convenient fiction. This is actually quite impressive. Nobody before Grothendieck proposed to think the continuum as an approximation to the discrete (In the letter Grothendieck attributes this idea to Riemann. However, it is not that clear that actually Riemann was thinking the continuum in this terms. We will back to this point later.). We find this idea very suggestive on its own (We discuss about this in (Cruz Morales 2017). Let us just mention that this idea is an inversion respect to the usual way of thinking the relation between the continuum and the discrete. Usually the continuum is seen approximate by the discrete via a saturation (limit) procedure. Grothendieck is proposing a radical, but we think very fruitful, change of perspective. It would be very interesting to explore in more detail this line of thought.). For Grothendieck, the continuous structure of the space can be thought as an artificial, but necessary and convenient, construction in order to understand its actual structure. However, the relation between the discrete and the

continuum is not that simple for Grothendieck. In fact, in (Grothendieck n.d.) he proposed that there could be three possibilities for the structure of the space. The space can be discrete, continuum, or mixed discrete/continuum structure. In Grothendieck's words:

> "*Les développements en mathématique des dernières décennies ont d'ailleurs montré une symbiose bien plus intime entre structures continues et discontinues, qu'un ne l'imaginait encore dans la première moitié de ce siècle. Toujours est-il que de trouver un modèle 'satisfaisant' (ou, au besoin, un ensemble de tels modèles, se 'raccordant' de fa'çon aussi satisfaisante que possible), que celui-ci soit 'continu', 'discret' ou de nature 'mixte' — un tel travail mettra en jeu sûrement une grande imagination conceptuelle, et un flair consommé pour appréhender et mettre à jour des structures mathématiques de type nouveau. Ce genre d'imagination ou de 'flair' me semble chose rare, non seulement parmi les physiciens (où Einstein et Schrödinger semblent avoirété parmi les rares exceptions), mais même parmi les mathématiciens (et là je parle en pleine connaissance de cause)*."

Again, it is clear that he was thinking not just in the mathematical space but also in the physical one. Unlike what he did in algebraic geometry and in tame topology, his new geometry of forms is not just a mathematical description of the space but it should be thought as a description useful for physical purposes. This is the reason, we believe, that the notion of form was so central in this new approach. A form has many manifestations in the world not only in mathematics, so any theory about the space focused on the idea of form should reflect those manifestations. We think Grothendieck had fully conscience about this.

4.1 Riemann's Heritage

An important point to analyze in Grothendieck's conception of the geometry of forms is Riemann's influence. Of course, Riemann has a big influence in the whole grothendieckean work, but in relation to the geometry of forms that influence is particularly interesting. In (Grothendieck n.d.) we find:

> "*Il doit y avoir déjà quinze ou vingt ans, en feuilletant le modeste volume constituant l'oeuvre complète de Riemann, j'avais été frappé par une remarque de lui 'en passant'. Il y fait observer qu'il se pourrait bien que la structure ultime de l'espace soit 'discrète', et que les représentations 'continues' que nous nous en faisons constituent peut-être une simplification (excessive peut-être, à la longue. . .) d'une réalité plus complexe; que pour l'esprit humain, 'le continu' était plus aisé à saisir que 'le discontinu', et qu'il nous sert, par suite, comme un 'approximation' pour appréhender le discontinu. C'est là une remarque d'une pénétration surprenante dans la bouche d'un mathématicien, à un moment où le modèle euclidien de l'espace physique n'avait jamais encore été mis en cause; au sens strictement logique, c'est plutôt le discontinu qui, traditionnellement, a servi comme mode d'approche technique vers le continu*."

Here Grothendieck again refers Riemann as the one who was the first in thinking the continuum as an approximation of the discrete, and in this sense the continuous structure of the space as an approximation of a more intricate discrete structure of it

(We already saw that Grothendieck also mentions Riemann in relation with this idea in his letter to Yamashita on June 9th, 1986.). What it is also important to consider is that Grothendieck pointed out that the relation between the continuum and the discrete can be determined by the fact that the continuum is easier to assimilate by the human mind (he uses the expression human spirit) than the discrete. This has deep philosophical implications in terms of the nature of space. The question is to some extent this idea is already in Riemann in the explicit way that Grothendieck is presenting it.

Riemmann's contribution started to be clarified in a letter to Yamashita on September 16th, 1986. There Grotehndieck wrote:

> "*When I wrote about the theory of mathematical continuum begin a kind of approximation and idealization of very large and complex discrete aggregates, I thought I was citing an idea of Riemann, from what I remembered from reading through his collected works, many years ago (presumably in his exposition of the idea of what is now called riemannian geometry). When I later looked for a precise reference in his work, I was rather surprised that I wasn't able to find it - and it seems you didn't find it either. So maybe, I found his idea only 'in between the lines' of his actual words*."

The "in between the lines" at the end is really remarkable. At the end of his Habilitation thesis (Riemann 1854), Riemann discussed about the possibility of a discrete structure for the space. That idea yielded a big impression on Grothendieck. After looking with more detail, he recognized that in an explicit way the idea is not in Riemann but he certainly believed that it can be extracted from what Riemann wrote. This is a risky implication. Thus, why does Grothendieck make that implication? We could find a partial answer again in the letter to Yamashita from September 16th:

> "*It would almost seem that there has been something like 'telepathic' communication (to give it a name) from Riemann's thought-world to mine. I am still convinced now, or rather, I well know, that I never discovered myself this idea (because this kind of discovery of some significant view, I mean the moment of such discovery, one never forgets I believe) but that I got it from Riemann, and with some surprise.*"

This passage is very beautiful. That "telepathic communication" between Grothendieck and Riemann, two of the biggest mathematical geniuses of all times, connected by their mathematical passion and the wish of understanding the notion of space, is one of the most remarkable examples of mathematical creativity. Despite of being separated by many years, both of them had the intention of unveiling the inner and deep structure of the space and they walked together going behind that goal. Thus, without doubt Grothedieck's ideas about the space, and in particular the geometry of forms he envisioned in the late 1980s, are Riemann's heritage.

5 The Notion of Space: Mathematics, Physics, and Philisophy

In the previous section we discussed how thanks to the formulation of his geometry of forms Grothendieck began interested in the relations between the physical and the mathematical space. It is also important to remark that in fact Grothedieck was

beyond mathematics and physics in his reflections and also consider the importance of philosophy in order to understand what a space is. This give us a new perspective about Grothedieck's thoughts in the late 1980s and also about his understanding of the notion of space as a whole. In particular, Grothendieck was interested in the role of mathematical models for the physical reality and the philosophical background of them. He thought that only putting together philosophy, physics, and mathematics we can find an adequate approach to the problem of space. Regarding this, in (Grothendieck n.d.) he wrote:

> "*Pour résumer, je prévois que le renouvellement attendu (s'il doit encore venir) viendra, plutôt d'un mathématicien dans l'âme, bien informé des grands problèmes de la physique, que d'un physicien, Mais surtout, il y faudra un homme ayant 'l'ouverture philosophique' pour saisir le noeud du problème. Celui-ci n'est nullement de nature technique, mais bien un problème fondamental de 'philosophie de la nature'.*"

And in a letter to AY on June 24th, 1991 he also pointed out:

> "*Je suspecte que les nouvelles structures à dégager seront beaucoup plus subtiles qu'un simple paraphrase de modèles continus connus en termes discretes. Et surtout, qu'avant toute tentaive de dégager des nouveaux modèles, présmés meilleurs que les anciens, il símpose de poursuivre une réflexion philosophico-mathématique très servie sur la notion même de 'modèle mathématique' de quelque aspect de la realité - sur son rôle, son utilité, et sus limites*."

It is beautiful to see how Grothendieck uses the expression natural philosophy. The problem of space is not just a mathematical, physical of philosophical one, it is a problem of natural philosophy in a very broad sense that includes mathematics, physics, and philosophy but also goes beyond them. This is evidence of how deep Grothendieck's thoughts about the space became. The philosophical side of Grothendieck's approach to the notion of space is, in general, unknown but for us it is very relevant and it should be studied in more detail. Here, we just mention this aspect to motivate the study and the debate about it.

Grothendieck pointed out that the problem of space is not only of technical nature. It is not enough to think of the problem in mathematical or physical terms. We also need a right conceptual approach and it is in this point where philosophy should enter. One of the aspects to be consider is how the mathematical models of the space that we can propose reflect the nature and the structure of the physical space and it is consistent philosophically. Of course, this is not an easy problem and Grothendieck had no answer. Again, the reflection on the role of mathematical models is of philosophical nature as Grothendieck saw (see Grothendieck n.d.):

> "*J'ai le sentiment que la r'eflexion fondamentale qui attend d'être entreprise, aura à se placer sur deux niveaux différents.*
> 1. *Une réflexion de nature 'philosophique', sur la notion même de 'modèle mathématique' pour une portion de la réalité. Depuis les succès de la théorie newtonienne, c'est devenu un axiome tacite du physicien qu'il existe un modèle mathématique (voire même, un modèle unique, ou 'le' modèle) pour exprimer la réalité physique de façon parfaite, sans 'd'ecollement' ni bavure. Ce consensus, qui fait loi depuis plus de deux siècles, est*

comme une sorte de vestige fossile de la vivante vision d'un Pythagore que 'Tout est nombre'. Peut-être est-ce là le nouveau 'cercle invisible', qui a remplacé les anciens cercles métaphysiques pour limiter l'Univers du physicien (alors que la race des 'philosophes de la nature' semble définitivement éteinte, supplantée haut-la-main par celle des ordinateurs). Pour peu qu'un veuille bien s'y arrêter ne fut-ce qu'un instant, il est bien clair pourtant que la validité de ce consensus-là n'a rien d'évident. Il y a même des raisons philosophiques très sérieuses, qui conduisent 'a le mettre en doute a priori, ou du moins,à prévoir à sa validité des limites très strictes. Ce serait le moment ou jamais de soumettre cet axiome à une critique serrée, et peut-être même, de 'd'emontrer', au delà de tout doute possible, qu'il n'est pas fondé: qu'il n'existe pas de modèle mathématique rigoureux unique, rendant compte de l'ensemble des phénomènes dits 'physiques' répertoriés jusqu'à présent.

Une fois cernée de façon satisfaisante la notion même de 'modèle mathematique', et celle de la 'validite' d'un tel modèle (dans la limite de telles 'marges d'erreur' admises dans les mesures faites), la question d'une 'théorie unitaire' ou tout au moins celle d'un 'modèle optimum' (en un sens à préciser) se trouvera enfin clairement posée. En même temps, on aura sans doute une idée plus claire aussi du degré d'arbitraire qui est attaché (par nécessité, peut-être) au choix d'un tel modèle.

2. *C'est apr'es une telle réflexion seulement, il me semble, que la question 'technique' de dégager un modèle explicite, plus satisfaisant que ses devanciers, prend tout son sens. Ce serait le moment alors, peut-être, de se dégager d'un deuxième axiome tacite du physicien, remontant à l'antiquité, lui, et profondément ancré dans notre mode de perception même de l'espace: c'est celui de la nature continue de l'espace et du temps (ou de l'espace-temps), du 'lieu' donc où se déroulent les 'phénomènes physiques*."

We include this extensive quote because it illustrates in Grothendieck's own words what we are trying to argue, i.e., that the technical description of the space only makes sense for Grothendieck after a conceptual approach. It took him around 40 years to reach this conclusion. This quote also shows the evolution of Grothendieck's ideas from his pure technical questions in his work in algebraic geometry to the more physical-philosophical questions in his geometry of forms. Thanks to this, it is very clear that notion of space is indeed a leitmotif in Grothedieck's mathematical life.

6 Conclusion

Along this essay we have shown how the idea of space in Grothedieck evolved from his generalization of the notion of (algebraic) variety via schemes to the formulation of a new approach for the foundations of topology in what he called geometry of forms, passing through the notion of topos and tame space. Of course, for Grothendieck the notion of variety and topological space (in the classical sense) were useful in mathematics but he realized of new situations where new ideas on space were needed. In particular, one of these situations is related to the notion of form and its role in topology and, in general, in mathematics.

Grothendieck wrote that there are three main areas of mathematical reflection, namely arithmetic (associated to the notion of integer numbers), analysis (associated to the notion of magnitude), and geometry (associated to the notion of form). From

the beginning he put the notion of form as belonging to the realm of geometry and more general topology, but even more, it is the central notion for geometry in the sense that geometry for him is the branch of mathematics in charge of studying the aspects of the universe, not just the mathematical universe of course, related to the form.

Following that line, for Grothendieck there are three different notions which are important from a mathematical point of view, namely number, magnitude, and form. In (Cruz Morales Submitted) we argue that number and magnitude are manifestations of a more general notion of form and in this sense number can be thought as an arithmetical form while magnitude can be thought as an analytic form. What Grothendieck called form should be interpreted as a geometrical form (or shape). What we have is a fundamental triad that gives the most important aspects to be considered in a mathematical approach to the notion of space, i.e., arithmetical, analytical, and geometrical aspects. We have argued that this was the quest of Grothendieck for more than 40 years.

On the other hand, as we have already mentioned before in this text, Zalamea's interpretation (Zalamea 2019) of the space-number in Grothendieck's work as analogue to the space-time in physics gives a nice way to understand Grothendieck's ideas. Zalamea proposes that we should think Grothendieck's work as an attempt to develop the notion of space-number as the conjunction of arithmetic and geometry in one entity. This analogy is very suggestive since, as we have shown, Grothendieck was interested in physical aspects of the space. For this reason, we find very plausible to think of his last works in terms of constructing a theory for a space-number. We think that this line of thought could be very helpful in order to understand the notion of space in Grothendieck and it is a future project try to follow it.

In (Bolzano 2004) Bolzano presents an idea of mathematics as a general theory of forms. In some sense Grothendieck's approach to the notion of space is a realization of Bolzano's idea, including physics and philosophy in the picture. We believe this should be a way to think of Grothendieck's contribution to the human knowledge. There are still many problems to be considered and many gaps to be filled in order to understand Grothendieck's ideas about the space. In this text we wanted to present a new perspective and contribute with some elements which, to the best of our knowledge, were not considered before in the literature. More than giving a definitive answer about the problem of space in Grothendieck's work we wanted to give a picture as complete as we could do it, in order to start the dialogue. A dialogue that follows Grothendieck's own advice will require a mathematical, physical, and philosophical way of thinking.

Acknowledgments The author wants to thank Fernando Zalamea for many illuminating discussions on Grothendieck's work over the years, his constant and warm support and having read a draft version of this essay. Of course, any mistake that remains is completely responsibility of the author. He also thanks Winfried Scharlau for providing him with Grothendieck's correspondence with Yamashita and Tsuji and for conversations on Grothendieck's life, and Yuri Manin for supporting the idea of writing this text. Finally, he thanks Pierre Deligne for many interesting emails and his patience in explaining his ideas.

References

Bolzano B (2004) On the concept of mathematics and its classification. In: The mathematical works of Bernard Bolzano (trans: Russ S). Oxford University Press

Cartier P (2001) A mad day's work: from Grothendieck to Connes and Kontsevich. The evolution of concepts of space and symmetry. Bull Am Math Soc 38:389–408

Connes A, Consani C (2014) The arithmetic site. arXiv:1405.4527

Cruz Morales JA (2017) Una visión del continuo en Grothendieck y el yoga de una "geometría de las formas": Apuntes para una reflexión. En: Grothendieck. Visiones sobre la multiplicidad de su obra, Editorial Nomos, Bogotá

Cruz Morales JA (2018) On "topology of the forms" of Grothendieck: a first view. Communication at International Congress of Mathematicians, Rio de Janeiro

Cruz Morales JA (Submitted) On the notion of form: a dialogue between Grothendieck and Thom

Deligne P (1998) Quelques idées maîtresses de l'oeuvre de A. Grothendieck. In: Matériaux pour l'histoire des mathématiques au XX siècle, Actes du colloqne à la mémoire de Jean Dieudonné (Nice, 1996). Séminaires et congrès. SMF

Deligne P (2020) Private communication

Grothendieck A (1958) The cohomology theory of abstract algebraic varieties. In: Proceedings international congress of mathematicians. Cambridge University Press, Cambridge

Grothendieck A (1986) Vers une géométrie des formes. Université de Montpellier. Archives Grothendieck. https://grothendieck.umontpellier.fr/

Grothendieck A (1997) Esquisse d'un programme. En: Geometric Galois actions. 1. Around Grothendieck's Esquisse d'un programme, London mathematical society lecture notes series 242. Cambridge university Press, Cambridge

Grothendieck A Recoltés et Semailles. Réflexions et témoignages sur un passé de mathématicien. Unpublished manuscript

Jackson A (2004) Comme Appelé du Néant— as if summoned from the void: the life of Alexandre Grothendieck 2. Not Am Math Soc 51(10):1196–1213

Lochak P, Schneps L (1997) Geometric Galois actions. 1. Around Grothendieck's Esquisse d'un programme, London Mathematical Society lecture notes series 242. Cambridge University Press, Cambridge

Riemann B (1854) On the hypotheses which lie at the foundations of geometry. Habilitation thesis

Ruelle D (2007) The Mathematician's brain: a personal tour through the essentials of mathematics and some of the great minds behind them. Princeton University Press, Princeton

Scharlau W (2010) Wer ist Alexander Grothendieck? Anarchie, Mathematik, Spiritualität, Einsamkeit. Eine Biographie. Teil 3: Spiritualität

Scharlau W (2020) Private communication 2016

van den Dries L (1998) Tame topology and -minimal structures, London Mathematical Society lecture notes series 248. Cambridge University Press, Cambridge

Zalamea F (2019) Grothendieck. Una guía a la obra matemática y filosófica. Edtorial Nomos, Bogotá

Geometry in the Twentieth Century: A Return to Euclid – The Work of Herbert Busemann

Athanase Papadopoulos

Contents

1 Introduction 1886
2 Leibniz, Busemann, and a Question of Euclidean Geometry 1887
3 Busemann's Doctoral Dissertation 1889
4 Classical Notions: Questions and Results 1891
5 The Sign of the Curvature, Angle Excess, and Area of Triangles: Replacing Equalities by Inequalities 1892
6 Convexity Theory 1894
7 Busemann's Work Recognition 1897
8 On the Work of A. D. Aleksandrov 1897
9 In Guise of a Conclusion 1899
References 1900

Abstract

This is an exposition of the work of Herbert Busemann (1905–1994), seen as a return to the geometry of Ancient Greece. The importance of this work, its recognition, and its relation with other works are discussed.

Keywords

Metric geometry · Metric spaces · G-spaces · Herbert Busemann

AMS Classification

01A60 · 53C70 · 53C45 · 54E35

A. Papadopoulos (✉)
Institut de Recherche Mathématique Avancée (IRMA) and Centre de Recherche et d'Expérimentation sur l'Acte Artistique (ITI CREAA), Université de Strasbourg and CNRS, Strasbourg Cedex, France
e-mail: athanase.papadopoulos@math.unistra.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_60

1 Introduction

The twentieth century saw the spectacular development of several fields in mathematics: algebraic topology, low-dimensional manifolds, algebraic geometry, number theory, dynamical systems, mathematical physics, to name a few. Among these fields, there is one which, in my opinion, is not sufficiently emphasized: metric geometry. In writing this, I am thinking of the work of Herbert Busemann, whose collected papers have been published recently in Springer's Collected Works Series (Busemann 2018a, b). Busemann's work goes against the tide, and I think that it is still under-evaluated. This is why I would like to highlight it here.

The path opened up by Busemann is in a sense a return to Euclid's geometry. I will explain in what sense, and in fact this is the theme of the whole present chapter. Let me first give some indications on what I mean by this return.

To start with, I would like to make it clear that in this chapter, when I talk about Euclid's geometry, I am not necessarily referring to the propositions found in the *Elements*, but rather to a geometry inspired by them and which is close to them in several respects. To explain this, let me start by recalling that an important (maybe the most important) part of the geometric books of Euclid's *Elements* is an exploration of the properties of triangles. Likewise, a significant part of Busemann's geometry includes a study of the geometry of triangles on surfaces (which are not necessarily the Euclidean plane): such a surface may be the sphere, or the hyperbolic plane, or a convex surface embedded in three-space, or any other two-dimensional manifold. In particular, classical "non-Euclidean" geometry, as a study of properties of non-Euclidean triangles, is part of Euclid's geometry as I intend it here. For example, Menelaus' *Spherics*, which I discuss in another chapter of this Handbook (see the chapter ► "Menelaus' Spherics in Greek and Arabic Mathematics"), being a treatise on the intrinsic properties of spherical triangles, is an example of what I call Euclid's geometry. Higher-dimensional hyperbolic and spherical geometry, in particular the study of non-Euclidean polyhedra, which generalize triangles, are also part of this geometry.

The geometry developed by Busemann is defined axiomatically – with a few number of axioms: these include those of a metric space, to which are added some axioms on the existence of shortest lines joining pairs of points. Thus, the notion of distance between two points is one of the basic notions in that geometry. I recall incidentally that Euclid's *Elements* do not use the notion of distance between two points. From the philosophical point of view, this is important and needs a thorough discussion, but for what concerns us here, this is not an issue. Furthermore, in Busemann's geometry, considerations of differentiability are excluded. Thus, the methods of differential calculus are not part of it. Consequently, Riemannian geometry is not part of this geometry either.

Why call such a geometry "Euclidean"? Because it deals with problems similar to those raised by Euclid's *Elements* or motivated by them. I shall start by explaining this in the next section, using a significant example, namely, a work of Busemann arising from a question raised by Leibniz. At the same time, the result I am about to describe will show how non-Euclidean geometry naturally fits into this framework.

2 Leibniz, Busemann, and a Question of Euclidean Geometry

Leibniz addressed the question of defining a plane. His work on this subject is reproduced in his *Mathematical writings* (Leibniz 1849, p. 166). The problems raised by such a definition are presented and commented on by Busemann in his article *On Leibniz's definition of planes* (Busemann 1941) and are taken up again in his book *The geometry of geodesics* (Busemann 1955, p. 46). The question of the definition of a plane can be traced back to Euclid, although, as Busemann remarks, the definitions of surface and line that are given in the *Elements* are difficult to work with. We recall that Euclid, in Book I of the *Elements*, defines a *line* as something "that has length and no width," but the terms length and width are not defined there. Likewise, he defines a *plane* as something "that has only length and width." Busemann notes that ancient authors already found these definitions objectionable, and that it is not possible, within the framework of Euclidean geometry, to characterize, based only on these definitions, straight lines among curves, and planes among surfaces. Thus, one cannot give an answer to Leibniz's question based only on Euclid's *Elements*. Archimedes went further than Euclid. In *The sphere and the cylinder*, following a definition of straight line due to Heron, he introduced as an axiom the fact that the straight line is a shortest curve among all curves joining two points. He also developed a method for calculating lengths of curves. All this is recalled in the same passage by Leibniz, who considers that this cannot be used for a definition of a straight line, and proposes to define a straight line as the set of points fixed by the rotation of some body – which assumes that the notion of body precedes that of straight line, and which gives rise to other difficulties. I refer here to Couturat's comments on Leibniz's work, see Chap. IX, of his treatise *The logic of Leibniz* (Couturat 1901), titled *Geometric calculation*. Regarding Leibniz's question, Busemann writes in the introduction to his article (Busemann 1941):

> The well-known unsatisfactory Euclidean definitions of straight lines and planes have evoked early attempts to replace them by better ones. Leibniz proposed *to define a plane as the locus of points which have equal distances from two given points*, and the straight line as locus of points having equal distance from three given points. Since the idea of metric space or anything which could be substituted for it had not yet been developed, this attempt of Leibniz did not grow into a well-founded theory. The present note tries to give a critical investigation of the implications of Leibniz's definition from the point of view of metric spaces.

To address such a problem, and similar ones, Busemann introduced the axioms of a G-space (the letter G stands for "geodesic"). In fact, these axioms are already present in Busemann's doctoral thesis (defended in 1931), and they are the same ones he used throughout the rest of his works. This is the moment to introduce these axioms:

1. The space is equipped with a distance function $d(x, y)$.
2. Every bounded infinite set has an accumulation point. (Busemann calls a space satisfying this property *finitely compact*.)

3. For any distinct points x, z, there is a point y such that $d(x, y) + d(y, z) = d(x, z)$. (One says that *$y$ lies between x and z*.)
4. For every point p, there exists a positive real number ρ_p such that for any distinct points x, y in the ball of center p and radius ρ_p, there exists a point z such that y is between x and z. (This is a property of local extendability.)
5. (Uniqueness of extension of geodesics) For any distinct points x and y, if y is between x and z_1 and between x and z_2, and if $d(y, z_1) = d(y, z_2)$ then $z_1 = z_2$.

We note that it is not assumed here that the metric space is a manifold. In fact, it is an important open problem, which is presently solved only under some special extra hypotheses, to know whether a G-space is a manifold; see the reviews by Andreev (Andreev 2018) and Berestovskii (Berestovskii 2018a) who both made important contributions to this problem.

The notion of G-space is used extensively in many of Busemann's writings, e.g., in his 1942 book *Metric Methods in Finsler Spaces and in the Foundations of Geometry* (Busemann 1942), in his 1943 paper *On spaces in which points determine a geodesic* (Busemann 1943), and in his 1955 book *The geometry of geodesics* (Busemann 1955).

Busemann and B. B. Phadke (one of his former students) write in (Busemann and Phadke 1993, p. 190): "One of the most attractive features of G-space theory is that it maintains a strong connection with the foundations of geometry."

With the above axioms in the background, Busemann can address Leibniz's question of defining a plane. He calls a subset of a G-space "flat" if, equipped with the induced metric, it is a G-space. In his paper (Busemann 1941) and in his book (Busemann 1955), §§46 and 47, he considers Leibniz's question in the setting of G-spaces to which a single hypothesis is added; it concerns *bisectors*, that is, loci of points equidistant from two distinct points. The hypothesis says the following:

> *Any bisector contains with any two points at least one segment connecting them.*

Busemann calls this property *flatness* of the bisectors. It is obviously satisfied in the case of the classical geometries of constant curvature. He proves the converse:

> *The three classical spaces of constant curvature (Euclidean, spherical, and hyperbolic) are the only G-spaces in which all bisectors are flat.*

Thus, motivated by Leibniz's question, Busemann deduces a remarkable (probably the simplest) characterization of the three classical geometries. He writes in (Busemann 1955, p. 307), considering such a characterization of the classical spaces: "Besides being of considerable historical interest, these results form perhaps the most direct and conceptually simplest access to the elementary geometries, and they have many important applications."

In his book *The geometry of geodesics* (Busemann 1955, p. 310), Busemann shows that with a slight modification of the notion of flatness, one gets a *local* characterization of the three classical geometries. In fact, he gives a characterization

of spaces that are *locally* the above spaces of constant curvature, based on a "local bisector theorem," a criterion involving a property of small enough open balls; see Theorem 46.1 of (Busemann 1955). It is interesting that Busemann obtains the fact that the metric he ends up with is either the Euclidean, hyperbolic or spherical by establishing the trigonometric formulae for right triangles in such a space, see (Busemann 1955, p. 320). Busemann calls the *Helmholtz–Lie problem* the question of characterizing the three elementary geometries. From the above bisector theorem characterizing the three classical geometries, he deduces several corollaries, in particular, statements concerning motions and the Helmholtz–Lie problem, of which the following very general statement is given (Busemann 1955):

> *If in a G-space, for any four points a, a', b, c satisfying $d(a, b) = d(a', b)$ and $d(a, c) = d(a', c)$, there exists a motion leaving b and c fixed and carrying a to a', then the space is either locally Euclidean, hyperbolic or spherical.*

Another application of the bisector theorem is the fact, obtained in (Busemann 1955), see also (Busemann 1947), that in dimension two, the three elementary geometries are the only geometries in which the area of a triangle is expressible in terms of its sides. More precisely, Busemann proves the following:

> *If every point of a G-surface has a neighborhood in which a notion of area for triangles is defined and has the same value for all isometric triples, then the surface is locally Euclidean, hyperbolic or spherical.*

3 Busemann's Doctoral Dissertation

In one of his last papers, written with Phadke and titled *Novel results in the geometry of geodesics* (Busemann and Phadke 1993), Busemann recalls that he already introduced the axioms of a G-space (without the name) in his doctoral dissertation, defended in 1931. The history of this dissertation is interesting. The thesis supervisor was Richard Courant. Busemann declares in his autobiographical memoirs that he wrote his thesis "against" Courant. He recounts the following:[1]

> Officially, I took my degree with Courant. This was only "officially," in the sense that I was really inspired by [Pavel Sergueevitch] Aleksandrov, who visited Göttingen regularly. He gave me the idea of the subject of the thesis. I wrote it, but of course he could not be the official reviewer of my thesis, so he was my co-referee. I must say that my thesis was partly in protest against Courant. I went to Courant originally, he gave me something and it turned

[1] Unless otherwise stated, Busemann's quotations are taken from the recording of an interview with Constance Reid, presumably made on April 22, 1973 and kept at the library of the University of Göttingen. I would like to thank again Herbert Goenner for providing me with a copy of this interview. I presented this recording in more detail in the biography I wrote on Busemann, included in the first volume of his *Collected works* (Busemann 2018a, p. 3–12).

> out to be much easier than what he thought. I did it and it became a small paper,[2] but not enough for a thesis. Then he gave me something else, I did this too, but then Kolmogorov came to Göttingen, Courant showed it to him, and Kolmogorov said: "Oh, that's all very fine, but it's well known." So I became really mad, and I went away to Rome. I was angry with Courant, and wrote my thesis there, on my own.

According to the paper by Busemann and Phadke (Busemann and Phadke 1993, p. 181), Busemann became convinced of the importance of non-Riemannian metrics after he read the beginning of Minkowski's *Geometrie der Zahlen* in 1926. At the same time, he heard a course on point set topology and learned about Fréchet's concept of metric spaces.[3] We read in that paper:

> The older generation ridiculed the idea of using these spaces as a way to obtain results of higher differential geometry. But it turned out that a few simple axioms on distance suffice to obtain many nontrivial results of Riemannian geometry and, in addition, many which are quite inaccessible to the classical methods.

Busemann wrote his thesis by himself, while he was visiting Rome, and he sent his work to Courant. The latter responded in a letter dated August 8, 1930 that this "geometric work" is sufficient for a "reasonable" dissertation.[4] Let me say a few words about the results of the thesis.

The subject of Busemann's thesis is the axiomatic characterization of Minkowski spaces, that is, metric spaces that underlie finite-dimensional vector spaces equipped with a norm.

The paper *Über die Geometrien, in denen die "Kreise mit unendlichem Radius" die kürzeten Linien sind* (On the geometries where circles of infinite radius are the shortest lines) is extracted from that thesis. An English translation of this paper is contained in Vol. I of Busemann's *Selected Works* (Busemann 1932). Here, the term "circle of infinite radius" means a horocycle, that is, the limit of a family of circles passing through a given point and whose center tends to infinity along a geodesic. The notion of horocycle (and its higher-dimensional analog, horosphere) was first used in hyperbolic geometry and it plays a crucial role there. In Euclidean geometry, the limit of such a family of circles is a straight line. The main result that Busemann proves in the paper (Busemann 1932) is that a two-dimensional space where horocycles are straight lines is isometric to a Minkowski space.

The metric characterization of Minkowski spaces remained one of Busemann's main objects of study. I refer the interested reader to the survey titled *Busemann's work on the characterization of Minkowski geometries* (Papadopoulos 2018a) which appeared in the section containing preliminary essays of Volume II of Busemann's *Selected Works* (Busemann 2018b).

[2] Busemann talks here about his paper (Busemann 1930).

[3] The concept of metric space as we intend it today, was first introduced (without the name) by M. Fréchet in his thesis, defended in 1906 (Fréchet 1906).

[4] An English translation of Courant's report on Busemann's dissertation is contained in Volume II of Busemann's Selected Works.

In the next section, I will mention other problems considered by Busemann and results he obtained in metric geometry.

4 Classical Notions: Questions and Results

The questions and results I mention in this section may seem unrelated but they all pertain to classical geometry, and many of them date back to the time of Euclid. They are all topics on which Busemann worked. His methods of proof, making use only of elementary geometry, are often ingenious.

1. A distance characterization of the three classical geometries. Since we already talked about the characterization of the three constant curvature geometries, I start by giving another local distance characterization of these three geometries (Busemann 1955, p. 47):

   If every point in a G-space has a neighborhood in which every quadruple of point is isometric to a quadruple of points in a Euclidean, hyperbolic or spherical space (the choice of such a space may a priori *depend on the point), then the space is locally Euclidean, hyperbolic or spherical.*
2. A theory of spaces with prescribed geodesics, and in particular, a thorough study of Hilbert's Problem IV. Roughly speaking, this problem asks for a determination of all metrics on subsets of Euclidean space whose geodesics are the (Euclidean) straight lines. There are several papers by Busemann on this problem, e.g., (Busemann 1939; Busemann 1976), see also the survey (Papadopoulos 2014). Busemann was the main promoter of Hilbert's Problem IV, and his work acted as a catalyst to the solution of that problem that was given by A. V. Pogorelov in 1973. The latter writes in the introduction to his monograph *Hilbert's fourth problem* (Pogorelov 1979): "The occasion for the present investigation is a remarkable idea due to Herbert Busemann, which I learned from his report to the International Congress of Mathematicians at Moscow in 1966."
3. A theory of perpendicularity in G-spaces and conditions under which this relation is symmetric (Busemann 1961).
4. The question of convexity of metric spheres (that is, sets of points whose distance to a given point is constant) in G-spaces (Busemann and Phadke 1980; Busemann and Phadke 1983; Busemann and Phadke 1993).
5. A theory of parallelism. In the paper (Busemann and Phadke 1993, p. 190), Busemann and Phadke write: "It is hard to understand that the theory of parallels was almost completely neglected in the classical theory after Lobachevsky's discoveries. We now explain the concepts upon which a beautiful theory of parallels can be built in a very general setting." In his 1933 paper *On spaces with convex spheres and the parallel postulate* (Busemann 1933), under the assumption of convexity of metric spheres and the fact that the parallel postulate is satisfied, Busemann studied conditions under which a space can be mapped onto a Euclidean space such that the geodesics are mapped to straight lines. Here, the

fact that the *parallel postulate* holds means that if g is an arbitrary straight line (that is, a geodesic isometric to the real line) and P a point not on g, then for every sequence of points x_n on g converging to infinity, the line Px_n converges to a unique geodesic. Parallelism is also the subject of Chap. III of *The geometry of geodesics* (Busemann 1955), titled *Perpendiculars and parallels*, see also the papers (Busemann 1979; Busemann and Phadke 1993) and the survey (Busemann 1961). A theory of parallelism in spaces with nonpositive curvature is developed in (Busemann 1955, p. 248).

6. Angle bisectors in triangles. In his paper *Planes with analogs to Euclidean angular bisectors* (Busemann 1975), Busemann gives a characterization of Minkowski planes based on a property of angular bisectors of triangles, namely, the fact that an angle in a triangle divides the opposite side in the ratio of the adjacent sides.
7. A theory of length-preserving maps. Here, Busemann's original aim is to determine all locally finite length-preserving self-maps of the Euclidean, hyperbolic, and spherical spaces (Busemann 1964). The first example that he gives of a length-preserving map is folding a piece of Euclidean paper. Among the results he proves is that any finitely compact space with an intrinsic metric and a pairwise transitive group of motions that possesses a proper length-preserving map has constant curvature.
8. The question of existence of axes of motions, the theory of axial motions, translation distances, asymptotes to axes, and parallelism of axes, see (Busemann 1955, 1970, Busemann and Phadke 1993).
9. A theory of horospheres (which he calls limit spheres) (Busemann and Phadke 1993).
10. A theory of asymptotes, and the relation between asymptotes and limit spheres (Busemann 1933, 1955, 1961). The asymptoticity relation is in general neither symmetric nor transitive. Symmetry and transitivity hold in spaces which Busemann calls *straight Desarguesian* (Busemann 1955, p. 139).
11. A study of the properties of fundamental groups of compact spaces of negative curvature (Busemann 1961).
12. The introduction of a theory of timelike spaces: a metric setting for the theory of general relativity, see (Busemann 1967).

The next notions need special attention, and I include them in separate sections.

5 The Sign of the Curvature, Angle Excess, and Area of Triangles: Replacing Equalities by Inequalities

Busemann introduced a notion of sign of the curvature (negative, nonpositive, vanishing, nonnegative, and positive) in metric spaces. This theory is developed in his book *The geometry of geodesics* (Busemann 1955, p. 236 ff.) and in other works. It is a generalization of the notion of sign of the curvature in Riemannian manifolds in the sense that it coincides with this notion when the underlying space is a

Riemannian manifold. Let us recall the definitions. In fact, the theory works for a class of metric spaces that are more general than G-spaces: for most purposes it suffices that the space is *midpoint convex* in the sense that for any two points x and y, there exists a point z whose distances from x and from y are both equal to half of the distance from x to y. Such a point z (which is generally not unique) is called a *midpoint* of x and y. A G-space is a special case of a midpoint convex space. A midpoint convex space which is complete is geodesic (any two points can be joined by a geodesic).

Busemann says that a midpoint convex metric space (M, d) is *nonpositively curved* if for any three points a, b, c in M, if a' is a midpoint of a and c and b' a midpoint of b and c, then

$$d(a', b') \leq 1/2(d(a, b)).$$

There are variations of this definition, namely, the space is called *negatively curved, nonnegatively curved, or positively curved* if the above inequality is replaced by a strict inequality or by the reverse large or strict inequality respectively. If the inequality is replaced by an equality, the space is said to be *zero curved*. These notions are used today under the name *Busemann nonpositive (resp. nonnegative, positive, negative, zero) curvature*.

The inequality expressing negative, nonpositive, nonnegative, positive curvature replaces an equality valid in Euclidean geometry by an inequality. Obtaining theorems in curved spaces from analogous theorems in Euclidean geometry by replacing equalities by inequalities, is known since antiquity. This is highlighted in my article (▶ "Menelaus' Spherics in Greek and Arabic Mathematics, this handbook) on the work of Menelaus of Alexandria (2nd c. AD). One proposition of Menelaus' *Spherics* says that the sphere is negatively curved in the above sense (see Proposition 27 of (Rashed and Papadopoulos 2017)), but there many other propositions involving this notion. In the *Geometry of geodesics*, Busemann proves that a G-space has nonpositive curvature in the above sense if and only if every point has a neighborhood in which all the triangles of sides α, β, γ and angles A, B, C opposite to these sides satisfy the *cosine inequality*

$$\gamma^2 \geq \alpha^2 + \beta^2 - 2\alpha\beta \cos C,$$

that is, if the usual Euclidean cosine formula holds, but where the equality has been replaced by an inequality (this is Proposition 41.3 of (Busemann 1955)). In the same book, Busemann studies several other metric notions of curvature, see in particular Chap. 5. One of them uses the notion of angle excess of triangles in G-spaces, see (Busemann 1955, p. 279ff). He also gives several statements that make relations between the theory of parallels and that of angle excess. Let me recall here that the fact that the angle excess of a spherical triangle is positive is also a result of Menelaus, see (Rashed and Papadopoulos 2017, Proposition 12). In hyperbolic geometry, angle excess is negative; this was known to J. H. Lambert, for whom the existence of hyperbolic geometry was hypothetical, see §81 of his *Theory of*

parallels (Lambert 1766), in which he defines the area of a triangle in the hyperbolic plane as the *angle defect* of that triangle. Building a theory of angle excess in G-spaces is one of the themes of Chap. V of Busemann's *Geometry of geodesics* (Busemann 1955), titled *The influence of the sign of the curvature on geodesics*.

Talking about area, let me quote another result of Busemann giving a characterization of Minkowski spaces in terms of a property of areas of its triangles. This is Theorem 51.5 of (Busemann 1955). The theorem concerns G-spaces that are *straight*, that is, in which the property of local extendability in the axioms of a G-space is replaced by a global one. The result says that in a straight space, the fact that the area of any triangle ABC depends only on the distances BC and the distance from A to the segment BC is equivalent to the fact that the space is a Minkowski plane with symmetric perpendicularity.

6 Convexity Theory

Among the introductory essays included in Volume II of the *Selected Works*, there is one titled *Busemann and Feller on curvature properties of convex surfaces* (A'Campo Neuen and Papadopoulos 2018), written by Annette A'Campo and the author of the present chapter. It is a commentary on the paper *Krümmungseigenschaften konvexer Flächen* (Curvature properties of convex surfaces), written in 1935, by Busemann and Feller (Busemann and Feller 1936).

This work of Busemann and Feller is strongly influenced by Hilbert's vision of mathematics. Both Busemann and Feller obtained their PhD in Göttingen, which, in those times, was one of the two most important mathematical centers of the world (the other one being Paris). Although their thesis advisor was Courant, both Busemann and Feller were shaped by Hilbert's view on the foundations of geometry. Hilbert retired in 1930 but continued to give a course on the philosophy of mathematics. Feller obtained his PhD in 1927, and Busemann in 1931. Before talking about their work, let me say a few words on some work of Hilbert related to the same topic.

In 1932, together with Stephan Cohn-Vossen,[5] Hilbert published his remarkable book *Geometry and the Imagination*[6] (Hilbert and Cohn-Vossen 1932) which is based on Hilbert's lectures held in 1920/21. The book contains a chapter on the geometry of surfaces, but the point of view is purely geometric, unlike the one of Euler, Monge, and most of the other mathematicians who followed, who studied surfaces and curves using analytic geometry. The methods used in the chapter on differential geometry of Hilbert and Cohn-Vossen's book constitute a return to the methods of the Greeks. For instance, their description of the osculating circle at a point of a plane curve starts with: "We draw a circle through P_1 and two neighboring

[5] For a comprehensive review on Cohn-Vossen's work, we refer the reader to the article (Aleksandrov 1947) by A. D. Aleksandrov.

[6] The original German title is: Anschauliche Geometrie.

points on the curve. If we let the two neighboring points approach P_1, the corresponding circle converges to a limit position [...]. This limit circle is called the osculating circle at P_1." No differentiation is involved. The discussion then is generalized to space curves. The osculating plane of a space curve at a point is constructed as the limit of the plane through the tangent at P and a neighboring point that approaches P (Hilbert and Cohn-Vossen 1932, p. 158). After a study of intrinsic properties that characterize the sphere, Hilbert and Cohn-Vossen in (Hilbert and Cohn-Vossen 1932), §32 study convex surfaces sharing the same properties.

Between the work of Hilbert and Cohn-Vossen and that of Busemann and Feller on convexity, I mention the work of Bonnesen and Fenchel on the same topic. They wrote a book titled *Theorie der konvexen Körper* (The theory of convex bodies) which was published in 1934, the year Werner Fenchel, who also belonged to the Göttingen school, moved to Denmark. This book, which is one of the founding texts of modern convexity theory, is also written in the tradition of Hilbert's foundations of geometry. Section 17 starts with the following: "Far-reaching assertions can be made about the curvature relations of convex curves without making any assumption of differentiability" (Bonnesen and Fenchel 1934, p. 153). Indeed, the authors, in that section and the next one, develop a geometric theory of curvature of convex curves and surfaces without differentiability conditions.

Busemann's works on convexity is in the same style. His first published work on the subject is the joint paper (Busemann and Feller 1936) with Feller.[7] In this paper, the authors study curvature properties of convex surfaces in three-space. Notions like tangent line, tangent plane, osculating circle, radius of curvature, etc. are studied using their original geometrical meaning rather than differential calculus. Busemann and Feller prove that a convex surface has almost everywhere second-order derivatives. (It was only known before that such a surface is almost everywhere differentiable, cf. (Reidemester 1921).) Among the other results, they obtain that there is a well-defined nearest point projection of curves lying in the complement of a convex body bounded by a convex surface onto this convex surface which is length nonincreasing. They also study geodesics on convex surfaces. They describe their goal in the introduction as follows: "The aim of this paper is to develop the elements of differential geometry of convex surfaces without making the usual regularity assumptions. Hereby necessarily, purely geometric considerations often replace the usual analytic arguments."[8] They show that without any differentiability hypothesis, one can define a notion of curvature and obtain the analogs of the classical theorems of Euler, Meusnier, and others on the differential geometry of surfaces. Their arguments are synthetic. The theorems of Meusnier and Euler are recurrent in

[7] William Feller (1906–1970) studied in Göttingen, obtaining his PhD with Courant, like Busemann. Feller obtained his PhD in 1927, and Busemann in 1931. They both emigrated to the USA after the Nazis formed a coalition in Germany. They wrote seven joint papers, published between 1934 and 1945. The reader interested in this turbulent period of the history of Europe in relation with mathematics may refer to Feller and Busemann's biographies contained in their Selected Works editions, (Feller 2015, Vol. 1) and (Busemann 2018a).

[8] Cf. the English translation in (Busemann, 2018a, p. 235).

Busemann's work. In a paper he wrote 15 years later (Busemann 1950), Busemann introduced a new notion of curvature in Minkowski spaces, and to justify the relevance of this notion, he showed that analogs of the theorems of Meusnier and Euler hold in this new setting as well. The study of the differential geometry of curves and surfaces, based on the notion of tangent, normal, order of contact, etc. of plane curves finds its origin in the work of Archimedes and Apollonius (third and second centuries B.C.) and of other Greek geometers.

This work of Busemann and Feller was the beginning of a lifelong program conducted by Busemann on the study of convexity theory using a minimal amount of differential geometry or differential calculus. He worked later on various geometric aspects of this subject, including volume, Brunn–Minkowski theory and convexity in general metric spaces. The subject of convexity dates back to Greek antiquity. Busemann, in his study of convex curves in G-spaces in his book *The geometry of geodesics* (Busemann 1955, p. 154), after he proves that a convex simple closed curve in a two-dimensional G-space has a supporting line at each point, writes that "it is remarkable that an exact proof of this proposition in the plane was already given by Archimedes." On the notion of angle, tangents, and contact in Greek mathematics, I refer the reader to the remarkable editions by R. Rashed, (Rashed 2008, 2009a, b, 2010, 2015; Apollonius 2008).

In 1956, together with C. M. Petty, Busemann formulated a set of problems on convex bodies, in a paper titled *Problems on convex bodies* (Busemann and Petty 1956). Most of these problems are still open, see the commentary by D. Burago (Burago 2018) in Volume II of Busemann's *Selected Works*. The latter, in the conclusion to his article, writes that "[he believes] that this circle of problems posed by Busemann opened a can of worms, a new direction or perhaps even a new area of research which is likely to be very active for many years to come."

In 1958, Busemann published a book titled *Convex surfaces* (Busemann 1958). In the Introduction, he declares that his purpose is to present a subject, convex surfaces, which "during the past 25 years has experienced a striking and beautiful development, principally in Russia, but has remained largely unknown, at least in the USA." Indeed, the book is a tribute to the work of A. D. Aleksandrov. The following paragraph gives us some information on how Busemann presented his book project to the publisher:

A letter dated July 23, 1956, from the editor-in-chief of Interscience Publishers Inc., addressed to Busemann, starts with:[9] "It is with very special pleasure indeed that we learn from Dr. Stocker of your willingness to present in our Tracts series an account of important recent development of the work of certain Russian geometers." The letter refers to the work of A. D. Aleksandrov and the school he founded in Russia. I am not going to attempt any review of the results of Aleksandrov and his school because this would lead us much too far, but I shall mention some of this work later in this chapter. It is interesting to know that Busemann's book on convexity which was initially planned to be a review of the work of the Russian school on this subject, was eventually translated into Russian in 1964, see

[9]The letter is part of Courant's correspondence kept at the Elmer Holmes Bobst Library in New York. Carbon copies of the letter were sent to Courant and to Bers.

(Busemann 1958). The book contains several important results on convex bodies. Among these results, let me mention a proof of a theorem in convexity theory that was independently obtained by Busemann and Fenchel and Jessen (Fenchel and Jessen 1938), stating that if two convex bodies in Euclidean n-space have equal p-th order area measure (we shall not go into the definition here) for some p satisfying $1 \leq p \leq n - 1$, then the two bodies are the same up to a translation.

7 Busemann's Work Recognition

The recognition of Busemann's work came from the Soviet Union. In 1985, he was awarded the Lobachevsky Prize "for his innovative book *The Geometry of geodesics.*" The book was written 30 years before. The Lobachevsky Prize (or Lobachevsky medal, depending on the epoch and on the institution which delivered it), is still a prestigious prize. Its first recipients were Sophus Lie in 1897, Wilhelm Killing in 1900, and David Hilbert in 1903. A. D. Aleksandrov obtained it in 1951. The list of the more recent recipients includes Weyl, Pontryagin, H. Hopf, P. S. Alexandroff, Efimov, Kolmogorov, Hirzebruch, Arnold, Margulis, Gromov, and Chern.

Busemann's work was not appreciated for its true value in the West. In the three-volume set titled *A Century of Mathematics in America* (Duren et al. 1988) published in 1988 by the AMS, there is no mention of Busemann's work. In a letter to V. Pambuccian, dated November 26, 1989 and published in Volume II of Busemann's *Selected Works* (Busemann 2018b, p. 64), Busemann writes: "... Your complaint regarding the neglect of synthetic geometry in the USA is unfortunately justified. *The Geometry of Geodesics* (GG) and *Recent synthetic Differential Geometry* are still available, however for ridiculously high prices. . . . The absence of interest in the USA is recognizable from the very fact that every year there are approximately twice as many copies sold abroad than in the USA. And the factor two is probably lower than the real one, for three of my books were translated in Russian, and appeared in well-printed, large, very cheap editions, e.g., GG costs 1.88 rubles."

Busemann's work started to be recognized in the West only in the 1980s, when metric geometry was revived by M. Gromov, who introduced the methods of synthetic global geometry. In fact, metric geometry became fashionable in France after Gromov started giving courses on this topic in Paris at the end of the 1970s. His 1979 course was published in several editions, see (Gromov 1981). Igor Belegradek, in his *Math Reviews* review of the first English edition of this book (1991), writes that it is "one of the most influential books in geometry in the last twenty years." This coincided with the last years of Busemann's career. Gromov is an heir of the Russian school of metric geometry founded by A. D. Aleksandrov. I will talk about the latter in the next section.

8 On the Work of A. D. Aleksandrov

A. D. Aleksandrov founded an important school on metric geometry, with a large number of students and collaborators. Like Busemann, he was interested in the most basic notions of geometry. Classical problems of convexity, isoperimetry, and

isoepiphany, on which he and the school he founded were working, became at the forefront of research. For him, like for Busemann, classical synthetic geometry and the old geometric problems that originate in Greek Antiquity were more important than Riemannian geometry based on linear algebra and tensor calculus. In some sense, his work was a return to Euclid and Archimedes.

Like Busemann, Aleksandrov's interest abides in the most basic notions of geometry. Aleksandrov's work was perpetuated by his students and his student's students. Yurii Burago, Yurii Borisov, Ilya Backelman, Alexei Pogrelov, Victor Zalgaller, Grigorii Perelman, Yurii Reshetnyak, Anton Petrunin, and Valerii Berestovskii are among his descendents – there are many others. Aleksandrov was an invited speaker at the 1960 ICM, where he delivered a talk titled *Modern development of surface theory* (Aleksandrov 1960) in which he presented an outline of his work together with that of his students, including the metric condition of the boundedness of curvature of triangles in terms of angle excess leading to this metric definition of curvature, and a variety of notions and results obtained under the condition of bounded curvature, including a definition of area, a study of convexity, embedding problems, parallel translation, and several others. We saw that in the paper (Busemann and Feller 1936), Busemann and Feller proved that a convex surface in three-dimensional space has a second order differential at almost every point, and that the Dupin indicatrix has the usual form.

In 1939, A. D. Aleksandrov extended this result to an arbitrary convex hypersurface in n-dimensional Euclidean space. I refer the reader to Aleksandrov's book *Intrinsic geometry of convex surfaces* (Aleksandrov 2005), in which he provided proofs of the theorems of Euler, Meusnier, and Rodrigues in the general setting of convex surfaces, attributing these generalizations to Busemann and Feller. In the same book, he proved that any convex two-dimensional manifold is nonnegatively curved in the sense he introduced and which we recalled above, and that every nonnegatively curved two-dimensional manifold homeomorphic to the sphere is isometric to some convex surface.

We mentioned Busemann's work on "timelike spaces," which provides a geometric setting to general relativity. Aleksandrov with his students developed a similar theory, which they called *chronogeometry*, see the review in (Papadopoulos 2018c) and the references there.

The metric notions of curvature introduced by Aleksandrov and by Busemann turned out to be at the bases of very important developments in geometry, and the terms "Busemann geometry" and "Aleksandrov geometry" refer to these works. Two AMS Mathematics Subject Classification codes carry the names of Busemann and Aleksandrov geometries[10] and there is a multitude of theorems and research problems that rely on their intrinsic definitions of the sign of the curvature.

In (Kutateladze et al. 2012), Kutateladze, Novikov, and Reshetnyak write:

[10]These are: 53C70: Direct methods (G-spaces of Busemann, etc.) and 53C45: Global surface theory (convex surfaces la A. D. Aleksandrov).

> It would be impossible to understand Aleksandrov's philosophy without turning to the roots of the science he loved best. In 1981 he wrote: "The spirit of modern mathematics lies in the tendency to return to the Greeks." His favorite motto was: "Back to Euclid!" Geometry was a part of the culture of the ancient world.

They also write:

> Aleksandrov steered mathematics toward the synthetic geometry of the ancients in a much subtler way than is now recognized. What is meant is not just a transition from smooth local geometry to geometry in the large without restrictions on differentiability. By enriching differential geometry with the tools of functional analysis and measure theory, he strove to regain the grandeur of mathematics in the times of Euclid. And by synthesizing geometry with other branches of twentieth-century mathematics, he aspired to the ancient ideal of a unified science.

See also the article (Katateladze 2012) with the significant title *Aleksandrov of Ancient Hellas*. Likewise, Zalgaller, in his paper (Zalgaller 2012) writes about Alexandrov: "When talking about methodology, he often pronounced the slogan: 'Retreat to Euclid!', fighting for visual methods."

9 In Guise of a Conclusion

Since his doctoral dissertation and until the end of his life, Busemann used to formulate problems and work on them without relying on the trends that were fashionable in his time. In an article that appeared in the Los Angeles Times on June 14, 1985, (Dembart 1985), on the occasion of the attribution of the Lobachevsky Prize to Busemann, the author writes: "Few mathematicians ever make it into public consciousness, but Busemann has had a hard time even within his own field, in part, at least, because he never worked on trendy problems and never followed the crowd." The same journalist quotes Busemann saying: "If I have a merit, it is that I am not influenced by what other people do."

A geometry freed from the dominion of differential calculus kept growing since the last quarter of the twentieth century, reinvigorated in the works of Thurston whose methods constitute a return to the synthetic methods of non-Euclidean geometry, and of Gromov, closer in spirit to the Busemann and Aleksandrov approaches. Problems on area and volume, isoperimetry, and convexity, became catalyzers of a new trend of mathematical research.

Finally, I started this article by mentioning Leibniz. I would like to end it by mentioning another mathematician-philosopher, René Thom.

Thom was one of the main founders of modern topology. His work on cobordism, for which he was awarded the Fields medal in 1958, inaugurated the combination of the methods of differential topology with those of algebraic topology. He writes in (Thom 1991, p. 10) that in his youth, "[he] had a certain taste for Euclidean geometry which immediately attracted [him] very much." He kept this taste for Euclidean geometry – the "old Euclidean geometry," as he used to call it, with its constructions,

criteria for triangle equality, etc. all his life, even though, in the mathematical milieu in which he was to evolve, it was against the grain. Having disappeared from the school and university curricula. Euclidean geometry became a field not only unknown, but, what is worse, he said, poorly understood. He writes in (Thom 2004): "I have kept a preference for Euclidean geometry that my colleagues and fellow mathematicians do not forgive me."

Busemann's work is still poorly known and greatly underrated. I find no good reason for that, except that this work lacks publicity, which, it seems, Busemann himself was reluctant to do. In his letter to V. Pambuccian, dated November 26, 1989, which I quoted above, he writes: "The reason for the neglect of synthetic geometry appears to me to lie first in the fact that every professor cares about seeing his ideas live on via his students. To admit that his methods use to a large extent unnecessary hypotheses is inconvenient and is simply swept under the rug. In fact, I do not know, except for some of my students, anybody in the USA who has actually read GG."

Acknowledgments This work is supported by the Interdisciplinary Thematic Institute CREAA, as part of the ITI 2021–2028 program of the Université de Strasbourg, the CNRS, and the Inserm (funded by IdEx Unistra ANR-10-IDEX-0002, and by SFRI-STRAT'US ANR-20-SFRI-0012 under the French Investments for the Future Program).

References

A'Campo Neuen A, Papadopoulos A (2018) Busemann and Feller on curvature properties of convex surfaces. In: Papadopoulos A (ed) Herbert Busemann, Selected works, Vol. II. Springer, Cham, pp 67–88

Aleksandrov AD (1939) Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. Leningrad State Univ Annals [Uchenye Zapiski] Math Ser 6:3–35

Aleksandrov AD (1947) On the works of S. E. Cohn-Vossen, Uspehi Matem. Nauk (N.S.) 2 3(19): 107–114. English trans: Iacob A

Aleksandrov AD (1960) Modern development of surface theory. Proc Int Congr Math 1958:3–18

Aleksandrov AD (2005) Intrinsic geometry of convex surfaces, Kutateladze SS (ed) (trans: Vakhrameyer, S). CRC Press, Chapman & Hall/CRC, Taylor & Francis group, Boca Raton, FL

Aleksandrov AD, Zalgaller VA (1967) Intrinsic geometry of surfaces, Academy Sc. USSR, Moscow, 1962 (English trans: Danskin JM). AMS Translations of Mathematical Monographs, Amercian Mathematical Society, Providence

Andreev PD (2018) Busemann's problems on G-spaces. In: Papadopoulos A (ed) H. Busemann, Selected works, Vol. II. Springer, Cham, pp 85–113

Apollonius (2008) Les Coniques, tome 3: Livre V, commentaire historique et mathématique, édition et traduction du texte arabe. de Gruyter, 550 p

Berestovskii V (1979) Busemann homogeneous G-spaces. Dokl Akad Nauk SSSR 247(3):525528

Berestovskii V (1982) Homogeneous Busemann G-spaces. Sibirsk Mat Zh 23(2) 3–15, 215

Berestovskii VN (2018a) Busemann's results, ideas, questions and locally compact homogeneous geodesic spaces. In: Papadopoulos A (ed) H. Busemann, Selected works, Vol. II. Springer, Cham, pp 41–84

Berestovskii VN (2018b) Herbert Busemann and convexity. In: Papadopoulos A (ed) H. Busemann, Selected works, Vol. II. Springer, Cham, pp 89–118

Bonnesen T, Fenchel W (1934) Theorie der konvexen Körper. Springer, Berlin. English translation: Theory of convex bodies, BCS Associates, Moscow, Idaho, 1987
Burago D (2018) On several problems posed by H. Busemann and related to geometry in normed spaces. In: Papadopoulos A (ed) H. Busemann, Selected works, Vol. II. Springer, Cham, pp 119–121
Busemann H (1930) Die Vollständigkeit der Minimalfolgen von Eigenwertproblemen. Nachrichten Göttingen 1930:295–307
Busemann H (1932) Über die Geometrien, in denen die "Kreise mit unendlichem Radius" die krzesten Linien sind. Math Ann 106(1):140–160. (English trans: A'Campo A) On the geometries where circles of infinite radius are the shortest lines, in: H. Busemann, Selected works, Vol. I, pp 155–172)
Busemann H (1933) Über Räume mit konvexen Kügeln und Parallelenaxiom. Nachrichten Göttingen 1933:116–140. (trans: A'Campo A) In H. Busemann, Selected works, Vol. I, A. Papadopoulos (ed.), Springer, Cham, 2018, pp 173–191
Busemann H (1939) Two-dimensional metric spaces with prescribed geodesics. Ann Math (2) 40:129–140
Busemann H (1941) On Leibniz's definition of planes. Am J Math 63:101–111
Busemann H (1942) Metric Methods in Finsler spaces and in the foundations of geometry. In: Annals of mathematics studies, no. 8. Princeton University Press, Princeton
Busemann H (1943) On spaces in which points determine a geodesic. Trans Am Math Soc 54: 171–184
Busemann H (1947) Two-dimensional geometries with elementary areas. Bull Amer Math Soc 53(4):402–407
Busemann H (1950) The foundations of Minkowskian geometry. Comment Math Helv 24:156–187
Busemann H (1955) The geometry of geodesics. Academic Press, Inc., New York. Reprinted by Dover in 2005. Russian translation, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1962, 504 p
Busemann H (1958) Convex surfaces, interscience tracts in pure and applied mathematics. Interscience Publishers, New York/London. Russian translation, Convex surfaces, Izdat. Nauka, Moscow, 1964
Busemann H (1961) The synthetic approach to Finsler spaces in the large, C.I.M.E. Summer Sch., 23, Geometria del Calcolo delle Variazioni. Springer, Heidelberg, pp 1–72
Busemann H (1964) Length-preserving maps. Pacific. J Math 14:457–477
Busemann H (1967) Timelike spaces. Diss Math 53, 52 p
Busemann H (1970) Recent synthetic differential geometry. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol 54. Springer-Verlag, Berlin/Heidelberg/New York
Busemann H (1975) Planes with analogues to Euclidean angular bisectors. Math Scand 36:5–11
Busemann H (1976) Problem IV: Desarguesian spaces. Math Dev Hilbert Probl, Proc Symp Pure Math 28, De Kalb 1974:131–141
Busemann H (1979) Minkowskian geometry, convexity conditions and the parallel axiom. J Geom 12(1):17–33
Busemann H (2018a) In: Papadopoulos A (ed) Selected works, Vol. I. Springer, Cham
Busemann H (2018b) In: Papadopoulos A (ed) Selected works, Vol. II. Springer, Cham
Busemann H, Feller W (1936) Krümmungseigenschaften Konvexer Flächen. Acta Math 66(1): 1–47. English translation by A. A'Campo Neuen: Curvature properties of convex surfaces. Selected works, Herbert Busemann, vol. II, p. 235–269
Busemann H, Petty CM (1956) Problems on convex bodies. Math Scand 4:88–94
Busemann H, Phadke BB (1980) Nonconvex spheres in G-spaces. J Indian Math Soc, New Ser 44: 39–50
Busemann H, Phadke BB (1983) Peakless and monotone functions on G-spaces Tsukuba. J Math 7(1):105–135
Busemann H, Phadke BB (1993) Novel results in the geometry of geodesics. Adv Math 101(2): 180–219
Couturat L (1901) La logique de Leibniz: d'après des documents inédits. Alcan, Paris

Dembart L (1985) An unsung geometer keeps his own plane, Los Angeles Times, June 14, 1985
Duren PL, Askey R, Merzbach UC, Edwards HM (eds) (1988) A century of mathematics in America. American Mathematical Society, Providence
Feller W (2015) Schilling IRL, Vondraek Z, Woyczynski WA (eds) Selected papers. Springer, Cham
Fenchel W, Jessen B (1938) Mengenfunktionen und konvexe Körper. Danske Vid Selskab Mat-Fys Medd 16(3):1–31
Fréchet M (1906) Sur quelques points du calcul fonctionnel, Thèse présentée á la faculté des Sciences de Paris pour obtenir le grade de Docteur ès Sciences, 1906, published in Rendiconti del Circolo Matematico di Palermo, vol 22, pp 1–72
Gromov M (1981) Structures métriques pour les variétés riemanniennes. In: Lafontaine J, Pansu P (eds) Textes Mathématiques, 1, 1st edn. Cedic/Fernand Nathan, Paris. vii, 152 p. English trans: Metric structures for Riemannian and non-Riemannian spaces (trans: Bates SM). With appendices by Katz M, Pansu P, and S. Semmes. 3rd printing, Modern Birkhaüser classics. Birkhaüser, Basel. 2007
Hilbert D, Cohn-Vossen S (1932) Anschauliche Geometrie. Springer, Berlin/Heidelberg
Kutateladze SS (2012) Aleksandrov of ancient Hellas. Sib Electron Math Rep 9:A6–A11
Kutateladze SS, Novikov SP, Reshetnyak YG (2012) Aleksandr Danilovich Aleksandrov (on the 100th anniversary of his birth). Russ Math Surv 67(5):959–965
Lambert JH (1766) Theorie der Parallellinien. French translation in: La thorie des lignes parallles de Johann Heinrich Lambert, ed. A. Papadopoulos and G. Thret, Collection Sciences dans l'Histoire, Librairie Scientifique et Technique Albert Blanchard, Paris, 2014
Leibniz GW (1849) Mathematische Schriften. zweite Abteilung I, Berlin
Papadopoulos A (2014) Hilbert's fourth problem. In: Papadopoulos A, Troyanov M (eds) Handbook of Hilbert geometry, IRMA lectures in mathematics and theoretical physics, vol 22. European Mathematical Society Publishing House, pp 391–431
Papadopoulos A (2018a) Busemanns work on the characterization of Minkowski geometries. In: Herbert Busemann, selected works, vol II. Springer Nature, Cham, pp 145–159
Papadopoulos A (2018b) Herbert Busemann. Not Am Math Soc 65(3):341–343
Papadopoulos A (2018c) Chronogeometry. In: Papadopoulos A (ed) Herbert Busemann, selected works, Vol. II. Springer, Cham, pp 133–141
Pogorelov AV (1979) Hilbert's fourth problem (trans: Silverman RA), edited by Kra I in cooperation with E. Zaustinskyi, V. H. Winston and Sons, Washington, DC. Original published by Nauka
Rashed R (2008) Apollonius: Les Coniques, tome 1.1: Livre I, commentaire historique et mathématique, édition et traduction du texte arabe. de Gruyter, 666 p
Rashed R (2009a) Apollonius: Les Coniques, tome 2.2: Livre IV, commentaire historique et mathématique, édition et traduction du texte arabe. de Gruyter, 319 p
Rashed R (2009b) Apollonius: Les Coniques, tome 4: Livres VI et VII, commentaire historique et mathématique, édition et traduction du texte arabe, Scientia Graeco-Arabica, vol 1.4. de Gruyter, 572 p
Rashed R (2010) Apollonius: Les Coniques, tome 2.1: Livres II et III, commentaire historique et mathématique, édition et traduction du texte arabe. de Gruyter, 682 p
Rashed R (2015) Angles et grandeur: D'Euclide à Kamal al-Din al-Farisi, Scientia Graeco-Arabica, vol 17. De Gruyter, Boston/Berlin
Rashed R, Papadopoulos A (2017) Menelaus' spherics: early translation and al-Māhānī/al-Harawī's version, Series: Scientia Graeco-Arabica, 21. De Gruyter, Berlin/Boston, 890 p
Reidemester K (1921) Über die singulären Randpunkte eines konvexen Körpers. Math Ann 83: 116–118
Thom R (1991) Prédire n'est pas expliquer, Entretiens avec Emile Noël. Eshel, coll. La Question, Champs/Paris
Thom R (2004) Exposé introductif. In: Logos et Théorie des Catastrophes, Petitot J (dir.), Actes du colloque international de Cerisy de 1982, Centre culturel international de Cerisy-la-Salle, Genève, Patiño, 1989. Publié aussi dans Gazette des mathématiciens, édition spéciale, pp 7–22
Zalgaller VA (2012) Memoirs of A. D. Aleksandrov and his Leningrad geometry seminar. Sib Electron Math Rep 9:A26–A80

Section VII

Logic

Bharath Sriraman

Logic in the History and Philosophy of Mathematical Practice

Valentina Harizanov

Contents

1 Introduction 1906
2 Cross-References 1920
References 1920

Abstract

Mathematical logic is the study of reasoning about mathematical objects and the degree to which mathematical and scientific reasoning can be formalized and mechanized. Logic provides the foundations of mathematics and of theoretical computer science. Classical logic defined truth, developed the theory of infinite numbers, resolved paradoxes of naive set theory, defined what an algorithm is, and established that certain mathematical principles are independent from the rest of mathematics. Modern logic contributes to current developments in mathematics and its foundations, and continues to closely interact with various areas of mathematics. Essential and surprising connections are often found in order to solve big problems. To bring various resources into a single argument, common methods, uniform terminology, and concisely stated theorems are needed. This interaction of disciplines has not only led to new approaches and results but also to new areas of mathematical research, such as computable structure theory, reverse mathematics, and the emerging quantum computing theory.

Keywords

Axiom · Arithmetic · Curve · Computable · Elliptic · Fermat · Model · Nonstandard · Peano · Proof · Theory · Turing · Ultrapower

V. Harizanov (✉)
Department of Mathematics, George Washington University, Washington, DC, USA
e-mail: harizanv@gwu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_118

1 Introduction

Mathematical logic is the study of mathematical and scientific reasoning and degree to which it can be formalized and made algorithmic. It also provides the foundations of mathematics. It studies the common language used in mathematics, how main mathematical objects are defined, and formalizes the laws of reasoning with these objects. Logic played a key role in developing the axiomatic systems for geometry, arithmetic, and analysis. It resolved foundational crisis of mathematics at the beginning of the twentieth century. This led to formalism championed by Hilbert. Hilbert's program for the foundations of mathematics involved formalization of classical mathematical theories based on consistent axiomatic systems, which are also complete. Work of Genzen, Gödel, Turing, Tarski, and others led to discovery of general phenomena about power and limitations of the axiomatic methods with deep mathematical and philosophical ramifications. In particular, Gödel's incompleteness theorem established that a consistent system with an algorithmic set of axioms, capable of expressing arithmetic, cannot be complete. Hence, there are statements that cannot be decided from the standard axioms, thus leading to nonstandard models of theories. A. Robinson developed precise model-theoretic machinery to extend the real numbers by infinite and infinitesimal numbers, allowing us to study calculus in the original framework of Newton and Leibniz.

Logic has been instrumental in solving some of the biggest mathematical problems, such as Hilbert's tenth problem and Dehn's word problem in combinatorial group theory. Hilbert's tenth problem asked whether there is an effective procedure to decide whether a given Diophantine equation has a solution in integers. Dehn's word problem was about groups presented by finitely many generators and relators and asked for an algorithm deciding whether two words in the generators represent the same element. Negative solutions to both problems were possible after decades of work by many mathematicians, but only after computability theory was established as an exact theory. Hilbert's tenth problem for rationals is currently one of the main open problems in mathematics. Computability theory also led to complexity theory in computer science where time and space resources for algorithms are taken into consideration. Complexity theory uses computable reductions to define the hardest problems within certain complexity classes. The question whether nondeterministic polynomial time algorithms can be simulated by deterministic polynomial time algorithms is one of the biggest open problems of our times.

Logic tries to explain essential phenomena of mathematical practice and to use these explanations to further contribute to mathematics. Set-theoretic developments in logic established that some principles mathematicians ordinarily use in their practice are neither provable nor refutable from the rest of mathematics, thus reflecting certain philosophical and practical preferences. In current mathematical practice, various areas of logic and, more broadly, mathematics continue to interact with each other in an essential way leading to new, more general discoveries. The chapters in this section present several areas of logic and mathematics and their interactions with mathematical practice. While the chapters survey some of the

fundamental notions and results of classical logic, they also show modern topics of current research interest.

The main areas of mathematical logic include model theory, set theory, computability theory, proof theory, and others. One of the common themes is the study of definability in various structures and how it captures relevant semantic notions. Work in set theory provided the foundations of mathematics by showing that a great deal of mathematics can be formalized in terms of sets and their axioms. More recently, work in the foundations of mathematics, such as in reverse mathematics, focuses on establishing parts of mathematics within particular formal systems. Reverse mathematics has become a very active area of research. Its goal is to determine which principles are required to prove theorems of ordinary mathematics and is usually carried out using subsystems of second-order arithmetic. Results in reverse mathematics are intricate and involve interplay of mathematical logic with other areas of mathematics.

In the chapter *Model Theory of Nonstandard Structures with Applications*, Kossak focuses on three results in model theory, which he establishes using nonstandard models. Along the way, he presents and surveys many model-theoretic concepts and ideas. Model theory is the study of the relationship between formal theories and algebraic structures. It is closely related to universal algebra, the study of general algebraic structures. Model theory is also related to algebraic geometry, which deals with geometric problems involving zeroes of polynomials of several variables, as well as other related problems.

A *structure* $\mathcal{M}$ consists of a nonempty set M, called the domain or the universe, and certain constants, relations, and functions on the domain. Every relation and every function (or operation) has assigned a length or arity. The corresponding constant, relation, and function symbols form the *language* or the *signature* of the structure. The usual rational, real, or complex numbers with addition and multiplication have an algebraic *field* structure. Both addition and multiplication are associative binary operations. Every number has an additive inverse with 0 being the additive identity. Hence, when restricted to addition, these numbers have a commutative *group* structure. Every nonzero number has a multiplicative inverse with 1 being the multiplicative identity. Hence, the nonzero numbers also have a commutative group structure under multiplication. Integers with addition and multiplication form a *ring* structure. The size of natural numbers, integers, or rational numbers is $\aleph_0$, the smallest infinite size. On the other hand, the size of reals or complex numbers is $2^{\aleph_0}$, which is the same as the size of the power set of the set N of natural numbers, which is the set of all subsets of N.

Here, we are focusing on the first-order language and formulas. This language is fairly expressive for everyday mathematics. Gödel discovered the completeness theorem for first-order logic, which established the correspondence between provability in a formal system and truth in a structure. A *theory* is a set of sentences (formulas with no free variables) in a given language L. The theory of a structure $\mathcal{M}$ is the set of all first-order sentences it satisfies. It is *complete* since for each sentence either the sentence or its negation is true in the model. Two structures for the same language are *elementarily equivalent* if they satisfy the same first-order sentences.

An *automorphism* of $\mathcal{M}$ is a permutation p of its domain, which preserves the structure. For example, if $u * v = w$ in $\mathcal{M}$, then also $p(u) * p(v) = p(w)$ in $\mathcal{M}$. Thus, we can view p as renaming of the elements. More generally, for two structures, a bijection of the domains that preserve the structure is called an *isomorphism*. It is one of the main concepts of mathematics. Isomorphic structures are elementarily equivalent, but the converse is not true.

A theory may have more than one model, as Keisler says (Chang and Keisler 1990): "The mathematical world was forced to observe that a theory may have more than one model in the 19th century, when Bolyai and Lobachevsky developed non-Euclidean geometry, and Riemann constructed a model in which the parallel postulate was false but all the other axioms were true. Later in the 19th century, Frege formally developed the predicate logic, and Cantor developed the intuitive set theory in which our models live."

Any model of some theory not isomorphic to the intended, standard model is called *nonstandard*. The set of natural numbers is $N = \{0, 1, 2, \ldots\}$. The idea of establishing an axiomatic system that captures the *standard model of arithmetic*, $\mathcal{N} = (N, +, \cdot, 0, 1, <)$, was developed in the late nineteenth century by Dedekind and Peano. One of these formal systems became known as Peano arithmetic. It has an infinite but computable set of axioms. Hilbert asked for a proof that such a system is complete. However, an algorithmic set of such axioms generates an incomplete theory, as established by Gödel. The first-order theory of $\mathcal{N}$ is known as the true arithmetic. In 1934, Skolem constructed a nonstandard model of true arithmetic, hence elementarily equivalent to $\mathcal{N}$.

While the notion of a substructure plays an important role in algebra and mathematics in general, model theory has a stronger notion of an elementary substructure. Let $n \geq 1$. For an n-tuple $\overline{a}$ of elements in the domain, the *type* of $\overline{a}$ is the set of all first-order formulas with n variables, which are satisfied by $\overline{a}$. Thus, the type captures all first-order properties that $\overline{a}$ has in $\mathcal{M}$. For example, the element 1 has different types in the structure $(N, <)$ and its substructure $(N - \{0\}, <)$, as witnessed by the formula $\theta(x) : \forall y \neg (y < x)$. If p is an automorphism of $\mathcal{M}$, then the type of $p(\overline{a})$ is the same as the type of $\overline{a}$.

For two structures $\mathcal{S}$ and $\mathcal{M}$ for the same language, $\mathcal{M}$ is an *elementary extension* of $\mathcal{S}$ (equivalently, $\mathcal{S}$ is an elementary substructure or submodel of $\mathcal{M}$) if $S \subseteq M$ and the types of the finite tuples from S are the same in $\mathcal{S}$ and $\mathcal{M}$. It follows that a structure and its elementary extension are elementarily equivalent. One of the main results that is often used is that every infinite structure $\mathcal{M}$ has a proper elementary extension $\mathcal{M}^*$. If $\mathcal{M}^*$ is not isomorphic to $\mathcal{M}$, then $\mathcal{M}^*$ can be considered to be a nonstandard extension of $\mathcal{M}$. Moreover, if the language is at most countable, then there is a proper elementary extension of the same cardinality. The standard model of calculus is the usual ordered field of real numbers. In 1960, A. Robinson preciscely established another approach to calculus and analysis – nonstandard analysis. He extended the standard model to a nonstandard one containing infinite numbers and infinitesimals.

The analysis and classification of definable sets is an important part of model theory and logic in general. For a structure $\mathcal{M}$, an n-ary relation R on the domain

M is *definable* if there is a formula with n free variables, possibly with parameters from M, satisfied in $\mathcal{M}$ exactly by the tuples in R. Clearly, every finite set is definable since its elements can be taken as parameters. Hence, co-finite sets are also definable. A set is called *arithmetical* if it is definable in the standard model of arithmetic $\mathcal{N}$. Parameters are even not needed here since all natural numbers are definable. An infinite structure where the only definable sets are either finite or co-finite is called *minimal*. The ordered set of integers is not minimal since the set of natural numbers is definable. An important example of a minimal structure is the field of complex numbers, a result due to Chevalley and Tarski. Moreover, the field of complex numbers is *strongly minimal* since every elementarily equivalent structure is also minimal.

A structure is called *ordered* if one of its relations (named in the language) is a linear ordering. Kossak uses elementary extensions and automorphisms to show that the ordered set of natural numbers $(N, <)$ is minimal, demonstrating the power and elegance of this method. The structure $(N, <)$ is not strongly minimal since every proper elementary extension of it contains a copy of the ordered set of integers. An ordered structure is o-*minimal* (order-minimal) if each definable set is the union of finitely many open intervals and points. The ordered field of real numbers is o-minimal, a result due to Seidenberg and Tarski. Kossak uses a similar argument as above to show that the ordered set of integers is o-minimal.

The second application of the elementary extension method of the chapter is to establish an infinite Ramsey theorem based on arithmetic sets, a result by Kossak and Schmerl. Here, Kossak uses a conservative proper elementary extension of the standard model of arithmetic. An elementary extension is *conservative* if every definable set in the extension, when intersected with the subdomain remains definable in the substructure.

In the case of the field $\mathcal{C}$ of complex numbers, a proper elementary extension of the same cardinality turns out to be isomorphic to $\mathcal{C}$. On the other hand, no proper elementary extension of the field $\mathcal{R}$ of real numbers is isomorphic to $\mathcal{R}$. Kossak uses properties of reals, complex numbers, and their elementary extensions to show that there is a nontrivial automorphism of $\mathcal{C}$, which is not the complex conjugation. Then he shows that there are $2^{2^{\aleph_0}}$ automorphisms of $\mathcal{C}$, which is as many as there are functions from the set of all complex numbers to itself.

Kossak's examples illustrate how the arguments in nonstandard elementary extensions lead to results about the standard structures.

The chapter *Computability Theory* by Harizanov, Srinivasan, and Verta explains mathematical theory of algorithms, its development, and some of the current trends. The topics include various formalisms, effective enumerations, relative computability, arithmetical and hyperarithmetical hierarchies, and computable structures. Although the algorithmic method has been pervasive in mathematical practice for centuries, it was not sufficiently rigorous to deal with negative solutions to algorithmic problems that are now known as undecidable problems. Modern computability theory started in the 1930s and was founded by Gödel, Church, Petér, Turing, Kleene, Post, and others. In his influential 1936 paper, Turing invented what we

now call a *Turing machine*, which is essentially an abstraction of a computing agent. He also introduced the universal machine, which can simulate any other Turing machine on any input, leading to the idea of stored-program computer. Turing also proved that there are undecidable problems. Gödel was very pleased with Turing's approach. He wrote: "That this really is the correct definition of mechanical computability was established beyond any doubt by Turing."

There were other formalisms for an intuitive concept of an effectively calculable function, which turned out to be equivalent to Turing computability, thus leading to the same class of formally computable functions. Hence, this theory of algorithms provided the foundation for mathematical theory of computability (also known as recursion theory) and the emerging science of computing. Turing's formalism was the most convincing one. Soare (2016) says: "It was Turing who saw the computer itself in the marble, a simple intuitive device equipped with only a finite program and using only a finite sequence of strokes at each stage in a finite computation, the vision closest to our modern computer." The development of computability theory led to the acceptance of the *Church-Turing thesis* which says that a function is intuitively effectively calculable exactly when it is formally computable. This thesis is somewhat unique in mathematics since it is not a statement with a mathematical proof but an assertion that can be explained, defended, or debated.

A Turing machine on a given input may halt or it may not halt, computing forever. This leads to the notion of a partial computable function. Without loss of generality, we can consider only functions and relations on natural numbers. The set of inputs for which a partial function f is defined is called the *domain* of f. A partial function f is *partial computable* if there is a Turing machine that on every input in the domain of f halts and outputs its value, while it does not halt on inputs that are not in the domain of f. *Computable* functions are partial computable functions that are total. Partial computable functions are the same as partial recursive functions defined by Kleene, starting with some basic functions and applying the operations of composition, primitive recursion, and unbounded search. Computable functions are also called *recursive functions* or general recursive functions for emphasis. An important proper subset of recursive functions, known as primitive recursive functions, was previously used by Gödel in his work on formal arithmetic.

A relation R of natural numbers is *computable* if there is a Turing machine that always halts and on the elements in R outputs 1, while it outputs 0 on the elements in the complement (negation) of R. Equivalently, the characteristic function of R is computable. Hence, the negation of a computable relation is also computable. Since each Turing machine is a finite list of instructions, Turing machines can be computably enumerated, even without repetitions. For every $n \geq 1$, this enumeration gives a computable enumeration of all n-ary partial computable functions, now each function in the listing appearing infinitely often since it can be computed by infinitely many Turing machines. For $n = 1$, we have a computable listing of all unary partial computable functions: $\varphi_0, \varphi_1, \varphi_2, \ldots$ For a function φ_e, we can think of its index e as its Gödel code and can use e as an input for another Turing machine. This gives rise to a universal Turing machine, which can run any Turing machine on any input.

Church and Kleene introduced computably enumerable sets in 1936. We say that a set S is *computably enumerable* (also called recursively enumerable) if and only if S is empty or the range of a computable function. For example, the theorems in a formal system with a computable set of axioms form a computably enumerable set. It can be proven that a set is computably enumerable if and only if it is the domain W_e of some partial computable function φ_e. Clearly, every computable set is computably enumerable since a decision algorithm can be transformed into an enumeration algorithm. The converse is not true. The diagonal *halting set* K consists of all inputs e on which the Turing machine with index e halts, i.e., $K = \{e : e \in W_e\}$. It is easy to show that K is computably enumerable. However, K is not computable. It turns out that computable sets are exactly computably enumerable sets that also have computably enumerable complements. Since K is not computable, the relation $\{(e, x) : x \in W_e\}$ is also not computable, hence the halting problem of determining whether e-th Turing machine halts on input x is undecidable.

Computably enumerable sets are closely related to Diophantine equations and Hilbert's tenth problem mentioned earlier, which was posed in 1900 at the International Congress of Mathematicians in Paris. Diophantine equations are polynomial equations in finitely many variables. The problem asked to devise a process that determines whether any given Diophantine equation has integer solutions. The algorithmic solvability of a Diophantine equation in the integers is equivalent to the algorithmic solvability of a Diophantine equation in the natural numbers. The negative solution to the problem was obtained in 1970 after many years of work of Davis, Putnam, J. Robinson, and Matiyasevich. It is not hard to see that a set of natural numbers defined using existential quantifiers and a polynomial equation is computably enumerable. Such sets are called *Diophantine sets*. The corresponding polynomial is called its Diophantine definition. For example, the formula $\theta(x)$: $\exists y \exists z(x = (y + 2)(z + 2))$ defines composite numbers. In 1950, Davis conjectured that every computably enumerable set is definable by an existential polynomial formula. His conjecture was finally proven in 1970 which, together with the existence of computably enumerable sets that are not computable, shows the undecidability of the Hilbert's tenth problem.

Computability theory developed the following classification of sets and problems they encode. A set is Σ_0^0 or Π_0^0 if it is computable. For $n > 0$, a set A is Σ_n^0 if there is a computable $(n + 1)$-ary relation R such that for every $a \in N$, we have $a \in A \Leftrightarrow (\exists x_1)(\forall x_2)\cdots(Qx_n)\ R(a, x_1, x_2, \ldots, x_n)$, where Q is a corresponding quantifier. Π_n^0 sets are defined similarly starting with a universal quantifier. They are the complements of Σ_n^0 sets. It can be shown that computably enumerable sets are exactly Σ_1^0 sets. This hierarchy of sets of natural numbers is a proper hierarchy since Σ_n^0 sets are strictly contained in Σ_{n+1}^0 sets. Clearly, every Σ_n^0 set is also Π_{n+1}^0. It follows from Gödel's work that a set is first-order definable in the standard model of arithmetic $\mathcal{N}$ if and only if it belongs to the union of all Σ_n^0 sets, which is the same as the union of all Π_n^0 sets. Thus, such sets are called arithmetical, and the hierarchy is called *arithmetical hierarchy*. There is also correspondence between the complexity of the sets in the arithmetical hierarchy and the complexity of the defining formulas. A set is Δ_n^0 if it is both Σ_n^0 and Π_n^0.

We can relativize some of the notions of computability theory by using Turing's notion of an *oracle* machine. A machine with an oracle for a set X is a Turing machine which carries out its computation with the additional resource of read-only access to the characteristic function of X. Hence, even if a set Y is not computable, it may be computable with an oracle for another set X. For example, a set is computable in its complement. If an oracle is computable, then it can be eliminated. When Y is computable in X, we also say that Y is Turing reducible to X, or that the Turing degree of Y is $\leq$ than the Turing degree of X: $\deg(Y) \leq \deg(X)$. Two sets have the same Turing degree if they can be computed from each other. All computable sets have the least Turing degree, denoted $\mathbf{0}$. The Turing degree of any noncomputable set is strictly greater than $\mathbf{0}$. Turing degrees are only partially ordered since two Turing degrees can be incomparable, as was first shown by Kleene and Post. There are $2^{\aleph_0}$ Turing degrees and they have a very interesting structure.

The halting set relative to X is defined as K above but replacing the machines with machines with oracle X. It is also called the *jump* of X and denoted X'. Hence, K is the jump of any computable set. The Turing degree of X' is strictly greater than the Turing degree of X. Hence, there is no greatest Turing degree. The jump operator can be iterated, starting with K as the first jump. The second jump is K' and so on. There is following correspondence between arithmetical hierarchy and Turing degrees. While Δ^0_1 sets are the computable sets, Δ^0_2 sets are exactly those computable in the halting set K, and Δ^0_3 sets are exactly those computable in K', and so on. Arithmetical hierarchy can be extended to hyperarithmetical hierarchy by extending the definition of Σ^0_n and Π^0_n sets to Π^0_α and Σ^0_α sets for infinite constructive ordinals α.

In current research, computability theory interacts with other areas of logic and mathematics outside of logic leading to new areas such as computable structure theory, computable analysis, algorithmic randomness, reverse mathematics, etc. The chapter presents some of the ideas of computable structure theory where algebraic structures and constructions are effective. Even in 1930, a famous algebraist Van der Waerden introduced the notion of an explicitly given field as one "the elements of which are uniquely represented by distinguishable symbols with which addition, subtraction, multiplication and division can be performed in a finite number of operations." This led to the notion of a computable structure. A structure for a finite language is *computable* if its domain is computable and its functions and relations are computable. In computable structures, quantifier-free formulas define computable relations, and first-order formulas in general define arithmetical relations. The standard model of arithmetic is a computable structure. In computable structures, computable infinitary formulas define hyperarithmetical relations.

In the chapter *Reverse Mathematics*, Stillwell presents the program of reverse mathematics. The goal of reverse mathematics is to determine what axioms of mathematics are necessary to prove specific mathematical theorems. Partial motivation for this program goes back to the old mathematical questions about Euclid's *Elements* and later about some of the principles of set theory. The program of modern reverse mathematics was introduced by H. Friedman in his lectures at the International Congress of Mathematicians in Vancouver in 1974. It was also further

developed by Simpson and his students who established many of its first results. Many other researchers joined later, making it a very active research area. The context is usually second-order arithmetic, working over a weak base theory. Some of the definitions and methods are inspired by constructive analysis and proof theory. This study has very close connections to computability, as the natural models involved are well-behaved collections of Turing degrees, and the set existence axioms are related to logical definability and computational strength. For example, the common base theory RCA_0 states, roughly, that computable objects exist, and that induction can be performed on computably enumerable sets. Friedman explains in his founding paper (Friedman 1975) two principal themes that emerge: "I. When the theorem is proved from the right axioms, the axioms can be proved from the theorem. II. Much more is needed to define explicitly a hard-to-define set of integers than merely to prove their existence."

Stillwell presents the historical development of mathematical practice that led to reverse mathematics, as well as its main systems that can prove much of mainstream mathematics. He begins with the subsystem of Euclid's geometry without the parallel axiom, which is also called the *neutral geometry*. In this system, there are many interesting equivalents of the *parallel axiom*, which states that for any line and a point outside it, there is exactly one line through the point that does not intersect the line. Eventually, it was shown that the parallel axiom is not a theorem of the neutral geometry. Stillwell presents Beltrami's model from 1868 of neutral geometry where the parallel axiom does not hold since there is more than one parallel line. This leads to non-Euclidean geometries where the parallel axiom and its equivalents do not hold. Thus, the parallel axiom is i*ndependent* from the axioms of neutral geometry. It can neither be proved nor refuted. In general, if a theory T and a statement (sentence) σ are such that $T \cup \{\neg\sigma\}$ is consistent (has a model), then σ cannot be proved from a theory T. Here, T could be a base theory.

Stillwell further presents set theory starting with Cantor's naive set theory from 1874 and discovery of various infinite sizes. Two sets have the same size (cardinality) if there is $1-1$ and onto correspondence (bijection) between them. The size of every infinite countable set is $\aleph_0$. Cantor's diagonal argument can be used to show that for every set of size κ, its power set has a strictly larger size 2^κ. The sets of real and complex numbers are of size $2^{\aleph_0}$, the same size as the set of all infinite binary sequences. Using properties freely to define sets created Russell's paradox. In particular, the property $x \notin x$ gives us the "set of all sets that do not contain themselves," which after all could not be a set. If it were some set D, then we will have a contradiction: $D \in D \Leftrightarrow D \notin D$.

A *well-ordering* of a set is a total (linear) ordering of its elements such that every nonempty subset has the least element. The ordered set of integers is not well-ordered because the subset of negative integers does not have the least element. The orderings of natural numbers ω and other ordinals $\omega + 1, \omega + 2, \ldots, \omega + \omega, \omega + \omega + 1, \ldots$ are well-orderings. Cantor's bold hypothesis that every set can be well-ordered by some ordering presented further challenges. This statement became known as the well-ordering principle. Zermelo showed that the well-ordering principle is equivalent to

the axiom of choice. The *axiom of choice* says that for every set of nonempty sets, there is a function that chooses an element in each set. If the union of the sets can be well-ordered, then the least elements in the sets can be chosen. In 1908, Zermelo gave axioms of set theory by building sets from the empty set by certain constructions such as pairing, union, power set, definable subsets. There is also an axiom of infinity, which says that an infinite set exists. His system avoided Russell's paradox and proved the equivalence of the axiom of choice with the well-ordering principle. Zermelo's axiomatic system was abbreviated by Z. In 1922, Fraenkel added the axiom of replacement, which says that the image of any set under any definable function is also a set. This gave us to a stronger system, abbreviated by ZF, which now serves as a standard base system of set theory. In this system, there are many interesting equivalents of the axiom of choice, which are important in various areas of mathematics. For example, the axiom of choice is equivalent to the existence of a basis in any vector space. It has been shown that the axiom of choice is independent of ZF. Namely, Gödel obtained a model of ZF where the axiom of choice holds, and later Cohen produced a model of ZF where the axiom of choice does not hold. To establish this result, Cohen discovered a new method known as the forcing method.

Stillwell further presents the arithmetization of geometry and introduction of Dedekind cuts which reduce the arithmetic of real numbers to the arithmetic of rational numbers. Cantor's pairing function provides coding of pairs of natural numbers by natural numbers and allows coding of other mathematical objects by natural numbers. This is what Poincarė called at the beginning of the twentieth century the arithmetization of mathematics.

Peano arithmetic mentioned earlier is an axiomatic system for natural numbers with the unary successor function S, constant 0, addition, and multiplication. While the axioms say that the successor function is 1–1, and 0 is not in its range, addition and multiplication are captured by their recursive definitions: $x + 0 = x$ and $x + S(y) = S(x + y)$, and similarly for multiplication. In the first-order Peano arithmetic, there is an induction axiom schema for every first-order formula $\theta(x)$:

$$(\theta(0) \wedge \forall x(\theta(x) \Rightarrow \theta(S(x)))) \Rightarrow \forall x(\theta(x)).$$

In the second-order Peano arithmetic, there are also set variables denoted by capital letters. The second-order induction axiom is:

$$(0 \in X \wedge \forall x(x \in X \Rightarrow S(x) \in X)) \Rightarrow \forall x(x \in X).$$

There is also a comprehension axiom schema for any formula $\theta(x)$:

$$\exists X \forall x[x \in X \Leftrightarrow \theta(x)],$$

where formula θ can contain free set variables different from X. A comprehension scheme states that any set of natural numbers definable by a formula exists.

Gödel introduced the arithmetization of computability. For $n > 0$, a first-order formula is Σ_n^0 if it is equivalent to a formula in the prenex normal form which begins

with a block of existential quantifiers and has n alternations of blocks of quantifiers. Π_n^0 formulas are defined similarly starting with a block of universal quantifiers. The negation of Σ_n^0 formula is a Π_n^0 formula and vice versa. In particular, computably enumerable relations are captured by Σ_1^0 formulas of Peano arithmetic. Furthermore, it can be shown that a property is computable (recursive) if it can be defined both by a Σ_1^0 and a Π_1^0 formula in arithmetic. A formula equivalent to both Σ_1^0 and a Π_1^0 formula is also called a Δ_1^0 formula.

This brings us to the base system for reverse mathematics known as RCA_0 standing for *recursive comprehension axiom* (with the subscript indicating a particular version of the system). In RCA_0, the induction schema is restricted to Σ_1^0 formulas and the set existence axiom schema to Δ_1^0 formulas. This comprehension schema is saying that for every computable property, there is a corresponding set X. More precisely, if there are Σ_1^0 formula γ and Π_1^0 formula δ such that

$$\forall x[(\theta(x) \Leftrightarrow \gamma(x)) \wedge (\theta(x) \Leftrightarrow \delta(x))], \text{then } \exists X \forall x[x \in X \Leftrightarrow \theta(x)].$$

The minimal model of RCA_0 is the set of natural numbers with computable sets. RCA_0 can be viewed as a formalization of constructive mathematics. While RCA_0 can prove few important theorems in analysis, Stillwell shows that it cannot prove even the weak form of Bolzano's principle – that a bounded nondecreasing sequence of rational numbers has a limit.

Many problems in mathematics can be viewed as choosing an infinite path in a tree. A subtree of the full binary tree is a subset of its nodes, which is closed under initial segments. Such a tree is computable if its set of nodes is computable. The full binary tree has $2^{\aleph_0}$ infinite paths. A new axiom which we will consider is called *weak König's lemma*. It states that an infinite binary tree has an infinite path. Weak König's lemma is a theorem in infinite combinatorics since we can choose an infinite path by extending a finite path by adding the leftmost node with an infinitely many successors. However, there are computable infinite binary trees without computable infinite paths, so RCA_0 cannot prove weak König's lemma. When added to RCA_0, it gives us a stronger system WKL_0, which is the right axiomatic system for many theorems. Weak König's lemma has many interesting equivalents in RCA_0. For example, Friedman established that it is equivalent to the Heine-Borel theorem that any covering of the interval $[0, 1]$ by an infinite sequence of open intervals contains a finite subcovering. The full König's lemma states that an infinite finitely branching tree has an infinite path. The full König's lemma is not provable in WKL_0.

In an even stronger system ACA_0, standing for *arithmetic comprehension axiom*, the second-order induction axiom scheme is allowed, and the comprehension axiom scheme is allowed for arbitrary formulas. This system implies the first-order induction for all formulas. ACA_0 can prove all axioms of the first-order Peano arithmetic. The minimal model of ACA_0 is the set of natural numbers with arithmetical sets. ACA_0 can prove a number of important theorems in analysis, the full König's lemma, and some theorems in infinite combinatorics. Friedman proved that ACA_0

is conservative over the first-order Peano arithmetic in the sense that the theorems of the first-order Peano arithmetic are exactly those of ACA_0 that involve only variables for natural numbers.

These three systems are sufficient to prove a great deal of mainstream mathematics, but there are two other well-known systems that capture some theorems beyond the reach of ACA_0. These systems are ATR_0 and Π^1_1-CA_0, standing for arithmetical transfinite recursion, and comprehension for formulas with the second-order universal quantifier, also known as Π^1_1 formulas. ATR_0 can be viewed as a formalization of *predicative* mathematics, a program advocated by Poincaré and Weyl in which a set cannot be defined in terms of a collection that includes the set itself. On the other hand, Π^1_1 comprehension allows nonpredicativity. Together, these five systems are known as "big five." Reverse mathematics always looks for minimal resources needed to prove a theorem, leading to new systems.

The chapter *Countable Nonstandard Models: Following Skolem's Approach* by Dimitrov and Harizanov goes back to Skolem's original construction of a *countable* nonstandard model of arithmetic. The chapter presents this construction and shows how it provides an idea for the construction of more effective nonstandard models, a topic of current research interest. The existence of a nonstandard model of arithmetic is now often proven using the compactness theorem for first-order logic. The compactness theorem says that a set of sentences has a model exactly when every finite subset of sentences has a model. The existence of nonstandard models can also be established by ultrapower constructions. The ultrapower construction is very elegant, but it depends on the axiom of choice and typically produces uncountable structures from infinite structures.

Going back to Skolem's construction from 1934 and in Skolem (1955) was very illuminating since it provided a direct connection with current investigation of effective ultrapowers. Skolem's original proof produces a countable nonstandard model in an explicit way. Skolem (1955) says: "More than 30 years ago I proved by use of a theorem of Löwenheim that a theory based on axioms formulated in the lower predicate calculus could always be satisfied in a denumerable infinite domain of objects. Later one has often expressed this by saying that a denumerable model exists for such a theory. Of particular interest was of course the application of this theorem to axiomatic set theory, showing that also for this an arithmetical model can be found." In 1959, Tennenbaum proved that there is no computable nonstandard model of arithmetic.

The ultraproduct construction is a major mathematical tool that is used in algebra, mathematical logic, and other areas. It allows us to construct new models of theories from old ones. The ultraproduct construction has some striking applications, such as a very elegant proof of the compactness theorem for first-order logic and the construction of a nonstandard model of analysis. The ultraproduct is a quotient of the direct product of structures modulo an equivalence relation determined by an ultrafilter. An ultrapower is an ultraproduct when there is only one structure. Ultrafilters are also used to determine the truth in the ultraproduct. To describe an ultrafilter, first start with a nonempty set I (for example, $I = N$) called the index set, which will be used to index structures. An *ultrafilter* over I is a nonempty set U of subsets of I such that U is closed

under the intersection of two sets and under the superset relation, and for every subset X of I, either X or its complement belongs to U. The construction of nontrivial ultraproducts requires nonprincipal ultrafilters. An ultrafilter is *nonprincipal* if it does not contain any singletons. The existence of nonprincipal ultrafilters, however, cannot be proven in ZF. However, the axiom of choice proves not just the existence of nonprincipal ultrafilter but a stronger statement called the ultrafilter lemma, which is strictly weaker than the axiom of choice. Łoś's Theorem, sometimes called the fundamental theorem of ultraproducts, establishes the satisfiability in the ultraproduct. An important consequence of Łoś's theorem is the following: if $\mathcal{B} = \prod_U \mathcal{A}$ is an ultrapower, then $\mathcal{A}$ and $\mathcal{B}$ are elementarily equivalent.

Dimitrov encountered the need for an effective ultrapower construction while investigating the automorphisms of the structure of computably enumerable vector spaces. It turned out that similar constructions had been used before to study models for fragments of arithmetic, first by Feferman, Scott, and Tennenbaum, and later by Lerman, Hirshfeld, and Wheeler. In 2009, Dimitrov introduced the notion of a cohesive power of computable structures. It is an effective ultrapower construction, which starting with an infinite computable structure produces an infinite countable structure. The role of an ultrafilter is played by a cohesive set. Cohesive sets are infinite sets of natural numbers, which are indecomposable by computably enumerable sets. More precisely, an infinite set C of natural numbers is *cohesive* if for every computably enumerable set W with complement $\overline{W}$, either the intersection $W \cap C$ or the intersection $\overline{W} \cap C$ is finite. Given a cohesive set C, we start with a computable structure and in its countable infinite power consider partial computable sequences modulo C. Hence, an effective ultrapower is always at most countable structure and it can be isomorphic to the original structure. However, in general, by Dimitrov's theorem, effective ultrapowers preserve only first-order properties expressed by sentences at lower levels of arithmetical hierarchy. It is possible for isomorphic computable structures to have non-elementarily equivalent effective ultrapowers over a fixed cohesive set.

We will now give a more precise definition of a cohesive power over a cohesive set C. Let us restrict ourselves to a finite language L. Let $\mathcal{A}$ be a computable structure for L with domain A contained in N. The cohesive power of $\mathcal{A}$ will be denoted by $\prod_C \mathcal{A}$. For convenience, let $\mathcal{B} = \prod_C \mathcal{A}$. Let D be the set of all partial computable sequences ψ with elements in A, which are defined on almost all (i.e., all but finitely many) elements of C. Two such sequences are equal modulo C if on almost all elements of C, they are defined and equal. Equivalence modulo C is an equivalence relation, and the elements of the cohesive power of $\mathcal{B}$ are the equivalence classes $[\psi]$ of partial functions ψ in D. Let R be a relation symbol in L, say a binary one. Then $R^{\mathcal{B}}([\psi], [\phi])$ if and only if the set of all i such that $\psi(i)$, $\phi(i)$ are defined and $R^{\mathcal{A}}(\psi(i), \phi(i))$ contains all but finitely many elements of C. The interpretation of a function symbol in $\mathcal{B}$ is defined pointwise. If c is a constant symbol in L, then the interpretation of c in $\mathcal{B}$ is the equivalence class of the constant sequence with value $c^{\mathcal{A}}$.

Dimitrov proved a variant of Łoś's theorem for cohesive powers: all Π_2^0 and Σ_2^0 sentences are true in $\mathcal{A}$ if and only if they are true in a cohesive power of $\mathcal{A}$. While Π_3^0

sentences true in a cohesive power of $\mathcal{A}$ are also true in $\mathcal{A}$, the converse does not hold. Feferman, Scott, and Tennenbaum found a Π^0_3 sentence that is true in the standard model of arithmetic but not in its cohesive power. Recently, simpler examples of such sentences have been produced for linear orders, the cohesive powers of which has been investigated by Dimitrov, Harizanov, Morozov, Shafer, A. Soskova, and Vatev. They also showed that if a structure $\mathcal{A}$ is n-decidable, then its cohesive power satisfies the same Π^0_{n+2} and Σ^0_{n+2} sentences as $\mathcal{A}$ does. The *elementary diagram* (or complete diagram) of a structure $\mathcal{A}$ is the set of all first-order sentences of L with added constants for the elements of the domain, which are true in $\mathcal{A}$. A Σ^0_n *diagram* of $\mathcal{A}$ is the set of all Σ^0_n sentences in the elementary diagram. A structure is *n-decidable* if its domain is computable and its Σ^0_n diagram is computable. A structure is *decidable* if its domain is computable and its elementary diagram is computable. Computable structures are the same as 0-decidable structures. Hence, for a decidable structure, the effective ultrapower is elementarily equivalent to the structure. Thus, more decidability in the structure leads to more satisfiability in the effective ultrapower.

The chapter *Fermat's Last Theorem* by McLarty presents the historical development of the ideas and knowledge needed for the solution of one of the most famous problems in mathematics, which has had great impact on several different areas. Fermat's last theorem states that no three nonzero integers a, b, and c satisfy the equation $a^n + b^n = c^n$ for any natural number $n > 2$. It has a very long history. Around 1637, Pierre de Fermat wrote on a margin of his copy of *Arithmetica* that he has a "marvelous proof" that there is no such solution, a remark discovered only after his death. For $n = 2$, there are infinitely many such solutions, known as Pythagorean triples. While Fermat himself had a proof for $n = 4$, no general proof of his statement was found, so the statement became known as Fermat's conjecture. For a long time, Fermat's conjecture seemed like an isolated statement of number theory without many connections to other areas of mathematics and without significant philosophical interest. In spite of efforts of many, the conjecture remained unresolved until 1994 when Andrew Wiles (with help of Richard Taylor) completed the proof he announced a year earlier, confirming Fermat's conjecture.

It can be shown that to prove that Fermat's equation has no solution for $n > 2$, it suffices to prove that it has no solution for $n = 4$ and for all odd primes n. The special cases for $n = 3$, 5, and 7 were resolved in the eighteenth and nineteenth century by Euler, Dirichlet and Legendre, and Lamé, respectively. Moreover, in the early nineteenth century, Germain developed several new approaches toward more general solution of the problem and established a theorem about the divisibility of solutions to Fermat's equation. In the mid-nineteenth century, Kummer obtained a proof for prime exponents known as regular primes.

The case $n = 3$ was first established by Euler in the 1770s. Although his proof had a part that needed to be fixed, McLarty argues that it was very clever and contributed to the idea of the emerging field of algebraic number theory. One of the central topics of algebraic number theory is the study of algebraic number fields. An algebraic number field is a finite-degree extension of the field $\mathbb{Q}$ of rational numbers, where degree means the dimension of the field as a vector space over $\mathbb{Q}$. For example, a

field can be extended by the adjunction of the k-th roots of elements of the field. Kummer's work on Fermat's last theorem contributed to the development of this theory. Algebraic number fields contain rings of integers, which may or may not have unique prime factorization, the topic which led to the development of the theory of ring ideals and class field theory. The need for such new concepts illustrates what philosophers call formation of mathematical concepts. While these developments produced many important results, the methods stemming from Kummer's work failed to completely resolve Fermat's conjecture.

It turned out that the solution to Fermat's conjecture was closely related to several other areas of mathematics, including number theory and geometry. McLarty talks about the *Langlands program* which is a broad project in modern mathematical research about various important and influential connections between number theory and geometry. In the 1970s, Hellegouarch and Demyanenko linked Fermat's conjecture to the arithmetic of elliptic curves. A *real elliptic curve* is defined by the cubic equation $y^2 = x^3 + ax + b$, where the coefficients a, b are real numbers and the roots of the right-hand side are all distinct. If a and b are rational numbers, such an elliptic curve is called an *arithmetic elliptic curve*. It is an amazing geometric fact that the intersecting lines with the cubic curve determine rational points on the curve that form a commutative group. A point is rational if both of its coordinates are rational numbers. A point r is torsion if $nr = 0$ for some positive natural number n. Hellegouarch and Demyanenko independently showed that if certain kinds of torsion points can occur on arithmetic elliptic curves, then Fermat's Conjecture will be refuted. However, Mazur proved that such points are not possible.

While quadratic curves have genus 0, and elliptic curves have genus 1, the curves of the Fermat's equations for $n > 3$ have genus >1. Mordell conjectured that every algebraic curve with genus >1 has finitely many rational points. In 1983, Faltings established Mordell's conjecture by proving that any equation in two variables corresponding to a curve of genus >1 has finitely many rational solutions. This led to the result marking a big progress toward establishing Fermat's conjecture, that for each exponent $n > 2$, Fermat's equation has at most finitely many integer solutions.

Frey associated to every integer solution of the Fermat's equation with exponent n an elliptic curve, which later became known as Frey-Hellegouarch curve. Proving Fermat's last theorem amounted to showing that such a curve does not exist. In 1982, Frey suggested that these curves cannot have certain property called modularity. Serre formulated the conjecture precisely, which was later proven by Ribet and published in 1990. The *modularity conjecture*, also known as Shimura-Taniyama-Weil conjecture, which states that every arithmetic elliptic curve is modular would establish Fermat's conjecture. Ribet's theorem showed that establishing the modularity conjecture for a class of elliptic curves called semistable elliptic curves would be sufficient for the proof of Fermat's last theorem. Wiles announced his proof in 1993, but there was a gap, which was circumvented with assistance from Taylor and completed in 1994. (Full modularity conjecture was proved in 2001.)

Ribet's proof builds on many earlier results. McLarty also argues that Ribet's and other proofs implicitly use the language of category theory. Wiles' proof is based on replacing the problems about elliptic curves with the problems of Galois

representations, which are algebraic in nature. He then uses Mazur's theory of deformations. The full proof is contained in two papers, one by Wiles and one written jointly by Wiles and Taylor, which together make up the whole of the May 1995 issue of the *Annals of Mathematics*. Wiles's proof is amazing, suggesting further applications and discoveries. In 1993, Mazur said (Kolata 1993): "A lot more is proved than Fermat's last theorem. One could envision a proof of a problem, no matter how celebrated, that had no implications. But this is just the reverse. This is the emergence of a technique that is visibly powerful. It is going to prove a lot more."

Now, the logic question arises whether Fermat's last theorem is provable in Peano arithmetic or some finitary fragment of Peano arithmetic. This is widely open. Peano arithmetic can interpret polynomials with integer coefficients and can deal with rational approximations to real numbers. McLarthy explains Macintyre's argument that the real challenge to proving Fermat's last theorem in Peano arithmetic is finding approximations to the modularity theorem. Bounded arithmetic with exponentiation is also called exponential function arithmetic and abbreviated by EFA. It is worthwhile to investigate whether Fermat's last theorem is provable in EFA.

McLarty shows how Fermat's last theorem demonstrates the unity of mathematics and interactions of different and seemingly disconnected areas. It inspired new areas to be created. It also demonstrated the need to state results precisely and in a common, uniform, and efficient language so that they are available to other disciplines if needed. It showed that concrete examples are often useful as guide and illustrations of more abstract and more general theories.

2 Cross-References

- ▶ Computability Theory
- ▶ Countable Nonstandard Models: Following Skolem's Approach
- ▶ Fermat's Last Theorem
- ▶ Model Theory of Nonstandard Structures with Applications
- ▶ Reverse Mathematics

Acknowledgments Harizanov gratefully acknowledges support of FRG NSF grant DMS-2152095.

References

Chang CC, Keisler HJ (1990) Model theory. North-Holland, Amsterdam

Friedman H (1975) Some systems of second order arithmetic and their use. In: Proceedings of the international congress of mathematicians. Vancouver 1974, vol 1, pp 235–242

Kolata G (1993) At last, shout of 'Eureka!' in age-old math mystery. New York Times

Skolem Th (1955) Peano's axioms and models of arithmetic. In: Mathematical interpretations of formal systems. North-Holland, Amsterdam, pp 1–14

Soare RI (2016) Turing computability. Theory and applications. Springer-Verlag, Berlin

Model Theory of Nonstandard Structures with Applications

Roman Kossak

Contents

1 Introduction 1922
2 Preliminaries 1923
3 Minimality and Order-Minimality 1925
4 Infinite Ramsey Theorem 1927
5 Automorphisms of the Field of Complex Numbers 1929
References 1930

Abstract

Every infinite mathematical structure $\mathcal{M}$ has an extension $\mathcal{M}^*$ that has the same first-order properties as $\mathcal{M}$, but is not isomorphic to $\mathcal{M}$. In this sense, $\mathcal{M}^*$ can be considered a nonstandard extension of $\mathcal{M}$. A short discussion of the idea of nonstandard models is followed by proofs of three easy standard results that use nonstandard extensions in essential ways. The aim is to explain basic model-theoretic concepts behind such proofs.

Keywords

Model theory · First-order logic · Nonstandard models · Automorphisms

R. Kossak (✉)
The Graduate Center, CUNY, New York, NY, USA
e-mail: rkossak@gc.cuny.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_71

1 Introduction

The set of natural numbers N with addition and multiplication is called the standard model of arithmetic. A model of arithmetic is the set of positive elements of a discretely ordered ring whose addition and multiplication satisfy a set of arithmetic axioms. A model of arithmetic is nonstandard if it is not isomorphic to the standard model.

True Arithmetic, TA, is the set of all first order sentences that hold in the standard model. In a 1934 paper, Thoralf Skolem gave a construction of a nonstandard model of TA. This was originally seen as a negative result: first-order logic is incapable of characterizing the standard model up to isomorphism. The nonstandard models showed up as uninvited guests obfuscating the foundational landscape of mathematics. Hence, perhaps slightly pejorative label "nonstandard."

In the development of mathematical logic in the second half of the 20th century, nonstandard models of arithmetic got vindicated. Here is a quotation from Gaifman (2004):

> Non-standard models are interesting mathematical structures, studied as such. Their interest is to a large extent logic oriented, stemming from the fact that the models share many properties with the standard one, but is very different structurally.

Nonstandardness in set theory has a different nature. In arithmetic, there is one intended object of study – the standard model. There is no intended model in set theory. Set-theorists talk about universes of sets, but what happens in those universes very much depends on the axioms that they satisfy, and there are many axiomatic systems to choose from.

There is no standard model of set theory, but still there is a notion of nonstandardness. The core of each model of set theory is its class of ordinal numbers. Each ordinal number α is a well-ordered set of ordinal numbers smaller than α. Every model of set theory thinks that its ordinals are well ordered. This means that each subset of α, that is an element of the model has a least element. However, α may have other subsets, and those subsets may not have least elements. If this happens, the model is considered nonstandard.

In the 1960s, Abraham Robinson, inspired by Skolem's discovery of nonstandard models of arithmetic, introduced a new approach to infinitesimal calculus – nonstandard analysis. He used model theory to move analysis from the (standard) ordered field of real numbers to its (nonstandard) extension that included infinite numbers and infinitesimals. The standard/nonstandard terminology applies mostly to the distinction between the standard real numbers and the infinite numbers and infinitesimals.

In nonstandard analysis, and in some applications of nonstandard models of arithmetic, one uses elementary extensions to formalize arguments involving infinitely large and infinitely small quantities. It matters that the usual arithmetic and algebraic arguments apply to such quantities as well. Structural, model-theoretic properties of the extension usually do not play a role.

I will present nonstandard proofs of three standard results that use model theory in a different way.

Every infinite structure has an elementary extension. There is an extensive knowledge on where such extensions come from and what properties they may have. If $\mathcal{N}$ is an elementary extension of $\mathcal{M}$, the key observation is that while $\mathcal{M}$ and $\mathcal{N}$ share the same first-order properties, their second-order properties may be very different, and that can be useful.

Here is a quick example. All details are explained in the following two sections. Let (N, <) be the ordered set of the natural numbers. If (N^*, <) is an elementary extension of (N, <) then $N^*\backslash N$ has no least element. While (N, <) is well-ordered, (N^*, <) is not. Being well-ordered is a second-order property. Its definition has a quantifier ranging over all subsets of the ordered set. We will take advantage of the difference between (N, <) and (N^*, <) to prove something about (N, <), namely, we will show that (N, <) is minimal, as defined in the third section. The result itself is just an observation, but its proof is remarkable. Since (N, <) is well ordered, it has no nontrivial automorphism. No elementary extension of (N, <) is well ordered, and all elementary extensions of (N, <) admit nontrivial automorphisms. We will use this structural property to prove the result.

In the fourth section, I will give a nonstandard proof of the infinite Ramsey theorem. The proof is only a modification of the standard proof, but I will also outline a proof of a stronger result that uses a specific result about conservative elementary extensions of model of arithmetic.

The last section is devoted to a proof that the field of complex numbers $\mathcal{C}$ has a nontrivial automorphisms other than complex conjugation. The result formulated this way is a bit unusual. Every permutation of a transcendence basis of $\mathcal{C}$ extends to an automorphism; hence, the automorphism group of the complex field is of cardinality $2^{2^{\aleph_0}}$. However, if we try to come up with an example of nontrivial automorphism that is not complex conjugation, none comes to mind. It is not hard to see that complex conjugation is the only nontrivial continuous automorphism. This suggests that other automorphisms are not easy to come by. In fact it has been shown that the statement that the complex field has only two automorphisms is consistent with the axioms of Zermelo-Fraenkel set theory (ZF) without the axiom of choice. The nonstandard argument, presented in the last section, allows to see that there must be automorphisms in question. The proof is generalized to show that the automorphism group of $\mathcal{C}$ has the maximum number $2^{2^{\aleph_0}}$ of conjugacy classes.

There are deep applications of nonstandard methods in algebra, analysis, topology, functional analysis, probability, stochastic analysis, and number theory. For examples, see Tao (2013) and Loeb and Wolff (2015). The results presented here are chosen for their simplicity. My main goal is to explain the elements of model theory behind them.

2 Preliminaries

This section is brief overview of basic concepts of model theory that will be used in this note.

In model theory, a structure is a set M, called the domain of the structure, together with sets of functions $\{f_i\}_{i\in I}$, relations $\{R_j\}_{j\in J}$, and constants $\{c_k\}_{k\in K}$. For each

i, j, and k, $f_i : M^{n_i} \to M$, $R_j \subseteq M^{m_j}$, and $c_k \in M$. The index sets are arbitrary, in particular, each of them could be empty.

We will use $\mathcal{M}$, $\mathcal{N}$ for structures and M, N, ... for their respective domains. The exceptions are N, Z, Q, R, and C, which stand for the sets of natural, integer, rational, real, and complex numbers.

The logical symbols of first-order logic are variables, Boolean connectives, existential and universal quantifiers, left and right brackets, and the equality symbol. The language of a structure $\mathcal{M}$ consists the set of function symbols f_i, relation, and constant symbols for the functions, relations, and constants of $\mathcal{M}$. The language of a structure is also called its signature.

By a formula of the language of a structure, we mean a first-order formula whose nonlogical symbols are in the language of the structure.

We will use the bar notation for tuples of elements and strings of variables. If $\varphi(\overline{x})$ is a formula in the language of a structure $\mathcal{M}$ and $\overline{a}$ is a tuple of elements of M, then we say that $\varphi(\overline{a})$ is satisfied in $\mathcal{M}$, written if $\mathcal{M} \models \varphi(\overline{a})$, if what $\varphi(\overline{x})$ says holds in $\mathcal{M}$. What I wrote here looks trivial. Just substitute "holds" for "is satisfied." Hidden under the rug is Tarski's formal definition of satisfaction that proceeds by induction on the logical complexity of formulas. What I mean by "$\varphi(\overline{a})$ holds in $\mathcal{M}$" refers to the intuitive understanding of first-order formulas that is taught in discrete mathematics courses. Tarski's definition can be found in any textbook on mathematical logic. For us here, the intuitive approach suffices.

A theory is a set of first-order sentences (formulas without free variables) in a given language. If T is a theory and $\mathcal{M}$ is a structure for the language of T, then $\mathcal{M}$ is a model of T, written $\mathcal{M} \models T$, if $\mathcal{M} \models \varphi$ for all φ in T.

A type of a tuple of elements from the domain of a structure $\mathcal{M}$, written $\mathrm{tp}^{\mathcal{M}}(\overline{a})$, is the set of formulas $\{\varphi(\overline{x}) : \mathcal{M} \models \varphi(\overline{a})\}$, where $\varphi(\overline{x})$ ranges over all formulas of the language of $\mathcal{M}$. In other words, $\mathrm{tp}^{\mathcal{M}}(\overline{a})$ is the set of all first-order properties that $\overline{a}$ has in $\mathcal{M}$.

If $\mathcal{M}$ and $\mathcal{N}$ are structures for the same language and $M \subseteq N$, then $\mathcal{N}$ is an elementary extension of $\mathcal{M}$ if for all tuples $\overline{a}$ from $\mathcal{M}$, $\mathrm{tp}^{\mathcal{M}}(\overline{a}) = \mathrm{tp}^{\mathcal{N}}(\overline{a})$. When we move from $\mathcal{M}$ to an elementary extension, the types of the tuples from $\mathcal{M}$ don't change. Below, when we say that $\mathcal{N}$ is an elementary extension of $\mathcal{M}$, we tacitly assume that M is a proper subset of N.

For an example of an extension that is not elementary, let $\mathcal{M} = (\mathrm{N} \setminus \{0\}, <)$ (0 is a natural number) and $\mathcal{N} = (\mathrm{N}, <)$. Let $\varphi(x)$ be $\forall y \neg (y < x)$. Then $\mathcal{M} \models \varphi(1)$ and $\mathcal{N} \models \neg\varphi(1)$; hence, $\mathrm{tp}^{\mathcal{M}}(1) \neq \mathrm{tp}^{\mathcal{N}}(1)$. In fact, for each $n > 0$, $\mathrm{tp}^{\mathcal{M}}(n) \neq \mathrm{tp}^{\mathcal{N}}(n)$, and that is because each element of either structure is uniquely determined by the number of elements separating it from the least element, and there is a first-order sentence expressing that number is n; one sentence for each natural number n.

The compactness theorem says that if every finite set of sentences of a theory T has a model, then T has a model. The following theorem is an easy consequence of the compactness theorem. It can also be proved directly by an ultrapower construction. Ultraproducts and ultrapowers are in the arsenal of nonstandard methods and are important for applications in algebra. For a full discussion and advanced applications in commutative algebra, see Schoutens (2010).

We say that a structure is infinite if its domain is infinite.

Theorem 1 *Every infinite structure has an elementary extension.*

Let $\mathcal{M}$ be (M, f, R, c), where $f : M \to M$, $R \subseteq M^2$, and $c \in M$. Let α be a permutation of M, then α is an automorphism of $\mathcal{M}$ if for all $a, b \in M$:

- $\alpha(f(a)) = f(\alpha(a))$;
- $(a, b) \in R$ iff $(\alpha(a), \alpha(b)) \in R$
- $\alpha(c) = c$.

A general definition for an arbitrary structure $\mathcal{M}$ has the same conditions for all functions, relations, and constants of $\mathcal{M}$.

The model-theoretic version of the second bulleted condition above is

$$\mathcal{M} \models R(a, b) \Leftrightarrow R(\alpha(a), \alpha(b)).$$

This is the base case in the proof by induction by complexity of formulas of the following very useful fact. Below, if $\overline{a} = a_1, \ldots, a_n$, then $\alpha(\overline{a})$ is $\alpha(a_1), \ldots, \alpha(a_n)$.

Theorem 2 *Let α be an automorphism of $\mathcal{M}$. Then, for all $\overline{a}$, $\mathrm{tp}^{\mathcal{M}}(\overline{a}) = \mathrm{tp}^{\mathcal{M}}(\alpha(\overline{a}))$.*

Let $\mathcal{M}$ be a structure. A set $A \subseteq M^n$, for $n > 0$, is definable in $\mathcal{M}$ if there are a formula $\varphi(\overline{x}, \overline{y})$ of the language of $\mathcal{M}$ a tuple $\overline{b}$ such that

$$A = \left\{\overline{a} : \mathcal{M} \models \varphi\left(\overline{a}, \overline{b}\right)\right\}.$$

3 Minimality and Order-Minimality

In every structure, all finite sets and their complements are definable. For example, the set $\{a, b\}$ is defined by $(x = a) \vee (x = b)$, and its complement is defined by $\neg[(x = a) \vee (x = b)]$. A structure $\mathcal{M}$ such that every definable subset of M is either finite or cofinite is called minimal.

The definition of minimality involves only definable subsets of the domain M. Every infinite structure has definable sets that are neither finite nor cofinite. For example, $\{(x, y) : \mathcal{M} \models x = y\}$.

The ordered set of integers $(Z, <)$ is not minimal. For example, the set of positive integers is definable and it is neither finite nor cofinite.

A prime example of a minimal structure is the field of complex numbers $\mathcal{C}$. The proof that $\mathcal{C}$ is minimal is nontrivial. The result is known as the Chevalley–Tarski theorem, and its equivalent version in algebraic geometry says that the projection of a constructible set is constructible. This formulation of the theorem reveals an important aspect of minimality. The projection of a definable set is definable. For example, if $A \subseteq M^2$ is defined by $\varphi(x, y)$, then its projection onto the first coordinate

is defined by $\exists y\varphi(x, y)$. It is hard to see directly what subsets of the domain of a structure are definable just by inspecting formulas with one free variable. Instead, it is often necessary to analyze definability of sets in higher dimensions, that is, sets defined by formulas with arbitrary numbers of free variables.

The rest of this section is devoted to the proof that $(\mathrm{N}, <)$ is minimal. It is not an important theorem; in Doets (1996), it is just an exercise that follows the argument given for the special case that even numbers are not definable in $(\mathrm{N}, <)$. The result can be proved by an analysis of what first-order formulas of the language with just one binary relation symbol $<$ can define in $(\mathrm{N}, <)$. The proof in Doets (1996) that we will follow here is different. We will use an elementary extension and an automorphism.

To show that $(\mathrm{N}, <)$ is minimal, suppose it is not. Then, there is a formula $\varphi(x)$ that defines a set that is neither finite nor cofinite. Let $(\mathrm{N}^*, <)$ be an elementary extension of $(\mathrm{N}, <)$. Because the extension is elementary, $(\mathrm{N}^*, <)$ is a linearly ordered set, and every element of N^* has an immediate successor, and every element other than 0 has an immediate predecessor.

In $(\mathrm{N}, <)$ every element is uniquely determined by its distance from 0. This cannot change in the elementary extension. Hence, for any standard n and any nonstandard c, $(\mathrm{N}^*, <) \models n < c$. Then, for every nonstandard c, $[c] = \{c + n : n \in \mathrm{Z}\} \subseteq \mathrm{N}^*$, where $c + n$ denotes either n the successor of c or $-n$ predecessor of c, depending on whether n is positive or negative. For the proof that $(\mathrm{N}, <)$ is minimal, we do not need anything more.

Let ψ be the sentence $\forall x\ \exists\ y[x < y \wedge \varphi(y) \wedge\ \neg\ \varphi(y + 1)]$.

Since the set defined by $\varphi(x)$ in $(\mathrm{N}, <)$ is neither finite nor cofinite, $(\mathrm{N}, <) \models \psi$. By elementarity of the extension, $(\mathrm{N}^*, <) \models \psi$. It follows that there is a nonstandard c such that $(\mathrm{N}^*, <) \models \varphi(c) \wedge\ \neg\ \varphi(c + 1)$. For such a c, let $\alpha(x) = x + 1$ if $x \in [c]$, otherwise $\alpha(x) = x$.

Because α is an order-preserving permutation of N^*, it is an automorphism of $(\mathrm{N}^*, <)$. By Theorem 2, c and $\alpha(c)$ have the same type in $(\mathrm{N}^*, <)$, but we also have $(\mathrm{N}^*, <) \models \varphi(c) \wedge\ \neg\ \varphi(\alpha(c))$ and this contradiction finishes the argument.

In the argument, nothing was mentioned about parameters from N that $\varphi(x)$ could have. That was unnecessary. In $\mathcal{N}$ all elements are definable; hence, all formulas have equivalent parameter-free versions, so without loss of generality we could assume that $\varphi(x)$ has no parameters.

Structures $\mathcal{M}$ and $\mathcal{N}$ for the same language are elementarily equivalent if for each first-order sentence φ, $\mathcal{M} \models \varphi$ if and only if $\mathcal{N} \models \varphi$.

Since the type of each element of a structure includes all sentences that hold in the structure, if $\mathcal{N}$ is an elementary extension of $\mathcal{M}$, then $\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent.

A structure $\mathcal{M}$ is strongly minimal if it is minimal, and every structure that is elementarily equivalent to it is also minimal. The field of complex numbers is strongly minimal, and so are most known minimal structures. The ordered set $(\mathrm{N}, <)$ is a rare exception. Every elementary extension of $(\mathrm{N}, <)$ contains a copy of $(\mathrm{Z}, <)$; hence, none of those extensions is minimal.

A structure is ordered if one of its relations linearly orders the domain. Any ordered structure that has a subset order-isomorphic to (Z, <) is not minimal, but some of them are o-minimal. A structure is o-minimal (short for order-minimal) if each definable subset of its domain is the union of finitely many open intervals, including the unbounded ones, and points. A prime example of an o-minimal structure is the ordered field of real numbers. This nontrivial result is known as the Seidenberg–Tarski theorem. In terms of real analytic geometry, it says that the projection of a semi-algebraic set is semi-algebraic.

A proof similar to the one for the minimality of (N, <) shows that (Z, <) is o-minimal. Unlike (N, <), (Z, <) has nontrivial automorphisms. They are all of the form $\alpha(x) = x + n$, for integer n; but none of them can be used in the proof, since we need automorphisms that fix any finite number of parameters from Z. For that reason, we move to an elementary extension (Z^*, <) and argue as in the case of (N, <).

All dense linear orderings are o-minimal. There is simple proof using automorphisms that the ordered set of the rational numbers (Q, <) is o-minimal. There is no need to move to an elementary extension. The automorphism group of (Q, <) is rich enough.

In conclusion, the point of this section was to show how certain results about an intended, standard structure can be obtained by moving the argument to its nonstandard elementary extension. This works when the extension has some useful second-order properties that the standard structure lacks. The example of (N, <) might be the simplest, but it is not isolated. The analysis and classification of definable sets and types of elements in a structure are often made easier by moving an elementary extensions that admit lots and lots of automorphism, and some very general results in model theory show that such extensions always exist.

4 Infinite Ramsey Theorem

For a set $X \subseteq \mathrm{N}$, $[X]^n$ is the set of all increasing n-tuples of elements of X. The infinite Ramsey theorem says that for all k, $n > 0$, for every function $f: [\mathrm{N}]^n \to \{1, \ldots, k\}$, there is an infinite $X \subseteq \mathrm{N}$ such that f is constant on $[X]^n$. We call X a homogeneous set for f.

We will only prove the theorem for $k = n = 2$. Increasing k only makes some formulas a bit longer. The proof for arbitrary n proceeds by induction. To prove the inductive step, one reduces the problem of finding a homogeneous set X for a given partition of $[\mathrm{N}]^{n+1}$ to the problem of finding a homogeneous set for a partition of $[\mathrm{N}]^n$. We will show how it is done in the case of $n = 2$. For $n = 1$, the theorem just says that in every partition of an infinite set into finitely many parts, at least one of the parts is infinite.

Let $f: [\mathrm{N}]^2 \to \{1, 2\}$ be given. Let $\mathcal{M}$ be $(\mathrm{N}, <, f)$ and $\mathcal{M}^* = (\mathrm{N}^*, <, f^*)$ be an elementary extension of $\mathcal{M}$.

Let c be an element of $\mathrm{N}^* \backslash \mathrm{N}$. As was shown in the previous section, for all standard n, $\mathcal{M}^* \models n < c$.

Since the extension is elementary, f^* extends to f and it is a function from $[\mathrm{N}^*]^2$ to $\{0, 1\}$.

We will define a sequence of natural numbers $\{b_n\}_{n\in\mathrm{N}}$ by recursion. Let b_0 be any natural number, and let $i_0 = f^*(b_0, c)$. We have

$$\mathcal{M}^* \models \exists x[b_0 < x \wedge f^*(b_0, x) = i_0].$$

Since b_0 and i_0 are in N, by elementarity

$$\mathcal{M} \models \exists x[b_0 < x \wedge f(b_0, x) = i_0].$$

Let b_1 be the least such x, and let $i_1 = f^*(b_1, c)$]. We have,

$$\mathcal{M}^* \models \exists x[b_1 < x \wedge f^*(b_0, x) = i_0 \wedge f^*(b_1, x) = i_1].$$

Arguing as above, let b_2 to be the least natural number such that

$$b_1 < b_2 \wedge f(b_0, b_2) = i_0 \wedge f(b_1, b_2) = i_1.$$

Now let $i_2 = f^*(b_2, c)$. . ..

Proceeding in this manner for each $n > 0$, we get an increasing sequence such that for all $j < n - 1$ and all $j < k < n$, $f(b_j, b_k) = i_j = f^*(b_j, c)$. It follows that $\{b_n\}_{n\in\mathrm{N}}$ is partitioned into

$$\{b_i : \forall k[i < k \to f(b_i, b_k) = 0]\} \cup \{b_j : \forall k[\, j < k \to f(b_j, b_k) = 1]\}.$$

At least one of these sets must be infinite, and we can take any of them that is infinite to be the set X required in the conclusion of the theorem.

The proof outlined above is not much different from the standard proof, but its details are simplified by the possibility of referring to the nonstandard c.

The sequence $\{b_n\}_{n\in\mathrm{N}}$ is determined by the values $f^*(n, c)$, for $n \in \mathrm{N}$. Since the extension $\mathcal{M}^*$ and c were chosen arbitrarily, we do not know what those values might be, so what the proof tells are very little about the homogeneous set X. We can get more information from a more advanced variant of the argument.

A set is arithmetic if it is definable in the standard model of arithmetic. The arithmetic version of Ramsey's theorem can be stated as follows.

Theorem 3 *For all k, $n > 0$, for every arithmetic partition $f : [\mathrm{N}]^n \to \{1, \ldots, k\}$ there is an arithmetic infinite homogeneous set.*

We finish this section with an outline of the proof. For details and a more general statement, see Kossak and Schmerl (2006, Theorem 2.2.16).

If $\mathcal{N}$ is an elementary extension of $\mathcal{M}$, then we say that the extension is conservative if for every definable set X in $\mathcal{N}$, $X \cap M$ is definable in $\mathcal{M}$. By a result of Phillips (1974), the standard model of arithmetic has a conservative extension.

Let an arithmetic $f : [\mathrm{N}]^2 \to \{0, 1\}$ be given. Let $\mathcal{M}$ to be the standard model of arithmetic, and let $\mathcal{M}^*$ its conservative elementary extension. Because f and the

ordering of the natural numbers have arithmetic definitions, there is no need to expand (N, +, ·) to (N, +, ·, <, f).

The recursive construction of $\{b_n\}_{n\in\mathrm{N}}$ can be carried out in $\mathcal{M}^*$ up to a nonstandard number d, that is, $\{b_n\}_{n\in\mathrm{N}}$ becomes the standard part of the sequence $\{b_n\}_{n<d}$. Since $\{b_n\}_{n<d}$ is definable in $\mathcal{M}^*$, by conservativity of the extension, $\{b_n\}_{n\in\mathrm{N}}$ is definable in $\mathcal{M}$. The homogeneous set X has a definition that uses $\{b_n\}_{n\in\mathrm{N}}$ as a set parameter; hence X is arithmetic.

5 Automorphisms of the Field of Complex Numbers

Let $\mathcal{C}$ be the field of complex numbers. Any two algebraically closed fields of the same characteristic, cardinality, and transcendence degree are isomorphic. Hence, up to isomorphism, $\mathcal{C}$ is the only algebraically closed field of characteristic 0 and cardinality $2^{\aleph_0}$.

In the case of finite or countable languages, Theorem 1 can be restated as: Every infinite structure has an elementary extension of the same cardinality. If $\mathcal{C}^*$ is such an extension of $\mathcal{C}$, then, by elementarity, it is also algebraically closed; hence, $\mathcal{C}^*$ is isomorphic to $\mathcal{C}$. In this sense, it is not really a nonstandard extension, nevertheless we will use the fact that it is a proper extension to prove something about $\mathcal{C}$, but first we need to represent $\mathcal{C}^*$ in a special way.

Let $\mathcal{R}$ be the field of real numbers. In contrast with $\mathcal{C}$, no elementary extension $\mathcal{R}^*$ is isomorphic to $\mathcal{R}$. It is easy to see that each such an extension has to have infinitely large numbers. It is a non-Archimedean ordered field.

We will prove that $\mathcal{C}$ has a nontrivial automorphism other than complex conjugation.

Let $\mathcal{R}^*$ be an elementary extension of $\mathcal{R}$ of cardinality $2^{\aleph_0}$, and let

$$C^* = \{a + bi : a, b \in R^*\}.$$

There are proofs of the fundamental theorem of algebra that can be adapted to show that $\mathcal{C}^*$, whose domain is C^*, and its addition and multiplication defined as in the field of complex numbers is an algebraically closed field, but there is also a short model-theoretic argument. The field of complex numbers can be represented as R^2 with addition and multiplication that are definable in $\mathcal{R}$. Moreover, the fact that, under this representation, $\mathcal{C}$ is an algebraically closed field is expressed by a set of first-order sentences comprised of the field axioms, and an infinite set of sentences expressing that all polynomial equations of degree n have solutions – one sentence for each n. Since $\mathcal{R}^*$ is an elementary extension of $\mathcal{R}$, all those sentences are true about $\mathcal{C}^*$, and that concludes the argument.

It is also true that $\mathcal{C}^*$ is an elementary extension of $\mathcal{C}$, but this not be used in the argument below.

Since the cardinality of C^* is $2^{\aleph_0}$, $\mathcal{C}^*$ is isomorphic to $\mathcal{C}$, let $F : C^* \to C$ be an isomorphism.

For $a,\ b \in \mathrm{C}^*$, let $\alpha(a+bi) = a - bi$ and let $\beta = F(\alpha)$. Then β is an automorphism of $\mathcal{C}$, and the set of fixed points of β is an ordered non-Archimedean subfield of $\mathcal{C}$. Hence, β is not conjugation in $\mathcal{C}$.

While the result presented above is more of an anecdote, with a bit more of model theory we can prove more.

A real closed field is a field that is elementarily equivalent to the ordered field of real numbers. In particular, all elementary extensions of $\mathcal{R}$ are real closed. There are $2^{2^{\aleph_0}}$ pairwise nonisomorphic real closed fields of the same size. This follows from general results about isomorphism types of ordered structures, but also can be proven in this special case in several different ways. For interesting recent results on model theory of real closed fields, their integer parts, and nonstandard models of arithmetic, see D'Aquino et al. (2010) or Dolich et al. (2015).

Let $\mathcal{G}$ be the automorphism group of a structure $\mathcal{M}$. Then f and g in $\mathcal{G}$ are conjugate if for some h in $\mathcal{G}, f = hgh^{-1}$. Model theoretically, f and g are conjugate if and only if the expanded structures $(\mathcal{M}, f)$ and $(\mathcal{M}, g)$ are isomorphic. The conjugacy class of f is the set of all automorphisms of $\mathcal{M}$ that are conjugate to f.

If $\mathcal{S}$ is a real closed field, then, by the same argument as the one we used at the beginning of this section, there is an automorphism $\alpha_{\mathcal{S}}$ of $\mathcal{C}$ such that the set of fixed points of $\alpha_{\mathcal{S}}$ is isomorphic to $\mathcal{S}$. If $\mathcal{T}$ is a real closed field that is not isomorphic to $\mathcal{S}$, then $\alpha_{\mathcal{S}}$ and $\alpha_{\mathcal{T}}$ are in different conjugacy classes. Thus, we have shown that the automorphism group of $\mathcal{C}$ is of size $2^{2^{\aleph_0}}$. Moreover, we proved the following theorem.

Theorem 4 The automorphism group of $\mathcal{C}$ has $2^{2^{\aleph_0}}$ conjugacy classes.

Another example of an application of a nonstandard extension of $\mathcal{C}$ is the proof of the fundamental theorem of algebra in Leibman (2005). The proof uses a small amount of model theory. It is a standard, nonstandard analysis style proof involving a fixed point lemma applied to a topology on a quotient space of $\mathcal{C}^*$.

References

D'Aquino P, Knight JF, Starchenko S (2010) Real closed fields and models of Peano arithmetic. J Symb Log 75(1):1–11. https://doi.org/10.2178/jsl/1264433906. https://doi-org.ezproxy.gc.cuny.edu/10.2178/jsl/1264433906

Doets K (1996) Basic model theory. Studies in logic, language and information. CSLI Publications, Stanford; FoLLI: European Association for Logic, Language and Information, Amsterdam

Dolich A, Knight JF, Lange K, Marker D (2015) Representing Scott sets in algebraic settings. Arch Math Log 54(5–6):631–637. https://doi.org/10.1007/s00153-015-0431-1. https://doi-org.ezproxy.gc.cuny.edu/10.1007/s00153-015-0431-1

Gaifman H (2004) Non-standard models in a broader perspective. In: Nonstandard models of arithmetic and set theory, Contemporary mathematics, vol 361. American Mathematical Society, Providence, pp 1–22. https://doi.org/10.1090/conm/361/06585. https://doi-org.ezproxy.gc.cuny.edu/10.1090/conm/361/06585

Kossak R, Schmerl JH (2006) The structure of models of Peano arithmetic, Oxford logic guides, vol 50. The Clarendon Press, Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/

9780198568278.001.0001. https://doi-org.ezproxy.gc.cuny.edu/10.1093/acprof:oso/9780198568278.001.0001. Oxford Science Publications

Leibman G (2005) A nonstandard proof of the fundamental theorem of algebra. Am Math Mon 112 (8):705–712. https://doi.org/10.2307/30037572. https://doi-org.ezproxy.gc.cuny.edu/10.2307/30037572

Loeb PA, Wolff MPH (eds) (2015) Nonstandard analysis for the working mathematician, 2nd edn. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-7327-0. https://doi-org.ezproxy.gc.cuny.edu/10.1007/978-94-017-7327-0

Phillips RG (1974) Omitting types in arithmetic and conservative extensions. In: Victoria symposium on nonstandard analysis (Univ. Victoria, Victoria, B.C., 1972), Lecture notes in mathematics, vol 369, pp 195–202

Schoutens H (2010) The use of ultraproducts in commutative algebra, Lecture notes in mathematics, vol 1999. Springer, Berlin. https://doi.org/10.1007/978-3-642-13368-8. https://doi-org.ezproxy.gc.cuny.edu/10.1007/978-3-642-13368-8

Tao T (2013) Compactness and contradiction. American Mathematical Society, Providence. https://doi.org/10.1090/mbk/081. https://doi-org.ezproxy.gc.cuny.edu/10.1090/mbk/081

Computability Theory

Valentina Harizanov, Keshav Srinivasan, and Dario Verta

Contents

1 Brief History of Computability Theory 1934
2 Turing Computable Functions and Sets 1938
3 Partial Recursive Functions and Church-Turing's Thesis 1939
4 Enumerations and Computably Enumerable Sets 1943
5 The Halting Set, Diophantine Sets and Hilbert's Tenth Problem 1946
6 Undecidable Problems and Arithmetical Hierarchy 1949
7 Oracles and Turing Degrees 1954
8 Computable Ordinals and Hyperarithmetical Hierarchy 1956
9 Computable Models and Computable Infinitary Formulas 1958
10 Cross-References 1960
References 1960

Abstract

Computability theory is the mathematical theory of algorithms, which explores the power and limitations of computation. Classical computability theory formalized the intuitive notion of an algorithm and provided a theoretical basis for digital computers. It also demonstrated the limitations of algorithms and showed that most sets of natural numbers and the problems they encode are not decidable (Turing computable). Important results of modern computability theory include the classification of the computational difficulty of sets and problems. Arithmetical and hyperarithmetical hierarchies and the Turing degrees provide important measures of the level of such difficulty. There are uncountably many Turing degrees and they are partially ordered. These classifications are used in contemporary research to investigate the complexity of problems in modern computable mathematics, for example, regarding structures and their properties.

V. Harizanov (✉) · K. Srinivasan · D. Verta
Department of Mathematics, George Washington University, Washington, DC, USA
e-mail: harizanv@gwu.edu; ksrinivasan@gwu.edu; dverta1@gwu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_101

Keywords

Turing machine · Partial recursive · Recursively enumerable · Diophantine · Undecidable problems · Arithmetical · Hyperarithmetical · Turing degree · Computable structure · Computable formulas

1 Brief History of Computability Theory

Computability theory is the mathematical theory of algorithms. The study of particular algorithms can be traced back to Euclid (330 BCE) and his famous algorithm for determining the greatest common divisor of a pair of integers. The word *algorithm* itself has roots in the name of the mathematician Muhammad ibn Musa al-Khwarizmi who developed methods of solving linear as well as quadratic equations (e.g., completing the square) in the beginning of the ninth century. A set is computable if there is a decision procedure for identifying its elements. Computable sets often encode decidable problems.

In the late 1800s, Frege tried to put arithmetic on solid foundations, but he encountered problems with consistency exemplified by Russell's paradox. Frege used quantifiers and developed predicate logic. During the same period, Cantor worked on set theory and studied ordinal and cardinal numbers. Perhaps most importantly from the perspective of later computability theory, Cantor introduced the diagonal argument to show that an uncountable size existed. This argument will be used crucially in Gödel's incompleteness theorem, as well as to show the undecidability of the halting problem.

Computability theory is closely related to Diophantine equations and Hilbert's Tenth Problem. Diophantine equations, named after Diophantus of Alexandria, are polynomial equations in finitely many variables for which integer solutions are studied in number theory. At the Second International Congress of Mathematicians in 1900, Hilbert posed 23 unsolved problems that would challenge mathematicians of the twentieth century. Hilbert's Tenth Problem asked to devise a process that determines whether any given Diophantine equation with integer coefficients has a solution in the integers. In modern terms, it was a decision problem: asking for an algorithm to determine which Diophantine equations have integer solutions. It was the only decision problem among Hilbert's problems. If such an algorithm existed, all that would be required to solve it would be to write the algorithm down. However, as it turned out, no such algorithm exists.

To demonstrate that no algorithmic solution exists would require considerably more theoretical work, as limitative theorems often do. It is easy to construct an equilateral triangle with straightedge and compass, but the theory of constructible real numbers was needed to show that squaring the circle is impossible. It is easy to solve a quadratic equation with roots, but the development of Galois theory was needed to show that quintic equations cannot be solved with roots. In the same way, it was only after computability theory was developed that it could be shown that no algorithm could ever solve this decision problem. The negative solution to Hilbert's Tenth Problem, that there is no algorithm that determines whether a Diophantine

equation has a solution in the integers, was finally obtained in 1970 by the work of Davis, Putnam, J. Robinson, and Matiyasevich.

At the same conference, Hilbert also famously asked for a proof that arithmetic is consistent. The answer to that problem turned out to be negative. In the 1920s, Hilbert wanted to establish his own program akin to that of Frege's to get a nice axiomatic system from which all of mathematics could be deduced, but this turned out to be impossible. Gödel's famous incompleteness theorem from 1931 (Gödel 1931) established that a consistent system with an algorithmic set of axioms, capable of expressing arithmetic, cannot be complete.

Modern computability theory started in 1936 with Turing's seminar paper (Turing 1936) on "computable numbers, with an application to the Entscheidungsproblem" (decision problem). The Entscheidungsproblem was a problem posed by Hilbert and Ackermann in 1928 to devise an algorithm that decides whether a first-order sentence is universally valid, i.e., true in every first-order structure. It turned out that there is no such algorithm. Turing's paper was recommended for publication by Church, who had himself just published a paper that reached the same conclusion as Turing's, although by a different method. Later that year, Turing moved to Princeton University to study for his PhD in mathematical logic under Church's direction, which he completed in 1938. In (Turing 1936), Turing introduced a Turing machine, or as he would call it an a-machine, which is essentially an abstraction of a computing human that can read and write on a paper line while being in a certain mind state. This formalism had profound significance for the emerging science of computing. The idea of the Turing machine provides a rigorous mathematical formalism to what was until then a rough concept of being effectively calculable or computable. The Church-Turing thesis relates precisely these two concepts, saying that a function that is effectively calculable by a human is so exactly when it is computable by a Turing machine. Turing machine formalism provides an exact account of what functions are not algorithmically computable, allowing for negative solutions to decision problems. At the same time, Post (Post 1936) also worked independently on a formalism for computation, although he did not justify his system as a model of computation as Turing did.

While Turing's work was going on, there was a parallel development in mathematical logic that provided another foundation for computability – the recursive functions. That definitions by recursion could be used as the building blocks of a certain class of functions was an old idea going back to Skolem, but Gödel made it precise in 1931, when he found that he needed to define what we now call the class of primitive recursive functions in order to prove his incompleteness theorems. That is, in order to encode the provability predicate in the language of arithmetic, he needed to show that primitive recursive functions in general were definable in arithmetic. Work by Herbrand, Gödel, and Kleene generalized this notion to what we now call general recursive functions. These functions turned out to be the same as Turing computable ones. Gödel (Gödel 1931) was the first to introduce the diagonal set the complement of which can be viewed as what we now call the halting set. The halting set captures the halting problem: whether the e-th Turing machine halts on input e, which turned out to be undecidable.

The subject of computability theory started as "recursive function theory," established and promulgated by Church and Kleene when talking about

Herbrand-Gödel recursive functions (despite the disapproval of Gödel himself). Davis pushed forward the notion of calling it "computability" as evidenced by his own usage within his book (Davis 1958). In the preface, he wrote "This book is an introduction to the theory of computability and noncomputability, usually referred to as the theory of recursive functions." It became more and more standard throughout the next decades (as it can be seen in Cooper 2004; Cutland 1997; Griffor 1999; Rogers 1967; Soare 1987, 2016; Weber 2012).

Kleene is also known for Kleene-Mostowski arithmetical hierarchy that relates definability to computability in an essential way. Furthermore, Kleene worked on more general computability based on relativization and, along with Post, on Turing degrees. Post investigated various computable reductions of problems to each other and, in particular, studied Turing degrees in the 1940s and 1950s. Post was also very interested in finding a computably enumerable set that is not computable but also not Turing equivalent to the halting set. Although he never succeeded with these constructions, his attempts led to many important notions of sets with various Post-type properties. A set A is computable in a set B, or Turing reducible to B, if given a decision algorithm for membership to B, we can find a decision algorithm for membership to A; equivalently, A can be computed by a Turing machine with oracle B. This idea for oracle Turing machines originated from Turing's PhD dissertation in 1938.

Kleene and Post (Kleene and Post 1954) constructed incomparable Turing degrees computable in the halting set, using a finite extension method (also, see Cooper 2004; Soare 1987). In this method, a binary sequence representing a set is constructed by extending its finite initial segments, while attempting to satisfy certain requirements. We first decompose the conditions of the construction into infinitely many requirements, assign priorities to the requirements by listing them in an infinite sequence, and then attempt to satisfy them in order of their priority. Once satisfied, a requirement will remain satisfied forever. Such construction can be viewed as a diagonal argument, which is sometimes also called a "wait and see argument" or "finite forcing argument." However, Kleene-Post's construction used the halting set as an oracle, so the sets it constructed are computable in the halting set.

Friedberg and Muchnik (Friedberg 1957; Muchnik 1956) independently discovered the priority method with injury in order to obtain two Turing incomparable computably enumerable sets. To obtain computably enumerable sets, the construction has to be computable, so it cannot depend on any non-computable oracle. Friedberg-Muchnik's argument led to the development of a more general method called finite injury priority method, a powerful and intricate mathematical technique unique to computability theory. The finite injury priority method tries to satisfy mutually conflicting requirements arranged in a priority list by fitting together opposite strategies. While attempting to satisfy the requirements in order of their priority, it can possibly injure some lower priority requirements that have already been satisfied. As they became unsatisfied, they will need to be satisfied again, but injury can happen only finitely often for each requirement, so in the limit every requirement will be satisfied. The priority methods can be further classified based on injuries of the requirements and how many limits are needed to eventually satisfy them (see Ash and Knight 2000; Soare 1987).

Computability theory has found many applications to other areas of mathematics. In the 1970s, Metakides and Nerode and other researchers in the West and independently in the Siberian school of constructive mathematics initiated a study of computability in mathematical structures by using modern computability-theoretic tools, such as the priority methods, leading to the development of computable algebra and computable model theory and more generally computable mathematics (see Fokina et al. 2014; Harizanov 1998; Millar 1999). Other interactions of computability theory with other areas of mathematics led to the development of generalized computability, computable analysis, algorithmic randomness, algorithmic learning theory, reverse mathematics, approximate computability, and other areas.

The theory of formal languages, of great importance in computer science, is closely related to computability. Chomsky, a recognized linguist, philosopher, and cognitive scientist, developed in the 1950s a hierarchy of formal grammars and languages. A Chomsky grammar consists of a finite set T of terminal symbols and a finite set V of variables, along with substitution rules of the form $\alpha \to \beta$ for words α, β over $T \cup V$, indicating that a word α can be replaced by a word β. These rules are called productions. We can think of terminal symbols as corresponding to words in a dictionary for a natural language and of variables as corresponding to grammatical categories such as ⟨noun phrase⟩ or ⟨verb⟩. Variables include a special starting symbol, sentential variable S. The formal language generated by a specific grammar is the set of all words over the terminal symbols that can be generated starting with sentential symbol S and using the production rules of the grammar. We can think of a formal language as the set of all grammatically correct words over terminal symbols, where a word in a formal language corresponds to a sentence in a natural language. The constraints that these formal grammars must satisfy determine where they lie in the hierarchy. There are broadly four types of grammars and languages they generate, ranging from the most restrictive to the least restrictive one: regular, context-free, context-sensitive, and phrase structure. These languages have corresponding computational models, which are essentially special kinds of Turing machines where memory may be limited.

Regular languages are exactly those that can be recognized by finite state automata, context-free languages are exactly those that can be recognized by pushdown automata, and context-sensitive languages are exactly those that can be recognized by linear bounded automata. Not every language generated by Chomsky's grammar is computable. That is, the existence of a generative algorithmic grammar does not guarantee the existence of a recognition algorithm for the correct sentences of the language. For example, in Chomsky's hierarchy of languages, regular, context-free, and context-sensitive languages are all computable, while some phrase structure languages may not be computable. One of the most remarkable results in computable mathematics is that Chomsky's phrase structure languages coincide with the recursively enumerable languages. It can be established that a set is computable exactly when both the set and its complement are computably enumerable. Hence, a phrase structure language is computable when its complement (the set of all grammatically incorrect statements) is also a phrase structure language. For more on Chomsky's languages, see (Chomsky 1959; Sipser 2013).

Recently, the authors in Ji et al. (2020) claim that the class of recursively enumerable languages coincides with an important class in quantum complexity theory, thus demonstrating deep connections among quantum physics, computer science, and computability theory.

2 Turing Computable Functions and Sets

In (Turing 1936), Turing introduced a universal Turing machine, one that is capable of copying any other Turing machine and running it on any input. This was instrumental for the development of digital computers.

Each Turing machine has an infinite tape divided into cells holding symbols and has a head that can read a single symbol, write another one, and move left or right one cell at the time. The finite set of states contains a starting state q_{start} and a halting state q_{end}. In addition to a finite set of states, each Turing machine has a finite alphabet, including a blank symbol B, and it computes with words over that alphabet. This action of a specific Turing machine is according to its program applied to a given input. The program consists of finitely many instructions given as tuples of the form

$$(p, a, q, b, X) \text{ or simply } p, a, q, b, X$$

where p, q are states, a, b are symbols from the alphabet, and X stands for L (go left) or R (go right). This instruction says that while in state p and reading symbol a, change to state q (if different from p), replace a by b (if different from b), and move one cell left if $X = L$ or right if $X = R$. The computation begins by placing the input word at the beginning of the tape, one symbol in each cell, while the rest of the tape is blank, and putting the machine head in the starting state over the left-most symbol. The computation then proceeds in steps according to the program. It stops when the program reaches a halting state; otherwise, it computes forever. If a computation reaches a halting state, then it outputs the word that is written on the tape at that final step.

While computer scientists work with various alphabets for the Turing machines, in mathematics we can code symbols and words by natural numbers, $\mathbb{N} = \{0, 1, 2, \ldots\}$, so we may assume that every Turing machine alphabet consists of a single symbol 1, in addition to the blank symbol. Every natural number x is represented in unary, by a string of $x + 1$ consecutive 1's. Hence 0 is represented as 1, while 1 is represented as 11, 2 as 111, etc.

For example, the program with the following two instructions, will compute a function $s(x) = x + 1$:

$$q_{start}, 1, q_{start,}1, R$$
$$q_{start}, B, q_{end}, 1, R$$

This program will scan a given input until it reaches a blank symbol; it will then convert the first blank symbol to 1 and halt.

A unary function g on natural numbers, $g : \mathbb{N} \to \mathbb{N}$, is called *computable* if there is a Turing machine that on every input x halts and outputs $g(x)$. Hence, the successor function above, $s(x)$, is computable. A Turing machine computes a binary function on natural numbers, say $h(x, y) = x + y$, if starting with input x, y, a pair written as x and y in unary separated by a blank symbol, ends the computation in the halting state outputting $x + y$ in unary. We similarly define computable n-ary functions $h : \mathbb{N}^n \to \mathbb{N}$, where the inputs are n-tuples of natural numbers.

We can also consider partial functions on natural numbers, such as $\psi(x) = x - 5$ which is defined only on numbers x such that $x \geq 5$. That is, the domain of ψ, $dom(\psi) = \{5, 6, 7, \ldots\}$. Such a function ψ is *partial computable* if there is a Turing machine M that on every input x in the domain halts and outputs the correct value $\psi(x)$, while it does not halt, computing forever, on the inputs x that are not in the domain. We can show that the function above is partial computable by devising a corresponding Turing machine. Such a Turing machine will compute forever on inputs 0, 1, 2, 3, and 4. We can also show that a binary partial function $\theta(x, y) = x - y$, which is defined only on pairs x, y for which $x \geq y$, is partial computable. Two partial functions are equal if they have the same domains, and for each input in the domain, they output the same value.

Every partial computable function has a corresponding Turing machine.

Conversely, given any Turing machine M and a number $n \geq 1$, there is a corresponding partial computable n-ary function $\varphi_M^{(n)}$ that is (partially) computed by M.

Partial functions include total functions, so partial computable functions include computable functions, which are sometimes called total computable for emphasis.

A set $A \subseteq \mathbb{N}$ of natural numbers is *computable* if there is a Turing machine that on an input x always halts, and outputs 1 if $x \in A$, and outputs 0 if $x \notin A$. This is equivalent to saying a set A is *computable* if its characteristic function c_A is computable, where

$$c_A(x) = \begin{cases} 1, & \text{if } x \in A; \\ 0, & \text{if } x \notin A. \end{cases}$$

It is easy to see that the complement of a computable set is computable. All finite sets and co-finite sets are computable. Many other sets are computable. A set of natural numbers is identified with a unary relation. We similarly define computable binary relations and computable relations of any arity.

3 Partial Recursive Functions and Church-Turing's Thesis

The notion of Turing machines was instrumental in the development of digital computers, since they provided a "machine" realization of the intuitive notion of computability, offering the most intuitive account of what functions are and are not algorithmically calculable. However, in some cases it takes considerable effort to write a Turing machine program, for example, a program that computes the function

$f(x, y) = x^y$. Exponentiation is a recursive operation, so the power of recursion could be exploited. Work by Herbrand and Gödel led to the notion to what we now call general recursive functions. Their original formulation was rather cumbersome though, so we will present here the more standard formulation of recursive functions owing to Kleene (see Kleene 1936; Cooper 2004; Soare 1987). We will start with some basic functions and then describe certain operations to obtain more functions from the existing ones. The defined class of functions will be the smallest class, with respect to inclusion, which can be obtained from the basic functions by a finite number of allowed operations. Thus, recursive functions are defined, fittingly, recursively. We start with three base cases.

The *zero function* $Z : \mathbb{N} \to \mathbb{N}$ is defined by $Z(x) = 0$.

The *successor function* $s : \mathbb{N} \to \mathbb{N}$ is defined by $s(x) = x + 1$.

For each $n \in \mathbb{N}$ and each $i \in \{1, \ldots, n\}$, the *projection function* $U_i^n : \mathbb{N}^n \to \mathbb{N}$ is defined by $U_i^n(x_1, \ldots, x_n) = x_i$.

Hence for every $n \geq 1$, there are n projection functions; and $U_1^1(x) = x$. The zero function, successor function, and projection functions are called *basic functions*.

We now have recursion steps that start with functions and define further functions. The first operation on functions is composition.

If $f : \mathbb{N}^k \to \mathbb{N}$ and $g_1, \ldots, g_k : \mathbb{N}^n \to \mathbb{N}$ are given functions, then the function c obtained by *composition* is defined by

$$c(x_1, \ldots, x_n) = f(g_1(x_1, \ldots, x_n), \ldots, g_k(x_1, \ldots, x_n)).$$

Hence, $c : \mathbb{N}^n \to \mathbb{N}$. Composition is also called substitution.

For example, for a function $Z_2 : \mathbb{N}^2 \to \mathbb{N}$ such that $Z_2(x, y) = 0$, we have

$$Z_2(x_1, x_2) = Z(x_1) = Z\big(U_1^2(x_1, x_2)\big),$$

so Z_2 is obtained by composition, according to the above definition, from functions $f = Z$ and $g_1 = U_1^2$ for $n = 2$ and $k = 1$.

Similarly, $s_3(x, y, z) = z + 1$ is obtained by composition as

$$s_3(x, y, z) = s\big(U_3^3(x, y, z)\big).$$

The next operation is called primitive recursion.

If $f : \mathbb{N}^n \to \mathbb{N}$ and $g : \mathbb{N}^{n+2} \to \mathbb{N}$ are given functions, then the function p obtained by *primitive recursion* is defined by

$$p(x_1, \ldots, x_n, 0) = f(x_1, \ldots, x_n) \text{ and}$$
$$p(x_1, \ldots, x_n, y + 1) = g(x_1, \ldots, x_n, y, p(x_1, \ldots, x_n, y)).$$

Hence, $p : \mathbb{N}^{n+1} \to \mathbb{N}$.

Note that the second part in the definition is not self-referential since we have $y + 1$ on the left-hand side and y on the right-hand side. Recursion here is on y, while

$x_1, \ldots, x_n$ can be viewed as parameters. For brevity, let us write $\vec{x} = x_1, \ldots, x_n$. Then, we have

$$p\left(\vec{x}, 0\right) = f\left(\vec{x}\right),$$
$$p\left(\vec{x}, 1\right) = g\left(\vec{x}, 0, p\left(\vec{x}, 0\right)\right), \text{etc.}$$

For example, for the addition function $p(x, y) = x + y$, we have

$$p(x, 0) = x = U_1^1(x),$$
$$p(x, y + 1) = p(x, y) + 1 = s(p(x, y)) = s_3(x, y, p(x, y))$$
$$\text{where } s_3(x, y, z) = z + 1.$$

The set of *primitive recursive functions* is the smallest set of functions containing the basic functions that is closed under composition and primitive recursion.

Hence, addition is primitive recursive. Similarly, we can show that multiplication $x \cdot y$ and exponentiation x^y are primitive recursive functions. A set or a relation is primitive recursive if its characteristic function is primitive recursive. For example, the set of even numbers, the set of prime numbers, and the usual order relation on natural numbers are primitive recursive. We can further show that some other kinds of recursion also lead to primitive recursive functions. For example, the function $f : \mathbb{N} \to \mathbb{N}$ defined by

$$f(0) = 1,\ f(1) = 2 \text{ and } f(x + 2) = f(x) + f(x + 1),$$

and known as the *Fibonacci sequence* 1, 2, 3, 5, 8, ... is primitive recursive.

While many important functions in elementary number theory are primitive recursive, and every primitive recursive function is intuitively computable, not every intuitively computable function is primitive recursive. One such function is the *Ackermann function* $\psi : \mathbb{N}^2 \to \mathbb{N}$ defined by double recursion as follows:

(i) $\psi(0, y) = y + 1$,
(ii) $\psi(x + 1, 0) = \psi(x, 1)$,
(iii) $\psi(x + 1, y + 1) = \psi(x, \psi\ (x + 1, y))$.

It can be shown that ψ is a well-defined total function that is intuitively computable. For example, let us compute $\psi(1, 1)$.

$$\psi(1, 1) = \psi(0, \psi(1, 0)) \text{ by (iii)}$$
$$= \psi(1, 0) + 1 \text{ by (i)}$$
$$= \psi(0, 1) + 1 \text{ by (ii)}$$
$$= 1 + 1 + 1 \text{ by (i)}$$
$$= 3$$

We can show that for every primitive recursive function $f : \mathbb{N}^n \to \mathbb{N}$, there is a constant a_f such that

$$\forall x_1 \ldots \forall x_n \big(f(x_1, \ldots, x_n) < \psi\big(a_f, x_1 + \cdots + x_n\big)\big).$$

Hence, if $\psi(x, y)$ is primitive recursive, then the unary function obtained by setting $x = y$, $g(x) = \psi(x, x)$, is also primitive recursive so for some fixed a we have

$$\forall x(\psi(x, x) < \psi(a, x)), \text{hence}$$

$\psi(a, a) < \psi(a, a)$, which is a contradiction.

The third operation that we will introduce on total or partial functions is called minimalization or unbounded search or μ-operator. Here, μy stands for the "least y such that." For a unary partial function g by $g(y)\downarrow$, we denote that $y \in dom\,(g)$ so g is defined on y. If we want to emphasize the value $v = g(y)$, we write $g(y)\downarrow = v$. By $g(y)\uparrow$, we denote that $y \notin dom\,(g)$ so $g(y)$ is undefined. In the following definition, we may have parameters $\vec{x} = x_1, \ldots, x_n$.

If $g : \mathbb{N}^{n+1} \to \mathbb{N}$ is a partial function, then an n-ary function h obtained from g by *minimalization* is defined as $h\left(\vec{x}\right)$ is the least y, if it exists, such that $g\left(\vec{x}, y\right)\downarrow = 0$ and $g\left(\vec{x}, z\right)\downarrow$ for all $z < y$. If such y does not exist, then $h\left(\vec{x}\right)\uparrow$.

We write $h\left(\vec{x}\right) = \mu y\left(g\left(\vec{x}, y\right) = 0\right)$. Even if g is a total function, h may not be total.

The set of *partial recursive functions* is the smallest set of functions containing the basic functions that is closed under composition, primitive recursion, and minimalization.

A *recursive function* is a partial recursive function which is total.

When composition or primitive recursion is applied to partial functions on an input and undefinability is encountered at any point, then the resulting function will not be defined either on that input. Recursive functions are also called general recursive.

For example, nowhere defined function ε, i.e., one with the empty domain, is partial recursive since

$$\varepsilon(x) = \mu y(y \neq y).$$

For a computable permutation of natural numbers f, the inverse function f^{-1} is also a computable permutation since

$$f^{-1}(x) = \mu y(f(y) = x).$$

On the other hand, it can be shown that the inverse function of a primitive recursive permutation is not always primitive recursive.

We can prove the following important result about the equivalence of two formal definitions of computability. Both total and partial recursive functions have Turing machine counterparts.

A total function is recursive if and only if it is computable by a Turing machine.

A partial function is partial recursive if and only if it is partial computable by a Turing machine.

Like Turing machines, recursive functions also capture our intuition of computability. After all, the three base case functions can be trivially calculated by hand, and the three recursion step operations can all be computed through straightforward algorithms, provided that unbounded search ends. It is a remarkable fact that a function is computable via a Turing machine if and only if it is recursive in the above sense. This fact lends credence to the Church-Turing thesis that intuitively computable functions (also called effectively computable functions) coincide with formally computable functions.

The minimalization operator only yields a recursive function when the function is total, and how do we know when such a function is total? The aforementioned equivalence makes this the same as asking, how do we know when a given Turing machine halts on all inputs? This is an algorithmically unsolvable (hence undecidable) problem. For partial recursive functions, we allow the minimalization operator to yield partial functions and not just total ones. It is just that we cannot algorithmically determine which partial recursive functions are general recursive.

Church-Turing's Thesis

(i) *A total function is intuitively computable iff it is formally computable, say by a Turing machine, or is a recursive function.*

(ii) *More generally, a partial function is intuitively partial computable iff it is partial recursive or partial computable by a Turing machine.*

Another reason to accept Church-Turing's thesis is that all other formalisms of algorithms lead to the same classes of computable functions and partial computable functions. Furthermore, we do not know of any example of an intuitively computable function that cannot be shown to be formally computable.

We often use Church-Turing's thesis when we want to show that we have an algorithm or partial algorithm by describing it informally, without writing the exact Turing machine program or a partial recursive function derivation. For example, a natural number is prime if it is divisible only by 1 and by itself; i.e., it has only trivial factors. The set of all prime numbers is computable since there is an algorithm that for a given natural number determines whether it has a nontrivial factor by checking all its possible divisors. For more on partial computable functions and Church-Turing's thesis, see (Cooper 2004; Davis 1958; Rogers 1967; Soare 1987).

4 Enumerations and Computably Enumerable Sets

We know that every finite set is computable. It is not hard to see that if an infinite set A is computable, then there is a computable function f that enumerates its elements in the increasing order:

$$A = \{a_0 < a_1 < a_2 < \cdots\},$$

where $a_n = f(n)$ for every n.

This leads to the following more general definition. Recall that for a partial function ψ, by $ran(\psi)$ we denote its range, the set of all outputs or values it produces, $\{\psi(n) : n \in dom(\psi)\}$.

A set A of natural numbers is *computably enumerable* if A is empty or there is a computable unary function f such that

$$A = ran(f) = \{f(0), f(1), \ldots\}.$$

For an infinite set, an enumeration $f(0), f(1), \ldots$ can be modified by crossing repetitions to be one-to-one. However, it does not have to be in an increasing order, so we cannot know at any stage in advance whether a number that has not been enumerated in A will later be enumerated or not. Assume that there is a computable enumeration in the increasing order:

$$f(0) < f(1) < f(2) < \cdots.$$

From such an enumeration, we can design a decision algorithm since given x we can find n such that $x < f(n)$ and then decide whether $x \in \{f(0), \ldots, f(n-1)\}$. If "yes," output 1, if "no," output 0. Hence, computable infinite sets are exactly those computably enumerable sets for which there is a computable enumeration in the increasing order.

The *partial characteristic function* of a set A is defined by

$$pc_A(x) = \begin{cases} 1, & \text{if } x \in A; \\ \text{undefined}, & \text{if } x \notin A. \end{cases}$$

We can show that A is computably enumerable if and only if the partial characteristic function of A is partial computable. It is easy to see that from a computable characteristic function we can construct a partial computable partial characteristic function. The converse is not always possible.

It can also be shown that a set A is computably enumerable if and only if it is the domain of a partial computable function ψ: $A = dom(\psi)$. Moreover, a set A is computably enumerable if and only there is a computable binary relation R such that for every x,

$$x \in A \Leftrightarrow (\exists y) R(x, y).$$

All pairs of natural numbers can be computably listed without repetitions, for example, by placing them first into infinitely many infinite lists as

$$\begin{aligned} &(0,0), (0,1), (0,2), (0,3), \ldots \\ &(1,0), (1,1), (1,2), \ldots \\ &(2,0), (2,1), \ldots \\ &\ldots \end{aligned}$$

and then systematically listing them all in a single infinite list, say by following finite diagonals, to obtain

$$(0,0);(1,0),(0,1);(2,0),(1,1),(0,2);\ldots$$

This method gives us computable bijection $f\colon \mathbb{N}^2 \to \mathbb{N}$:

$$f(x,y) = \frac{1}{2}(x+y+1)(x+y)+y.$$

That is, because (x, y) is the y^{th} entry in its diagonal and the previous diagonals contain $1 + 2 + \cdots + (x + y)$ entries. Sometimes, instead of $f(x, y)$, we write $\langle x, y\rangle$ emphasizing that this is a computable pairing function that codes a pair of natural numbers by a single one.

Similarly, we can computably enumerate all k-tuples of natural numbers for $k > 2$, as well as the set of all finite tuples of natural numbers. These computable enumerations or numberings can be viewed as algorithmically coding various mathematical objects by natural numbers. This idea goes back to Gödel, so we often call the object's position in the list its Gödel number. In particular, all computably enumerable sets can be simultaneously computably listed by first listing in a computable manner all Turing machines. Each Turing machine is a finite list of instructions that consist of finite sequences of symbols. Let

$$M_0, M_1, M_2, \ldots$$

be a fixed computable enumeration of all Turing machines. We can assure that there are no repetitions in this enumeration, i.e., that the enumeration is one-to-one. For a Turing machine M_e, we will denote the unary partial computable function it computes by φ_e and, for $n > 1$, an n-ary partial computable function it computes by $\varphi_e^{(n)}$. Hence,

$$\varphi_0, \varphi_1, \varphi_2, \ldots$$

is a computable enumeration of all unary partial computable functions. For each partial computable function ψ, there are infinitely many Gödel codes e such that $\psi = \varphi_e$. This statement is known as the *Padding Lemma*. This is true since to the Turing machine M_e that computes φ_e, we can add extraneous instructions to obtain new programs computing the same function.

Kleene has established a normal form for partial computable functions, which shows that every partial recursive function can be obtained from primitive recursive functions by using at most one application of the μ-operator.

Kleene Normal Form Theorem

There is a ternary primitive recursive predicate T and a primitive recursive unary function U such that for every Gödel code e and input x, we have

$$\varphi_e(x) = U(\mu z T(e, x, z)).$$

This normal form generalizes to n-ary partial computable functions with the corresponding Kleene's predicate $T_n(e, x_1, \ldots, x_n, z)$.

For φ_e, we denote its domain $dom(\varphi_e)$ by W_e.

A set A is computably enumerable if and only if $A = W_e$ for some e.

Thus,

$$W_0, W_1, W_2, \ldots$$

is a computable enumeration of all computably enumerable sets. Hence, there are countably many computably enumerable sets. They are partially ordered by set-theoretic inclusion $\subseteq$ and are closed under intersection $\cap$ and union $\cup$, so they form a lattice. A *lattice* is a partially ordered set in which any two elements have a least upper bound (also called supremum) and a greatest lower bound (also called infimum). Here, given computably enumerable sets U and V, the union $U \cup V$ is their supremum, and the intersection $U \cap V$ is their infimum. The least element in this lattice is the empty set ø, and the greatest element is $\mathbb{N}$. For more on the lattice of computably enumerable sets, see (Odifreddi 1989; Soare 1987).

5 The Halting Set, Diophantine Sets and Hilbert's Tenth Problem

There are computably enumerable sets with complements that are not computably enumerable. These are exactly the noncomputable computably enumerable sets. For example, the halting set K is computably enumerable, but not computable. The set K consists of all inputs e on which the Turing machine with index e halts. That is,

$$K = \{e : e \in W_e\} = \{e : \varphi_e(e)\downarrow\}.$$

The set K is computably enumerable because it can be algorithmically listed by the procedure that simultaneously runs

$$\varphi_0(0), \varphi_1(1), \ldots, \varphi_e(e), \ldots$$

and enumerates those e for which $\varphi_e(e)$ converges, as soon as the convergence occurs. Here, simultaneously means that we start with the first Turing machine and run it for one step, and then at each step we add a new Turing machine program and also run all activated machines for an additional computational step.

We will now prove that K is not a computable set.

The proof is by contradiction, so we will assume that K is computable. The complement of K is the *divergence set*:

$$\overline{K} = \{e : \varphi_e(e) \uparrow\}.$$

Since we assumed that K is computable, $\overline{K}$ would also be computable and hence computably enumerable. If $\overline{K}$ were computably enumerable, then for some i we have $\overline{K} = W_i$. Then,

$$i \in K \Leftrightarrow i \in W_i \text{ (by definition of } K)$$

$$\Leftrightarrow i \in \overline{K} \left(\text{since } \overline{K} = W_i\right),$$

which is a contradiction.

Another proof that K is not computable is a diagonal argument that uses Church-Turing's thesis. Assuming that K is computable, define the following total function

$$g(x) = \begin{cases} \varphi_x(x) + 1 & \text{if } x \in K \\ 0 & \text{if } x \notin K \end{cases}$$

By Church-Turing's thesis, g is a computable function; hence $g = \varphi_i$ for some i. Since g is total, $i \in K$. Then,

$$g(i) = \varphi_i(i) + 1 = \varphi_i(i),$$

which is a contradiction.

Noncomputability of the halting set K implies the undecidability of the *halting problem* given as a set of pairs

$$H = \{(x, e) : \varphi_e(x) \downarrow\}.$$

If H were computable, then K would be computable since

$$e \in K \Leftrightarrow (e, e) \in H.$$

However, H is still a computably enumerable binary relation.

There are many characterizations of a computably enumerable set. Another equivalent notion of a computably enumerable set arose from number theory and involved Diophantine equations. A Diophantine equation is an equation of the form

$$p(x_0, \ldots, x_{m-1}) = 0,$$

where p is a polynomial in the variables $x_0, \ldots, x_{m-1}$ with integer coefficients. Diophantine equations are named after Diophantus of Alexandria, who wrote *Arithmetica* in 13 books in the third century.

When Hilbert proposed his list of unsolved problems in mathematics, the Tenth Problem was the following:

> "Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined in a finite number of operations whether the equation is solvable in rational integers."

Solving Hilbert's Tenth Problem was by no means easy even after the advent of computability theory. It required decades of groundbreaking work, starting with the development of a robust theory of Diophantine sets. In 1970, Matiyasevich added the final link to the proof chain forged over many years by Davis, Putnam, and J. Robinson. Matiyasevich used interesting facts about Fibonacci numbers to get the negative result. He established that there is no algorithm that determines whether a Diophantine equation has a solution in the integers.

Lagrange's four-square theorem from 1772 establishes that every natural number can be expressed as the sum of squares of four integers. Hence, the algorithmic solvability of a Diophantine equation in the integers is equivalent to the algorithmic solvability of a Diophantine equation in the natural numbers. Moreover, by Skolem's result of 1934, it is sufficient to consider only Diophantine equations of total degree four.

Thus, without loss of generality, we will only consider natural number solutions to Diophantine equations. We say that an n-ary relation R on natural numbers is *Diophantine* if there exists a polynomial $p(x_1, \ldots, x_m, y_1, \ldots, y_n)$ such that for every n-tuple of natural numbers $a_1, \ldots, a_n$, we have

$$R(a_1, \ldots, a_n) \Leftrightarrow \exists x_1 \exists x_2 \cdots \exists x_m p(x_1, x_2, \ldots, x_m, a_1, \ldots, a_n) = 0.$$

Note that $a_1, \ldots, a_n$ can be viewed as parameters for which the corresponding equations in variables $x_1, \ldots, x_m$ have or do not have solutions as determined by relation R.

We say that a k-ary function f is a *Diophantine function* if its graph G_f is a $(k+1)$-ary Diophantine relation, where

$$G_f = \{(a_1, \ldots, a_k, b) : f(a_1, \ldots, a_k) = b\}.$$

Davis, Putnam, and J. Robinson showed that the set of Diophantine sets is closed under unions and intersections, but not under complements. They further showed that under the assumption that exponentiation is a Diophantine relation, relations defined through existential quantifiers and bounded universal quantifiers are Diophantine relations, from which it follows that Hilbert's Tenth Problem is algorithmically unsolvable, as we shall see shortly. Matiyasevich completed their work by showing that the exponentiation is, in fact, Diophantine (see Matiyasevich 1993).

Here is the main thrust of the proof that Diophantine functions are exactly recursive functions. It is clear that any Diophantine function is a computable function; it is simply a matter of running a search for the quantified variables along with the output variable. We show the converse by induction on the complexity of recursive functions. To start with the base cases, $Z(a) = b$ if and only if $b = 0$, so Z is Diophantine. Clearly, $s(a) = b$ if and only if $b - a - 1 = 0$, so s is Diophantine; and $U_i^n(a_1, \ldots, a_n) = b$ if and only if $a_i - b = 0$, so U_i^n is Diophantine. Now there are

three induction steps. For primitive recursion, we use the fact that exponentiation is Diophantine, which allows us to show that Gödel's coding β function is Diophantine. For minimalization, we also use the fact that exponentiation is Diophantine, which allows us to use bounded universal quantifiers in addition to existential quantifiers to define Diophantine relations. Therefore, we can show by induction on the complexity of recursive functions that every recursive function is a Diophantine function.

It is not hard to see that every Diophantine set is computably enumerable. Say we have that

$$a \in A \text{ if and only if } \exists x_1 \exists x_2 \cdots \exists x_m p(x_1, \ldots, x_m, a) = 0.$$

We can produce an effective listing of all $(m + 1)$-tuples of integers, evaluate the polynomial at each $(m + 1)$-tuple, and, every time the polynomial is equal to 0, output the value a.

The following key result is due to Davis, Putnam, J. Robinson, Matiyasevich:

Every computably enumerable set of integers is Diophantine.

Since not all computably enumerable sets are computable, it follows that there exist Diophantine sets that are not computable. The last statement immediately implies a negative answer to Hilbert's question. Indeed, the halting set K is a noncomputable Diophantine set, and let $p(x_1, \ldots, x_m, y)$ be its Diophantine definition. If we had an algorithm to decide when polynomial equations have solutions in $\mathbb{N}$, thinking of the value of a as a coefficient of the polynomial equation $p(x_1, \ldots, x_m, a) = 0$, we would have an algorithm to determine for which values a of y there exist $x_1, \ldots, x_m \in \mathbb{N}$ such that the polynomial equation is satisfied. Thus, we would have an algorithm to decide membership in K, which would make K computable.

It should be noted that Hilbert's Tenth Problem is still open for solutions in the field of rational numbers. We do not know much about Hilbert's Tenth Problem over the rational numbers. It is easy to show that if it were undecidable, it would then follow that it is undecidable over integers. At the same time, attempts to approach Hilbert's Tenth Problem over rationals from different perspectives have produced many interesting developments and results (see Shlapentokh 2007).

6 Undecidable Problems and Arithmetical Hierarchy

Kleene established the following result.

Fixed Point Theorem (also called *Recursion Theorem*)

For a standard enumeration of all partial computable functions

$$\varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_e, \ldots$$

if a (total) computable function f (say a virus) changes all indices by producing

$$\varphi_{f(0)}, \varphi_{f(1)}, \varphi_{f(2)}, \ldots, \varphi_{f(e)}, \ldots$$

then it will fail in the following sense. For some e we will have

$$\varphi_{f(e)} = \varphi_e.$$

The last equality is the equality of partial functions, and e is called a fixed point. In fact, f will have infinitely many such fixed points.

The following result due to Rice provides many examples of undecidable problems given as index sets. A set of unary partial recursive functions is nontrivial if it is nonempty and does not contain all functions.

Rice's Theorem

Let $\mathcal{X}$ be a nontrivial set of unary partial recursive functions. Consider its index set defined as

$$X = \{i : \varphi_i \in \mathcal{X}\}.$$

Then, X is noncomputable.

To prove this theorem by contradiction, assume otherwise, that X is computable. Since $\mathcal{X}$ is nontrivial, fix some a and b such that $\varphi_a \in \mathcal{X}$ and $\varphi_b \notin \mathcal{X}$. Hence, $a \in X$ and $b \notin X$. Define the following total function:

$$f(x) = \begin{cases} a & \text{if } x \notin X \\ b & \text{if } x \in X. \end{cases}$$

It follows by definition of X that

$$\forall x (f(x) \in X \Leftrightarrow x \notin X).$$

The function f is computable by Church-Turing's thesis since we assumed that X is computable. We apply Kleene's fixed point theorem to f to obtain a fixed point e. Then,

$$\varphi_{f(e)} = \varphi_e,$$

but we have

$$f(e) \in X \Leftrightarrow e \notin X,$$

contradicting the fact that X is an index set.

For example, let $\mathcal{X}$ be the set of all unary partial computable functions with non-empty domain. It is a nontrivial set of functions, so its index set $\{i : W_i \neq \emptyset\}$ is not computable. If other words, the problem of determining whether W_i is a non-empty or empty set is undecidable.

We can show that the halting set K is not an index set as defined above. On the other hand, we can use Rice's theorem to show that the following decision problems are undecidable:

whether W_e is finite or infinite,
whether W_e is computable or not,
whether φ_e is total (equivalently, recursive or $W_e = \mathbb{N}$) or not,
whether φ_e is a constant function,
whether φ_e is a zero function Z,
whether W_e is cofinite (i.e., its complement is finite).

Also, the problem of deciding whether two partial computable functions are equal, given by $\{(x, y)$: $\varphi_x = \varphi_y\}$, is undecidable since otherwise its characteristic function $c(x, y)$ is computable, so if we fix an index a for a zero function, $Z = \varphi_a$, we get that the unary function $c(x, a)$ is computable, so the problem of determining whether φ_x is a zero function Z would be decidable, which is a contradiction.

We can use a diagonal argument to show that there is partial recursive function that cannot be extended to a (total) recursive function. Let a partial function g be defined as follows:

$$g(x) = \begin{cases} \varphi_x(x) + 1 & \text{if } \varphi_x(x) \downarrow \\ \text{undefined} & \text{if } \varphi_x(x) \uparrow . \end{cases}$$

We can use Church-Turing's thesis to show that g is partial recursive: on input x, computably find Turing machine M_x and run M_x on x. If it halts and outputs y, halt and output $y + 1$; otherwise, compute forever. We will show by contradiction that g cannot be extended to a (total) recursive function. Assume otherwise, that g can be extended to some recursive function f. Fix e such that $f = \varphi_e$.

Then, $\varphi_e(x)$ is defined for every x.

$$\begin{aligned} g(e) &= \varphi_e(e) + 1 \text{ (since } \varphi_e(e) \downarrow) \\ &= f(e) \text{ (since } f \text{ extends } g) \\ &= \varphi_e(e) \text{ (since } f = \varphi_e), \end{aligned}$$

which is a contradiction.

Thus, we can use Rice's theorem to show that the problem of determining whether φ_e is extendible or not to a (total) recursive function is undecidable.

We would like to further classify the undecidable problems. The Kleene-Mostowski *arithmetical hierarchy* provides framework for classification of some of these problems. It relates quantifier complexity of relations with their computability in an important way. We also have the following classification of sets of

natural numbers. A set A is $\Sigma_0^0 = \Pi_0^0$ if it is computable. Let n be a positive natural number. A set A is Σ_n^0 if there is a computable $(n+1)$-ary relation R such that for every $a \in \mathbb{N}$,

$$a \in A \Leftrightarrow (\exists m_1)(\forall m_2)(\exists m_3)\cdots(Qm_n)R(a, m_1, m_2, m_3, \ldots, m_n),$$

where Q is $\exists$ if n is an odd number, and Q is $\forall$ if n is an even number. The existential quantifier here can be viewed as a projection. The negation corresponds to complementation, and the universal quantifier can be expressed using the existential quantifier and negation.

A set B is Π_n^0 if and only if its complement $\overline{B}$ is Σ_n^0.

For a partial computable function φ_e with Gödel code (index) e, we write $\varphi_{e,s}(x) \downarrow$ if e, $x < s$, and there is an output $< s$ of $\varphi_e(x)$ after $< s$ in the corresponding computation, i.e., this computation halts by stage s. If the output is y, we also write $\varphi_{e,s}(x) \downarrow y$. The relations

$$R(e, x, s) \Leftrightarrow \varphi_{e,s}(x) \downarrow \text{ and}$$
$$P(e, x, s, y) \Leftrightarrow \varphi_{e,s}(x) \downarrow y$$

are computable relations since there is a bound on the number of computation steps.

The predicate defined as

$$T(e, x, z) \Leftrightarrow (P(e, x, s, y) \wedge z = \langle s, y \rangle)$$

is Kleene's T predicate from Kleene's normal form.

The relation $\exists s(\varphi_{e,s}(x)\downarrow)$ is a Σ_1^0 binary relation which says that the computation for φ_e on input x halts at some stage.

Thus, the relation $\forall s(\varphi_{e,s}(x)\uparrow)$ is a Π_1^0 binary relation which says that the computation for φ_e on input x never halts.

The above hierarchy of sets of natural numbers is a proper hierarchy since Σ_n^0 sets are strictly contained in Σ_{n+1}^0 sets, and Π_n^0 sets are strictly contained in Π_{n+1}^0 sets. Clearly, every Σ_n^0 set is also Π_{n+1}^0, and every Π_n^0 set is also Σ_{n+1}^0. It then follows that the union of all Σ_n^0 sets is the same as the union of all Π_n^0 sets. A set is called *arithmetical* if it belongs to this union.

A set A is arithmetical iff A is Σ_k^0 or Π_k^0 for some $k \geq 0$.

Equivalently, a set $\mathcal{A}$ is arithmetical if it can be obtained from a computable relation by finitely many projections and complementations.

We say that a set is Δ_n^0 if it is both Σ_n^0 and Π_n^0. We have that Δ_n^0 sets are strictly contained in Π_n^0 (or Σ_n^0) sets. By definition, Δ_0^0 sets are the computable sets. The following theorem is due to Post.

If a set A is Δ_1^0, then A is computable.

If a set A is Δ_1^0, then A is both Σ_1^0 and Π_1^0. Then both A and $\overline{A}$ are computably enumerable. By computably enumerating them simultaneously, we can devise a decision algorithm for A since every natural number will appear either in A or $\overline{A}$.

Let $\mathcal{N} = (\mathbb{N}, +, \cdot, S, 0)$ be the *standard model of arithmetic* in the language of arithmetic, $L = \{+, \cdot, S, 0\}$ denoting plus, times, successor, and zero. The name comes from the fact that a relation is arithmetical if and only if it is definable in $\mathcal{N}$ by a first-order formula. Gödel established that all computable relations are definable in $\mathcal{N}$, and hence all arithmetical relations are definable in $\mathcal{N}$. Hence for every arithmetical set A, there is a first-order formula $\alpha(x)$ such that

$$A = \{n : \mathcal{N} \vDash \alpha(n)\}.$$

This can be generalized to all arithmetical relations. Recall the classification of first-order formulas. The upper subscript 0 signifies that the quantification is only over variables that stand for individual elements. A formula is a $\Sigma_0^0 = \Pi_0^0$ formula if it contains no quantifiers, so it is quantifier-free. For $n > 0$, a formula is a Σ_n^0 (Π_n^0, respectively) formula if it is equivalent to a formula in the prenex normal form which begins with a block of existential (universal, respectively) quantifiers and has n alternations of blocks of like quantifiers.

A relation definable in $\mathcal{N}$ by a quantifier-free formula is computable. It follows that a relation definable in $\mathcal{N}$ by a Σ_n^0 formula (Π_n^0 formula, respectively) is a Σ_n^0 relation (Π_n^0 relation, respectively).

For any computable relation, there are two natural defining formulas: one with a block of existential quantifiers followed by a formula with only bounded quantifiers, $\forall x < y$ and $\exists x < y$, and the other one with a block of universal quantifiers followed by a formula with only bounded quantifiers. A block of existential (universal) quantifiers can be replaced by a single existential (universal) quantifier by coding tuples of natural numbers by a single natural number. Matiyasevich proved that bounded quantifiers can be eliminated from the above formulas, so a computable set is definable in $\mathcal{N}$ both by a Σ_1^0 and a Π_1^0 formula (see A.7 of Appendix in Ash and Knight 2000). Moreover, a set that is Σ_n^0 (Π_n^0, respectively) is definable in $\mathcal{N}$ by a Σ_n^0 (Π_n^0, respectively) first-order formula.

Given a set complexity class $\mathfrak{C}$, we say that a set A is $\mathfrak{C}$-*complete* if A is in $\mathfrak{C}$, and A is $\mathfrak{C}$-hard, that is, every set X in $\mathfrak{C}$ can be reduced to A. The reduction means that there is a computable function $f : \mathbb{N} \to \mathbb{N}$ such that for every n we have

$$n \in X \Leftrightarrow f(n) \in A.$$

The reduction function f does not have to be one-one, it can be many-one, so the above reduction is called an m-reduction. We can show that the halting set K is complete for the set of all computably enumerable sets.

Furthermore, we can, for example, show that the set $FIN = \{e: W_e \text{ is finite}\}$ is Σ_2^0 as follows:

$$\begin{aligned} e \in FIN &\Leftrightarrow \exists s\big(dom(\varphi_{e,s}) = dom(\varphi_e)\big) \\ &\Leftrightarrow \exists s \forall t \forall x\big(t > s \Rightarrow \big(\varphi_{e,t}(x) \downarrow \Rightarrow \varphi_{e,s}(x) \downarrow\big)\big) \end{aligned}$$

We can further show that every Σ_2^0 set can be reduced to *FIN* via a computable function. Hence, *FIN* is Σ_2^0-complete. The complement set $INF = \{e\text{: } W_e \text{ is infinite}\}$ will be Π_2^0-complete.

For more on arithmetical hierarchy and complexity of problems, see (Fokina et al. 2014; Montalbán 2021; Shoenfield 1993; Soare 1987).

7 Oracles and Turing Degrees

In his 1938 PhD dissertation (Turing 1938), Turing briefly wrote about a Turing machine with an oracle. An oracle will supply additional information presented on a read-only oracle tape about the characteristic function of some set of natural numbers. Unlike a standard Turing machine that computes functions autonomously, an oracle Turing machine at some steps must query the outside world, i.e., oracle, before continuing. The oracle contains information about membership of elements in some set. If the oracle set is $Y \subseteq \mathbb{N}$, then the corresponding oracle Turing machines will compute *Y*-partial computable functions. Hence, we will have an effective enumeration of all unary *Y*-partial computable functions:

$$\varphi_0^Y, \varphi_1^Y, \varphi_2^Y, \ldots$$

If the oracle *Y* is computable, then it is not needed since each answer given by its characteristic function $c_Y(n)$ can be replaced by a program that computes $c_Y(n)$, so we can omit the superscript *Y*. Oracle Turing machines lead to relative computability and the notion of Turing reducibility. Post developed the notion of Turing reducibility in his paper (Post 1944). For sets *X* and *Y*, we say that *X* is *Turing reducible* to *Y*, or *X* is Turing computable in *Y*, or *X* is *Y*-computable if $c_X = \varphi_e^Y$ for some *e*. We denote this relative computability by $X \leq_T Y$. The relation $\leq_T$ is a partial order relation on the subsets of the set of natural numbers. Every set is computable in itself, so $X \leq_T X$. Furthermore, the reducibility relation is transitive since $X \leq_T Y$ and $Y \leq_T Z$ implies $X \leq_T Z$.

By $X \equiv_T Y$, we denote that *X* is Turing equivalent to *Y*, i.e., $X \leq_T Y$ and $Y \leq_T X$. We also say that *X* and *Y* have the same *Turing degree* and write $\deg(X) = \deg(Y)$. Post referred to Turing degrees as the degrees of unsolvability. Every set *X* and its complement $\overline{X}$ have the same Turing degree. To prove that $X \leq_T \overline{X}$, for a given *x*, we ask oracle $\overline{X}$ whether *x* belongs to $\overline{X}$ and change "yes" to "no" (i.e., $c_{\overline{X}}(x) = 1$ to $c_X(x) = 0$) and vice versa. Hence, we also have that $\overline{X} \leq_T X$.

There are countably many sets in each Turing degree since there are countably many *Y*-computable functions for every *Y*. By $X <_T Y$, we denote that $X \leq_T Y$ but $Y \not\leq_T X$. For the order on degrees, we use the usual notation $\leq$, $<$, and $=$. We write

$$\deg(X) \leq \deg(Y) \text{ if } X \leq_T Y.$$

All computable sets are Turing equivalent, so they have the computable Turing degree, which we denote by boldface $\mathbf{0}$.

The computable degree **0** *is the least Turing degree.*

Every Turing degree is an equivalence class containing all sets that are computable in each other. Each equivalence class is countable, but there are continuum many sets of natural numbers. Hence:

There are continuum many Turing degrees. They are partially ordered.

The partially ordered set of Turing degrees forms an upper semi-lattice since there is always supremum. The supremum (least upper bound) of Turing degrees deg(X) and deg(Y) is the Turing degree of the set

$$X \oplus Y = \{2n : n \in X\} \cup \{2n+1 : n \in Y\}.$$

Hence, every finite set of Turing degrees has supremum. It can be shown that the infimum (greatest lower bound) of two Turing degrees may not exist.

For a set Y, the Turing *jump of* Y is defined as follows:

$$Y' = \left\{e : \varphi_e^Y(e) \text{ halts}\right\} = \left\{e : \varphi_e^Y(e)\downarrow\right\}.$$

The set $K = \{e\text{: } \varphi_e(e)\downarrow\}$ is also denoted by $\emptyset'$. We showed that K is non-computable; hence $\emptyset <_T K = \emptyset'$. Similarly, we can show that

$$Y <_T Y'$$

Hence:

There is no greatest Turing degree.

The jump operator can be iterated:

$$Y^{(k+1)} = \left(Y^{(k)}\right)' = \left\{e : \varphi_e^{Y^{(k)}}(e)\downarrow\right\}.$$

We say that for $n \geq 1$, $Y^{(n)}$ is the n-th Turing jump of Y. Let $\mathbf{y}^{(n)} = \deg(Y^{(n)})$. Hence, for every Turing degree $\boldsymbol{y}$ we have an ascending sequence of Turing degrees:

$$\mathbf{y} < \mathbf{y}' < \mathbf{y}'' < \mathbf{y}''' < \cdots.$$

We define the *ω-jump* of Y by

$$Y^{(\omega)} = \left\{\langle x, n\rangle : x \in Y^{(n)}\right\},$$

and let $\mathbf{y}^{(\omega)} = \deg(Y^{(\omega)})$. The degree $\mathbf{y}^{(\omega)}$ is a natural upper bound for the sequence above, but we can prove that no infinite ascending sequence of Turing degrees has a supremum.

The following facts show the connection between arithmetical hierarchy and Turing degrees. Let $n \geq 2$.

A set A is Σ_n^0 *iff A is computably enumerable relative to* $\emptyset^{(n-1)}$.
A set A is Δ_n^0 *iff A is computable relative to* $\emptyset^{(n-1)}$*; that is,* $A \leq_T \emptyset^{(n-1)}$.

Kleene and Post (1954) proved that there are incomparable Turing degrees, $\boldsymbol{a}$ and $\boldsymbol{b}$, by constructing sets A and B such that neither $A \leq_T B$ nor $B \leq_T A$. Moreover, they showed that the sets A and B are Turing reducible to K.

In fact, we can show that there are continuum many pairwise Turing incomparable degrees. Furthermore, for any given non-computable Turing degree $\boldsymbol{d}$, we can construct a Turing degree that is incomparable with it. A Turing degree $\boldsymbol{a}$ is called *minimal* if it is non-computable, i.e., $\boldsymbol{a} > \mathbf{0}$ and there are no Turing degrees strictly between $\mathbf{0}$ and $\boldsymbol{a}$. Spector showed that such degrees exist. In fact, there are continuum many minimal degrees, and Sacks established that there is such a degree strictly below the degree of the halting set. For more on Turing degrees, see (Cooper 2004; Lerman 1983; Odifreddi 1989; Soare 1987).

8 Computable Ordinals and Hyperarithmetical Hierarchy

One development in computability theory that has found applications in many different areas of mathematical logic is the notion of computable ordinals. To review terminology, we say that two ordered sets have the same order type if there exists an order isomorphism between them. Ordinal numbers are order types of well-ordered sets. A well-ordering of a set is a total (linear) ordering of its elements such that every nonempty subset has a least element. For finite well-ordered sets, the ordinals look like the natural numbers, but then they go far beyond them for infinite well-ordered sets:

$$0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \omega + 3, \ldots, \omega + \omega, \ldots, \omega + \omega + \omega, \ldots,$$
$$\omega^2, \ldots, \omega^3, \ldots, \omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots, \epsilon_0, \ldots$$

The ordinals form a well-ordered proper class, which enables us to do transfinite induction and transfinite recursion with them.

Now, a *computable ordinal* is the order type of a computable well-ordering, i.e., the order type of a well-ordered set (X, R) where X is a computable set and R is a computable relation. Since there are only countably many Turing machines, it follows that there are only countably many computable well-orderings and thus only countably many computable ordinals. Yet there are uncountably many countable ordinals, so the fact that the class of ordinals is well-ordered implies that there exists a least non-computable ordinal, and it is easy to see that this must be a countable ordinal, i.e., the order type of a countable well-ordering. We call this

ordinal ω_1^{CK}, in analogy to ω_1, the least uncountable ordinal. The CK stands for Alonzo Church and Stephen Kleene, who arrived at this ordinal through very different means.

Kleene was not studying computable well-orderings, but rather the seemingly unrelated problem of giving names to ordinals. Kleene proposed that we use natural numbers to denote ordinals. We define this notation system by transfinite recursion.

Let 1 be a notation for 0.

If a is a notation for α, then let 2^a be a notation for $\alpha + 1$.

If the e^{th} Turing machine halts on all inputs and outputs only notations for strictly increasing ordinals with limit β, then let $3 \cdot 5^e$ be a notation for β.

The set of all natural numbers which are the notations for some ordinals is called Kleene's $\mathcal{O}$, and we say that an ordinal is *constructive* if it has a notation in $\mathcal{O}$. Since $\mathcal{O}$ is countable, it follows that there are only countably many constructive ordinals and thus there exists a least non-constructive ordinal, and it is easy to see that this ordinal must be countable. In 1955, Spector surprisingly showed that this ordinal is actually equal to ω_1^{CK} and more than that, the set of computable ordinals is the same as the set of constructive ordinals. This result allows us to encode transfinite induction on countable ordinals in natural numbers via Kleene's ordinal notation. This will prove critical in defining the hyperarithmetical hierarchy.

The *hyperarithmetical hierarchy* will allow us to iterate Turing jumps through the computable ordinals. We use transfinite recursion on Kleene's ordinal notation.

Let $H(1)$ be the empty set.

Let $H(2^a)$ be the Turing jump of $H(a)$.

If $3 \cdot 5^e$ is a notation for β, let $H(3 \cdot 5^e)$ be the set of all ordered pairs (u, v) such that u is a notation for some $\alpha < \beta$ and $v \in H(u)$.

Now, it is not in general the case that $H(a) = H(b)$ if a and b are notations for the same ordinal α, but $H(a)$ and $H(b)$ have the same Turing degree, which is what ultimately matters.

In analogy to the arithmetical hierarchy, we can now define Σ^0_α, Π^0_α, and Δ^0_α relations for arbitrary computable ordinals α. Let a be a notation for an infinite α. Then we say that a relation is Σ^0_α if it is computably enumerable relative to $H(a)$, it is Π^0_α if its complement is computably enumerable relative to $H(a)$, and is Δ^0_α if it is computable relative to $H(a)$. (This is well-defined by the property above about different notations for the same ordinal yielding the same Turing degree.) And we say that a relation is *hyperarithmetical* if it is Δ^0_α for some computable ordinal α.

This hyperarithmetical hierarchy is only one way to generalize the arithmetical hierarchy, however. Another, seemingly unrelated, generalization is the analytical hierarchy, which is a way of classifying relations definable in the language of second-order logic. Arithmetical relations are defined in the language of first-order logic, where we only allow quantifiers ranging over individual elements. Second-order logic permits quantifiers ranging over arbitrary sets and relations, but for our purpose it suffices to restrict quantifiers to functions. We say that a relation is $\Sigma^1_0 = \Pi^1_0 = \Delta^1_0$ if it is arithmetical, noting that it is permitted to carry free function variables. If R is a Π^1_n relation, then

$$\exists f_1 \exists f_2 \cdots \exists f_k R$$

is Σ^1_{n+1}. If R is a Σ^1_n relation, then

$$\forall f_1 \forall f_2 \cdots \forall f_k R$$

is Π^1_{n+1}. We say that a relation is Δ^1_n if it is both Σ^1_n and Π^1_n. We say that a relation is *analytical* if it is Δ^1_n for some natural number n.

Kleene found a surprising connection between these two generalizations of the arithmetical hierarchy. He showed that every hyperarithmetical set is analytical, and more specifically that

A set is hyperarithmetical if and only if it is Δ^1_1.

An example of a hyperarithmetical set that is not arithmetical is the set of all Gödel codes of formulas of Peano arithmetic that are true in the standard model of arithmetic. That is because $\emptyset^{(\omega)}$ is a hyperarithmetical set that is not arithmetical. Kleene's O is not even hyperarithmetical; it is a Π^1_1-complete set. Here are a few amazing results that involve hyperarithmetical sets.

Kleene constructed a computable tree that is non-finitely branching and has infinite paths but has no infinite hyperarithmetical path. Hence, we have a computable structure with very noncomputable properties.

On the other hand, Spector obtained a surprising result that if a well-ordering has a hyperarithmetical copy, then it has a computable copy.

For more on hyperarithmetical hierarchy, see (Ash and Knight 2000; Sacks 1990).

9 Computable Models and Computable Infinitary Formulas

We consider countable structures for computable languages. A computable language is a countable language with algorithmically presented set of symbols and their arities. The universe A of an infinite countable structure $\mathcal{A}$ can be identified with the set of natural numbers. If L is the language of $\mathcal{A}$, then L_A is the language L expanded by adding a constant symbol for every $a \in A$, and $\mathcal{A}_A = (\mathcal{A}, a)_{a \in A}$ is the corresponding expansion of $\mathcal{A}$ to L_A. The *open diagram* of a structure $\mathcal{A}$, $D(\mathcal{A})$, is the set of all quantifier-free sentences of L_A true in $\mathcal{A}_A$. A structure is *computable* if its open diagram is computable. The *Turing degree of* $\mathcal{A}$ is the Turing degree of the atomic diagram of $\mathcal{A}$. Hence a structure is computable if its Turing degree is $\mathbf{0}$. If a structure has a finite language, then it is computable if its domain is computable and its relations and functions are computable.

For example, consider the following linear ordering $\prec$ on the set of natural numbers:

$$0 \prec 2 \prec 4 \prec \cdots \prec 5 \prec 3 \prec 1.$$

We say that this ordering has order type $\omega + \omega^*$ where order type ω^* stands for the usual ordering of negative integers, which is the reverse order type of the usual ordering for the natural numbers. The structure $\mathcal{A} = (\mathbb{N}, \prec)$ is computable because its domain is computable and its only relation $\prec$ is computable. Specifically, there is a procedure to decide for every pair of elements from the domain, natural numbers a, b, whether $a \prec b$ or $b \prec a$. First, we determine whether the numbers are even or odd (every even number is smaller than any odd number), and if the numbers are of the same parity, their natural ordering determines their $\prec$ ordering (the same for even numbers and reverse for odd numbers).

The *elementary diagram* of $\mathcal{A}$, denoted by $D^c(\mathcal{A})$, is the set of all first-order sentences of L_A that are true in $\mathcal{A}_A$. Henkin's construction of a model for a consistent complete decidable theory is effective and produces a structure $\mathcal{A}$ with a computable domain such that the elementary diagram of $\mathcal{A}$ is computable. A structure $\mathcal{A}$ is called *decidable* if its elementary diagram $D^c(\mathcal{A})$ is computable. It is easy to see that not every computable structure is decidable. For example, the *standard model of arithmetic*, $\mathcal{N} = (\mathbb{N}, +, \cdot, S, 0)$, is computable but not decidable. For more on computable and decidable models, see (Ershov 1980; Ershov and Goncharov 2000; Harizanov 1998; Millar 1999).

We can use computable ordinals to define computable infinitary formulas. The computable Σ_0 and Π_0 formulas are the finitary quantifier-free formulas. The computable $\Sigma_{\alpha+1}$ formulas are of the form

$$\bigvee_{n \in W_e} \exists \vec{y}_n \psi_n \left(\vec{x}, \vec{y}_n \right),$$

where for $n \in W_e$, ψ_n is a Π_α formula indexed by its Gödel number, and $\exists \vec{y}_n$ is a finite block of existential quantifiers. That is, $\Sigma_{\alpha+1}$ formulas are computably enumerable disjunctions of $\exists \Pi_a$ formulas. Similarly, $\Pi_{\alpha+1}$ formulas are computably enumerable conjunctions of $\forall \Sigma_\alpha$ formulas. It can be shown that a computable Σ_1 formula is of the form

$$\bigvee_{n \in \mathbb{N}} \exists \vec{y}_n \theta_n \left(\vec{x}, \vec{y}_n \right),$$

where $\left(\theta_n \left(\vec{x}, \vec{y}_n \right) \right)_{n \in \mathbb{N}}$ is a computable sequence of quantifier-free formulas. If α is a limit ordinal, then Σ_α (Π_α, respectively) formulas are of the form $\bigvee_{n \in W_e} \psi_n$ ($\bigwedge_{n \in W_e} \psi_n$, respectively), such that there is a sequence $(\alpha_n)_{n \in W_e}$ of ordinals having limit α, given by the ordinal notation for α, and every ψ_n is a Σ_{α_n} (Π_{α_n}, respectively) formula.

There is a version of *compactness theorem* due to Kreisel and Barwise, which states that if Γ is a Π^1_1 set of computable infinitary sentences such that every Δ^1_1 subset of Γ has a model, then Γ has a model. As a corollary, we obtain that if Γ is a Π^1_1 set of computable (infinitary) sentences, and if every Δ^1_1 subset of Γ has a computable model, then Γ has a computable model (see Ash and Knight 2000).

Ash established that in a computable structure, a computable infinitary Σ_α formula defines a Σ^0_α relation in the hyperarithmetical hierarchy, and similarly a computable infinitary Π_α formula defines a Π^0_α relation.

For example, consider a computable relation R on the above introduced computable linear ordering $\mathcal{A}$ of order type $\omega + \omega^*$, which is its initial segment of type ω; that is, $R = 2\mathbb{N}$. This relation can be defined in $\mathcal{A}$ by a computable Σ_2 formula:

$$x \in R \Leftrightarrow \bigvee_{n\in\mathbb{N}} \exists x_0 \cdots \exists x_n [x_0 \prec x_1 \prec \cdots \prec x_n \wedge x = x_n \wedge$$
$$\forall y[\neg(y \prec x_0) \wedge \neg(x_0 \prec y \prec x_1) \wedge \cdots \wedge \neg(x_{n-1} \prec y \prec x_n)]].$$

Similarly, the complement of R, the final segment of $\mathcal{A}$ of order type ω^*, can be defined by the following computable infinitary Σ_2 formula:

$$x \in \neg R \Leftrightarrow \bigvee_{n\in\mathbb{N}} \exists x_0 \cdots \exists x_n [x_0 \succ x_1 \succ \cdots \succ x_n \wedge x = x_n \wedge$$
$$\forall y[\neg(y \succ x_0) \wedge \neg(x_0 \succ y \succ x_1) \wedge \cdots \wedge \neg(x_{n-1} \succ y \succ x_n)]].$$

Tennenbaum and Denisov showed independently that there is an isomorphic computable copy $\mathcal{B}$ of $\mathcal{A}$, a computable linear ordering of order type $\omega + \omega^*$ in which the initial segment of type ω is not computable. It can be shown that it does not even have to be computably enumerable. However, because of definability of R and $\neg R$, by Ash's result, such a relation must be Δ^0_2 on $\mathcal{B}$. For more on computable formulas and their connections with computability, see (Ash and Knight 2000; Fokina et al. 2014; Ha et al. 2021).

10 Cross-References

▶ Reverse Mathematics

Acknowledgment Harizanov and Srinivasan gratefully acknowledges support of FRG NSF grant DMS-2152095, and Verta support of NSF graduate research fellowship.

References

Ash CJ, Knight JF (2000) Computable structures and the hyperarithmetical hierarchy. Elsevier, Amsterdam

Chomsky N (1959) On certain formal properties of grammars. Inf Control 2:137–167

Cooper SB (2004) Computability theory. Chapman & Hall, Boca Raton

Cutland NJ (1997) Computability. An introduction to recursive function theory. Cambridge University Press, Cambridge

Davis M (1958) Computability and unsolvability. McGraw-Hill, New York

Ershov YuL (1980) Decidability problems and constructive models. Nauka, Moscow (Russian)

Ershov YuL and Goncharov SS 2000 Constructive models. Siberian School of Algebra and Logic, Kluwer Academic/Plenum Press (English translation)
Fokina EB, Harizanov V, Melnikov A (2014) Computable model theory. In: Downey R (ed) Turing's legacy. Developments from Turing's ideas in logic, Lecture notes in logic. Cambridge University Press, Cambridge, pp 124–194
Friedberg RM (1957) Two recursively enumerable sets of incomparable degrees of unsolvability (Solution of Post's problem 1944). Proc Natl Acad Sci 43:236–238
Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für Mathematik und Physik 38(1931):173–198 (German)
Griffor ER (ed) (1999) Handbook of computability theory. Elsevier, Amsterdam
Ha T, Harizanov V, Marshall L, Walker H (2021) Computability and definability. In: Cenzer D, Porter C, Zapletal J (eds) Structure and randomness in computability and set theory. World Scientific, Singapore, pp 285–355
Harizanov VS (1998) Pure computable model theory. In: Ershov YuL, Goncharov SS, Nerode A, Remmel JB (eds) Handbook of recursive mathematics, vol 1. Elsevier, Amsterdam, pp 3–114
Ji Z, Natarajan A, Vidick T, Wright J, Yuen H (2020) MIP*=RE. https://arxiv.org/abs/2001.04383
Kleene SC (1936) General recursive functions of natural numbers. Math Ann 112:727–742
Kleene SC, Post EL (1954) The upper semi-lattice of degrees of recursive unsolvability. Annals of Mathematics 59:379–407
Lerman M (1983) Degrees of unsolvability. Local and global theory. Springer-Verlag, Berlin
Matiyasevich YuV (1993) Hilbert's tenth problem. MIT Press, Cambridge, MA
Millar TS (1999) Pure recursive model theory. In: Griffor ER (ed) Handbook of computability theory. Elsevier, Amsterdam, pp 507–532
Montalbán A (2021) Computable structure theory. Within the arithmetic, Perspectives in logic. Cambridge University Press, Cambridge
Muchnik AA (1956) On the unsolvability of the problem of reducibility in the theory of algorithms. Dokl Akad Nauk SSSR 108:194–197 (Russian)
Odifreddi PG (1989, 1999) Classical recursion theory, vol I–II. Elsevier, Amsterdam
Post EL (1936) Finite combinatory processes – formulation 1. J Symb Log 1:103–105
Post EL (1944) Recursively enumerable sets of positive integers and their decision problems. Bull Am Math Soc 50:284–316
Rogers H (1967) Theory of recursive functions and effective computability. McGraw-Hill, New York
Sacks GE (1990) Higher recursion theory. Springer-Verlag, Berlin
Shlapentokh A (2007) Hilbert's tenth problem. Diophantine classes and extensions to global fields. Cambridge University Press, Cambridge
Shoenfield JR (1993) Recursion theory, Lecture notes in logic. A K: Peters, Natick
Sipser M (2013) Introduction to the theory of computation. Cengage Learning, Boston
Soare RI (1987) Recursively enumerable sets and degrees. A study of computable functions and computably generated sets. Springer-Verlag, Berlin
Soare RI (2016) Turing computability. Theory and applications. Springer-Verlag, Berlin
Turing AM (1936) On computable numbers, with an application to the Entscheidungsproblem. Proc Lond Math Soc 42:230–265
Turing AM (1938) Systems of logic based on ordinals. PhD dissertation, Princeton University
Weber R (2012) Computability theory. American Mathematical Society, Providence

Reverse Mathematics

John Stillwell

Contents

1 The Parallel Axiom 1964
2 The Axiom of Choice 1966
3 Arithmetization of Analysis 1969
4 Axioms for Numbers and Sets of Numbers 1972
5 Computability 1973
6 Arithmetization of Computation 1975
7 Theorems and Nontheorems of RCA_0 1976
8 Trees in Analysis 1977
9 Theorems and Nontheorems of WKL_0 1979
10 Arithmetic Comprehension 1981
11 Theorems and Nontheorems of ACA_0 1982
12 The Big Five 1983
13 Conclusions 1984
14 Cross-References 1986
References 1986

Abstract

Reverse mathematics is a new take on an old idea: asking which axioms are necessary to prove a given theorem. This question was first asked about the parallel axiom in Euclid's geometry and later about the axiom of choice in set theory. Obviously, such questions can be asked in many fields of mathematics, but in recent decades, it has proved fruitful to focus on subsystems of *second-order arithmetic*, where much of mainstream mathematics resides. It has been found that many basic theorems of analysis and topology, as well as certain parts of infinite algebra and combinatorics, can be proved in such systems. And, remarkably, almost all the basic theorems fall into one of five particular systems:

J. Stillwell (✉)
University of San Francisco, San Francisco, CA, USA
e-mail: stillwell@usfca.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_43

a base system RCA_0 of "constructive mathematics" and four others defined by certain axioms about real numbers. Moreover, many of the theorems not provable in RCA_0 turn out to be *equivalent to* one of these defining axioms, so we know precisely which axiom is needed to prove them. Thus, after some motivational remarks about the parallel axiom and the axiom of choice, we will concentrate on the study of RCA_0 and its extensions, which is what "reverse mathematics" is generally taken to mean today.

Keywords

Reverse mathematics · Logic · Computability · Arithmetization · Analysis

1 The Parallel Axiom

The role of the parallel axiom in geometry has been questioned since the time of Euclid himself, if not earlier. In Book I of his *Elements*, Euclid needs the parallel axiom to prove the Pythagorean theorem but first he proves many theorems without it, such as his Proposition 5: in a triangle with two equal sides, the angles opposite those sides are equal. The subsystem of Euclid's geometry without the parallel axiom is known as *neutral geometry*, and we may call theorems A and B in Euclid's geometry *equivalent* if the implications $A \Rightarrow B$ and $B \Rightarrow A$ are provable in neutral geometry.

The parallel axiom can be stated in many equivalent ways, but perhaps the simplest was given by Playfair (1795), asserting the existence of exactly one parallel to a given line through a given point:

Parallel Axiom For any line $\mathcal{L}$ and a point P outside it, there is exactly one line $\mathcal{M}$ through P that does not meet $\mathcal{L}$.

It has been observed since ancient times that many consequences of the parallel axiom are actually equivalent to it. Among them are:

- The angle sum of any triangle is π.
- The points at constant distance from a given line (on one side) form a line.
- Similar triangles of different sizes exist.
- The Pythagorean theorem.

In fact, many equivalents of the parallel axiom arose from attempts to *prove* the parallel axiom in neutral geometry, by searching for absurd consequences of its negation. For example, if the parallel axiom is false then the angle sum of some triangle is not π. This may seem unlikely, but it is not clearly absurd. In particular, it is not clearly absurd for the angle sum of a triangle to be less than π, as would happen under the hypothesis that there is *more than one parallel* $\mathcal{M}$ to a given line $\mathcal{L}$ through a point P outside $\mathcal{L}$.

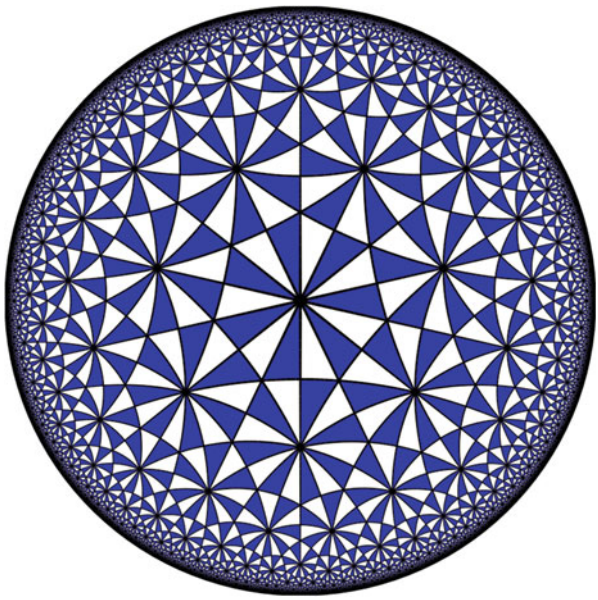


Fig. 1 The conformal disk model

The search for absurd consequences of the latter hypothesis ended when Beltrami (1868) found *models* of neutral geometry plus the hypothesis of more than one parallel, thus establishing that:

- The hypothesis of more than one parallel is consistent with neutral geometry.
- Therefore, the parallel axiom is not provable in neutral geometry.

Perhaps the clearest of Beltrami's models is the *conformal disk model* (also called the Poincaré disk, because Poincaré rediscovered it in 1882), shown in Fig. 1. The figure shows an open disk (the "plane" of the model) tiled with curvilinear triangles, each with angles $\pi/2$, $\pi/3$, and $\pi/7$. (Image from Wikimedia Commons, due to Claudio Rocchini.)

The "points" of this model are points inside the unit circle, its "lines" are circular arcs orthogonal to the unit circle, and its "angles" are the angles between these circular arcs. "Distance" is not ordinary distance in the plane; it is given by a formula that makes all the triangles of equal size. The distance function was rediscovered by Poincaré (1882) as an invariant of the mappings $z \mapsto \frac{az+b}{\bar{b}z+\bar{a}}$ of the unit disk $\{z: | z| <1\}$ in the plane $\mathbb{C}$ of complex numbers, where $|a|^2 - |b|^2 = 1$. These mappings are the *isometries* of the model and they include mappings sending any triangle in the tiling to any other, thereby showing that all these triangles are "congruent" in the sense of the model.

Since there are infinitely many triangles between any "point" and the unit circle, each "line" has infinite length. It can be checked, using the distance formula, that all the axioms of neutral geometry are satisfied in this model. It is also clear that the parallel axiom is *not* satisfied, because one can find a "line" $\mathcal{L}$, a "point" P outside $\mathcal{L}$, and several "lines" $\mathcal{M}$ through P that do not meet $\mathcal{L}$.

Thus Beltrami's model shows that the parallel axiom is *not* a theorem of neutral geometry, so its many equivalents are not theorems of neutral geometry either. On the contrary, there is an equally valid *non-Euclidean geometry* in which the parallel axiom and its equivalents are false. To prove any of these equivalents, it is *necessary* to add the parallel axiom to neutral geometry.

The train of thought we followed to arrive at this conclusion is the prototype of reverse mathematics. It involves:

- A weak "base theory" T (in this case, neutral geometry).
- An axiom A, seemingly not provable in T, but with many interesting equivalents.
- Construction of a *model* of T plus the negation $\neg A$ of A, thus showing that A and its equivalents are not provable in the base theory T.

The crucial (and typically most difficult) step is the construction of the model. In the case of Euclid's geometry, there was an interval of over 2000 years between the question of proving the parallel axiom and its negative answer by means of the Beltrami's models. In the next case, history speeds up, but the models are still formidably difficult to construct.

2 The Axiom of Choice

A host of new axiomatic issues in mathematics arose from the discovery of Cantor (1874) that the infinite sets $\mathbb{N}$ of natural numbers and $\mathbb{R}$ represent *different kinds of infinity*. Specifically, $\mathbb{R}$ is *uncountable*, in the sense that there is no bijection between $\mathbb{R}$ and $\mathbb{N}$. The easiest proof of uncountability is the one found by Cantor (1891) for $\mathcal{P}(\mathbb{N})$, the set of subsets of $\mathbb{N}$. Namely, any countable collection S_0, S_1, S_2, ... of subsets of $\mathbb{N}$ is not all of $\mathcal{P}(\mathbb{N})$, because it omits the "diagonal" set $S = \{n : n \notin S_n\}$. The uncountability of $\mathbb{R}$ follows by pairing subsets of $\mathbb{N}$ with real numbers.

Before the discovery of uncountable sets, infinity was thought to be manageable by the proof method of *induction*, whereby one proves that a property P holds for all members of $\mathbb{N}$ by proving first that P holds for 0 and then proving that P holds for $n + 1$ whenever it holds for n. The method of induction rests on the *well-ordering* property of $\mathbb{N}$, according to which each subset of $\mathbb{N}$ has a least member, and it can be used for any infinite set in bijection with $\mathbb{N}$; that is, any *countable* set.

Cantor (1883) boldly hypothesized, as a "law of thought," that every set S can be well-ordered. That is, there is relation $\prec$ on S with the properties:

- For all $a \in S$, $a \not\prec a$
- For all $a, b \in S$, $a \prec b$, or $b \prec a$
- For all $a, b, c \in S$, if $a \prec b$ and $b \prec c$ then $a \prec c$

which define a *linear order*, together with the property:

- For each set $T \subseteq S$, there is a $\prec$ least member of T

which defines a *well-order*.

Cantor's hypothesis brings the advantage of inductive proofs to uncountable sets but seems doubtful even for $\mathbb{R}$. The usual ordering $<$ on $\mathbb{R}$ is a linear ordering, but certainly not a well-ordering – for example, the set of positive numbers has no least member – and no way to define a well-ordering seems apparent. Thus, the existence of a well-ordering for each infinite set seems at best to be an unprovable axiom.

This raises the question: what axioms are suitable for set theory and, in particular, what is a base theory against which axioms such as well-ordering might be studied? Cantor himself did not propose axioms for set theory. With the exception of his well-ordering hypothesis, he used plausible but informal reasoning about infinite sets. However, even plausible reasoning about infinite sets can lead to contradictions.

For example, it seems plausible that every property $P(x)$ of sets x corresponds to a set S; namely, $S = \{x : P(x)\}$. But, as Russell (1902) famously pointed out, what if $P(x)$ is the property $x \notin x$? In this case, if $S \notin S$ then $S \in S$ by definition of S, and likewise $S \in S$ implies $S \notin S$. This contradiction became known as the *Russell paradox*. To avoid it, axioms for set theory must avoid identifying sets with properties – sets must be produced in some other manner. This in fact is not hard to do, but it was not done until after Zermelo (1904) had wrestled with the problem of proving Cantor's well-ordering hypothesis.

With remarkable ingenuity, Zermelo succeeded in proving well-ordering from an apparently much more plausible hypothesis we now call the *axiom of choice*.

Axiom of Choice Any set S of nonempty sets x has a *choice function*; that is, a function f with domain S such that $f(x) \in x$ for each $x \in S$.

It is easy to prove, conversely, that existence of a well-ordering for every set implies the axiom of choice. Suppose that S is a set of nonempty sets x, and consider the union, U, of all the sets x in S. Then a well-ordering $\prec$ of U gives a choice function f for S; namely, let $f(x) =$ - least member of x in U.

Zermelo (1908) gave the first set of axioms for set theory, simultaneously avoiding the Russell paradox and providing a base theory in which the axiom of choice could be proved equivalent to Cantor's well-ordering hypothesis. The idea of Zermelo (and probably of Cantor before him) was that sets should arise from the empty set by certain set construction operations, much as the natural numbers

arise from 0 by the successor operation. Zermelo's set construction operations are:

- **Pairing.** From any sets a and b form the set $\{a, b\}$.
- **Union.** From any set S form the union of all sets x in S.
- **Power Set.** From any set x form the set $\mathcal{P}(x) = \{y : y \subseteq x\}$ of all subsets of x.
- **Definable Subset.** For any set S and any formula $P(x)$ in the language of set theory form the subset $\{x \in S : P(x)\}$. (The language of set theory has symbols only for the relations $\in$ and $=$ plus parentheses and symbols for logical terms such as "and," "or," "not," "implies," "for all," and "there exists.")

In addition, there are axioms stating the existence of the empty set and an infinite set (essentially $\mathbb{N}$).

Zermelo's axioms suffice to prove the equivalence of the axiom of choice and well-ordering, and they avoid the Russell paradox by disallowing the collection of *all* sets x with a given property $P(x)$ into a set. The definable subset axiom allows collection of sets with property $P(x)$ only within a set S already constructed. Indeed, in Zermelo's set theory, there is no largest set because, for any set S, $\mathcal{P}(S)$ is larger, as was proved by Cantor (1891). This is not unlike the fact that there is no largest natural number, since every natural number has a successor.

The Zermelo axiom system Z is a candidate for a base system in which to study the axiom of choice and its equivalents; however, it is more usual to use a somewhat stronger system ZF, with a set construction axiom called *replacement* instead of Zermelo's definable subset axiom. Replacement was introduced by Fraenkel (1922), and it can be expressed as follows:

Replacement For any function definition $f(x)$ in the language of set theory, and any set y, form the range of f on y; that is, the set $\{f(x) : x \in y\}$.

ZF is stronger than Z because it allows the construction of sets such as:

$$\{\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}\,\mathcal{P}(\mathbb{N}), \ldots\},$$

which is the range of the function f on $\mathbb{N}$ defined inductively by $f(0) = \mathbb{N}$ and $f(n+1) = \mathcal{P}(\,f(n))$. Because of its strength, ZF has become the standard system of set theory in mathematics, and it can still serve as a base for studying the axiom of choice, because the axiom of choice is not provable in ZF, and nor is its negation.

We know this by the construction of models. Assuming the consistency of ZF, Gödel (1938) found a model of ZF plus the axiom of choice, thereby showing that the axiom of choice is not *dis*provable in ZF. Cohen (1963) found a model of ZF in which the axiom of choice does not hold, thereby showing that the axiom of choice is not provable in ZF. Thus the axiom of choice, and all of its equivalents,

are *independent* of ZF, much as the parallel axiom is independent of neutral geometry.

In another echo of the situation in geometry, there are many interesting equivalents of the axiom of choice. One example from mainstream mathematics is *existence of a basis for any vector space*. Hamel (1905) essentially proved that the axiom of choice implies existence of a basis; the reverse was proved by Blass (1984).

3 Arithmetization of Analysis

The discovery of non-Euclidean geometry raised doubts about the value of Euclid's *Elements* as a foundation of mathematics and as a guide to the mathematical study of space. At the same time, the use of geometric intuition to study lines and curves was confronted by the difficulty of proving the fundamental theorem of algebra. The proof of Gauss (1816), in particular, relied on a special case of the *intermediate value theorem*, according to which a continuous real function that passes from negative to positive values somewhere takes the value zero.

Bolzano (1817) noticed this difficulty and made a start on proving the intermediate value theorem by giving essentially the modern ε-δ definition of continuity. But he could not complete the proof without making a new assumption about the real numbers. He had to assume the *least upper bound* property, which states that any bounded set of real numbers has a least upper bound.

The way out of this difficulty, ironically, was found by looking more closely at the *Elements*, particularly Book V. The solution is to *arithmetize* the line, by viewing some of its points as rational numbers. This was first done by Dedekind in 1858 and published in his book Dedekind (1872). Book V of the *Elements* observes that any point on the line is determined by rational points above and below it, because if points p and q are different, there is a rational point r between them. Dedekind tweaked this idea by considering the *sets* L and U of rational numbers respectively less than, or greater than or equal to, a given point p. Then p corresponds to the pair $\langle L, U\rangle$, where L and U are *any* nonempty subsets of $\mathbb{Q}$ with the properties:

- Each member of L is less than every member of U.
- L has no greatest member.
- $L \cup U = \mathbb{Q}$.

Such a pair is called a *Dedekind cut,* and L (which by itself suffices to determine the point p) is called a *lower Dedekind cut*.

The Dedekind cut idea shows that each real number p may be encoded by a set L of rational numbers, and hence ultimately by a set of natural numbers, as we will see below. Also, the sum and product of real numbers, and their algebraic properties, reduce to corresponding properties of the rational numbers. For example, if:

$L_1 =$ lower Dedekind cut for p_1,
$L_2 =$ lower Dedekind cut for p_2,
then $L_1 + L_2 = \{r_1 + r_2 : r_1 \in L_1, r_2 \in L_2\} =$ lower Dedekind cut for $p_1 + p_2$.

We can then prove that the sum $L_1 + L_2$ of lower Dedekind cuts is commutative, associative, and so on, directly from the corresponding properties of the sum $r_1 + r_2$ of rational numbers. It is similar, with the appropriate modifications to deal with negative numbers, to prove the algebraic properties of product for real numbers. One can also give rigorous proofs of results such as $\sqrt{2}\sqrt{3} = \sqrt{6}$, which Dedekind believed had never been done before.

Thus Dedekind cuts reduce the *arithmetic* of real numbers to the arithmetic of rational numbers. They also prove Bolzano's least upper bound property, and very easily. Given any set of real numbers p_i, with corresponding lower Dedekind cuts L_i, the lower Dedekind cut for the least upper bound for the p_i is simply the *union* of the L_i. This follows immediately from the fact that $p_i \leq p_j$ just in case $L_i \subseteq L_j$.

With the real numbers "arithmetized" by reducing them to rational numbers and their arithmetic, it became possible to describe the line, the plane, and all manner of geometric figures in terms of rational numbers and sets of rational numbers, so arithmetization extended to essentially all of analysis and geometry. At the International Congress of Mathematicians in 1900, Poincaré expressed the state of affairs at the end of the nineteenth century:

> Today in analysis there remain only natural numbers or finite or infinite systems of natural numbers ... Mathematics, as one says, has been arithmetized. Poincaré (1902), p. 120.

He was able to say "natural numbers" instead of "rational numbers" because rational numbers m/n can be encoded by ordered pairs $\langle m, n \rangle$ of integers, integers m can be encoded by ordered pairs $\langle i, j \rangle$ of natural numbers with $i - j = m$, and ordered pairs of natural numbers can be encoded by single natural numbers. Perhaps the most obvious way to number the pairs $\langle i, j \rangle$ of natural numbers is by their position in Fig. 2, where

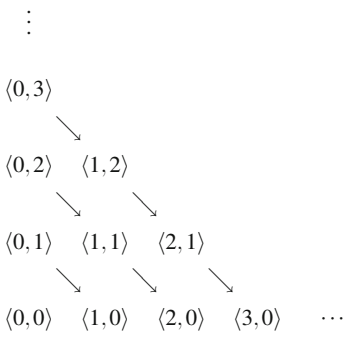


Fig. 2 Listing the ordered pairs of natural numbers

the zeroth pair is in the bottom left corner and any other pair is reached by following the arrows on successive diagonals in the left-to-right order.

In this ordering, the pair $\langle i, j\rangle$ is numbered by the *Cantor pairing function*:

$$P(i,j) = i + \frac{(i+j)(i+j+1)}{2},$$

so-called because it is essentially due to Cantor (1895). A more direct encoding of rational numbers by natural numbers can be seen in the list:

$$0;\ \frac{1}{1}, -\frac{1}{1};\quad \frac{1}{2},\quad -\frac{1}{2}, \frac{2}{1}, -\frac{2}{1};\quad \frac{1}{3}, -\frac{1}{3}, \frac{3}{1}, -\frac{3}{1};\quad \frac{1}{4}, -\frac{1}{4}, \frac{2}{3}, -\frac{2}{3}, \frac{3}{2},\ -\frac{3}{2}, \frac{4}{1}, -\frac{4}{1};\quad \ldots\ ,$$

which includes all the fractions $\pm m/n$, in lowest terms, in order of the sums $m + n$.

Rather surprisingly, it is not only the line and plane (and higher-dimensional spaces) that can be arithmetized, but any continuous real function $f : \mathbb{R} \to \mathbb{R}$. Such a function can in fact be encoded by a *single real number*, and hence by a single subset of $\mathbb{N}$. This result is implicit in results of Cantor from the 1880s, but it seems to have been first stated explicitly by Borel (1898), p. 109.

To see why a continuous $f : \mathbb{R} \to \mathbb{R}$ can be encoded by a real number, first note that f is determined by its values on the rational numbers, since any value $f(x)$ is the limit of values $f(r)$, for rational r, as r tends to x. Then, since there are countably many rational numbers $r_0, r_1, r_2, \ldots$, as the list above shows, f is determined by the countably many real numbers $x_0 = f(r_0)$, $x_1 = f(r_1)$, $x_2 = f(r_2), \ldots$. Finally, countably many real numbers $x_0, x_1, x_2, \ldots$ can be encoded by a single real number, in various ways.

For example, if we assume for simplicity that the reals are in the open interval $(0, 1)$ (which can be achieved by mapping $\mathbb{R}$ into this interval), then the decimal expansions of $x_0, x_1, x_2, \ldots$ can be woven into a single decimal expansion x as shown in Fig. 3.

Similar ideas can be used to encode points of $\mathbb{R}^2$, or more generally of any $\mathbb{R}^m$, as real numbers. This paves the way to encode any continuous $f : \mathbb{R}^m \to \mathbb{R}^n$ by a single real number and hence by a set of natural numbers. To sum up: *classical analysis, at least as far as it concerns continuous functions on $\mathbb{R}^m$, can be written in a language with variables ranging over natural numbers and sets of natural numbers*.

This is what it meant to say, in 1900, that "mathematics has been arithmetized."

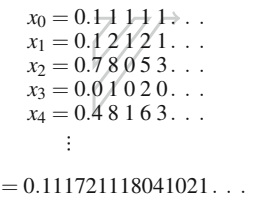


Fig. 3 Weaving a sequence of decimals into a single decimal

4 Axioms for Numbers and Sets of Numbers

Sets of natural numbers play a role not only in analysis, but even in arithmetic, as was evident from the first axiom system for arithmetic, proposed by Peano (1889). The Peano axioms take the *successor function* S as given, subject only to the conditions that 0 is not a successor, and that different numbers have different successors. Following an idea of Grassmann (1861), the axioms also define the sum and product functions *inductively* in terms of successor, and ensure the validity of induction in general by an *induction axiom*. Peano's statement of the induction axiom involves arbitrary sets of natural numbers, so two sorts of variable are required: one sort for natural numbers and the other sort for sets of natural numbers. We distinguish the two sorts by lower- and upper-case letters. The axioms are:

1. For all n, $0 \neq S(n)$.
2. For all m and n, if $m \neq n$ then $S(m) \neq S(n)$.
3. For all m and n, $m + 0 = m$ and $m + S(n) = S(m + n)$ (definition of sum).
4. For all m and n, $m \cdot 0 = 0$ and $m \cdot S(n) = (m \cdot n) + m$ (definition of product).
5. For all X, if $0 \in X$ and, for all n, $S(n) \in X$ when $n \in X$, then $X = \mathbb{N}$. (Induction.)

These axioms are sometimes called the *second-order* Peano axioms, because they involve *sets of* natural numbers as well as natural numbers. There are also *first-order* Peano axioms, which avoid the set variable X at the expense of replacing the single induction axiom by an infinite family of axioms, one for each property $\varphi(n)$ that can be written in the language with number variables, constant 0, function symbol S, relation symbol $=$, and the usual parentheses and logic symbols. In the first-order Peano axiom system, axiom 5 above is replaced by the *induction axiom schema*:

$5'$. If $\varphi(0)$ holds and, for all n, $\varphi(S(n))$ holds when $\varphi(n)$ holds, then $\varphi(n)$ holds for all n.

The induction axiom schema overcomes a certain vagueness in the second-order induction axiom, which does not tell us what it would mean to be "given" a set X. By writing down all properties φ that might define X – the so-called *arithmetically definable* properties – the first-order induction schema covers all uses of induction likely to occur in practice.

In fact, some believe that the only meaningful properties $\varphi(n)$ are those for which there is an algorithm that decides, for each n, whether $\varphi(n)$ holds or not. These "decidable properties" are a small subset of the arithmetically definable properties, as we will see when we formalize the concept of algorithm in the next section. This makes them suitable as part of a base system for analysis, in which only such "decidable" subsets of $\mathbb{N}$ are assumed to exist. It turns out that only a few of the basic theorems of analysis can be proved in this system, so one needs to add stronger set existence assumptions to prove the rest. What is more, most of the basic theorems of analysis are *equivalent* to one of a small number of set existence axioms. So set existence axioms play a role in analysis not unlike that played by the parallel axiom in geometry and the axiom of choice in set theory.

5 Computability

The origins of the computability concept are in logic, going back to a speculation of Leibniz that the validity of arguments might someday be tested by calculation. The first important step towards this goal was made by Boole (1847), who showed how to test validity in propositional logic by computation in what we now call Boolean algebra. Such computation is conceptually simple (essentially arithmetic mod 2) but propositional logic is unfortunately *too* simple to express most mathematical arguments. A logic strong enough for mathematics was introduced by Frege (1879). This *predicate logic,* as it is called, extends propositional logic by symbols for predicates (properties and relations between objects), variables x, y, z, ... for objects, and the quantifiers $\forall x$ and $\exists x$. Frege generated the valid formulas of predicate logic from axioms, by certain rules, in the manner of theorems in Euclid's *Elements*, rather than by computation as it was then understood. Moreover, it was not clear that Frege's axioms and rules were adequate to generate all logically valid formulas.

Much later, Gödel (1930) showed that Frege's system does indeed generate all logically valid formulas. At the same time, the concept of computation was being transformed, in imitation of logic itself.

Between the discoveries of Frege and Gödel, the most influential publication in logic was the *Principia Mathematica* of Whitehead and Russell (1910). Their three volumes, building on the work of Frege and Peano, gave axioms for logic and mathematics and derived many basic theorems – enough, they thought, to demonstrate that their system could prove all mathematical truths. They used a symbolic language in which conclusions were drawn from previously derived formulas in a "mechanical" manner, without needing to consider the meaning of the symbols. For example, the classical rule of *modus ponens* is of this kind: from A and $A \Rightarrow B$ derive B. We can apply this rule without knowing the meaning of A, B, or $\Rightarrow$. In the early 1920s, Post reduced the whole *Principia* system to a set of purely "mechanical" rules of this kind, each operating on strings of symbols (the "hypotheses") to produce a new string (the "conclusion").

Post's goal was to simplify the rules, hopefully to the point where one could determine whether a given string was producible from the axioms, and thus to determine whether a given mathematical sentence is true.

At first, this idea seemed promising. Post was able to reduce all derivation rules to the very simple type

from gA derive Ag', written $gA \rightarrow Ag'$ for short,

where A is an arbitrary string of symbols, and $\langle g, g' \rangle$ is a fixed pair of strings. He called a finite set of such rules a *normal system* and found that a certain normal system, applied to the axioms of *Principia*, could generate precisely its theorems. But, even with simple rules, it may be hard to foresee which strings they produce. Post was stymied by the rules below, on strings of 0s and 1s. The rules may be concisely described by saying: if a string has 0 on the left, attach 00 on the right and remove the first three symbols on the left; if the string has 1 on the left, attach 1101 on the right and remove the first three symbols on the left.

$$
\begin{array}{lll}
000A & \text{produces} & A00 \\
001A & \text{produces} & A00 \\
010A & \text{produces} & A00 \\
011A & \text{produces} & A00 \\
100A & \text{produces} & A1101 \\
101A & \text{produces} & A1101 \\
110A & \text{produces} & A1101 \\
111A & \text{produces} & A1101
\end{array}
$$

In all the cases Post tried, the string either dwindled to the empty string or else became periodic. However, he was unable to prove this, or to give an algorithm to decide whether a given string occurs as an outcome. To this day, the problem has not been solved!

At this point, as he recounted in Post (1941), Post's whole train of thought went into reverse. If there is no *apparent* algorithm to decide which strings are output by an arbitrary normal system, perhaps there is no such algorithm! If, as he thought, normal systems generate *all* sets that can be computably listed, then a set that is not the output of a normal system cannot be listed by any algorithm at all. One such "unlistable" set is a set of numbers *not* equal to the output set of any normal system. And we find such a set by the Cantor (1891) "diagonal" argument that $\mathcal{P}(\mathbb{N})$ is larger than $\mathbb{N}$.

In more detail: if $S_0, S_1, S_2, \ldots \subseteq \mathbb{N}$, then $D = \{n : n \notin S_n\}$ is a subset of $\mathbb{N}$ not among the S_n, because D differs from S_n with respect to the number N. Post's idea was to consider the subsets of $\mathbb{N}$ output by normal systems, which he called the *recursively enumerable* subsets of $\mathbb{N}$ (today, often called *computably* enumerable). The normal systems, being finite strings of symbols, can be enumerated and hence there are countably many recursively enumerable subsets of $\mathbb{N}$, which Post (1944) denoted by $W_0, W_1, W_2, \ldots$. Now, by Cantor's diagonal argument:

$$D = \{n : n \notin W_n\}$$

is *not* a recursively enumerable set, because it differs from each W_n with respect to the number n. But, interestingly, the *complement* of D,

$$\mathbb{N} - D = \{n : n \in W_n\}$$

is recursively enumerable. Informally speaking, we can list each member of $\mathbb{N} - D$ in a computation that proceeds by stages: at stage s do s steps in the generation of W_0, $W_1, W_2, \ldots, W_s$ and list any $n \leq s$ that occurs in the output of W_n.

It then follows, since D is *not* recursively enumerable, that there is no algorithm to decide membership in $\mathbb{N} - D$—otherwise, by applying the algorithm to 0,1,2, ... in turn we could recursively enumerate both D and $\mathbb{N} - D$.

To sum up: *assuming that any set whose members can be computably listed is recursively enumerable*, we can find a *normal system* (whose output is $\mathbb{N} - D$) *for*

which there is no algorithm to decide membership in its output set. Such a set is called *nonrecursive*.

Post discovered this nonrecursive recursively enumerable set $\mathbb{N} - D$ around 1921 but did not publish it at the time because he was troubled by the assumption that normal systems precisely capture the concept of computable enumeration. He felt, not unreasonably, that this was not a mathematical definition but more like a "law of nature," in need of eternal testing. It was not until 1936 that this remarkable assumption became generally accepted, thanks mainly to the paper of Turing (1936) which presented a more intuitively convincing model of computation.

Gödel, in particular, was not convinced until he saw Turing's paper. Later, in Gödel (1946), he pronounced it "a kind of miracle" that computability has a precise mathematical definition.

6 Arithmetization of Computation

Even before Gödel accepted that computability was a precise mathematical concept, he made a crucial contribution to computability theory by *arithmetizing* it. His famous proof of the incompleteness of *Principia Mathematica*, Gödel (1931), depends on numbering formulas and translating deductions into arithmetical processes, ultimately definable in Peano arithmetic. Gödel's arithmetization carries over to normal systems, with the result that any sentence of the form $n \in W$, where W is a recursively enumerable set, is equivalent to a sentence of Peano arithmetic of the form:

$$\exists x_1 \exists x_2 \ldots \exists x_k R(x_1, x_2, \ldots, x_k, n)$$

where R is quantifier-free. Such a formula is called a Σ_1^0 formula. Since R involves only the successor, sum, and product functions, equality, and symbols of propositional logic, we can decide, for any particular values of $x_1, x_2, \ldots, x_k, n$ whether $R(x_1, x_2, \ldots, x_k, n)$ holds. (In fact it was later shown, by Matijasevič (1971), that $R(x_1, x_2, \ldots, x_k, n)$ can be of the form $p(x_1, x_2, \ldots, x_k, n) = 0$, for where p is a polynomial with integer coefficients.)

Conversely, for any quantifier-free formula $R(x_1, x_2, \ldots, x_k, n)$ of Peano arithmetic, the set

$$W = \{n : \exists x_1 \exists x_2 \ldots \exists x_k R(x_1, x_2, \ldots, x_k, n)\}$$

is recursively enumerable. To enumerate the numbers n in this set, we enumerate the $(k+1)$ tuples of natural numbers $\langle x_1, x_2, \ldots, x_k, n\rangle$, and for each one add n to the list of members of W if $R(x_1, x_2, \ldots, x_k, n)$ holds.

Thus, Gödel's arithmetization shows that recursive enumerability is captured by the Σ_1^0 formulas of Peano arithmetic. This allows the concept of computability to be imported into Peano arithmetic, and finally brings us to the base system for reverse

mathematics: RCA_0. RCA stands for "recursive comprehension axiom," and the subscript denotes a refinement to the system originally proposed.

RCA_0 extends Peano arithmetic by a set existence axiom saying that, for any recursive property $\varphi(n)$, there is a set $X \subseteq \mathbb{N}$ such that:

$$n \in X \Leftrightarrow \varphi(n).$$

To express this axiom in the language of Peano arithmetic, we use the fact that a recursive property is one whose positive and negative cases are both recursively enumerable (the reason for this was given in the previous section, when we argued that an algorithm for membership in D would allow us to recursively enumerate both D and $\mathbb{N} - D$). Thus a recursive property $\varphi(n)$ is one for which there are Σ_1^0 formulas $\chi(n)$ and $\psi(n)$ such that:

$$\varphi(n) \Leftrightarrow \chi(n) \quad \text{and} \quad \chi(n) \Leftrightarrow \neg\psi(n).$$

Given such Σ_1^0 formulas, we can state the set existence axiom (schema) of RCA_0 as:

$$\exists X \forall n[n \in X \Leftrightarrow \varphi(n)] \tag{RCAx}$$

The other axioms are those of first-order Peano arithmetic, except that the φ in the induction axioms $5'$ are restricted to Σ_1^0 formulas. This is in keeping with the intention of keeping proofs in RCA_0 "constructive." It should also be mentioned that the formula $\varphi(n)$ can contain free set variables $Y, Z, \ldots$ other than X, so that X can be a set *constructed from* the sets $Y, Z, \ldots$. For example, for any set $Y \subseteq \mathbb{N}$, we can assert that the set X of even members of Y exists.

The existence of nonrecursive sets is the starting point of reverse mathematics, because it enables the construction of models that separate "constructive" theorems of analysis from "nonconstructive" ones. An obvious model of RCA_0 consists of the set $\mathbb{N}$ of natural numbers together with all its recursive subsets. We will call this the *minimal model,* because any model of RCA_0 necessarily include all recursive sets. It follows that anything RCA_0 can prove must hold in the minimal model. This creates opportunities for computability theory, where situations that give rise to *non*-recursive sets have been studied since the 1930s.

7 Theorems and Nontheorems of RCA_0

RCA_0 turns out to be a good base theory because it is strong enough to prove equivalences between statements, even though it cannot prove many of them outright. For example, in RCA_0 one cannot prove even the following weak form of Bolzano's least upper bound principle: *each bounded nondecreasing sequence of rational numbers has a limit.*

To see why, let f be a computable function that enumerates the members of a recursively enumerable but nonrecursive set W. Thus

$$W = \{\, f(n) : n \in \mathbb{N}\}$$

but the characteristic function of W,

$$\chi_W(m) = \begin{cases} 1 & \text{if } m \in W \\ 0 & \text{if } m \notin W, \end{cases}$$

is not computable.

Now consider the sequence $a_0, a_1, a_2, \ldots$, where

$$a_m = \sum_{i=1}^{m} 2^{-f(i)}.$$

This is a nondecreasing sequence of rational numbers, bounded by 1, whose limit a has binary expansion essentially the same as χ_W, because it has 1 in the mth place if $m \in W$, and 0 in the mth place otherwise. Thus the real number a is *not* computable and hence does not belong to the computable model of RCA_0.

On the other hand, the sequence $a_0, a_1, a_2, \ldots$ is the set of ordered pairs $\langle m, a_m \rangle$ which is recursive and hence *does* belong to the computable model. (To decide whether $\langle m, n \rangle$ belongs to the sequence, compute $f(0), f(1), \ldots, f(m)$ and see whether $n = \sum_{i=1}^{m} 2^{-f(i)}$.) Thus there is a sequence in the model whose limit is not in the model, which shows that RCA_0 cannot prove the existence of the limit.

Nevertheless, there are a few important theorems of analysis that RCA_0 *can* prove, because it is possible to construct the objects asserted to exist. One of these is the intermediate value theorem. Given a continuous function f (encoded as a set of natural numbers) on an interval $[a, b]$, with $f(a) < 0$ and $f(b) > 0$, one can compute a value c where $f(c) = 0$ by successive approximations. The idea is to repeatedly bisect the interval and to evaluate f at the bisection point. Either we will find c as a bisection point or else we can compute a nested sequence of intervals $[a_n, b_n]$ such that $f(a_n) < 0$ and $f(b_n) > 0$. Then c is the common point of these intervals, and the successive bisections eventually give us successive binary digits of c.

The axiom RCAx then allows us to assert the existence of c and hence prove the intermediate value theorem in RCA_0. We also get the fundamental theorem of algebra, by plugging the intermediate value theorem into the Gauss (1816) proof.

8 Trees in Analysis

The process of repeated bisection just described in the proof of the intermediate value theorem reappears in several well-known proofs in basic analysis. Here is another important example.

Bounded Value Theorem *If f is a continuous function on a closed interval $[a, b]$ then f is bounded on $[a, b]$.*

Proof. Suppose on the contrary that f is *not* bounded on $[a, b]$. Then f is unbounded on at least one half of $[a, b]$. Let $[a_1, b_1]$ be the leftmost half of $[a, b]$ on which f is unbounded, and repeat the argument on $[a_1, b_1]$. That is, f must be unbounded on at least one half of $[a_1, b_1]$; let $[a_2, b_2]$ be the leftmost such half, and so on.

Proceeding in this way, we obtain a sequence of intervals:

$$[a, b] \supseteq [a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots,$$

on which f is unbounded, and with a common point p whose binary expansion is computable from the sequence. Also, since f is continuous, $f(p)$ is defined and, in a sufficiently small interval containing p, the value of $f(x)$ differs from $f(p)$ by at most ε. But this contradicts f being unbounded on each $[a_n, b_n]$, since these intervals become arbitrarily small □.

Although it looks quite similar, this proof differs from the proof of the intermediate value theorem in one important respect. While the intermediate value c is computable from the given continuous function f, there is no apparent way to compute the common point p from the given function f. Finding the intervals $[a_n, b_n]$ depends on knowing all values of f over a given interval, and there would seem to be no finite computation that finds this information. We need a set existence axiom *stronger* than RCAx in order to claim the existence of the sequence of intervals $[a_n, b_n]$. It is convenient to take this axiom to be about *infinite paths in a binary tree*.

A binary tree underlies the proofs above. Its vertices are the interval $[a, b]$ and the infinitely many subintervals obtained from $[a, b]$ by halving. From each vertex, which is an interval I, there are edges to the left and right halves of I. This binary tree is shown in Fig. 4, with the vertex $[a, b]$ at the top.

The tree in the figure is called the *full* binary tree because there are two edges below every vertex v. In general, there may be only one edge below a vertex v, or none (in which case none of the vertices below v is in the tree). The sequence $[a, b] \supseteq [a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots$ found in the proof above is an example of an *infinite path* in the full binary tree, because each vertex $[a_{n+1}, b_{n+1}]$ is a half of the

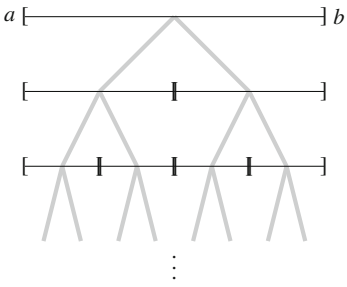


Fig. 4 A full binary tree

vertex $[a_n, b_n]$ and hence is connected to it by an edge. However, since we were not able to compute this path, we cannot necessarily prove, in RCA_0, that it exists. To guarantee its existence, we appeal to an axiom called the *weak Kőnig lemma,* which gets its name from a slightly stronger lemma due to Kőnig (1927).

The weak Kőnig lemma can be stated in terms of sets of natural numbers, but it is clearer to view it as a theorem of infinite graph theory. In this context, the lemma says: *if T is an infinite subtree of the full binary three, then T has an infinite path.* Its proof echoes the construction of the infinite path in the proof above. Start at the top vertex v_0 of *T*. Since *T* has infinitely many vertices, one of the (at most two) edges out of v_0 must have infinitely many vertices of *T* below it. Take the leftmost vertex, v_1, with infinitely many vertices below it, and repeat the argument. We can proceed in this way indefinitely, and hence find an infinite sequence of vertices, $v_0, v_1, v_2, \ldots$, each connected to the one before by an edge. This is an infinite path.

Although this argument is just as convincing as the proof above, it does not *compute* the path, so it is not expressible in RCA_0. Indeed, Kreisel (1953) found a recursive infinite binary tree *T* with *no* recursive infinite path. So, in the minimal model of RCA_0, *T* exists but its infinite paths do not, showing that RCA_0 cannot prove the weak König lemma. This makes the lemma a candidate for a new axiom. The resulting system RCA_0+weak König lemma is called WKL_0, and it can prove the bounded value theorem for continuous functions. RCA_0 cannot prove this theorem, but it can prove it *equivalent* to the weak König lemma. Thus, the weak König lemma is not just a convenient axiom to prove the theorem; it is in some sense the *right* axiom to prove it.

Similar discoveries about the weak König lemma, and other set existence axioms, led Friedman (1975) to view reverse mathematics as a program for finding "right" axioms.

> When the theorem is proved from the right axioms, the axioms can be proved from the theorem. Friedman (1975), p. 235.

9 Theorems and Nontheorems of WKL_0

As mentioned in the previous section, WKL_0 can prove the boundedness of continuous functions on closed intervals. The latter theorem is the first step towards the *extreme value theorem,* which says that a continuous function *f* on a closed interval attains a maximum and minimum. The classical proof of the extreme value theorem – if no maximum, let *l* be the least upper bound of *f*, so $1/(l - f(x))$ is continuous but unbounded – does not carry over directly to WKL_0. Nevertheless, the extreme value theorem can be proved in WKL_0, and RCA_0 proves it equivalent to the weak König lemma.

Another classical theorem provable by the weak König lemma is the *Heine-Borel theorem* for sequences, which says that any covering of [0,1] by open intervals

$(c_0, d_0), (c_1, d_1), (c_2, d_2), \ldots$ contains a finite subcovering $(c_0, d_0), \ldots, (c_k, d_k)$. (The full Heine-Borel theorem is for *sets* of open intervals, but in the language of arithmetic, we can only talk about countable sets, and this is usually all we need.) This theorem is also provable in WKL_0, as are the following consequences:

- Any function continuous on a closed interval is uniformly continuous.
- Any function continuous on a closed interval is Riemann integrable there.

Moreover, all of these theorems are provably equivalent to the weak König lemma in RCA_0. The equivalence of the weak König lemma and Heine-Borel for sequences was discovered by Friedman (1975). For proofs of all these results, and many others, see Simpson (2009).

The weak König lemma can also be used to prove the *Brouwer fixed-point theorem* in topology, which says that any continuous map of the n-dimensional ball into itself has a fixed point. This theorem is also equivalent to the weak König lemma in RCA_0. Moreover, Takayuki Kihara has recently shown that the same is true of two other celebrated topological theorems due to Brouwer:

- Invariance of domain: A continuous injective map of an open subset of $\mathbb{R}^m$ into $\mathbb{R}^m$ has an open image.
- Invariance of dimension: $\mathbb{R}^m$ and $\mathbb{R}^n$ are homeomorphic only if $m = n$.

Since these theorems are equivalent to the weak König lemma, they share in its "nonconstructive" character. Brouwer, after he made his name by his discoveries in topology, became a strong campaigner against nonconstructive methods, and in fact he rejected his own fixed-point theorem in a lecture in Berlin in 1927 (see van Dalen 2013, pp. 382 and 503). To my knowledge, he never rejected the invariance of domain and invariance of dimension, though we can now say that, relative to RCA_0, they are precisely as nonconstructive as the fixed-point theorem.

We have been speaking until now of the *weak* König lemma, which came to light because of its many interesting equivalents. The original lemma of Kőnig (1927) was not about binary trees, but the more general case of *finitely branching trees*; that is, trees in which each vertex belongs to finitely many edges. The full *König lemma* states that such an infinite tree T contains an infinite branch, and the proof is almost identical with that of the weak Kőnig lemma. Namely, pick a vertex v_0 and consider the finitely many edges out of v_0. One of them leads into an infinite subtree of T, so we let v_1 be such a vertex (say, the one with least index in the list of vertices of T). We repeat the argument with v_1 and continue indefinitely, thus defining an infinite branch.

Despite its similarity to the weak Kőnig lemma, the full Kőnig lemma is *not* provable in WKL_0. In WKL_0 we also lack guarantees for the convergence of sequences, already noticed as a failing of RCA_0. For both these things, we need a stronger set existence axiom.

10 Arithmetic Comprehension

A natural set existence axiom in second-order arithmetic, which is considerably easier to motivate than the recursive comprehension axiom RCAx or the weak König lemma, is the *arithmetic comprehension axiom* (schema), which simply removes the restriction on the formula φ in RCAx:

$$\exists X \forall n[n \in X \Leftrightarrow \varphi(n)], \qquad \text{(ACAx)}$$

where $\varphi(n)$ is now any formula in the language of Peano arithmetic, possibly with free set variables (but not including X). Arithmetic comprehension is stronger than recursive comprehension or the weak Kőnig lemma, which is why we have delayed its introduction until now: having seen the capabilities of recursive comprehension and the weak König lemma, we are in a better position to appreciate the need for extra strength.

When ACAx is the set existence axiom, it is also natural to remove the restriction on the induction axiom, and to bring back Peano's second-order induction axiom,

5. For all X, if $0 \in X$ and, for all n, $S(n) \in X$ when $n \in X$, then $X = \mathbb{N}$.

The system consisting of the Peano axioms plus ACAx is called ACA_0.

Since ACA_0 allows X to be defined by any arithmetic formula φ, Peano's induction axiom implies the first-order induction schema:

5′. If $\varphi(0)$ holds and, for all n, $\varphi(S(n))$ holds when $\varphi(n)$ holds, then $\varphi(n)$ holds for all n.

Thus ACA_0 can prove all the axioms of first-order Peano arithmetic, and hence all of its theorems. Rather surprisingly, the theorems of first-order Peano arithmetic are *precisely* the "purely arithmetic" theorems of ACA_0; that is, those that involve only variables for natural numbers. This was discovered by Friedman (1976), and we say that ACA_0 is *conservative over* first-order Peano arithmetic.

It means that the ability to reason about sets of natural numbers (or, equivalently, about real numbers and continuous functions) in ACA_0 does not help to prove purely arithmetic theorems. This is surprising in view of the common belief that analysis is needed to prove certain theorems of number theory. We will see shortly that quite a lot of analysis can be developed in ACA_0 – but Friedman's theorem shows that it gives no purely arithmetic theorems beyond those already be proved in first-order Peano arithmetic.

Like RCA_0, ACA_0 has a rather obvious minimal model; namely, $\mathbb{N}$ together with all its arithmetically definable subsets. As one would expect, there are models of WKL_0 that lie between the minimal models of RCA_0 and ACA_0, but describing them (and consequently showing that WKL_0 has strength strictly between RCA_0 and ACA_0) requires some deeper results from computability theory that are beyond the scope of this chapter. For more information on models of WKL_0, see Simpson (2009), pp. 318–319.

11 Theorems and Nontheorems of ACA_0

ACA_0 can prove all the theorems of WKL_0, because in any infinite arithmetically definable tree, the leftmost infinite path is arithmetically definable, and hence the weak König lemma follows by the axiom ACAx. Less obviously, ACA_0 can prove several theorems of analysis that WKL_0 cannot, such as:

- The sequential Bolzano-Weierstrass theorem: *any bounded infinite sequence of real numbers has a convergent subsequence*.
- The sequential least upper bound principle: any bounded sequence of real numbers has a least upper bound.
- Cauchy convergence criterion: a sequence $x_0, x_1, x_2, \ldots$ is convergent if, for any $\varepsilon > 0$, there is an n such that $|x_m - x_n| < \varepsilon$ for all $m > n$.
- Monotone convergence theorem: *any bounded monotonic sequence is convergent*.

Moreover, these theorems are all equivalent to each other, and to the axiom ACAx, over RCA_0.

Outside of analysis, ACA_0 also proves some interesting theorems of infinite combinatorics, many of them equivalent to ACAx. One interesting equivalent that we have already mentioned is the full König lemma.

Another interesting part of combinatorics around the level of ACA_0 is *Ramsey theory,* which originates in the paper Ramsey (1930). Ramsey theory typically concerns "coloring" the k-element subsets of natural numbers by finitely many colors, then showing that $\mathbb{N}$ contains an infinite subset M that is "monochromatic" in the sense that all k-element subsets of M have the same color. This result may be called RT(k) for short. For example, to prove RT(1), we have $k = 1$ so we color the members of $\mathbb{N}$ by finitely many colors. It is clear that infinitely many members of $\mathbb{N}$ get the same color. This almost obvious result RT(1) (also known as a "pigeonhole principle") can be leveraged to prove the less obvious RT(2), which in turn can be used to prove RT(3), and so on.

All of these theorems RT(k) can be proved in ACA_0, and in fact for $k > 2$, they are each *equivalent* to ACAx over RCA_0. However, RT(2) does *not* imply ACAx, and the sentence $\forall k$ RT(k), known as the infinite Ramsey theorem, is *not* provable in ACA_0. One interesting reason for its unprovability is that it is known to imply a purely arithmetic statement not provable in first-order Peano arithmetic, and hence not in ACA_0. This is the famous *Paris-Harrington theorem* of Paris and Harrington (1977). For more about Ramsey theory and ACA_0, see Hirschfeldt (2015).

This example, incidentally, helps to explain the subscript 0 in ACA_0. Friedman (1975) originally proposed a system ACA with a stronger comprehension axiom in place of ACAx, in which the formula $\varphi(n)$ defining membership in a set X was allowed to include quantified set variables. The infinite Ramsey theorem and the Paris-Harrington theorem are provable in ACA, so the purely arithmetic theorems of ACA go beyond first-order Peano arithmetic. Disallowing quantified set variables creates the system ACA_0 with exactly the same first-order strength as Peano

arithmetic, but with second-order theorems that are important for analysis. The subscripts 0 in RCA_0 and WKL_0 arise from similar restrictions on the systems originally proposed in Friedman (1975). Friedman (1976) himself proposed these restrictions along with his proof that ACA_0 is conservative over first-order Peano arithmetic.

12 The Big Five

The systems RCA_0, WKL_0, and ACA_0 are the most interesting from the viewpoint of mainstream mathematics, but there are two higher systems that cover parts of analysis with some flavor of set theory. Altogether, the five systems are known as "big five." From the historical viewpoint, the two higher systems round off the story of analysis and set theory nicely, by ending where Cantor began – with the introduction of countable ordinals and the analysis of closed subsets of $\mathbb{R}$. The two higher systems of the "big five" are:

- ATR_0, which stands for "arithmetical transfinite recursion," has the ability to handle countable ordinals. ATR_0 is strong enough to prove the infinite Ramsey theorem that lies beyond ACA_0. We will not go into detail about countable ordinals and ATR_0; for more see Simpson (2009), pp. 173–199.
- Π^1_1-CA_0, which stands for "Π^1_1 comprehension." The symbol Π^1_1 refers to a defining property $\varphi(n)$ of the form $\forall Y R(Y, n)$, where R is a formula of second-order arithmetic with no bound set variables other than Y. Π^1_1-CA_0 is the only one of the big five systems that allows a set-defining property $\varphi(n)$ with a set quantifier in its set existence axiom,

$$\exists X \forall n [n \in X \Leftrightarrow \varphi(n)].$$

Thus Π^1_1 comprehension takes us into the territory logicians call *impredicative*, where sets X are defined in terms of a totality that includes themselves.

The set existence axiom of ATR_0 is equivalent, over RCA_0, to Cantor's *perfect set theorem,* according to which any uncountable closed set contains a perfect set; that is, one that is closed and has no isolated point. For a proof, see Simpson (2009), pp. 185–188.

The set existence axiom of Π^1_1-CA_0 is equivalent, over RCA_0, to the *Cantor-Bendixson theorem,* according to which any uncountable closed subset of $\mathbb{R}$ is the union of a countable set and a perfect set. Again, we refer to Simpson (2009), particularly pp. 217–220. The ideas behind this equivalence go back to the first published paper on these systems of analysis, Friedman (1969), which was actually about a slightly stronger system Π^1_1-CA. Thus the history of reverse mathematics really begins at the level where analysis shades into set theory, as is also evident in Friedman's thesis, Friedman (1967). Only later, in Friedman (1975), was it noticed that weaker systems nicely capture more basic theorems of analysis. Then Friedman

(1976) introduced the slightly weaker "subscript 0" systems that make up the "big five" today.

For reasons we do not completely understand, RCA_0 seems to be an ideal base system, because it pulls noteworthy theorems into the logical equivalence classes of the set existence axioms for the other four members of the "big five." These include not only theorems of analysis, topology, and combinatorics – as we have seen – but also many theorems of countable algebra, which may be seen in Simpson (2009).

Nevertheless, as we have come to expect since the incompleteness theorem of Gödel (1931), some theorems remain out of reach.

In particular, there are two combinatorial theorems of the form "infinite sequence contains an infinite monotonic subsequence," modelled on the following consequence of the Bolzano-Weierstrass theorem, but more difficult to prove.

Ascending/Descending Sequence Principle (ADS) *Any infinite sequence of rational numbers contains an infinite monotonic subsequence.*

ADS easily follows from Bolzano-Weierstrass, hence it is provable in ACA_0. However, it is not provable in RCA_0, or in WKL_0. The complicated known facts about ADS are described in Hirschfeldt (2015), pp. 130–149. Perhaps the situation with the following analogues of ADS is just as complicated, but what is remarkable is their lofty position among the big five systems.

Kruskal's Tree Theorem *Any infinite sequence of finite trees contains an infinite subsequence that is increasing under the "embeds in" relation.*

Graph Minor Theorem *Any infinite sequence of finite graphs contains an infinite subsequence that is increasing under the graph minor relation.*

The tree theorem is due to Kruskal (1960). A proof may be found in Diestel (2010), along with a definition of the graph minor relation. The graph minor theorem was proved by Robertson and Seymour (2004) (the last in a series of papers that spanned 20 years!). Friedman showed that Kruskal's theorem is not provable in ATR_0. With Robertson and Seymour, he topped this result, in Friedman et al. (1987), by showing that the graph minor theorem is not provable in Π^1_1-CA_0.

13 Conclusions

While the idea of reverse mathematics applies in principle to many fields of mathematics, it has been particularly fruitful in subsystems of second-order arithmetic. This is where analysis and computability theory come together, thanks to their common foundations in arithmetic, and it becomes possible to exploit the wealth of results about computable and noncomputable structures that have been developed since the 1930s.

In particular, the "big five" systems give a precise analysis of the assumptions about infinity needed to prove basic theorems of analysis. The weakest system, RCA_0, in which only computable infinite objects are assumed to exist, may be viewed as a formalization of constructive mathematics. It is strong enough to prove the intermediate value theorem and the fundamental theorem of algebra, but not much else. However, RCA_0 is strong enough to prove many *equivalences* between statements, and it turns out that just four of these equivalence classes capture many basic theorems – not only in analysis but in topology, infinite combinatorics, and countable algebra. These four classes lie in four nonconstructive extensions of RCA_0: in order of inclusion, WKL_0, ACA_0, ATR_0, and Π^1_1-CA_0.

The defining axioms of WKL_0 and ACA_0 reveal principles not previously viewed as fundamental in analysis. In WKL_0, the defining axiom is the *weak König lemma*, which says that any infinite binary tree contains an infinite branch. This is a nonconstructive assumption, since there exists a computable infinite tree with no computable infinite branch. But its nonconstructive content is, in a sense that can be made precise, of "low degree." Despite its low nonconstructive content, the weak König lemma is equivalent to a version of the Heine-Borel theorem and the related properties of continuous functions: their uniform continuity and Riemann integrability, and the extreme value theorem. The weak König lemma is also equivalent to Brouwer's theorems in topology – the fixed-point theorem and the invariance of domain and dimension – so it pinpoints the nonconstructive content of these famous theorems.

The defining axiom of ACA_0 is *arithmetical comprehension*, which states the existence of sets defined in the language of arithmetic. Arithmetical comprehension includes both RCA_0 and WKL_0, since computability is definable in the language of arithmetic, but it goes further. In particular, arithmetical comprehension is equivalent to various sentences about convergence that WKL_0 cannot prove. Arithmetical comprehension is also equivalent to various statements of infinite combinatorics, such as the full König lemma (which is about finitely branching trees, not just binary trees) and various statements of Ramsey theory. Another measure of the strength of ACA_0 is its ability to prove all the theorems of first-order Peano arithmetic, whereas WKL_0 can match Peano arithmetic only for theorems with a limited number of quantifiers.

Nevertheless, there are theorems beyond the reach of ACA_0, and the systems ATR_0 and Π^1_1-CA_0 capture some of the important ones. They include some of the stronger theorems of Ramsey theory, and also theorems of analysis brought to light by Cantor, whose proofs involve countable ordinals. Just as RCA_0 may be viewed as a formalization of constructive mathematics, ATR_0 has been viewed as a formalization of *predicative* mathematics, in which set definitions are extended as far as possible as long as a set is not defined in terms of a collection that includes itself. Thus ATR_0 and Π^1_1-CA_0 allow us to pinpoint the position of certain theorems at or beyond the boundary of predicativity: the perfect set theorem belongs to predicative mathematics, but the Cantor-Bendixson theorem does not.

Thus reverse mathematics, even when restricted to the "big five" systems, illuminates the mathematics of infinite processes from the constructive to the predicative, and beyond.

14 Cross-References

► The Absolute Arithmetic Continuum and its Geometric Counterpart

References

Beltrami E (1868) Teoria fondamentale degli spazii di curvatura costante. Annali di Matematica Pura ed Applicata, ser 2, 2:232–255, in his *Opere Matematiche* 1:406–429, English translation in Stillwell (1996)

Blass A (1984) Existence of bases implies the axiom of choice. In: Axiomatic Set Theory (Boulder, Colorado, 1983), Contemporary Mathematics, vol 31. American Mathematical Society, Providence, pp 31–33

Bolzano B (1817) Rein analytischer Beweis des Lehrsatzes dass zwischen je zwey Werthen, die ein entgegengesetzes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege. Ostwald's Klassiker, vol. 153. Engelmann, Leipzig, 1905. English translation in Russ (2004), 251–277

Boole G (1847) Mathematical analysis of logic. Reprinted by Basil Blackwell, London, 1948

Borel É (1898) Leçons sur la théorie des fonctions. Gauthier-Villars, Paris

Cantor G (1874) Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. Journal für reine und angewandte Mathematik 77:258–262, in his *Gesammelte Abhandlungen*, 145–148. English translation by W. Ewald in Ewald (1996), Vol. II, 840–843

Cantor G (1883) Über unendliche, lineare Punktmannigfaltigkeiten. Mathematische Annalen 21:545–591, English translation by William Ewald in Ewald (1996), volume II, pp. 881–920

Cantor G (1891) Über eine elementare Frage der Mannigfaltigkeitslehre. Jahresbericht deutschen Mathematiker-Vereinigung 1:75–78, English translation by W. Ewald in Ewald (1996), Vol. II, 920–922

Cantor G (1895) Beiträge zur Begründung der transfiniten Mengenlehre. Mathematische Annalen 46(4):481–512, English translation by P. E. B. Jourdain in Cantor (1952)

Cantor G (1952) Contributions to the founding of the theory of transfinite numbers. Dover Publications, Inc., New York, N. Y., translated, and provided with an introduction and notes, by Philip E. B. Jourdain

Cohen P (1963) The independence of the continuum hypothesis I, II. Proc Natl Acad Sci 50, 51:1143–1148, 105–110

Davis M (ed) (2004) The undecidable. Dover Publications Inc., Mineola, corrected reprint of the 1965 original [Raven Press, Hewlett, NY]

Dedekind R (1872) Stetigkeit und irrationale Zahlen. Vieweg und Sohn, Braunschweig (english translation in: Essays on the theory of numbers, Dover, New York, 1963)

Diestel R (2010) Graph theory, graduate texts in mathematics, vol 173, 4th edn. Springer, Heidelberg. https://doi.org/10.1007/978-3-642-14279-6

Ewald W (1996) From Kant to Hilbert: a source book in the foundations of mathematics. Vol. I, II. The Clarendon Press, Oxford University Press, New York

Fraenkel A (1922) Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre. Mathematische Annalen 86:230–237. https://doi.org/10.1007/BF01457986

Frege G (1879) Begriffschrift. English translation in van Heijenoort (1967), pp 5–82 [van Heijenoort J (ed) (1967) From Frege to Gödel. A source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA]

Friedman H (1967) Subsystems of set theory and analysis. PhD thesis, MIT Department of Mathematics, Cambridge, MA

Friedman H (1969) Bar induction and $\pi_1^1 - CA$. J Symb Logic 34:353–362

Friedman H (1975) Some systems of second order arithmetic and their use. In: Proceedings of the international congress of mathematicians (Vancouver, BC, 1974), vol 1, pp 235–242

Friedman H (1976) Systems of second order arithmetic with restricted induction I, II. J Symb Logic 41:557–559

Friedman H, Robertson N, Seymour P (1987) The meta mathematics of the graph minor theorem, Logic and combinatorics. American Mathematical Society, pp 229–261

Gauss CF (1816) Demonstratio nova altera theorematis omnem functionem algebraicum rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse. Commentationes societas regiae scientiarum Gottingensis recentiores 3:107–142, in his Werke 3:31–56

Gödel K (1930) Die Vollständigkeit der Axiome des logischen Funktionenkalküls. Monatshefte für Mathematik und Physik 37:349–360

Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. I. Monatshefte für Mathematik und Physik 38:173–198, English translation in van Heijenoort (1967), 596–616 [van Heijenoort J (ed) (1967) From Frege to Gödel. A source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA]

Gödel K (1938) The consistency of the axiom of choice and the generalized continuum hypothesis. Proc Natl Acad Sci 25:220–224

Gödel K (1946) Remarks before the Princeton bicentennial conference on problems in mathematics. In Davis M (ed) (2004) The undecidable. Dover Publications Inc., Mineola, corrected reprint of the 1965 original [Raven Press, Hewlett, NY]

Grassmann H (1861) Lehrbuch der Arithmetic. Enslin, Berlin

Hamel G (1905) Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x+y) = f(x) + f(y)$. Mathematische Annalen 60:459–462

Hirschfeldt DR (2015) Slicing the truth, Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore, vol 28. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, on the computable and reverse mathematics of combinatorial principles, Edited and with a foreword by Chitat Chong, Qi Feng, Theodore A. Slaman, W. Hugh Woodin and Yue Yang

König D (1927) Über eine Schlussweise aus dem Endlichen ins Unendliche. Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, sectio scientiarum mathematicarum 3:121–130

Kreisel G (1953) A variant to Hilbert's theory of the foundations of arithmetic. British J Philos Sci 4:107–129 errata and corrigenda, 357 (1954)

Kruskal JB (1960) Well-quasi-ordering, the Tree Theorem, and Vazsonyi's conjecture. Trans Am Math Soc 95:210–225

Matijasevič YV (1971) Diophantine representation of recursively enumerable predicates. In: Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 1. Gauthier-Villars, Paris, pp 235–238

Paris J, Harrington L (1977) A mathematical incompleteness in Peano arithmetic. In: Barwise J (ed) Handbook of mathematical logic. North-Holland, Amsterdam

Peano G (1889) Arithmetices principia. Bocca, Torino

Playfair J (1795) Elements of geometry. Bell and Bradfute, Edinburgh

Poincaré H (1882) Théorie des groupes fuchsiens. Acta Math 1:1–62, in his *Œuvres* 2: 108–168. English translation in Poincaré (1985), 55–127

Poincaré H (1902) Du rôle de l'intuition et de la logique en mathématiques. Compte Rendu du Deuxième Congrès International des Mathématiciens, Paris

Poincaré H (1985) Papers on Fuchsian Functions. Springer, New York, translated from the French and with an introduction by John Stillwell

Post EL (1941) Absolutely unsolvable problems and relatively undecidable propositions – an account of an anticipation. In: Davis M (ed) (2004) The undecidable. Dover Publications Inc., Mineola, corrected reprint of the 1965 original [Raven Press, Hewlett, NY], pp 338–433

Post EL (1944) Recursively enumerable sets of positive integers and their decision problems. Bull Am Math Soc 50:284–316

Ramsey FP (1930) On a problem of formal logic. Proc London Math Soc 30:264–286

Robertson N, Seymour PD (2004) Graph minors. XX. Wagner's conjecture. J Comb Theory Ser B 92(2):325–357

Russ S (2004) The mathematical works of Bernard Bolzano. Oxford University Press, Oxford

Russell B (1902) Letter to Frege, pp 124–125, in van Heijenoort (1967), pp 124–125 [van Heijenoort J (ed) (1967) From Frege to Gödel. A source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA]

Simpson SG (2009) Subsystems of Second Order Arithmetic, 2nd edn. Perspectives in logic. Cambridge University Press, Cambridge/Association for Symbolic Logic, Poughkeepsie

Stillwell J (1996) Sources of hyperbolic geometry. American Mathematical Society, Providence

Turing A (1936) On computable numbers, with an application to the Entscheidungsproblem. Proc London Math Soc 42:230–265

van Dalen D (2013) L. E. J. Brouwer—topologist, intuitionist, philosopher. Springer, London

van Heijenoort J (1967) From Frege to Gödel. A source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA

Whitehead AN, Russell B (1910) Principia Mathematica. Cambridge University Press, Cambridge, 3 vols. 1910, 1912, 1913

Zermelo E (1904) Beweis dass jede Menge wohlgeordnet werden kann. Mathematische Annalen 59:514–516, English translation in van Heijenoort (1967), pp 139–141 [van Heijenoort J (ed) (1967) From Frege to Gödel. A source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA]

Zermelo E (1908) Untersuchungen über die Grundlagen der Mengenlehre I. Mathematische Annalen 65:261–281, English translation in van Heijenoort (1967), pp 200–215 [In: van Heijenoort J (ed) (1967) From Frege to Gödel. A source book in mathematical logic, 1879–1931. Harvard University Press, Cambridge, MA]

Countable Nonstandard Models: Following Skolem's Approach

Rumen D. Dimitrov and Valentina Harizanov

Contents

1 Models and Theories 1990
2 Constructing Models via Ultraproducts 1992
2.1 Ultrafilters 1992
2.2 Ultraproducts of Structures 1994
2.3 Compactness Theorem 1996
3 Computability Theory and Skolem's Model of Arithmetic 1997
3.1 Notes on Skolem's 1955 Paper 2000
3.2 Computability and Models of Fragments of Arithmetic 2003
4 Cohesive Sets 2003
4.1 Feferman-Scott-Tennenbaum Model 2005
4.2 Cohesive Powers of Computable Structures 2006
4.3 Concluding Remarks 2007
References 2008

Abstract

In 1934, Skolem gave a remarkable construction of a countable nonstandard model of arithmetic. His construction contains ideas of the ultrapower construction which was introduced in model theory 20 years later. However, typical ultrapower constructions produce uncountable models. Skolem's construction can also be connected with ideas from computability theory, formalized by Turing and others in 1936. The proof of one of Skolem's key statements can be interpreted using computability-theoretic notions of indecomposable sets developed in the late 1950s. We favor the computability-theoretic approach and

R. D. Dimitrov
Department of Mathematics & Philosophy, Western Illinois University, Macomb, IL, USA
e-mail: rd-dimitrov@wiu.edu

V. Harizanov (✉)
Department of Mathematics, The George Washington University, Washington, DC, USA
e-mail: harizanv@gwu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_21

describe the connections of Skolem's ideas with our current study of cohesive powers of models in computable model theory.

Keywords

Model theory · Computability · Countable models · Cohesive ultrapower

1 Models and Theories

Model theory studies in a rigorous and general way the intuitive notion of a mathematical model, which is used in all the sciences. In the abstract language of model theory, a *model* (also called *structure*) is a nonempty set of elements, called the domain, with certain relations, functions (operations), and constants on the domain. The names for these objects are called relation symbols, function symbols, and constant symbols. These symbols constitute the *language* (also called *signature*) $\mathcal{L}$ of the model. Every relation symbol or a function symbol in $\mathcal{L}$ has arity (length) and is interpreted in a model as a relation or a function of the same arity. We consider first-order formulas with equality, which allow existential ($\exists$) and universal ($\forall$) quantification only over first-order variables (standing for elements of the domain of the structure). A structure is countable if its domain is countable.

A *magma* is a nonempty set with a binary operation, $(M, *)$. If the operation is associative, it is called a *semigroup.* A semigroup with a unit element is called a *monoid,* and a monoid where every element has an inverse is called a *group.* A *field* is a structure with two binary operations and two constants, $(F, +, \cdot, 0, 1)$, such that $(F, +, 0)$ and $(F - \{0\}, \cdot, 1)$ are groups. Examples of familiar countable structures are the ordered sets of natural numbers $(N, <)$, the ordered set of integers $(Z, <)$, and the ordered set of rational numbers $(Q, <)$; the monoid $(N, +, 0)$ of the natural numbers with addition, the additive group of integers $(Z, +, 0)$, and the field of rational numbers $\mathcal{Q} = (Q, +, \cdot, 0, 1)$.

A *sentence* (or a statement) in $\mathcal{L}$ is a formula where all variables are quantified. Every sentence in $\mathcal{L}$ can be assigned a truth value in a structure for the same language $\mathcal{L}$. A *theory* is a (consistent) set of sentences. The theory of a structure $\mathcal{A}$, in symbols $Th(\mathcal{A})$, is the set of all first-order sentences true in $\mathcal{A}$. Let n be a positive integer. We classify a formula in prenex normal form as Σ_n (or Π_n) if it has n blocks of like quantifiers, beginning with a block of $\exists$ (or $\forall$) quantifiers. For a formula α, in prenex normal form, we let $neg(\alpha)$ denote the dual formula that is logically equivalent to the negation $\neg\alpha$. If α is Σ_n, then $neg(\alpha)$ is Π_n, and vice versa. In other words, Σ_1 formulas are existential formulas and Π_1 formulas are universal formulas, Π_2 formulas are $\forall\exists$ formulas, Σ_2 formulas are $\exists\forall$ formulas, Π_3 formulas are $\forall\exists\forall$ formulas, etc. The quantifier-free formulas are considered to be both Σ_0 and Π_0.

Peano arithmetic, introduced by Giuseppe Peano in 1889, is a theory axiomatized by a finite set of axioms plus an induction axiom scheme, thus by countably many axioms. Peano arithmetic is intended to be an explicit axiomatic description of what we call the familiar structure of *natural numbers*,

$$\mathbb{N} = \left(N; S^{\mathbb{N}}, +^{\mathbb{N}}, \cdot^{\mathbb{N}}, 0^{\mathbb{N}}, =^{\mathbb{N}}\right),$$

which we study since elementary school, with the domain $N = \{0, 1, 2, \ldots\}$, with the usual successor function, addition and multiplication functions (operations), the constant zero, and the equality relation, in the language $\mathcal{L} = \{S, +, \cdot, 0, =\}$. Note that we often do not include equality as part of $\mathcal{L}$ since it is assumed. Here, we do include equality because of a historical reference (Skolem 1955) to Skolem's presentation. We often do not write the superscript $\mathbb{N}$ when it is clear from the context what the intended interpretation of the symbol is in the structure, thus using the same notation for the relations, operations, and constants in the structure and their names in the language. Here are the axioms of Peano arithmetic.

1. $S(x) \neq 0$,
2. $S(x) = S(y) \Rightarrow x = y$,
3. $x + 0 = x$,
4. $x + S(y) = S(x + y)$,
5. $x \cdot 0 = 0$,
6. $x \cdot S(y) = x \cdot y + x$,
7. Induction scheme: for every first-order formula $\alpha(x, \overline{y})$ with a free variable x and a finite sequence of parameters $\overline{y}$, we have

$$(\alpha(0, \overline{y}) \& (\forall z)(\alpha(z, \overline{y}) \Rightarrow \alpha(s(z), \overline{y}))) \Rightarrow (\forall x)\alpha(x, \overline{y}).$$

Axioms 1–2 state that 0 is not a successor and that the successor function is one-to-one; axioms 3–4 give a recursive definition of addition, and axioms 5–6 give a recursive definition of multiplication. *Peano arithmetic* (abbreviated as PA) is the set of all sentences that can be deduced from the axioms of PA using the inference rules of a formal system:

$$\mathrm{PA} = \{\sigma : \mathit{Axioms}(\mathrm{PA}) \vdash \sigma\}.$$

Other familiar relations on N can be defined. For example, the relation $\leq$ can be defined as:

$$a \leq b \Leftrightarrow (\exists c)(a + c = b).$$

We say that a set is *countable* if it is infinite but its elements can be enumerated in a sequence indexed by natural numbers. The enumeration does not have to be algorithmic. There are countably many $\mathcal{L}$-sentences in PA. Its set of axioms is not logically equivalent to a finite set of axioms as shown in 1952 by Ryll-Nardzewski (1952) (see also Mostowski 1952). Whether a sentence in the language of arithmetic is an axiom can be algorithmically verified.

A consistent theory T is complete if for every first-order sentence θ in the language of T we have $T \vdash \theta$ or $T \vdash \neg\theta$. It is a consequence of the Gödel incompleteness theorem that PA is incomplete. The incompleteness of PA means

that there is a sentence γ in $\mathcal{L}$ such that neither γ nor $\neg\gamma$ is in PA. The collection of all $\mathcal{L}$-sentences true in $\mathbb{N}$, the theory of $Th(\mathbb{N})$, is called *True arithmetic.* That is,

$$Th(\mathbb{N}) = \{\theta : \theta \text{ is a first-order sentence } \& \ \mathbb{N} \models \theta\}.$$

The theory of any model is complete, and hence $Th(\mathbb{N})$ is complete. Note that for the sentence γ (as above), either γ or $\neg\gamma$ is in $Th(\mathbb{N})$ and hence PA $\subsetneq Th(\mathbb{N})$.

Structures $\mathcal{A}$ and $\mathcal{B}$ in the same language $\mathcal{L}$ are *elementarily equivalent* if they satisfy the same first-order sentences, i.e., they have the same first-order theory. Such two structures are *isomorphic* if the elements of $\mathcal{A}$ can be "renamed" so the structure $\mathcal{A}$ becomes essentially the same as the structure $\mathcal{B}$. More precisely, there is a bijection h (one-to-one and onto function), mapping the domain of $\mathcal{A}$ to the domain of $\mathcal{B}$ so that for each relation symbol R, each function symbol f, and each constant symbol c, we have:

$$\begin{aligned} &R^{\mathcal{A}}(\overline{a}) \Leftrightarrow R^{\mathcal{B}}(h(\overline{a})), \\ &f^{\mathcal{A}}(\overline{a}) = b \Leftrightarrow f^{\mathcal{B}}(h(\overline{a})) = h(b), \\ &h(c^{\mathcal{A}}) = c^{\mathcal{B}}. \end{aligned}$$

In particular, isomorphic structures have the same cardinality (size) since there is a bijection between their domains. While isomorphic structures are elementarily equivalent, the converse does not hold for infinite structures. Thus, nonisomorphic structures may be the models of the same first-order theory. Intended models of such theory (as determined by mathematical practice) as well as their isomorphic copies are called *standard,* while all other models of the theory are considered to be *nonstandard.*

Neither PA nor $Th(\mathbb{N})$ fully "captures" the structure $\mathbb{N} = (N; S, +, \cdot, 0, =)$. There are models of PA (and of $Th(\mathbb{N})$) that are not isomorphic to $\mathbb{N}$ and are thus nonstandard. Both the ultrapower construction and the compactness theorem from model theory provide such nonstandard models. For more on model theory see (Chang and Keisler 1990; Bell and Slomson 1974), and on models of PA see (Kaye 1991; Kossak and Schmerl 2006).

2 Constructing Models via Ultraproducts

2.1 Ultrafilters

Various product constructions have been one of the major tools for building models of theories. These constructions often produce nonstandard models. We will now define a family of sets that will be used to determine truth in one type of products, called ultraproduct. Consider a nonempty set I, which will be used to index structures. We will call it an index set. A *filter* over I is a nonempty set D of subsets of I such that D is closed under the finite intersections of sets and under the superset relation. That is,

(1) if X and Y belong to D, then $X \cap Y$ belongs to D, and
(2) if V belongs to D and $V \subseteq Z \subseteq I$, then Z belongs to D.

The notion of a filter was introduced by Cartan in 1937 and it can be defined more generally with respect to some partially ordered set. In our case, the subsets of I are ordered by the partial relation of set-theoretic inclusion $\subseteq$, and a filter includes some subsets that are large enough to satisfy properties (1) and (2).

The set of all subsets of I is called the power set of I and is denoted by $\mathcal{P}(I)$. Hence a filter is a subset of $\mathcal{P}(I)$. An important filter is a Fréchet filter, which consists of all co-finite subsets of I, i.e., those subsets of I with finite complements. A proper filter U is called an *ultrafilter* over I if, in addition to properties (1) and (2), we have

(3) for every subset X of I, either X or $I - X$ belongs to U.

It can be shown that for any $i \in I$, the set of all $X \subseteq I$ such that $i \in X$ forms an ultrafilter. Such an ultrafilter is called *principal.* Hence an ultrafilter U is *nonprincipal* if for every $i \in I$, we have that $\{i\} \notin U$.

A set of sets $\mathcal{K}$ has the *finite intersection property* if no intersection of finitely many elements of $\mathcal{K}$ is empty. It can be shown that a set U of subsets of I is an ultrafilter over I exactly when U has the finite intersection property and is a maximal such family with respect to inclusion $\subseteq$. Moreover, if there is a nonprincipal ultrafilter U over I, then I must be infinite. That is, if, to the contrary, $I = \{a_1, \ldots, a_n\}$, then $\{a_i\} \notin U$ for every i with $1 \leq i \leq n$, so $I - \{a_i\} \in U$. Hence,

$$(I - \{a_1\}) \cap \cdots \cap (I - \{a_n\}) = \emptyset,$$

which is a contradiction.

Another way to describe an ultrafilter U over I is to replace property (3) with the property:

(3a) if U contains the union of two sets, then it contains at least one of them.

The last property generalizes to finitely many sets, so if an ultrafilter contains a finite set of elements, it must contain at least one of these elements. Hence if U is a nonprincipal ultrafilter over I, then U contains all co-finite subsets of I. Thus, assuming that I is infinite, is there a nonprincipal ultrafilter over I? We can show that we can establish this assuming Zorn's lemma, also called Kuratowski-Zorn lemma. Zorn's lemma is equivalent to the axiom of choice. For partial order of set inclusion, Zorn's lemma states that if $\mathcal{K}$ is a nonempty set of sets that is closed under unions of nonempty chains then $\mathcal{K}$ has a maximal element.

Theorem 1 *(assuming Zorn's lemma) Let I be any infinite set. Then there is a nonprincipal ultrafilter over I.*

For proof see (Bell and Slomson 1974).

2.2 Ultraproducts of Structures

The notion of the ultraproduct of first-order structures in general has been introduced by Łoś (1955) in 1955. We will now give a series of definitions and results related to the ultraproduct construction.

Let I be an infinite set, and let $\mathcal{A}_i$ for $i \in I$ be a family of structures for the same language $\mathcal{L}$ with domains A_i. Let U be an ultrafilter over I. Then the *ultraproduct* of $\mathcal{A}_i$ over U, denoted by $\mathcal{B} = \Pi_U \mathcal{A}_i$, is a structure for $\mathcal{L}$ defined as follows.

(1) The domain of the ultraproduct consists of the equivalence classes of functions f such that $f(i) \in A_i$ for every $i \in I$, modulo the equivalence relation $=_U$ defined below. We can write f as a sequence $(f(i))_{i \in I}$ choosing for each i an element of A_i. The *equality modulo U* relation $=_U$ is defined as follows:

$$f =_U g \text{ iff } \{i : f(i) = g(i)\} \in U.$$

We think of elements of U as "large sets," so the agreement of f and g is on a large set. We can easily verify that $=_U$ is an equivalence relation (reflexive, symmetric, and transitive binary relation, such as equality). The equivalence class of a function f is denoted by $[f]$. The equivalence classes are the elements of the domain $\Pi_U A_i$ of the ultraproduct.

(2) Let R be an n-ary relation symbol in $\mathcal{L}$. We define its interpretation in $\mathcal{B}$:

$R^{\mathcal{B}}([f_1], \ldots, [f_n])$, holds iff

$$\{i \in I : R^{\mathcal{A}_i}(f_1(i), \ldots, f_n(i))\} \in U.$$

We can show that this relation is well defined.

(3) Let F be an n-ary function symbol in $\mathcal{L}$. We define its interpretation in $\mathcal{B}$:

$F^{\mathcal{B}}([f_1], \ldots, [f_n]) = [g]$, where for every $i \in I$, we have

$$g(i) = F^{\mathcal{A}_i}(f_1(i), \ldots, f_n(i)).$$

We can show that this function is well defined.

(4) Let c be a constant symbol in $\mathcal{L}$. We define its interpretation in $\mathcal{B}$:

$c^{\mathcal{B}} = [f]$ where for every $i \in I$,

$$f(i) = c^{\mathcal{A}_i}.$$

The following theorem, sometimes called the fundamental theorem of ultraproducts, establishes the satisfiability in the ultraproduct.

Theorem 2 *(Łoś 1955)*

If $\alpha(x_1, \ldots, x_n)$ *is a formula in language* $\mathcal{L}$*, then*

$$\mathcal{B} \models \alpha([f_1], \ldots, [f_n]) \text{ iff } \{i : \mathcal{A}_i \models \alpha(f_1(i), \ldots, f_n(i))\} \in U.$$

Hence if σ *is a sentence, then*

$$\mathcal{B} \models \sigma \text{ iff } \{i : \mathcal{A}_i \models \sigma\} \in U.$$

If all structures $\mathcal{A}_i$ are the same structure $\mathcal{A}$, then the ultraproduct is called an *ultrapower* and denoted by $\Pi_U \mathcal{A}$. An important consequence of Łoś's theorem is the following:

Corollary 1 *If* $\mathcal{B} = \Pi_U \mathcal{A}$ *is an ultrapower, then* $\mathcal{A}$ *and* $\mathcal{B}$ *are elementarily equivalent.*

Example: Let $\mathcal{A} = (N, +)$, the additive semigroup (in fact, monoid) of natural numbers. Additive number theory, also called Presburger-Skolem arithmetic, is the set of all first-order sentences true in $\mathcal{A}$ (see Smoryński 1991). A standard model of this theory is any model isomorphic to $\mathcal{A}$. Let $I = N$, and let U be a nonprincipal ultrafilter over I. Consider the ultrapower $\mathcal{B} = \Pi_U \mathcal{A}$ with domain B. Notice that for $a = [(a_0, a_1, a_2, \ldots)] \in B$ and $b = [(b_0, b_1, b_2, \ldots)] \in B$, we have

$$a + b = [(a_0 + b_0, a_1 + b_1, a_2 + b_2 \ldots)],$$

and

$$a + [(0, 0, 0, \ldots)] = a.$$

Łoś's theorem implies that $\mathcal{A}$ and $\mathcal{B}$ satisfy the same first-order sentences. Hence $\mathcal{B}$, although uncountable, is also a model of additive number theory. However, it has some very different properties from $(N, +)$. For example, consider $a = [(a_0, a_1, a_2, \ldots)] \in B$, where $a_0 = 0$, $a_1 = 1$, $a_2 = 1 \cdot 2$, $a_3 = 1 \cdot 2 \cdot 3, \ldots$ Then we can show that a has the property that for every natural number $n > 0$, there is a b such that:

$$a = b + \cdots + b \ (n \text{ times})$$

So, in the structure $\mathcal{B}$, a is divisible by every natural number $n > 0$.

Let U be a principal ultrafilter over N. If $U = \{X : X \subseteq N \,\&\, m \in X\}$, then $\Pi_U \mathcal{A}_i$ is isomorphic to $\mathcal{A}_m$.

Definition 1 A filter D is *countably incomplete* if it is not closed under countable intersections.

Let U be a nonprincipal ultrafilter over N. Clearly, N can be partitioned into countably many finite sets, so their co-finite complements are elements of U, and form a countable collection of sets with empty intersection. Hence any nonprincipal ultrafilter U over N is countably incomplete. Frayne, Morel, and Scott (1962) showed that for a countably incomplete ultrafilter U, the ultraproduct $\Pi_U \mathcal{A}_i$ is either finite or of cardinality at least continuum. Continuum is the cardinality of the power set of N, which is the same as the cardinality of the set of real numbers. Thus, such an ultraproduct $\Pi_U \mathcal{A}_i$ is never countably infinite.

Hence any ultrapower $\Pi_U \mathbb{N}$ of the standard model of arithmetic, $\mathbb{N} = (N; S, +, \cdot, 0, =)$, over a nonprincipal ultrafilter U has the same first-order theory as $\mathbb{N}$, but is not isomorphic to $\mathbb{N}$ since it is uncountable.

2.3 Compactness Theorem

Compactness theorem is another convenient tool for building nonstandard models. Compactness is a property of topological spaces and was introduced by Fréchet in 1906. In our case it means that satisfiability of an infinite theory is guaranteed by the satisfiability of all of its finite subtheories. Compactness theorem is often proved as a consequence of the Gödel completeness theorem. It can also be proved using the ultraproduct construction (see Frayne et al. 1962).

Theorem 3 *Let Σ be an infinite first-order theory. Assume that every finite subset of Σ has a model. Then Σ has a model.*

Proof idea Let the index set I be the set of all finite subsets of Σ. For every $i \in I$, let $\mathcal{A}_i$ be a model for i. Then a model $\mathcal{B}$ of Σ is produced as an ultraproduct $\Pi_U \mathcal{A}_i$ of the models $\mathcal{A}_i$ over a suitably defined ultrafilter U on I.

We can use compactness theorem to produce models with nonstandard elements. We first expand the original language with names for nonstandard elements and describe their properties using an infinite set of sentences in the expanded language.

Theorem 4 *There is a model of Th*$(\mathbb{N})$ *with nonstandard numbers*.

Proof Let $\mathcal{L}$ be the language of arithmetic. Let c be a new constant, and $\mathcal{L}' = \mathcal{L} \cup \{c\}$. Define the following set of sentences in $\mathcal{L}'$:

$$\Sigma = Th(\mathbb{N}) \cup \{c \neq 0, c \neq 1, c \neq 1+1, c \neq 1+1+1, \ldots\}.$$

We will apply compactness theorem to Σ. Let Σ' be a finite subset of Σ and let n be the largest number such that the following sentence

$$c \neq 1 + 1 + \cdots + 1(n \text{ ones})$$

belongs to Σ'. Then the standard model $\mathbb{N}$, with the constant c interpreted as the number $n+1$, is a model of Σ'.

By compactness theorem, Σ has a model $\mathcal{A}$. The reduct of $\mathcal{A}$ to $\mathcal{L}$ is a model of $Th(\mathbb{N})$ and it cannot be isomorphic to $\mathbb{N}$ because the interpretation of c in $\mathcal{A}$ is one of its nonstandard elements.

The ultraproduct proof of the compactness theorem may produce uncountable models. A different proof of the compactness theorem produces a countable model of $Th(\mathbb{N})$. Namely, it is easy to see that in any formal system a set of sentences Σ is consistent if and only if every finite subset of Σ is consistent. Assume that every finite subset of Σ is satisfiable (and thus consistent). Then Σ itself is consistent. Assuming that the language of Σ is finite or countable, we can use the proof of Gödel completeness theorem due to Henkin (in 1949) to construct a countable model of Σ. Thus, $Th(\mathbb{N})$ also has a countable nonstandard model.

By PA^- we denote essentially PA without the induction axiom scheme. Gödel and Rosser showed that if T is a theory in the language of arithmetic with a computable set of axioms, which extends PA^-, then there is a sentence (in fact, Π_1 sentence) ψ such that neither ψ nor $\neg\psi$ can be deduced from T. Thus both $T \cup \{\neg\psi\}$ and $T \cup \{\psi\}$ are consistent. Continuing this process we can show that there are continuum many countable nonstandard models of T (and hence of PA). Regarding the order type of such nonstandard models of arithmetic, Henkin showed that they are order isomorphic to a copy of the natural numbers followed by countably densely many (without ends) copies of integers.

3 Computability Theory and Skolem's Model of Arithmetic

Computability theory, the exact theory of algorithms, studies the power and inherent limitations of algorithmic methods. The most fundamental concept in computability theory is that of a computable set. A set, or more generally a relation, is *computable* if there is a decision procedure for determining in finitely many steps whether a given tuple of elements belongs to that relation.

A set of natural numbers is *computably enumerable* if there is an algorithm that enumerates (lists) its elements. Computably enumerable sets are often abbreviated by c.e. All c.e. sets can be simultaneously algorithmically listed by systematically coding all Turing machine programs

$$P_0, P_1, P_2, \ldots, P_e, \ldots$$

Let

$$\varphi_0, \varphi_2, \varphi_2, \ldots, \varphi_e, \ldots$$

be an algorithmic enumeration of all unary partial computable functions on natural numbers (computed by the corresponding Turing machine programs). Some of these functions are computable. A function is *computable* if it is partial computable and defined on all of N. However, a diagonal argument can be used to show that there is no algorithmic enumeration of all computable functions.

A set X is c.e. if there is a computable function f that enumerates its elements. That is, the set X is the range (the set of all values) of f:

$$X = rng(f) = \{f(n) : n \in N\}.$$

It can be shown that a set is c.e. if and only if it is the domain of a partial computable function. The domain of a partial computable function φ_e is denoted by W_e. We say that an index e is the Gödel code of φ_e. Each partial computable function has infinitely many such codes. Hence

$$W_0, W_1, W_2, \ldots, W_e, \ldots$$

is an algorithmic enumeration of all c.e. sets.

It can be shown that a set is computable if and only if both the set and its complement are c.e. There are c.e. sets the complements of which are not c.e. Namely, there are noncomputable c.e. sets. For example, the Turing *halting set*

$$K = \{e : \varphi_e(e) \text{ halts}\} = \{e : e \in W_e\}$$

is c.e., but not computable. The set K is c.e. because it is enumerated by an algorithm that simultaneously runs, for more and more computational steps, $\varphi_0(0)$, $\varphi_1(1)$, $\varphi_2(2)$, ... and enumerates those e for which $\varphi_e(e)$ converges, as soon as the convergence occurs. We can show that the complement $\overline{K}$ is not c.e. Assume otherwise. Then $\overline{K} = W_i$ for some i and

$$i \in \overline{K} \Leftrightarrow i \in W_i \Leftrightarrow i \in K,$$

which is a contradiction. Infinite computable sets can be viewed as c.e. sets the elements of which are enumerated in strictly increasing order.

Computable sets describe decidable problems. Computably enumerable sets describe partial decidable problems. There are continuum many non-c.e. sets that are further classified into an infinite hierarchy according to their computability-theoretic complexity.

A set X of natural numbers is called *arithmetical* if and only if it is definable in the standard model of arithmetic. That is, X is definable in $\mathbb{N}$ if there is a first-order

formula $\alpha(x)$ in the language of arithmetic with a free variable x, so that X is the set of all elements n in N such that $\alpha(n)$ is true in $\mathbb{N}$. There is a similar definition for an n-ary relation P on natural numbers. Such an arithmetical relation P is classified as being Σ_n^0 (Π_n^0, respectively) if and only if it is definable by a Σ_n (Π_n, respectively) formula in $\mathbb{N}$. Moreover, as Gödel and Matiyasevich established, $\Sigma_1^0 \cap \Pi_1^0$ sets are exactly the computable sets, Σ_1^0 sets are exactly the c.e. sets, Π_1^0 sets are exactly the complements of c.e. sets, etc. Any set of sentences in a computable language (including any finite language) can be represented as a subset of N by identifying sentences with their Gödel codes. It turns out that Peano arithmetic (PA) is a c.e. theory (Σ_1^0 set), while $Th(\mathbb{N})$ is not Σ_n^0 for any n.

We say that a countable structure for a computable language $\mathcal{A}$ is *computable* if its domain A is a computable set and its relations and functions are uniformly computable. In computable structures, we have a decision algorithm not only for quantifier-free sentences, but also for quantifier-free sentences with constant symbols added as names for all elements of the domain. Examples of computable structures include the ordered sets of natural numbers, integers, rational numbers, the additive group of integers, and the standard model of arithmetic. Note that computable structures may have complicated theories such as $Th(\mathbb{N})$.

It is much harder to think of a countable noncomputable structure. In the following example of such structure, let X be any noncomputable set. Define a linear order

$$\mathcal{A} = (N, \prec)$$

with (computable) domain N and a special order relation $\prec$, which is the same as the usual order relation $<$ except that the order is switched on each pair $(2n, 2n+1)$ if $n \notin X$. That is:

$$2n \prec 2n+1 \text{ if } n \in X,$$
$$2n+1 \prec 2n \text{ if } n \notin X.$$

Note that $\mathcal{A}$ is not computable as a structure, because the membership in the set X can be computed from the order relation $\prec$ of $\mathcal{A}$. Moreover, $\mathcal{A}$ is isomorphic to the ordered set of natural numbers $(N, <)$.

The standard model $\mathbb{N}$ of arithmetic is computable, while, as Tennenbaum (1959) showed, nonstandard models are not computable. Namely, if $\mathcal{M}$ is a nonstandard model of PA with domain M, then $(M, +, 0)$ is not computable. For proof and ramifications of Tennenbaum's theorem see (Kaye 1991; Schmerl 2011; Ash and Knight 2000).

For more on computability theory see (Soare 1987; Rogers 1967; Smoryński 1991) and on computable model theory see (Ash and Knight 2000; Fokina et al. 2014).

3.1 Notes on Skolem's 1955 Paper

In 1934, Skolem (1934) gave a remarkable construction of a countable nonstandard model of arithmetic. In 1955, he (Skolem 1955) presented the same result using more contemporary notation and terminology. His model can be described using notions from computability theory. In the introduction to his 1970 paper, Lerman (1970) gave the following description of Skolem's model:

"In 1934, Skolem proved the nonaxiomatizability of arithmetic by constructing nonstandard models for arithmetic. He did so by taking all arithmetic functions of one variable, and factoring out by the equivalence relation: $f \sim g$ if and only if $f(x) = g(x)$ almost everywhere (a.e.) in A, i.e., for all but finitely many $x \in A$, where A is any arithmetically indecomposable set. The model thus obtained satisfies exactly the same first-order sentences as the nonnegative integers."

The elements of Skolem's model of arithmetic are equivalence classes of functions of arbitrary arithmetic complexity. There have been studies of more effective models of fragments of arithmetic by Feferman, Scott, and Tennenbaum (1959), Hirshfeld, Lerman, McLaughlin, and Wheeler (Lerman 1970; Hirshfeld 1975; Hirshfeld and Wheeler 1975; McLaughlin 1989, 1990, 2007). Recursive ultrafilter construction has also been studied by Nelson (1992). More recently, cohesive powers of computable structures have been studied by Dimitrov, Harizanov, Miller, Mourad, Shafer, A. Soskova, and Vatev (2008, 2009a, 2014, n.d.). In this section, we give a more detailed analysis of Skolem's construction and its connection with effective ultrapowers of computable structures, known as cohesive powers (see Dimitrov 2009a).

The language $\mathcal{L} = \{', +, \cdot, 1, =\}$ used by Skolem is essentially our language $\{S, +, \cdot, 0, =\}$, where $'$ is the successor function S. Skolem uses a standard axiomatization of arithmetic, similar to PA but excluding zero. Since $'$ is definable in $\mathcal{L}_1 = \{+, \cdot, 1, =\}$, Skolem's formulas are essentially the first-order formulas in $\mathcal{L}_1$. The formulas are built "by use of quantifiers on an elementary expression, namely built by use of &, $\vee$ and $\neg$ on equations with polynomial terms on both sides." In other words, formulas are assumed to be in the prenex normal form. For $\mathcal{L}_1$-formulas, Skolem first replaces $y \neq x$, that is, $\neg(y = x)$ with

$$\exists z((x = y + z) \vee (y = x + z)). \tag{1}$$

Assuming that the sentences are in the prenex normal form, Skolem proceeds to omit the other propositional connectives. He replaces

$$(x = y) \vee (z = t)$$

with

$$xt + yz = xz + yt. \tag{2}$$

He replaces

$$(x = y)\&(z = t)$$

with

$$x^2 + y^2 + z^2 + t^2 = 2xy + 2zt. \tag{3}$$

Thus, every sentence is equivalent to a single equation involving polynomials on both sides preceded by a sequence of quantifiers. Skolem then considers "the true" sentences of type

$$\forall x_1 \ldots \forall x_n \exists y \Phi(x_1, \ldots, x_n, y),$$

where the arithmetic complexity of Φ is unrestricted. According to Skolem "the word true may mean provable in the system or that the statement is assumed as a further axiom." So we can assume that $\forall x_1 \ldots \forall x_n \exists y \Phi(x_1, \ldots, x_n, y)$ is an element of $Th(\mathbb{N})$. Such sentences define the arithmetical functions (now also called Skolem functions) as follows:

$$f(x_1, \ldots, x_n) = y \Leftrightarrow \Phi(x_1, \ldots, x_n, y)\&(\forall z)(\neg\Phi(x_1, \ldots, x_n, z) \vee (y \leq z)). \tag{4}$$

In the expression above we assume that $\leq$ is defined in $\mathcal{L}_1$ as usual. The set of all such arithmetical functions $f(x_1, \ldots, x_n)$ is denoted by F. Skolem then performs what we now call Skolemization where we repeatedly replace the first existentially bound variable y that is in the scope of the universal quantifiers $\forall x_1 \ldots \forall x_n$ by the function $f(x_1, \ldots, x_n) \in F$ as described in (4). In the example given by Skolem,

$$(\forall x)(\forall y)(\exists z)((x = y) \vee (x = y + z) \vee (y = x + z))$$

is replaced by

$$(\forall x)(\forall y)((x = y) \vee (x = y + f(x, y)) \vee (y = x + f(x, y))),$$

where in this case the Skolem function $f(x, y)$ is the "distance" function,

$$f(x, y) = \mid x - y \mid.$$

Note that we leave the occurrences of $\vee$ in the sentence above for the clarity of presentation.

We can assume that all sentences in $Th(\mathbb{N})$ are in the *Skolem normal form* (a block of universal quantifiers followed by a quantifier-free formula). Since the set F of arithmetical functions is closed under function composition (also called nesting or substitution), each sentence in $Th(\mathbb{N})$ is the universal closure of a single equation involving elements of F. Before defining a countable nonstandard model $\mathbb{N}^*$ of $Th(\mathbb{N})$, Skolem proves an important lemma.

Lemma 1 *(Skolem, Satz 1 in (Skolem 1934), Arithmetical Lemma in (Skolem 1955))*

Let $F_1 = \{f_1(t), f_2(t), \ldots\}$ be the set of arithmetical functions of one variable. There is a function $g(t)$ in F_1 such that for all nonequal $i, j \geq 1$, exactly one of the following holds:

$$\forall t > \max(i,j)\Big(f_i(g(t)) < f_j(g(t))\Big),$$

$$\forall t > \max(i,j)\Big(f_i(g(t)) = f_j(g(t))\Big),$$

$$\forall t > \max(i,j)\Big(f_i(g(t)) > f_j(g(t))\Big).$$

Remark 1 We will not provide a proof of Skolem's lemma here. However, we will later modify a proof of the existence of a cohesive set, a result by Dekker and Myhill (see Theorem VI in 12.3 of (Rogers 1967)), in order to show its similarity to Lemma 1.

Next, fix a function g the existence of which is given by Lemma 1, and let

$$f_i =^g f_j \text{ iff } \Big\{t : f_i(g(t)) = f_j(g(t))\Big\} = N \text{ almost everywhere.}$$

It is easy to see that $=^g$ is an equivalence relation. For any $f \in F_1$, let $[f]$ denote the equivalence class of f:

$$[f] = \{h \in F_1 : h =^g f\}.$$

Let the domain N^* be the set of all equivalence classes:

$$N^* = \{[f] : f \in F_1\}.$$

Let $\mathbb{N}^*$ be the structure in the language $\{+, \cdot, 1, =\}$ with domain N^* where $=_{\mathbb{N}^*}$, $1_{\mathbb{N}^*}$, $+_{\mathbb{N}^*}$, and $\cdot_{\mathbb{N}^*}$ are defined as follows. The equality $=_{\mathbb{N}^*}$ is the equality between equivalence classes (the elements of N^*). The constant $1_{\mathbb{N}^*}$ is the equivalence class of the function with constant value 1, denoted by [1], and

$$[f] +_{\mathbb{N}^*} [h] = [f + h],$$
$$[f] \cdot_{\mathbb{N}^*} [h] = [f \cdot h].$$

Skolem then verifies that the sentences in $Th(\mathbb{N})$ hold in $\mathbb{N}^*$. His verification can be viewed in comparison with a proof of Łoś's Theorem for ultrapowers. The fact that $\mathbb{N}^*$ is a nonstandard model of arithmetic which has $\mathbb{N}$ as its submodel is verified as follows. For $n \in N$, let $[n] \in \mathbb{N}^*$ be the equivalence class of the function with constant value n. The set $\{[n]: n \in N\}$ of equivalence classes of the constant

functions provide a copy of $\mathbb{N}$ in $\mathbb{N}^*$. The sentence $\forall x \exists y(x = y)$ defines the identity function

$$id(x) = x.$$

It is obvious that $[id] \in \mathbb{N}^*$ and $[id] \neq [n]$ for any $n \in N$, and hence $\mathbb{N}^*$ cannot be isomorphic to $\mathbb{N}$:

$$\mathbb{N}^* \ncong \mathbb{N}.$$

The class of arithmetically definable functions is countable because the language is countable. The model $\mathbb{N}^*$ is countable because its universe consists of equivalence classes of an equivalence relation $=^g$ defined on the countable family F_1.

A theory is *countably categorical* if it has exactly one countable model, up to isomorphism. Skolem's result shows that neither PA nor $Th(\mathbb{N})$ are countably categorical. Hence, the hope that $Th(\mathbb{N})$ (or even PA) has only one countable model up to isomorphism, the familiar structure $\mathbb{N}$, is shattered. However, $\mathbb{N}$ is the only computable such model.

In the introduction to his 1955 paper, Skolem refers to the Löwenheim-Skolem theorem and Skolem's paradox. For him "the notion set and particularly the notion subset in the case of infinite sets can only be asserted to exist in a relative sense." Since an axiomatization of the natural number series exists within set theory, the axiomatization of the number series also cannot be expected to be absolute. The model $\mathbb{N}^*$ constructed by Skolem is remarkable also because it can be described using computability-theoretic notions, which were developed much later than Skolem's original 1934 paper.

3.2 Computability and Models of Fragments of Arithmetic

In this section we will exploit notions of modern computability theory to build nonstandard models of theories. We will start by comparing Skolem's nonstandard model $\mathbb{N}^*$ with the following models:

(i) A structure denoted by $\mathcal{R}/_{\sim A}$, which was proposed by Feferman, Scott, and Tennenbaum (1959) as a nonstandard model of arithmetic,
(ii) A structure denoted by $\Pi_C\mathbb{N}$, called a cohesive power of $\mathbb{N}$, which was introduced by Dimitrov (2009a).

4 Cohesive Sets

The two approaches (i)–(ii) use various notions of algorithmically indecomposable sets. Skolem's Arithmetical Lemma 1 also implicitly uses this notion. We will now give relevant definitions and results from computability theory.

Definition 2

(1) An infinite set C of natural numbers is called *cohesive* if for every c.e. set W, either $W \cap C$ or $(N - W) \cap C$ is finite. That is, a set is cohesive if it is indecomposable into two infinite parts by a c.e. set.
(2) Some cohesive sets have c.e. complements called maximal sets. That is, a set M is *maximal* if M is c.e. and its complement $(N - M)$ is cohesive.

The notion of cohesiveness was introduced by Rose and Ullian (1963) in 1963, and is related to the notion of indecomposability introduced by Dekker and Myhill (1960) in 1960. Friedberg (1958), answering a question of Myhill from 1956, constructed a maximal set in 1958 without using an explicit reference to the notion of a cohesive set.

Similarly, r-cohesive and r-maximal sets are defined, where r stands for recursive, which is a synonym for computable.

Definition 3

(1) An infinite set $C \subseteq N$ is *r-cohesive* if for every computable set W, either $W \cap C$ or $(N - W) \cap C$ is finite.
(2) A set M is *r-maximal* if M is c.e. and $(N - M)$ is *r-cohesive.*

Clearly, maximal sets are r-maximal since computable sets are c.e., but the converse does not hold.

Theorem 5 *There is a function g: $N \to N$ such that its range rng(g) is a cohesive set.*

Proof Let $E_0, E_1, E_2, \ldots$ be any list of all c.e. sets.

We simultaneously define by induction an infinite sequence of sets

$$C_0 \supseteq C_1 \supseteq C_2 \supseteq \cdots$$

and a function g.

Let $C_0 = N$. Let $g(0) = 0$.

Assume that we have defined C_n and $g(n)$.

$$\text{Let } C_{n+1} = \begin{cases} C_n \cap E_n & \text{if } C_n \cap E_n \text{ is infinite,} \\ C_n \cap (N - E_n) & \text{otherwise.} \end{cases}$$

Let $g(n+1)$ be the least $c \in C_{n+1}$ such that $c > g(n)$.

Remark 2 The proof of Theorem 5 has been presented as a construction of a function g: $N \to N$ such that $rng(g)$ is a cohesive set. The original theorem of Dekker

and Myhill establishes that every infinite set has a cohesive subset. While there are countably many maximal sets, there are continuum many cohesive sets.

We can prove Skolem's Lemma 1 using the procedure outlined in the proof of Theorem 5. Using Skolem's list of arithmetical functions we define the following sets for $i, j \geq 1$ and $i \neq j$:

$$A_{ij1} = \Big\{n : f_i(n) < f_j(n)\Big\},$$
$$A_{ij2} = \Big\{n : f_i(n) = f_j(n)\Big\},$$
$$A_{ij3} = \{n : f_i(n) > f_i(n)\}.$$

While the sets A_{ijk} may not be c.e. (although they are arithmetical), we can construct a function g such that $rng(g)$ is indecomposable for the sequence A_{ijk} where $k = 1, 2, 3$ and $i, j \geq 1$ and $i \neq j$. We can call $rng(g)$ an arithmetically indecomposable set.

4.1 Feferman-Scott-Tennenbaum Model

Following Skolem's construction of a countable nonstandard model of arithmetic, Tennenbaum asked whether starting with computable functions and an r-cohesive set C, we could produce a nonstandard model of arithmetic. Below we give a description of such structure and the negative answer to Tennenbaum's question provided in (Feferman et al. 1959) by Feferman, Scott, and Tennenbaum. The structure they built is a model of only a fragment of arithmetic but is very interesting due to its constructive properties.

Let $\mathcal{R}$ be the set of all computable (recursive) functions, and let C be a fixed r-cohesive set. Feferman, Scott, and Tennenbaum constructed a structure, denoted by $\mathcal{R}/\sim_C$, consisting of equivalence classes of computable functions modulo an equivalence relation $\sim_C$ on $\mathcal{R}$. More specifically, for computable unary functions f and g let

$$f \sim_C g \text{ iff } C \subseteq^* \{n : f(n) = g(n)\}.$$

We write $X \subseteq^* Y$ if all but finitely many elements of X are also elements of Y. The domain of the structure $\mathcal{R}/\sim_C$ consists of the equivalence classes of computable functions under $\sim_C$. The arithmetic operations of addition and multiplication are defined naturally (poinwise). Feferman, Scott, and Tennenbaum proved that $\mathcal{R}/\sim_C$ is a model of only a fragment of arithmetic. They constructed a particular Π_3 sentence σ such that

$$\mathbb{N} \models \sigma \text{ but } \mathcal{R}/\sim_C \models \neg\sigma.$$

4.2 Cohesive Powers of Computable Structures

Recent developments in the area of computability-theoretic ultrapowers of structures arose from Dimitrov's study (2004, 2008, 2009b) of c.e. vector spaces over computable fields. The study of the computable infinite-dimensional vector space V_∞ has been ongoing since Metakides and Nerode (1977) introduced it in 1977. The vectors in the space V_∞ are the finitely nonzero infinite sequences of elements from a computable field F. The most natural infinite computable field is the field $\mathcal{Q}$ of rational numbers. In (Dimitrov 2008), we discovered an interesting class of nonstandard fields, later called cohesive powers of $\mathcal{Q}$, and denoted by $\Pi_C \mathcal{Q}$, where C is a cohesive set. The field $\Pi_C \mathcal{Q}$ has certain similarities with the classical ultrapower construction. In (Dimitrov 2009a), we introduced the notion of cohesive powers of computable structures in general. While cohesive powers arose from investigating lattices of c.e. vector spaces (Dimitrov and Harizanov 2016, 2017, n.d.), it became clear that $\Pi_C \mathbb{N}$ was an interesting structure in itself. We later discovered the exact connection between $\Pi_C \mathbb{N}$ and the structure $\mathcal{R}/_{\sim C}$ of Feferman, Scott, and Tennenbaum.

A partial computable function φ can be thought of as having a partial algorithm that takes elements $n \in N$ as inputs and either halts and outputs the value $\varphi(n)$, or computes forever in which case $\varphi(n)$ is undefined. We denote the halting computation by $\varphi(n)\downarrow$ and computing forever by $\varphi(n)\uparrow$. The domain of φ, the set of all inputs on which the partial algorithm halts, is a c.e. set. Two partial functions φ and ψ are equal, denoted by $\varphi \simeq \psi$, if on every input either both are undefined, or both are defined and have the same value, denoted by $\varphi(n)\downarrow = \psi(n)\downarrow$. Determining whether a partial computable function is total (hence computable) is an undecidable problem.

Definition 4 (Dimitrov 2009a) Let $\mathcal{L}$ be a finite language and $\mathcal{A}$ be a computable $\mathcal{L}$-structure with domain $A \subseteq N$. Let $C \subseteq N$ be cohesive. The *cohesive power of* $\mathcal{A}$ *over* C, denoted by $\Pi_C A$, is the $\mathcal{L}$-structure $\mathcal{B}$ defined as follows:

- Let $D = \{\varphi: (\varphi: N \to A)$ is partial computable with $C \subseteq^* dom(\varphi)\}$.
- For $\varphi, \psi \in D$, let $\varphi =_C \psi$ iff $C \subseteq^* \{i: \varphi(i)\downarrow = \psi(i)\downarrow\}$. The relation $=_C$ is an equivalence relation on D. Let $[\varphi]$ be the equivalence class of $\varphi \in D$ with respect to $=_C$.
- The domain of $\mathcal{B}$ is the set $B = \{[\varphi]: \varphi \in D\}$.
- Let R be an n-ary predicate symbol of $\mathcal{L}$. For $[\varphi_1], \ldots, [\varphi_n] \in B$, define $R^{\mathcal{B}}([\varphi_1]), \ldots, [\varphi_n]$ by

$$R^{\mathcal{B}}([\varphi_1], \ldots, [\varphi_n]) \Leftrightarrow C \subseteq^* \left\{ i : \bigwedge_{1 \le m \le n} \varphi_m(i)\downarrow \wedge R^{\mathcal{A}}(\varphi_1(i), \ldots, \varphi_n(i)) \right\}.$$

- Let F be an n-ary function symbol of $\mathcal{L}$. For $[\varphi_1], \ldots, [\varphi_n] \in B$, let ψ be the partial computable function defined by

$$\psi(i) \simeq F^{\mathcal{A}}(\varphi_1(i), \ldots, \varphi_n(i)).$$

Define

$$F^{\mathcal{B}}([\varphi_1], \ldots, [\varphi_n]) = [\psi].$$

- Let c be a constant symbol of $\mathcal{L}$. Then $c^{\mathcal{B}}$ is the equivalence class of the total computable function with constant value $c^{\mathcal{A}}$.

The cohesive power of a finite structure is isomorphic to the structure. The cohesive power of an infinite computable (hence countable) structure produces an infinite countable structure.

In (Dimitrov et al. 2014), the authors demonstrated that $\mathcal{R}/_{\sim C}$ and $\Pi_C\mathbb{N}$ are isomorphic when the complement of the set C is a maximal set. In that case for every partial computable function $\varphi \in D$ there is a total computable function f such that $\varphi =_C f$, hence $[\varphi] = [f]$.

The following theorem is a part of a more general result in (Dimitrov 2009a).

Theorem 6 (1) *If σ is a Π_2 sentence or a Σ_2 sentence, then $\Pi_C\mathcal{A} \models \sigma$ if and only if $\mathcal{A} \models \sigma$.*

(2) *If σ is a Π_3 sentence and $\Pi_C\mathcal{A} \models \sigma$, then $\mathcal{A} \models \sigma$.*

The converse of (2) does not hold. In (Dimitrov et al. n.d.), the authors produced natural counterexamples for linear orders. Using the previous theorem it cannot be deduced that $\Pi_C\mathbb{N}$ is a model of the full theory $Th(\mathbb{N})$, but just of a fragment of $Th(\mathbb{N})$. As mentioned before, the cohesive power of a computable field is a field. This is because the field axioms are of complexity at most Π_2. This is the case with many other familiar structures. In (Dimitrov et al. n.d.), we have also investigated the relationship between additional algorithmic properties of a computable structure and the isomorphism type of its cohesive power.

4.3 Concluding Remarks

In this chapter we focused on the existence of various nonstandard models. The algorithmic complexity of these models varies. Uncountable models are considered noncomputable because of their size. Even all countable nonstandard models of arithmetic are noncomputable. The elements of Skolem's countable model of arithmetic are equivalence classes of functions of arbitrary arithmetic complexity. It is not a computable model but we can measure the complexity of the elements of the model. The elements of the Feferman-Scott-Tennenbaum structure, which is only a model of a fragment of arithmetic, are equivalence classes of computable functions. Unfortunately, there is no algorithmic enumeration of all computable functions. The elements of cohesive powers of a computable structure are equivalence classes of partial computable functions. In general, the preservation of satisfaction of sentences

in cohesive powers, as well as in the Feferman-Scott-Tennenbaum structure, is limited to lower levels of arithmetical hierarchy because it depends on the existence of uniform procedures for finding witnesses of existential quantifiers. Compared to Skolem's model these structures favor a more constructivist approach at the expense of satisfiability of sentences at higher levels of the arithmetical hierarchy.

Acknowledgments The first author would like to thank F. Stephan and D. Skordev for useful discussion regarding Skolem's 1934 paper. The second author was partially supported by the Simons Foundation collaboration grant 429466 and GW Dean's Research Chair award. The authors are listed lexicographically, indicating equal contribution to the chapter.

References

Ash CJ, Knight JF (2000) Computable structures and the hyperarithmetical hierarchy. North Holland, Amsterdam

Chang CC, Keisler HJ (1990) Model theory. North-Holland, Amsterdam

Bell JL, Slomson AB (1974) Models and ultraproducts: an introduction, 3rd revised printing, North-Holland, Amsterdam

Dekker JCE, Myhill J (1960) Recursive equivalence types. University of California Publications in Mathematics, N.S., vol 3, pp 67–213

Dimitrov R, Harizanov V, Miller R, Mourad KJ (2014) Isomorphisms of non-standard fields and Ash's conjecture, 10th Conference on Computability in Europe. In: Beckmann A, Csuhaj-Varjú E, Meer K (eds) Language, life, limits, Lecture Notes in Computer Science 8493. Springer, pp 143–152

Dimitrov R, Harizanov V, Morozov A, Shafer P, Soskova AA, Vatev SV (n.d.) On cohesive powers of linear orders, submitted. https://arxiv.org/abs/2009.00340

Dimitrov R, Harizanov V (2016) Orbits of maximal vector spaces. Algebra and Logic 54:440–477

Dimitrov R, Harizanov V (2017) The lattice of computably enumerable vector spaces. In: Day A, Fellows M, Greenberg N, Khoussainov B, Melnikov A, Rosamond F (eds) Computability and complexity. Springer, pp 366–393

Dimitrov RD, Harizanov V (n.d.) Effective ultraproducts and applications, accepted for publication in Aspects of computation, Greenberg N, Ng KM, Wu G, Yang Y (eds), IMS, National University of Singapore, World Scientific

Dimitrov R (2004) Quasimaximality and principal filters isomorphism between ε^* and $\mathcal{L}^*(V_\infty)$. Arch Math Log 43:415–424

Dimitrov R (2008) A class of Σ^0_3 modular lattices embeddable as principal filters in $\mathcal{L}^*(V_\infty)$. Arch Math Log 47:111–132

Dimitrov R (2009a) Cohesive powers of computable structures, Annuaire de l'Université de Sofia "St. Kliment Ohridski." Fac Math Inf, Tome 99:193–201

Dimitrov R (2009b) Extensions of certain partial automorphisms of $\mathcal{L}^*(V_\infty)$, Annuaire de l'Université de Sofia "St. Kliment Ohridski." Fac Math Inf, Tome 99:183–191

Feferman S, Scott DS, Tennenbaum S (1959) Models of arithmetic through function rings, Notices of the American Mathematical Society 6:173–174. Abstract #556–31

Fokina E, Harizanov V, Melnikov A (2014) Computable model theory. In: Downey R (ed) Turing's Legacy: Developments from Turing Ideas in Logic. Cambridge University Press/ASL, pp 124–194

Frayne T, Morel AC, Scott DS (1962) Reduced direct products. Fundam Math 51:195–228

Friedberg RM (1958) Three theorems on recursive enumeration. I. Decomposition. II. Maximal set. III. Enumeration without duplication. Journal of Symbolic Logic 23:309–316

Hirshfeld J (1975) Models of arithmetic and recursive functions. Israel Journal of Mathematics 20: 111–126
Hirshfeld J, Wheeler WH (1975) Forcing, arithmetic, division rings. Lecture Notes in Mathematics 454. Springer, Berlin
Kaye R (1991) Models of Peano arithmetic. Oxford University Press, New York
Kossak R, Schmerl JH (2006) The structure of models of Peano arithmetic. Oxford University Press, New York
Lerman M (1970) Recursive functions modulo co-*r*-maximal sets. Transactions of the American Mathematical Society 148:429–444
Łoś J (1955) Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. Mathematical Interpretation of Formal Systems, North-Holland, Amsterdam, pp 98–113
McLaughlin TG (1989) Some extension and rearrangement theorems for Nerode semirings. Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 35:197–209
McLaughlin TG (1990) Sub-arithmetical ultrapowers: a survey. Annals of Pure and Applied Logic 49:143–191
McLaughlin TG (2007) Δ_1 ultrapowers are totally rigid. Archive for Mathematical Logic 46:379–384
Metakides G, Nerode A (1977) Recursively enumerable vector spaces. Annals of Mathematical Logic 11:147–171
Mostowski A (1952) On models of axiomatic systems. Fundam Math 39:133–158
Nelson GC (1992) Constructive ultraproducts and isomorphisms of recursively saturated ultrapowers. Notre Dame Journal of Formal Logic 33:433–441
Peano G (1889) Arithmetices Principia, Nova Methodo Exposita
Rogers H (1967) Theory of recursive functions and effective computability. McGraw-Hill Book Company, New York
Rose GF, Ullian JS (1963) Approximations of functions on the integers. Pacific Journal of Mathematics 13:693–701
Ryll-Nardzewski C (1952) The role of the axiom of induction in elementary arithmetic. Fundam Math 39:239–263
Schmerl JH (2011) Tennenbaum's theorem and recursive reducts. In: Kennedy J, Kossak R (eds) Set Theory, Arithmetic, and Foundations of Mathematics – Theorems, Philosophies. Lecture Notes in Logic 36, Association for Symbolic Logic, La Jolla, pp 112–149
Skolem Th (1934) Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen. Fundam Math 23: 150–161
Skolem Th (1955) Peano's axioms and models of arithmetic. Mathematical Interpretation of Formal Systems, North-Holland, Amsterdam, pp 1–14
Smoryński C (1991) Logical number theory I: an introduction. Springer-Verlag, Berlin
Soare RI (1987) Recursively enumerable sets and degrees. A study of computable functions and computably generated sets. Springer-Verlag, Berlin
Tennenbaum S (1959) Non-Archimedean models for arithmetic. Notices of the American Mathematical Society 6:270. Abstract #556-42

Fermat's Last Theorem

Colin McLarty

Contents

1 Introductory Overview 2012
2 Basics 2013
3 A Proof for $n = 3$ and a Valuable Mistake 2014
3.1 Proofs by Factorization 2015
3.2 Further on Factorization in Rings 2016
4 Poincaré, Mordell, and Weil 2017
4.1 Rational Roots of Cubic Polynomials 2017
4.2 The Addition Law on an Elliptic Curve 2018
4.3 Rational Points on Arithmetic Elliptic Curves 2019
5 The First Deep Geometric Result on FLT 2020
5.1 Curves as Surfaces 2021
5.2 Mordell's Conjecture, Faltings' Theorem 2022
6 Elliptic Curves Meet FLT 2023
6.1 Frey-Hellegouarch Curves 2023
6.2 Modularity 2024
7 The Ideas and Methods of Ribet's and Wiles's Proofs 2025
7.1 Ribet 2026
7.2 Wiles 2027
8 The Quondam Problem of Grothendieck Universes 2028
9 Peano Arithmetic 2028
10 The Wide-Open Question 2030
11 Conclusion 2031
References 2031

Abstract

For 300 years, Fermat's Last Theorem seemed to be pure arithmetic little connected even to other problems in arithmetic. But the last decades of the twentieth century saw the discovery of very special cubic curves, and the rise

C. McLarty (✉)
Case Western Reserve University, Cleveland, OH, USA
e-mail: colin.mclarty@case.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_44

of the huge theoretical Langlands Program. The Langlands perspective showed those curves are so special they cannot exist, and thus proved Fermat's Last Theorem. With many great contributors, the proof ended in a deep and widely applicable geometric result relating nice curves in rational coordinates to very nice surfaces in complex coordinates. This geometry is the only proof strategy for FLT yet known. Outstanding but not atypical of current mathematics, the proof challenges common philosophic views on abstraction and structure and raises still-open logical questions in arithmetic.

Keywords

Structuralism · Fermat's Last Theorem · Geometry · Peano Arithmetic

1 Introductory Overview

For most of its 350-year history, Fermat's Last Theorem (FLT) was intriguing arithmetic of no special philosophic interest. Even in arithmetic, it had little relation to other theorems. A generalized "prime factorization" misused in a fallacious proof of FLT has been a driving force in number theory to this day. That has gotten philosophers' and historian's attention as an example of mathematical concept formation, see Sect. 3.2. But that strategy did not prove FLT (yet).

The proof of FLT in 1995 had two beginnings in the late 1960s. It involves much other mathematics. And it challenges philosophic views of modern mathematics.

One beginning was a surprising specific geometric observation on FLT. The other was the vast theoretical *Langlands Program* in arithmetic algebraic geometry, see Sect. 6. These two had no connection at first except that both descended from André Weil's lifelong project to geometrize number theory. Weil sketched in his dissertation (1928), and explained at length in later works (1987, 1991), how he built that project on his youthful reading of Diophantus, Fermat, and Poincaré.

The geometric observation came when Yves Hellegouarch and Vadim Andreevich Demyanenko independently noticed that if any of the curves defined by certain cubic polynomials had certain odd geometric properties, then FLT was false. This was not the most exciting direction of inference: No one knew how to find curves with those properties, and people rather expected there were none, since people expected FLT was true. And in the other direction, a proof that no such curves exist would only rule out one kind of counterexample to FLT. It would not prove FLT.

Still, it was a striking new connection for FLT.

Poincaré (1901) related arithmetic to the geometry of cubic curves following leads in Fermat among others. Poincaré's work was developed notably by Louis Mordell and André Weil and the subject was well established by the 1970s. Intriguingly, though, Hellegouarch and Demyanenko used a less known aspect of it, namely *rational torsion points* on *elliptic curves.* Jean-Pierre Serre at the famed Séminaire Bourbaki had to say one knows very little about them (1971, p. 282). Within a few

years, Barry Mazur (1977a, 1978) changed that with deep results on torsion and on *modularity:*

> Although it was not realized at the time, the chain of ideas that was to lead to a proof of Fermats Last theorem had already been set in motion by Barry Mazur in [these papers]. (Darmon et al. 1997, p. 8)

Mazur was not aiming at FLT per se. He was interested in geometrizing arithmetic in general.

Gerhard Frey (1986), with crucial input from Serre, gave a strategy for FLT using modularity: Show any counterexample to FLT would give an elliptic curve that is *not* modular. But also prove the *Modularity Theorem* (MT), saying every relevant elliptic curve *is* modular. Number theorists had known for years that MT would be powerfully useful if it was true. Where FLT was easy to state but had few consequences, MT was hard to state but had rich consequences.

Up to this point, all the arguments were virtually pure arithmetic. They made subtle calculations using sophisticated geometric concepts in arithmetic but used no higher order logic, except some real and complex analysis so concretely calculational that it should be conservative over (first order) Peano Arithmetic (PA). The arguments were basically expressed in PA except that the authors were not interested in formalization.

Experts found Frey's strategy persuasive and Ken Ribet (1990), drawing on ideas from Serre and Mazur, proved counterexamples to FLT would produce non-modular elliptic curves. Few believed the other step, MT, was within reach. Andrew Wiles pursued it in near total secrecy for 6 years, though, and shocked the world in 1993 by proving enough of MT to give FLT as corollary (with some corrections over the next 2 years). Ribet's and Wiles' proofs too are led by directly arithmetical ideas but both invoke higher level apparatus. Within 5 years, Wiles' method of *modularity lifting* plus a lot of skilled calculating led to a proof of the full Modularity Theorem and made modularity even more useful than had been known. While the proof of MT is extremely valuable, it was FLT that inspired Wiles since age 10 and FLT made headlines around the world and Fermat's own statement of it is quoted in Latin at the start of Wiles (1995a).

2 Basics

A *Pythagorean triple* is any triple $\langle a, b, c\rangle$ of nonzero integers with

$$a^2 + b^2 = c^2.$$

Some sources require a, b, c positive. This makes little difference since $x^2 = (-x)^2$.

Euclid's *Elements* Book 10 proposition 19 shows how to generate infinitely many of these. For any distinct, nonzero integers m, n take

$$a = 2mn, b = m^2 - n^2, c = m^2 + n^2. \quad (1)$$

Routine calculation shows this gives Pythagorean triples.

Conversely, every Pythagorean triple is a multiple of one formed by Eq. 1. There are several interestingly different proofs of this: by using ordinary prime factorization, or by factorization with complex Gaussian integers $a + b\sqrt{-1}$, or by the geometry of rational points on a circle. Each of these proofs is common in textbooks and they are all in the Wikipedia article *Pythagorean triples*.

Sometime around 1630 Fermat wrote in the margin of a book that he had a "marvelous proof" that there are no nonzero integer solutions for cubes or fourth powers or any higher order.

$$a^3 + b^3 = c^3 a^4 + b^4 = c^4 a^n + b^n = c^n \, n \geq 5.$$

Essentially Fermat's own proof for fourth powers is common in textbooks today. No proof by him of the cubic case is known. He may have believed for the rest of his life that he had a proof of the cubic case, but it seems he fairly soon came to believe he had no proof for higher powers. See Edwards (1977, p. 2) and Weil (1987, p. 104).

For ease of reference, define F_n to be the Fermat equation with exponent n including the requirement that none of *a, b, c* are 0:

$$a^n + b^n = c^n \text{ with integers } abc \neq 0.$$

Notice F_n has rational solutions if and only if it has integer solutions. An integer solution is rational. And conversely any rational solution can be expressed by fractions with a common denominator, and then

$$(a/d)^n + (b/d)^n = (c/d)^n \text{ is equivalent to } a^n + b^n = c^n.$$

So it does not matter whether FLT is stated for integers or for rational numbers.

Also, if F_n has a solution, and n factors as $n = k \cdot m$, then F_m has a solution:

$$a^{k \cdot m} + b^{k \cdot m} = c^{k \cdot m} \text{ is equivalent to } \left(a^k\right)^m + \left(b^k\right)^m = \left(c^k\right)^m. \quad (2)$$

Every integer $n > 1$ has prime factors, and Fermat proved F_4 has no solutions. So if FLT is not true, there must be some odd prime p such that F_p has solutions.

3 A Proof for $n = 3$ and a Valuable Mistake

In 1770, Leonhard Euler published a brilliant but unclear proof of the cubic case, saying F_3 has no integer solutions. It centered on factoring numbers of the form $a^2 + 3b^2$ where *a, b* are integers. Some steps were not justified clearly.

Edwards (1977, pp. 43ff.) relates Euler's proof to using complex numbers of the form $a + b\sqrt{-3}$ with ordinary integers *a, b.* These numbers add and multiply, keeping this form:

$$\left(a + b\sqrt{-3}\right) + \left(c + d\sqrt{-3}\right) = (a + c) + (b + d)\sqrt{-3} \tag{3}$$

$$\left(a + b\sqrt{-3}\right) \cdot \left(c + d\sqrt{-3}\right) = (ac - 3bd) + (ad + bc)\sqrt{-3}. \tag{4}$$

This connects to Euler's argument by a special case of Eq. 4:

$$\left(a + b\sqrt{-3}\right) \cdot \left(a - b\sqrt{-3}\right) = a^2 + 3b^2.$$

Euler's unwarranted assumption could be justified in these terms if unique prime factorization generalizes from the ordinary integers $\mathbb{Z}$ to these numbers $a + b\sqrt{-3}$, with the primes existing in this new context. And it sort of does. Only it requires using terms $\left(1 + \sqrt{-3}\right)/2$ with denominator 2 as well as integer multiples of $\sqrt{-3}$. This indeed is one way to prove F_3 has no solutions.

In modern terms, the ring called $\mathbb{Z}\left[\left(1 + \sqrt{-3}\right)/2\right]$ has unique prime factorization:

$$\mathbb{Z}\left[\frac{1 + \sqrt{-3}}{2}\right] = \left\{a + b \cdot \frac{1 + \sqrt{-3}}{2} \middle| a, b \in \mathbb{Z}\right\}.$$

In this context, for example, 7 is no longer a prime since it has smaller factors

$$7 = 2^2 + 3 \cdot 1^2 = \left(2 + \sqrt{-3}\right)\left(2 - \sqrt{-3}\right).$$

In the ring $\mathbb{Z}\left[\left(1 + \sqrt{-3}\right)/2\right]$, it turns out $2 + \sqrt{-3}$ and $2 - \sqrt{-3}$ are prime.

Rings like $\mathbb{Z}\left[\left(1 + \sqrt{-3}\right)/2\right]$ are today called *rings of integers of algebraic number fields.* They do not always have unique prime factorization and that has been the topic of a great deal of number theory. See many algebraic number theory textbooks, including, for example, Stewart and Tall (2001, Chap. 3), and see Sect. 3.2 below.

3.1 Proofs by Factorization

Over the next century, Euler's proof for F_3 was perfected, others used similar arguments to prove F_5 has no solutions, and Gabriel Lamé did it for F_7. Then in 1847, Lamé unified those proofs into a purported proof of all of FLT. He expressed the idea quite clearly and so he was clearly wrong. But he was not wholly wrong. The historic events are well known and vividly described by Edwards (1977, pp. 76ff.) and Stewart and Tall (2001, pp. 185ff.).

Because people knew our Eq. 2, Lamé just had to show F_p has no solution with odd prime p. For that case, Lamé would take a *primitive p-th root of unity.* That is a complex number ζ_p such that $\zeta_p^p = 1$ while $\zeta_p \neq 1$. These always exist. They are those roots of $X^p - 1$ that are not roots of $X - 1$, so long division of polynomials shows they are exactly the solutions to

$$X^{p-1} + X^{p-2} + \cdots + X + 1 = 0. \tag{5}$$

To verify this, just multiply that polynomial by $X - 1$.

Using Eq. 5, simple calculation (try it by hand for $p = 3$) shows for any integers *a, b* there is a factorization

$$a^p + b^p = (a + b)(a + b\zeta_p)\left(a + b\zeta_p^2\right)\cdots\left(a + b\zeta_p^{p-1}\right).$$

Thus any solution $\langle a, b, c\rangle$ to F_p has a factorization

$$(a + b)(a + b\zeta_p)\left(a + b\zeta_p^2\right)\cdots\left(a + b\zeta_p^{p-1}\right) = c^p. \tag{6}$$

This is all correct. Then Lamé proved Eq. 6 is impossible by using further steps which, as his colleague Joseph Liouville quickly pointed out, amounted to an unjustified assumption of unique prime factorization for this sort of number.

Liouville knew this was a major gap. Then Liouville got a letter from Ernst Kummer showing unique factorization actually fails for $p = 23$. In fact, it fails for all primes $p \geq 23$ (see the Wikipedia article *Cyclotomic fields* and references there).

Kummer extended this kind of argument to give some results even in contexts where unique factorization failed. Further work on FLT in this line became extremely sophisticated and unavoidably technical. It is well covered in Edwards (1977). This strategy remained the main engine for proving new cases of FLT right up until the 1980s. It did not lead to the current proof of FLT.

3.2 Further on Factorization in Rings

Fermat's Last Theorem was not the first reason mathematicians studied unique prime factorization in these algebraic contexts. But it was one. And the tools created to study factorization pervade all of algebraic geometry and algebraic number theory today, naturally including the proof of FLT. So a few more words on the topic are appropriate though it is not a specific theme leading to Wiles' proof of FLT.

Mathematicians made two related responses to failure of unique prime factorization in a ring R. First was to define new *ideal divisors* for R, which are not elements of R but serve some of the same purposes as ordinary prime factorization in the integers $\mathbb{Z}$. Second was to measure "how close" R is to having unique prime factorization, which became one source of the far-reaching subject of *Class Field Theory* (Cox 1989; Kato et al. 2011). Kummer himself took early steps on both of these Mazur (1977b).

Today *ideals of rings* are bread and butter for algebraists, and often get into sophomore algebra courses. Philosophers and historians have taken this as a major example of mathematical concept formation. See, for example, Avigad (2006) and Boniface (2016). Class Field Theory is so imposing philosophers have barely studied it, and then indirectly through another of its sources the Quadratic Reciprocity Theorem. See, for example, Tappenden (2008), Yap (2011, 2020), and D'Alessandro (2020).

4 Poincaré, Mordell, and Weil

> What we possess of [Fermat's] methods for dealing with curves of genus 1 [i.e. elliptic curves] is remarkably coherent; it is still the foundation for the modern theory of such curves. It naturally falls into two parts; the first one, directly inspired by Diophantus, may conveniently be termed a method of ascent, in contrast with the descent which is rightly regarded as Fermat's own. (Weil 1987, p. 104)

Weil (1987, pp. 104–18) features the rich history of connections between arithmetic and cubic curves. See also Stillwell (1995). Whether Diophantus or Fermat themselves thought of their calculations in geometric terms is not clear. Poincaré (1901) did explicitly when he systematized Fermat's methods, related them to Kronecker's and Weierstrass's complex analysis, and expanded them to new problems.

4.1 Rational Roots of Cubic Polynomials

This computation was obvious to Fermat:

Theorem 1 *If a cubic polynomial P(X) in one variable with rational coefficients has two rational roots, or one rational root of multiplicity 2, then all its roots are rational.*

Proof. If α_1 is a rational root of $P(X)$, then $X - \alpha_1$ divides $P(X)$ and the quotient $Q(X)$ is quadratic with rational coefficients. If α_2 is a distinct rational root of $P(X)$, then $X - \alpha_2$ divides $Q(X)$. If α_1 is a double root of $P(X)$, then $X - \alpha_1$ divides $Q(X)$. Either way leaves a rational linear polynomial which necessarily has a rational root, and that is all the roots of $P(X)$. QED.

Notice this is pure arithmetic calculation, requiring no general notion of real or complex numbers let alone, say, the fundamental theorem of algebra.

Moving to two variables, and thus to the geometry of curves in the plane, does not change the idea.

Definition 1 A *rational point* in the coordinate plane $\mathbb{R}^2$ is a point with both coordinates rational numbers.

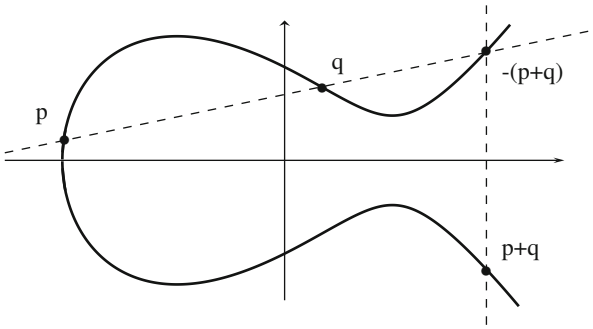


Fig. 1 The addition law on the elliptic curve defined by $Y^2 = X^3 - X + 1$

The curve in Fig. 1 is defined by a cubic equation in two variables

$$Y^2 = X^3 - X + 1. \tag{7}$$

Theorem 1 easily implies: If a straight line meets that curve in three points, and two of them are rational, then so is the third. And a line tangent to that curve at a rational point meets the curve in another rational point. To make those true even for vertical lines, we speak of a rational point on the curve at vertical infinity.

Equation 7 has six obvious rational (indeed integral) solutions: namely $\langle 0, \pm 1\rangle$, and $\langle -1, \pm 1\rangle$, and $\langle 1, \pm 1\rangle$. Take one of them, say $r = \langle 0, 1\rangle$. Routine calculation gives the tangent line to the curve at r, and by Theorem 1, it meets the curve in a new rational point, say s. Then take the tangent to the curve at s to find another rational point, and so on. Diophantus used this kind of algebra to find nontrivial rational solutions to suitable problems by starting from trivial ones.

Two questions naturally arise:

1. Is every rational solution to Eq. 7 generated this way starting from those six points? If not, does some other finite set of points generate them all?
2. Starting from a given one, say r, will this method continue generating new rational points? Or will it eventually return to r and start repeating?

The pursuit of those questions, and several other questions from number theory and from complex analysis, led to the idea of addition laws on elliptic curves.

4.2 The Addition Law on an Elliptic Curve

A wondrous geometric fact is that intersecting lines with a cubic curve produces an *addition law* so that the points of the curve form an Abelian group.

Definition 2 A *real elliptic curve* is a curve defined by an equation

$$Y^2 = X^3 + AX + B \tag{8}$$

where *A, B* are real numbers and the roots of $X^3 + AX + B$ are all distinct.

The addition law is uniquely determined by two rules:

1. The point at vertical infinity will be the 0 element.
2. For any three collinear points *p, q, r* on the curve, $p + q + r = 0$. (If the tangent to the curve at *p* also meets the curve at q, then $2p + q = 0$. And if *p* is an inflexion point then $3p = 0$.)

By these rules, if the line between points *r, s* on the curve is vertical, then

$$r + s + 0 = 0.$$

So define $-r = s$. Then look at the example in Fig. 1. The line connecting points *p, q* meets the curve in a third point and rule 2 says these three points add up to 0, so call that third point $-(p + q)$. Now the vertical line through $-(p + q)$ meets the curve in the negative of $-(p + q)$, thus $p + q$.

Simple reasoning left to the reader will verify most of the Abelian group axioms:

$$p + q = q + p, p + 0 = p, \text{for every } p \text{ there is a } q \text{ with } p + q = 0.$$

The associative law $p + (q + r) = (p + q) + r$ is not easy but it does follow by a classical geometric property of algebraic curves. See, e.g., Reid (2013, Chap. 2) or Stewart and Tall (2001, p. 229).

So each elliptic curve *E* has a group of real points $E(\mathbb{R})$ which is Abelian and contains at least the point at vertical infinity as 0 element. The same algebra gives an Abelian group $E(\mathbb{C})$ of the complex points on *E*. These geometric facts were important in early nineteenth century complex analysis, long before they were cast in group-theoretic form (McKean and Moll 1999, Chap. 2).

4.3 Rational Points on Arithmetic Elliptic Curves

Another wondrous fact is that the rationality considerations of Sect. 4.1 and the geometry and analysis of 1.4.2 fit right together to put Fermat's ascent and descent into modern geometric form.

Definition 3 An *arithmetic elliptic curve* is an elliptic curve defined by $Y^2 = X^3 + AX + B$ with rational coefficients *A, B*.

Theorem 1 shows the addition law on an arithmetic elliptic curve takes rational points to rational points, and so gives an Abelian group $E(\mathbb{Q})$ of rational points.

In this notation, the Fermat method of ascent by repeated tangent lines, described just after Eq. 7, takes a point r, then forms $s = -2r$, and then $-2\ s = 4r$, then $-8r$, $16r$, ... multiplying by -2 each time. The modern approach gives more points by repeated addition. Starting from one rational point r form the series:

$$r, r + r = 2r, 2r + r = 3r, 3r + r = 4r, 4r + r = 5r, \ldots.$$

A crucial concept is *torsion*. Point r is a torsion point if $n \cdot r = 0$ for some positive integer n, so that $(n + 1)\ r$ is just r again. Notice 0 is a torsion point trivially. And the sum of torsion points is torsion because $n \cdot r = 0$ and $m \cdot s = 0$ implies $mn \cdot (r + s) = 0$.

So, for every arithmetic elliptic curve E, the Abelian group $E(\mathbb{Q})$ of rational points has a subgroup $E_{tor}(\mathbb{Q})$ of rational torsion points. The questions at the end of Sect. 4.1 take this form:

1. Is the Abelian group $E(\mathbb{Q})$ generated by some finite set of its points?
2. Does $E(\mathbb{Q})$ have nonzero torsion points? If so, what are they, and what is the structure of $E_{tor}(\mathbb{Q})$?

Mordell proved the answer to the first is yes. Then Weil used a more transparent and systematic method extending this to a far more general result (Cassels 1973, p. 500). The question of how many elements it takes to generate a given $E(\mathbb{Q})$ drew a lot of attention, produced great results, and remains open. See the Wikipedia article *Birch and Swinnerton-Dyer conjecture*. For a recent Fields Medalist's contribution stressing elementary arithmetic aspects, see Bhargava and Shankar (2015).

While there had been notable work on rational torsion points, for example, by Beppo Levi (Schappacher and Schoof 1996), the torsion groups $E_{tor}(\mathbb{Q})$ were not well understood until Mazur's papers in the 1970s. A key consequence of Mazur (1978) shows how specific the arithmetic gets, and it will reappear in the noted "3–5 trick" in Wiles' proof (Sect. 7.2). For any point r on any arithmetic elliptic curve E, the smallest nonzero $n \in \mathbb{N}$ with $nr = 0$ in $E(\mathbb{Q})$ has to be either in the range 1–10, or else 12. Wiles (1995a, p. 542) uses the fact that it cannot be 15.

However, Mordell also built on Poincaré (1901) to pose a conjecture for other shapes of algebraic curves, which Weil could not settle.

5 The First Deep Geometric Result on FLT

> I asked myself whether many problems of indeterminate Analysis [in arithmetic] could not be connected to each other by a new classification of higher degree polynomials. . .. Poincaré (1901, p. 21)

The most obvious classification of polynomials is by degree. These two equations both have degree 3:

$$Y^2 = X^3 - X + 1 \tag{9}$$

$$Y^2 = X^3 - 3X + 2 \tag{10}$$

But Hilbert and Hurwitz (1890) already knew that in terms of number theory, Eq. 10 does not act like a cubic, it acts like a quadratic. It acts like a degree 2 equation. Hilbert, Hurwitz, and Poincaré all knew the right explanation is the geometry of the curves defined by these equations. The difference shows in plots of their real points. Compare Figs. 1 and 2. The curve defined by Eq. 10 crosses itself. Any cubic curve that crosses itself like that will behave arithmetically like a quadratic curve.

Specifically, all those mathematicians knew the important classification of polynomials for number theory is not by degree but by the *genus* of their curves. Here "curves" means spaces of complex number solutions, including points at infinity. Hilbert, Hurwitz, and Poincaré drew on Riemann's classification of what are now called Riemann surfaces (Riemann 1851; Pont 1974).

5.1 Curves as Surfaces

This idea is needed to state Mordell's conjecture, and again to describe modularity in Sect. 6.2. The conceptual point is to show these theorems involve perfectly concrete visual geometry.

Analysts talk about the *complex plane* while algebraic geometers call that same thing the *complex line.* More generally, the analyst's *Riemann surfaces* can be the geometer's *algebraic curves.* Of course, analysts and geometers are often the same people and the dual terminology has even been used as a textbook title: *Algebraic Curves and Riemann Surfaces* (Miranda 1995).

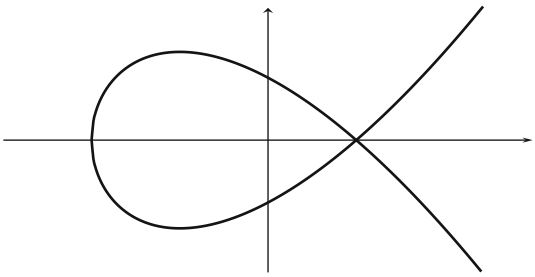


Fig. 2 The cubic curve defined by $Y^2 = X^3 - 3X + 2$

It is a choice between emphasizing topology or algebra. Topologically, the complex numbers form a plane with a real axis and imaginary axis, and small regions in a Riemann surface look like small regions in that plane. Algebraically, it is intuitive to say one equation in two variables defines a curve, and it is technically correct to say the complex numbers form a 1-dimensional complex vector space.

Visually, the curves in Figs. 1 and 2 formed by real solutions to Eqs. 9 and 10 are real cross-sections of complex curves – understanding that a complex curve is topologically a surface. Both curves go on to infinity so neither surface can be drawn whole to scale. But topologically, each curve expanding towards a point at infinity can be contracted down to put that point at a finite distance as in Fig. 3. Then the complex solutions to Eq. 9 form a torus. Those to Eq. 10 form a pinched torus where one part of the tube is pinched to a point. That pinch point is where the curve of real solutions in Fig. 2 crosses itself.

For more on genus, see numerous graphics in McKean and Moll (1999, Chap. 1) or the concise discussion in terms of algebra, topology, and differential geometry by Reid (2013, pp. 51ff.).

5.2 Mordell's Conjecture, Faltings' Theorem

In fact, quadratic curves (and self-crossing cubic curves) have genus 0, elliptic curves have genus 1, and when $n > 3$, then the curve of the Fermat equation F_n has genus greater than 1. Mordell conjectured that every algebraic curve with genus greater than 1 has at most finitely many rational points.

In 1983, Gerd Faltings proved this and much more (Mazur 1987). This was far-reaching arithmetic quite apart from FLT. It was beautiful classical arithmetic proved by insightful geometric means using the latest technology, including Grothendieck's *étale cohomology* (Faltings 1983). It was huge progress on FLT: No Fermat equation F_n with $n > 2$ has more than finitely many solutions! And several people gave explicit calculations based on this, concluding that F_n has no solutions, for the great majority of all $n > 2$ (Heath-Brown 1984). However, this proof did not identify which F_n have no solutions.

This result has not (yet) been the basis for any full proof of FLT.

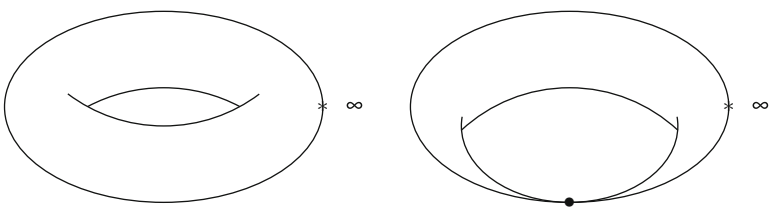


Fig. 3 The topology of a torus and a pinched torus, showing points at infinity

6 Elliptic Curves Meet FLT

Hellegouarch (1971) and Demyanenko (1971) independently used equations related to Eq. 11 below, to show that if certain kinds of torsion points can occur on elliptic curves, then FLT is false. It might naturally have occurred to one or both of them to turn this around and consider Eq. 11 for any hypothetical solution $\langle a, b, c\rangle$ to F_p for any odd prime p, but this never happened in print. Hellegouarch (2000) documents that he spoke on elliptic curves and FLT in the 1960s, but gives no details of what he said except that Serre corrected some errors.

At any rate, Hellegouarch and Demyanenko linked FLT to the arithmetic of elliptic curves. Mazur (1978) showed that Hellegouarch-Demyanenko torsion cannot occur on arithmetic elliptic curves; so these considerations would not refute FLT. But that did not prove FLT.

On the other hand, the beginnings of the Langlands Program are well documented, because in 1967, Robert Langlands wrote to Weil, to describe a huge unifying project (Mueller 2018). There has been great progress while a great deal remains to do.

> The Langlands Program is now a vast subject. There is a large community of people working on it in different fields: number theory, harmonic analysis, geometry, representation theory, mathematical physics. Although they work with very different objects, they are all observing similar phenomena. (Frenkel 2013, p. 77)

Wiles (1995a) is itself, technically, a step in the Langlands Program. But there is a persuasive argument that the overall concept of the Program mattered more in the process of proving FLT than the technical connection:

> Rather the crux of the Program is two pronged: its overall *vision* relating motives of all kinds to automorphic representations, and its *methods* which push representation theory to the forefront, and infuse the subject with a seemingly endless string of challenging problems. It is these aspects of the Langlands Program which (albeit indirectly) play a role in the proof of Fermat's Last Theorem. (Gelbart 1997, p. 191)

6.1 Frey-Hellegouarch Curves

Frey (1982) was apparently the first to associate with *every* integer solution to F_n

$$a^n + b^n = c^n$$

the elliptic curve defined by (Since Eq. 11 has nonzero term in x^2, it is not in the form of Eq. 8 used to define elliptic curves. But a linear change of variable $X' = X - ((a^n - b^n)/3)$ puts it in that form. Such issues often go without comment in the literature):

$$Y^2 = X(X - a^n)(X + b^n). \tag{11}$$

This is the Frey curve or Frey-Hellegouarch curve for the solution $\langle a, b, c\rangle$.

The goal would be to show this curve is impossible, and conclude there are no solutions to F_n. Frey suggested in a privately circulated note this curve could not be *modular,* as discussed in Sect. 6.2. A much fuller discussion was published with input from Serre (Frey 1986). See also Schappacher (1993/1994, p. 41).

This curve is better tied to F_n than appears at first. Obviously the roots of the righthand side are 0, and a^n and $-b^n$. But look at the differences between them:

$$a^n = a^n - 0b^n = 0 - (-b^n)c^n = a^n - (-b^n).$$

These differences are more intrinsic to the geometry of the curve than the roots are. A mere shift in the X-coordinate does not change the geometry. It changes the roots but does not change their differences. The most important classical invariant of a polynomial is its *discriminant,* defined by multiplying together the differences between its roots. The most obvious fact about the discriminant is that it equals 0 if and only if the polynomial has at least one multiple root.

6.1.1 Some Details of Arithmetic and Geometry

First, given $a^n + b^n = c^n$ with $abc \neq 0$, the discriminant of $X(X - a^n)(X + b^n)$ *is* $-(abc)^{2n}$ and thus nonzero. So the Frey-Hellegouarch curve is elliptic. Its plot over the real numbers $\mathbb{R}$ has no self-crossings. It cannot look like Fig. 2.

Second, if a and b are divisible by a prime number ℓ, then $a^n + b^n = c^n$ implies c is too. And any common factor of a, b, c can be divided out to give a smaller solution to F_n.

$$(k \cdot h)^n + (k \cdot i)^n = (k \cdot j)^n \text{ implies } h^n + i^n = j^n \text{ when } k \neq 0.$$

So, dividing out common factors, we can assume no prime divides both a and b. Such a solution to F_n is called *primitive* since it is not a multiple of any smaller solution.

Work on FLT requires calculating modulo various primes ℓ. So it is helpful that, for a primitive solution $\langle a, b, c\rangle$, the roots a^n and $-b^n$ of the righthand polynomial in Eq. 11 cannot both equal the root 0 modulo one prime ℓ. Intuitively, doing algebraic geometry modulo ℓ, the Frey-Hellegouarch curve looks at worst like Fig. 2. It cannot degenerate beyond one self-crossing. In the standard terminology, when $n > 3$, the Frey-Hellegouarch curve for a primitive solution to F_n is *semistable.*

6.2 Modularity

> I want to spend a few minutes considering one example (a conjecture, in fact) which shows how Number Theory can sometimes contrive to be a helpful, and possibly inspirational, goad to the rest of the Mathematical Sciences. (Mazur 1991, p. 593).

There is an obvious kinship between the geometry of the complex plane $\mathbb{C}$ and of the Euclidean plane $\mathbb{R}^2$ The absolute value of a complex number $a + b\sqrt{-1}$ equals the Euclidean norm of the vector $\langle a, b\rangle \in \mathbb{R}^2$:

$$| a + b\sqrt{-1} | = \left(a^2 + b^2\right)^{\frac{1}{2}} = \left\| \langle a, b\rangle \right\|.$$

By mid-nineteenth century, a similar link was known between the upper half of the complex plane on one hand, and the hyperbolic plane of non-Euclidean geometry on the other (McKean and Moll 1999, Sect. 9).

$$\mathbb{H} = \left\{ a + b\sqrt{-1} \in \mathbb{C} | b > 0 \right\}.$$

Essentially every account of elliptic curves describes how they can be covered by the complex numbers so that the curve gets a locally Euclidean structure. This has been central to the subject since the nineteenth century.

Mazur (1991) explains a late twentieth century discovery by Gennadii Vladimirovich Belyi: An elliptic curve can also get a locally hyperbolic structure from the upper half complex plane $\mathbb{H}$, if and only if it is defined by Eq. 8 with algebraic numbers *A, B* as coefficients (i.e., *A, B* are themselves the roots of integer polynomials).

A local Euclidean structure plus a local hyperbolic structure on one curve can have interesting interactions. But Mazur (1991) describes how much more happens when the local hyperbolic structure has a particular arithmetic form. Elliptic curves meeting that condition are called *modular* elliptic curves. Frey had suggested that Frey-Hellegouarch curves cannot be modular. That suggestion was proved by Ribet (1990).

The conjecture described in the quote of Mazur was called the conjecture of Taniyama-Shimura-Weil. Now it is the *Modularity Thesis* or *Modularity Theorem* (MT). It says every arithmetic elliptic curve is modular! Given Ribet's proof, a proof of MT would prove FLT. Indeed, by the considerations described in Sect. 6.1.1, FLT would follow just from proving every semistable elliptic curve is modular. That seemed out of reach to most experts in 1991, but Wiles (1995a) would prove it.

7 The Ideas and Methods of Ribet's and Wiles's Proofs

The proofs discussed up to here used extremely sophisticated conceptions and calculations but – with the notable exception of Faltings (1983) – did not use machinery so advanced that a graduate student in number theory would not normally know it. The success of Faltings (1983) and the progress through Frey (1986), plus the whole tendency of the Langlands Program, brought such machinery to the fore.

7.1 Ribet

Ribet (1990) proved an arithmetic conjecture by Serre. It turned out the most useful calculational handle on modular covers of elliptic curves, for this purpose, is certain technical witnesses to such covers. These witnesses are called *newforms.* Every modular cover of a given elliptic curve will be witnessed by many different newforms. Ribet showed that if a Frey-Hellegouarch curve did admit a hyperbolic structure making it modular, then any newform witnessing it could be reduced to one on a lower *level,* and then a lower one, to the point of impossibility. So Frey-Hellegouarch curves cannot have such witnesses and cannot be modular.

The essence of the argument is geometrically motivated calculation. The steps of the proof are largely calculational, albeit the calculations are often organized in terms of uncountable structures like the real, complex, and p-adic numbers, as well as the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ named in the paper's title.

As is typical of major proofs today, simply doing the calculations does not suffice to prove they work. Ribet uses considerable resources over several sections. He names just the chief ones in his summary:

> In §2, we recall material due to Raynaud [24] concerning Neron models of Jacobians. . .. In the next two §§, we recall the work of Deligne and Rapoport [4] and Cerednik-Drinfeld ([3], [7]) on the bad reduction of classical modular curves. . .. [To prove] the Main Theorem [w]e begin, in §6, with Mazur's result [20]. . .. (Ribet 1990, p. 433)

It should be obvious on its face that to bring all these resources into one comprehensible argument, the sources themselves must give concisely quotable theorems in relatively uniform terminology. This has to be done well to make proofs like Ribet's feasible to discover, and to make them reliable once given.

7.1.1 Structural Mathematics

Indeed the working methods are *structural* in that they define their objects just up to isomorphism, as philosophers have been remarking for some time. However, this has different goals and means in mathematics than philosophers usually discuss. Benacerraf (1965) depicts two young mathematicians using different "set theoretic" definitions of the natural numbers just because they were taught to. Benacerraf notes the two definitions work equally well because they produce isomorphic "structures," and proposes philosophers explore this as an issue for the ontology of mathematics. But it is also practical mathematics.

Generally, in practice, different constructions of one structure do not "work equally well." Different ones work best depending on context. And more:

1. A result proved in one context may be needed in another.
2. Efficiency may demand using a (reliably proven) result without lingering to specify any one construction or proof or context for its proof.

So mathematicians already have efficient, rigorous, routine ways to describe common structure behind different constructions.

The concept of structure that works for Ribet and his sources is the language of exact sequences and functors – in short categorical language. Some mathematicians will stress that Ribet and the others are not doing category theory. They are not lingering on categorical issues. But they use that language. That is how it is done.

7.2 Wiles

> To set the stage for the proof in (Wiles 1995a) one begins by replacing the problem on elliptic curves with a problem on Galois representations. (Wiles 1995b, p. 244)

After Ribet's proof, it remained to show all elliptic curves are modular – or at least all semistable ones – and that would prove FLT. Instead of working directly with hyperbolic covers of elliptic curves, Wiles used *Galois representations* which are algebraic constructs associated to elliptic curves. Wiles found modular elliptic curves can be recognized by their associated Galois representations, so these ones can be called *modular* representations. Then he used Mazur's theory of *deformations* of Galois representations (Mazur 1997).

Through all the technicalities, Wiles' proof has a clear conceptual outline often called an "$R = T$ strategy." In Wiles' case:

1. R captures a wide sense of "deformations" of modular representations of semistable elliptic curves, which reaches all Galois representations of those curves.
2. T captures a narrow sense of "deformations" of modular representations of semistable elliptic curves so the results of deformation remain modular.

The expression "$R = T$" is shorthand for proving the reach of T actually covers the reach of R (Saito 2013, Chap. 5). All the relevant Galois representations are reached from modular ones by deformations that preserve modularity. So all semistable elliptic curves are modular.

Of course Wiles' conceptual strategy did not erase concrete problems. Just the contrary. It pointed out certain concrete problems as decisive. The most famous example from arithmetic is the "3–5 trick," the topic of (Saito 2013, Chap. 4) and of many other publications and webpages. Wiles makes sophisticated use of a concrete fact of arithmetic on elliptic curves: the Abelian group $E(\mathbb{Q})$ of an arithmetic elliptic curve E cannot have both a point $r \neq 0$ with $3r = 0$ and a point $s \neq 0$ with $5\ s = 0$. It might have either one but cannot have both. This is not easy to prove. It was well known to Wiles since one proof of it is Mazur (1978). See Wiles (1995a, p. 542) or Saito's Proof of Proposition 4.3.

Wiles solved massive specific problems of many kinds to define the R and T he needed, and prove their relation. Wiles (1995a) gives 100 pages of detail, and that is dauntingly concise for all but leading experts. This has led to many more results on modularity and applications in number theory. And it finished the proof of FLT.

Here the numerous outside results needed, and the Grothendieckian functorial-categorical nature of the arguments, are even more prominent than in Ribet's case.

8 The Quondam Problem of Grothendieck Universes

> To avoid certain logical difficulties, we will accept the notion of a Universe, which is a set 'large enough' that the habitual operations of set theory do not go outside of it. (Grothendieck 1971, p. 146)

Whether this avoided difficulties or caused them has been a matter of controversy. Some mathematicians explicitly accept the large sets so as to use Grothendieck's large structure tools. Others reject the tools, or at least avoid using them in proofs, because of set theory. Everyone involved knew very well from the start how to give the specific number theoretic or geometric applications without these large structure tools. But the tools themselves were described, notably in Artin et al. (1972), using large cardinals which create a stronger set theory than ZFC.

The proofs bearing on FLT, including Ribet (1990) and Wiles (1995a), never mention Grothendieck universes. They have no reason to. But they refer to work in the Grothendieck school that does use them (McLarty 2010).

The set theoretic issue is moot now since even Grothendieck's largest structure tools turn out to be available in *finite order arithmetic,* far weaker than ZFC. Set theory at this level has been reinvented in many guises. It can take the form of Simple Type Theory (with arithmetic), or the Elementary Theory of the Category of Sets, or Zermelo set theory with Bounded Separation. Conservative extensions of any of these set theories can formalize Grothendieck's large structure tools by essentially the same definitions, and with the same theorems and proofs, as Grothendieck originally did (McLarty 2020).

Since this has been proved, no one who accepts anything like ZFC need feel any logical qualms about using these tools.

Intuitively, finite order arithmetic posits a set $\mathbb{N}$ of natural numbers, and a power set $P(\mathbb{N})$ containing all sets of natural numbers, and a power set $P^2(\mathbb{N})$ containing all sets of sets of natural numbers, and so on up through any finite level $P^n(\mathbb{N})$. This is a natural foundation for free and easy work in classical and basic abstract mathematics. It is too weak for some well-known purposes (Mathias 2001). But only very low levels of $P^n(\mathbb{N})$ are mentioned much in existing proofs of FLT. Likely no power sets are needed in principle, as discussed in the next section.

9 Peano Arithmetic

> $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}_p$, and $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and the Tate modules are all completions of finitist arithmetical structures. A substantial amount of their first-order theories admits natural finite approximations that do not quantify over second-order entities. Putting proofs of MT into PA will

> involve finding finite approximations of MT and of all the other principles that go into its proof. (Macintyre 2011, p. 15)

Peano Arithmetic (PA) is an axiomatic theory in first order logic, widely described in logic texts and online. On its face, PA deals with the natural numbers

$$\mathbb{N} = \{0, 1, 2, 3 \ldots\}.$$

Methods already familiar to Leopold Kronecker (1882) allow PA to interpret the integers $\mathbb{Z}$, rational numbers $\mathbb{Q}$, and even algebraic numbers $\mathbb{A}$. The algebraic numbers A are all the real or complex roots of polynomials with integer coefficients. Finite lists of numbers can be expressed in PA. So, for example, PA can interpret polynomials with integer coefficients by using their lists of coefficients.

While PA cannot talk about arbitrary real, or complex, or p-adic numbers, it can talk about and prove things about rational approximations to all these. This is crucial to discussing proofs of FLT in PA. Of course PA cannot prove as much about approximations to those sorts of numbers as even Second Order Arithmetic can, let alone stronger set theories. As a rule of thumb, algebraic calculations with these kinds of numbers can be approximated in PA. Arguments that depend on the entirety of some infinite series (or an exact limit to it) are a problem for PA.

Also crucially, PA can define some sets of numbers though these sets will not exist in PA. For example, PA can define prime numbers.

Definition 4 A natural number $n > 1$ is *prime* if and only if no natural numbers *k, m* both >1 have $n = km$.

So the set of prime numbers is called *definable in* PA. Notably, PA can state and prove that every natural number is the product of some finite list of prime numbers (unique up to reordering the list).

For any arithmetic elliptic curve E, PA can define the groups $E(\mathbb{Q})$ and $E_{tor}(\mathbb{Q})$. But when E($\mathbb{Q}$) is infinite, it will not itself exist in PA.

All of the work described above through Sect. 6 fits easily into PA in these ways (except Faltings', which is not used in the proof of FLT). The same is likely true of Ribet (1990) described in Sect. 7. But Ribet deals with the absolute Galois group $\mathrm{Gal}\left(\overline{\mathbb{Q}}/\mathbb{Q}\right)$. Neither that group, nor even its elements, can be defined in PA. Its elements can only be approximated in PA. Ribet gives a sustained argument absorbing much previous work that should be better explored in logic than it is yet.

The Macintyre quote opening this section well indicates the real challenge for proving FLT in PA: finding approximations for the Modularity Theorem. He says

> Spelling out the approximations would be a lengthy enterprise, and here it would be useful to have some metatheorems to apply. (Macintyre 2011, pp. 15 f.)

He points to a wide array of results already known for some areas of the proof, or close to areas of the proof. Number theorists often seek such approximations for

practical use quite apart from logic. They can be hard to find, and informative when found. As to logic:

> [There are] a number of distinct regions of [Wiles'] proof, almost all hitherto unvisited by proof theorists. . .. Some but certainly not all this territory is known to some model theorists. . .. Macintyre (2011, p. 8)

The project has not been completed. And even just to show that FLT is provable in PA in principle, there is no means currently known except to show at least in principle that such approximations can be given.

Then there are other important methods in current number theory, notably étale cohomology, not used in proving FLT and not much explored by logicians. These may well turn out to be easier than MT to get into PA. The language of MT is analytic geometry over the complex numbers, which can go beyond PA. Étale cohomology is closer to arithmetic in the first place.

10 The Wide-Open Question

> I want to stress that it is by no means clear to me right now how much induction is needed for [. . .] a proof of Wiles' theorem in PA. In particular, I think there is little evidence that bounded arithmetic plus the totality of exponentiation would suffice. . .. [O]ne has so little experience of more modern mathematics in the system that one should be cautious. (Macintyre 2011, p. 9)

Bounds in arithmetic are akin to approximations. Proving approximations to MT in bounded arithmetic (with any given kind of bounds, such as exponential) is akin to approximating MT in PA to begin with. There is little evidence on what bounds might work for MT because little is known of approximations to it in PA at all.

As a general fact of logic any theorem provable in PA is provable in some strictly weaker fragment of PA. This is because a proof of a theorem uses only finitely many axioms, and PA proves the consistency of all its finitely axiomatized subtheories (Hájek and Pudlák 1993, p. 168).

But it is a striking fact of proof theoretic experience that many interesting theorems are provable in the specific agreeably weak and conceptually attractive fragment that Macintyre calls bounded arithmetic with exponentiation. It is also called EFA, abbreviating two of its common names *Elementary Function Arithmetic* and *Exponential Function Arithmetic.* See Avigad (2003).

Since Wiles' proof was literally published in the *Annals of Mathematics,* it is a natural test case for Harvey Friedman's Grand Conjecture:

> Every theorem published in the *Annals of Mathematics* whose statement involves only finitary mathematical objects (i.e., what logicians call an arithmetical statement) can be proved in EFA. (Avigad 2003, p. 258)

No clear example is yet known to refute this claim (Early issues of the *Annals* had light articles on superexponential functions in arithmetic but these are not seriously to the point of Friedman's conjecture.)

Whether MT might refute Friedman's Conjecture remains to be known.

The real value of pursuing provability of MT, or FLT, in PA is likely to be not just securing the answer "yes it is provable in PA" or "no it is not provable in PA," but learning what resources from PA those theorems actually use. What fragment of induction suffices? To lift words from the quote of Macintyre above, the value would be to add some experience of more modern mathematics to proof theory.

11 Conclusion

The proof of FLT raises conceptual points listed here with the most relevant sections. Section 6.2 on modular curves, and Sect. 7 on the ideas of Ribet's and Wiles' proofs, are pivotal to the whole argument:

1. It showcases the geometric unity of current mathematics. It refutes persistent rumors that mathematics today is fragmented into specialities and that geometry was neglected in the twentieth century. Sections 4, 5, and 6.
2. It typifies current major proofs as key steps are long and rely on many prior advanced theorems. To make this work, theorems must be stated concisely and precisely, by uniform means. This is the practical reason for *structuralist* mathematics. It is not like the versions of structuralism most philosophers of mathematics consider. Section 7.
3. Far from displacing concrete arithmetic (as critics of "modern abstraction" feared), the modern techniques mobilize it. No one could find such long intricate proofs as these *without* concrete arithmetic as a guide. (The entire article.)
4. Some structural tools originally relied on large sets, though practitioners knew how to cut those tools out of the proof if desired. Now it is known these tools themselves can be rigorously founded without large sets. Section 8.
5. It is overwhelmingly likely that FLT is provable in PA but Macintyre (2011) documents points where this has not been demonstrated. Section 9.
6. The wide-open logical issue is that the leading ideas of the proof are so explicit that FLT is probably provable in some truly finitary fragment of PA. We are far from verifying what fragment. Section 10.

References

Artin M, Grothendieck A, Verdier JL (1972) Theorie des Topos et Cohomologie Etale des Schémas. Séminaire de géométrie algébrique du Bois-Marie, 4. Springer, three volumes, cited as SGA 4, New York

Avigad J (2003) Number theory and elementary arithmetic. Philos Math 11:257–284

Avigad J (2006) Methodology and metaphysics in the development of Dedekind's theory of ideals. In: Gray J, Ferreirós J (eds) The architecture of modern mathematics: essays in history and philosophy. Oxford University Press, Oxford UK, pp 159–186
Benacerraf P (1965) What numbers could not be. Philos Rev 74:47–73
Bhargava M, Shankar A (2015) Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves. Ann Math 181:191–242
Boniface J (2016) A process of generalization: Kummers creation of ideal numbers. In: Chemla K, Chorlay R, Rabouin D (eds) The Oxford handbook of generality in mathematics and the sciences. Oxford University Press, Oxford, UK, pp 483–500
Cassels JWS (1973) Louis Joel Mordell. 1888–1972. Biograph Memoirs Fellows R Soc 19:493–520
Cox D (1989) Primes of the form $x^2 + ny^2$. Wiley Interscience, New York
D'Alessandro W (2020) Proving quadratic reciprocity: explanation, disagreement, transparency and depth. Synthese:1–44. https://link.springer.com/article/10.1007/s11229-020-02591-6
Darmon H, Diamond F, Taylor R (1997) Fermat's last theorem. In: Coates J, Yau ST (eds) Elliptic curves, modular forms & Fermat's last theorem. International Press, Somerville, pp 2–140. Reprinted in several books and available on several websites
Demyanenko V (1971) The points of finite order of elliptic curves. Acta Arith 19:185–194
Edwards H (1977) Fermat's last theorem: a genetic introduction to algebraic number theory. Springer, New York
Faltings G (1983) Endlichkeitssatze für abelsche Varietäten über Zahlkörpern. Invent Math 73:349–366
Frenkel E (2013) Love and math: the heart of hidden reality. Basic Books, New York
Frey G (1982) Rationale Punkte auf Fermatkurven und getwisteten Modulkurven. J für die reine und angewandte Mathematik 331:185–191
Frey G (1986) Links between stable elliptic curves and certain Diophantine equations. No. 1 in Annales Universitatis Saraviensis. Series Mathematicae. Saarlande University, Saarbrücken
Gelbart S (1997) Three lectures on the modularity of $\overline{\rho}_{e,3}$ and the Langlands reciprocity conjecture. In: Cornell G, Silverman J, Stevens G (eds) Modular forms and Fermat's last theorem. Springer, New York, pp 155–191
Grothendieck A (1971) Revêtements Étales et Groupe Fondamental. Séminaire de géométrie algébrique du Bois-Marie, 1, Springer, New York, generally cited as SGA 1
Hájek P, Pudlák P (1993) Metamathematics of first order arithmetic. Springer, Berlin
Heath-Brown R (1984) Fermat's last theorem for "almost all" exponents. Bull Lond Math Soc 17(1):15–16
Hellegouarch Y (1971) Points d'ordre fini sur les courbes elliptiques. CRAS Paris 273:54–43
Hellegouarch Y (2000) Rectificatif àl'article de H. Darmon. Gazette des Mathématiciens 83, found at https://nitaj.users.lmno.cnrs.fr/hellegouarch.html
Hilbert D, Hurwitz A (1890) Über die diophantischen Gleichungen vom Geschlecht Null. Acta Math 14:217–224
Kato K, Kurokawa N, Saito T (2011) Number theory 2: introduction to class field theory. American Mathematical Society, Providence
Kronecker L (1882) Grundzüge einer arithmetischen theorie der algebraischen grössen. Crelle. J für die reine und angewandte Mathematik XCII:1–122
Macintyre A (2011) The impact of Gödel's incompleteness theorems on mathematics. In: Kurt Gödel and the foundations of mathematics: horizons of truth. Cambridge University Press, pp 3–25. Proceedings of Gödel Centenary, Vienna, Cambridge UK, 2006
Mathias ARD (2001) The strength of Mac Lane set theory. Ann Pure Appl Logic 110:107–234
Mazur B (1977a) Modular curves and the Eisenstein ideal. Publ Math IHES 47:133–186
Mazur B (1977b) Review of Ernst Edward Kummer, collected papers. Bull Am Math Soc 83:976–988
Mazur B (1978) Rational isogenies of prime degree. Invent Math 44:129–162

Mazur B (1987) On some of the mathematical contributions of Gerd Faltings. In: Gleason A (ed) Proceedings of the international congress of mathematicians, Berkeley, 1986, vol 1. American Mathematical Society, pp 7–12
Mazur B (1991) Number theory as gadfly. Am Math Mon 98:593–610
Mazur B (1997) Introduction to the deformation theory of Galois representations. In: Cornell G, Silverman J, Stevens S (eds) Modular forms and Fermat's last theorem. Springer, pp 243–312
McKean H, Moll V (1999) Elliptic curves: function theory, geometry, arithmetic. Cambridge University Press, Cambridge UK
McLarty C (2010) What does it take to prove Fermat's last theorem? Bull Symb Log 16:359–377
McLarty C (2020) The large structures of Grothendieck founded on finite order arithmetic. Rev Symb Log 13(2):296–325
Miranda R (1995) Algebraic Curves and Riemann Surfaces. American Mathemtical Society, Providence
Mueller J (2018) On the genesis of Robert P Langlands conjectures and his letter to André Weil. Bull Am Math Soc 55:493–528
Poincaré H (1901) Sur les propriétés arithmétiques des courbes algébriques. J des Mathematiques 7: 161–233
Pont JC (1974) Latopologie algébrique des origines à Poincaré. Presses Universitaires de France, Paris
Reid M (2013) Undergraduate algebraic geometry. Cambridge University Press, Cambridge, UK
Ribet K (1990) On modular representations of $\mathrm{Gal}\left(\overline{\mathbf{Q}}/\mathbf{Q}\right)$ arising from modular forms. Invent Math 100(2):431–476
Riemann B (1851) Grundlagen für eine allgemeine theorie der functionen einer veränderlichen complexen grösse. Inauguraldissertation, Universität Göttingen. In: Dedekind R, Weber H (eds) BernhardRiemann's Gesammelte mathematische Werke, vol 1876. B.G. Teubner, Leipzig, pp 3–45
Saito T (2013) Fermat's last theorem: basic tools. American Mathematical Society, Providence
Schappacher N (1993/1994) Les 350 ans du "Grand Theoreme de Fermat". L'Ouvert 73 and 75:1–9 and 32–44
Schappacher N, Schoof R (1996) Beppo Levi and the arithmetic of elliptic curves. Math Intell 18: 57–69
Serre JP (1971) *p*-torsion des courbes elliptiques, exposé 380. In: Séminaire Bourbaki. Secrétariat mathématique, Université Paris, Paris
Stewart I, Tall D (2001) Algebraic number theory and Fermat's last theorem. A. K. Peters Ltd, Natick, Massachusetts
Stillwell J (1995) Elliptic curves. Am Math Mon 102(9):831–836
Tappenden J (2008) Mathematical concepts and definitions. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford UK, pp 256–275
Weil A (1928) L'arithmétique sur les courbes algébriques. Acta Math 52:1–35
Weil A (1987) Number theory: an approach through history from Hammurapi to Legendre, Birkhäuser, New York
Weil A (1991) Souvenirs d'apprentissage. Birkhauser, New York
Wiles A (1995a) Modular elliptic curves and Fermat's last theorem. Ann Math 141:443–551
Wiles A (1995b) Modular forms, elliptic curves, and Fermat's last theorem. In: Proceedings of the International Congress of Mathematicians Zurich 1994, vol 1. Birkhauser, Zürich, pp 243–245
Yap A (2011) Gauss' quadratic reciprocity theorem and mathematical fruitfulness. Stud Hist Phil Sci 42(3):410–415
Yap A (2020) Noether as mathematical structuralist. In: Reck EH, Schiemer G (eds) Prehistory of mathematical structuralism. Oxford University Press, Oxford UK, pp 166–186

Section VIII

Proof

Bart Van Kerkhove and Joachim Frans

With immense sadness, we had to say a final goodbye to Joachim Frans (1989–2023), co-editor of this section. Joachim died on the morning of 7 December 2023 in a tragic traffic accident. He will be sorely missed by his colleagues and loved ones. We dedicate this section to his memory as a token of our deepest sympathy and appreciation.

Proof in the History and Philosophy of Mathematical Practice: An Introduction

Joachim Frans and Bart Van Kerkhove

Contents

1 Introduction 2038
2 The Chapters in This Section 2039
2.1 Valeria Giardino: Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research 2039
2.2 Line Edslev Andersen: The Social Epistemology of Mathematical Proof 2040
2.3 Rebecca Lea Morris: The Values of Mathematical Proofs 2040
2.4 Danielle Macbeth: Formal Proofs in Mathematical Practice 2041
2.5 Karine Chemla: The Dark Side of the History of Proof: Proving the Correctness of Algorithms in the Ancient World 2042
3 Conclusion 2042
4 Cross-References 2042
References 2043

Abstract

This introductory chapter sets the stage for an engaging exploration of the multifaceted concept of proof in the philosophy of mathematical practice. As a fundamental pillar of mathematics, proof has long been a subject of intense scrutiny for mathematicians and philosophers alike. Traditionally, proofs have been perceived as rigorous and deductive arguments, and this analysis was directed towards the notion of formal proof. However, recent developments have challenged this traditional view, highlighting the dynamic and evolving nature of mathematical proofs. In this section, we bring together diverse perspectives from leading scholars to delve into the philosophical significance and complexity of mathematical proof.

J. Frans (✉) · B. Van Kerkhove (✉)
Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium
e-mail: joachim.frans@vub.be; bart.van.kerkhove@vub.be

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_124

Keywords

Mathematical proof · Informal proof · Mathematical practice · Proof values · Diagrams · Social epistemology

1 Introduction

The fundamental aspect of proof in mathematics has long been emphasized by mathematicians and philosophers throughout history. Initially, proofs were perceived as rigorous arguments that logically derived new knowledge from accepted statements through precise step-by-step reasoning. This traditional view received significant philosophical attention, focusing on formal proof, justification, reliability, and truth.

However, upon closer examination of actual mathematical practices, a multitude of other questions and challenges emerge. These questions arise from questioning the belief that proofs are flawless deductive arguments, as informal arguments are also accepted as proofs in practice. This prompts exploration into the nature and persuasiveness of informal proofs. Additionally, understanding the process behind proofs and the reasoning patterns leading to their formation becomes a significant inquiry. The role of proofs in mathematical practice is also explored, since proofs can be valued for their elegance, beauty and various types of intellectual satisfaction.

Various researchers have paved the way toward a new attitude regarding the philosophical significance of mathematical proof. Imre Lakatos (1976), inspired by George Pólya (1945), highlights the heuristics and problem-solving strategies underpinning mathematical proofs and underscores the fallible and revisable nature of the proving process. Thomas Tymoczko (1979) encourages a reflection on the evolving nature of mathematical proof, especially in light of technological advancements and the upcoming role of computers in mathematical research. Philip Kitcher (1984) addresses the dynamic nature of mathematical discovery and the importance of looking at the historical development of mathematics. Yehuda Rav (1999) has made notable contributions to discussing the importance of proofs beyond their justificatory power and the relationship between formal and informal proofs.

While this introductory section cannot encompass the full scope of debates surrounding the role of proof in mathematical practice, it emphasizes its enduring significance in the philosophy of mathematical practice. As a pervasive theme, proof connects to various debates and approaches in the field. Attention toward questions concerning proof will always be present in each influential work in the philosophy of mathematical practice, such as Paolo Mancosu's (2008) introduction to different topics of philosophy of mathematical practice, William Aspray and Philip Kitcher (1988) and José Ferreirós and Jeremy Gray's (2006) volumes on the relationship between the history and philosophy of mathematics, and Benedikt Löwe and Thomas Müller's (2010) collection on the social dimensions of mathematics. This is not any different in this book, and the attention to mathematical proof will surpass the chapters within this section. Therefore, before presenting the contributions of this

specific section, we refer readers to other chapters that explore the multifaceted nature of mathematical proofs.

For instance, Otavio Bueno's chapter on heuristics and mathematical practice includes an insightful analysis of different types of mathematical proofs and the roles they can fulfill. Fenner Stanley Tanswell and Matthew Inglis' chapter on the language of proofs presents an empirical study on the use of imperatives in mathematical proofs, which they connect to the recipe model of proofs in their analysis. Jody Azzouni's contribution to this book offers another perspective on mathematical proofs through the algorithmic-device view. Giovanni Ferraro's chapter provides a detailed narrative of Euler's working method, shedding light on how he approached proofs. This list is not exhaustive, and we encourage every reader interested in the diverse philosophical aspects of mathematical proof to delve into this book and be captivated by its wealth of insights.

2 The Chapters in This Section

2.1 Valeria Giardino: Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research

Valeria Giardino delves into a widely discussed topic in the philosophy of mathematical practice, namely, the role of diagrams in mathematics. Diagrams are pervasive in various mathematical contexts, including recreational, educational, and research settings. Regardless of whether one deals with elementary or advanced mathematics, using diagrams to introduce concepts, convey understanding, and even prove mathematical results is ubiquitous.

This observation gives rise to numerous philosophical questions, with a central inquiry that has captivated philosophers of mathematical practice: What is the role, if any, of diagrams in mathematical proofs? Giardino highlights two distinct waves in the philosophical exploration of this question.

The first wave approaches the topic by examining and discussing specific case studies. These case studies often draw from Greek mathematics, given the prevalence of diagrams in this mathematical tradition. Through these case studies, philosophers can develop accounts of the precise role that diagrams play in mathematical proofs. However, as Giardino demonstrates, the first wave also goes beyond Greek mathematics and extends its investigations to contemporary mathematical practices, further evaluating the role of diagrams in proofs.

The second wave involves philosophical accounts that aim to provide a broader perspective on mathematics, and in doing so reevaluate our understanding of the role of diagrams in proofs. Rather than focusing solely on concrete cases, these approaches tackle questions about the nature of visual reasoning, the role of heterogeneity in mathematics, the reassessment of what constitutes a mathematical proof, and the conceptualization of the various roles that diagrams can assume.

By examining both specific case studies and overarching philosophical accounts, Giardino offers a comprehensive exploration of the role of diagrams in mathematical

proofs, shedding light on their significance in mathematical practice and prompting a reexamination of our understanding of proofs in mathematics.

2.2 Line Edslev Andersen: The Social Epistemology of Mathematical Proof

Line Edslev Andersen starts from the social epistemology of mathematics. Traditionally, mathematics has been regarded as an autonomous discipline, impervious to social and cultural influences, where results are deemed valid through objective and deductive reasoning processes. However, this viewpoint has encountered substantial challenges within the field of philosophy of mathematical practice. The central premise of social epistemology posits that mathematical knowledge is developed and refined through interactions and collaborations among mathematicians. Andersen's chapter focuses on two primary themes emerging from this endeavor.

The first theme consists of shedding light on the norms employed by the mathematical community to establish distinct standards for deductive proofs and other forms of mathematical justification. Andersen seeks to offer valuable insights into the (un)willingness to accept diverse mathematical validation methods.

The second theme pertains to the process of proof checking and the attendant issues concerning the authority and trustworthiness of mathematical knowledge. Given the impracticality of every mathematician thoroughly scrutinizing every result upon which their work relies, and the limitations imposed by the length or specialized expertise required for evaluating new proofs in their entirety, reliance on experts becomes necessary to discern what qualifies as a legitimate proof. Andersen provides an overview of various approaches that address how and when the mathematical community can accept an argument as valid proof.

While examining these two themes, Andersen presents challenging cases that underscore the fact that not all instances align with the existing framework offered by social epistemology. Consequently, the future holds intriguing and complex challenges yet to be fully addressed by the field.

2.3 Rebecca Lea Morris: The Values of Mathematical Proofs

Given the fundamental role of mathematical proofs in establishing the validity of mathematical theorems, it is unsurprising that philosophers have devoted significant attention to understanding the nature of mathematical justification. However, a closer examination of mathematical practice reveals that mathematicians are not solely concerned with proofs for their role in justification. Instances such as the practice of reproving theorems demonstrate that mathematicians attribute value to proofs beyond their justificatory function, and mathematicians frequently judge mathematical proofs using values that are not reducible to correctness.

Rebecca Lea Morris explores a range of proof values, distinguishing between practical values and abstract values. Practical value arises when proofs enhance the accessibility of mathematical knowledge, stimulate new discoveries, or foster innovative modes of thinking. While some practical values are intrinsic to all proofs, others are specific to certain proofs. Abstract values, on the other hand, extend beyond pragmatic objectives. Proofs possessing abstract value exhibit inherently valuable qualities. Morris delves into philosophical analyses of four abstract values: explanation, beauty, fit, and motivation.

In the concluding section of her chapter, Morris highlights the potential relationships between abstract and practical values, as well as the interconnections among abstract values themselves. This underscores the notion that comprehending the nature of a particular value represents merely the initial step in these philosophical inquiries, as the interplay between values opens up new and captivating avenues for research.

2.4 Danielle Macbeth: Formal Proofs in Mathematical Practice

The interest of the philosophy of mathematical practice in mathematical proofs is intricately linked to the recurring tension between fully formalized proofs and the proofs encountered in research practices. Proofs in actual research practices often exhibit gaps or exceed pure deductive reasoning and are commonly referred to as informal proofs. This tension becomes even more pronounced when considering recent advancements in formal proof checking and contemplating the relationship between informal proofs and their formal counterparts. Danielle Macbeth identifies two different positions concerning the relation between formal and informal proofs. On the one hand, she describes formalists who believe that informal proofs are in fact sketches of their formalized counterparts. On the other hand, you have the non-formalist position, which rejects the idea that formal proofs have an essential role to play in mathematical practice. This is closely related to the discussion of what these different types of proofs have to offer. There are various perspectives regarding the advantages or disadvantages of transitioning between these two types of proofs. Intuitively one might expect that formalized proofs possess persuasive strengths, while informal proofs contribute to insightful understanding.

Macbeth highlights the need for critical investigation of these perspectives and defends the view that the distinction does not simply lie in whether or not a proof is gap-free and whether or not the proof is infused with meaning. To accomplish this, she draws upon the historical contributions of Descartes, Leibniz, and Frege. By examining these figures' work on mathematical language, and emphasizing the importance of the distinction between different types of language, Macbeth argues that there is still a meaningful difference between rigorous and gap-free mathematical proofs and formal proofs, in the sense that the former can be seen as contentful and the latter not.

2.5 Karine Chemla: The Dark Side of the History of Proof: Proving the Correctness of Algorithms in the Ancient World

Karine Chemla directs our attention to the contrast between computation and proof. One possible understanding of this distinction is to view computation as a process of rule-following, where the content does not hold an essential role, while proof is connected to questioning and establishing relationships between statements. However, as Chemla's analysis reveals, this perspective does not withstand a more profound analysis.

What is particularly intriguing about Chemla's analysis is her departure from the standard narrative, often entailing an interpretation rooted in the Western history of mathematics. Instead, she challenges this view regarding proof and computation by exploring the history of verifying algorithm correctness in Chinese mathematics.

By examining the historical context of algorithmic validation in Chinese mathematics, Chemla offers a fresh perspective that challenges the traditional understanding of the relationship between proof and computation as well as to what is "proper" mathematics. In doing so, valuable insights are provided into alternative approaches of this central matter, enriching the discourse surrounding the interplay between proof and computation.

3 Conclusion

This section focuses on the philosophy of mathematical proof, which is essential in understanding mathematical practices. The following chapters explore various perspectives on the significance and intricacies of mathematical proof, providing readers with a comprehensive understanding of the subject.

Our hope is that this collection of chapters not only informs readers about the current state of research but also acts as a catalyst for future investigations. We believe that the insights and discussions presented here will serve as a solid foundation for further explorations. As the philosophy of mathematical proof continues to evolve, our goal is to inspire additional inquiry and encourage scholars to delve deeper into this captivating area of study.

4 Cross-References

- ▶ Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research
- ▶ Formal Proofs in Mathematical Practice
- ▶ Heuristics and Mathematical Practice
- ▶ Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice
- ▶ The Algorithmic-Device View of Informal Rigorous Mathematical Proof
- ▶ The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts

► The Social Epistemology of Mathematical Proof
► The Values of Mathematical Proofs

References

Aspray W, Kitcher P (eds) (1988) History and philosophy of modern mathematics. University of Minnesota Press, Minneapolis
Ferreirós J, Gray JJ (eds) (2006) The architecture of modern mathematics: essays in history and philosophy. Oxford University Press, Oxford
Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press, Oxford
Lakatos I (1976) Proofs and refutations. Cambridge University Press, Cambridge
Löwe B, Müller T (eds) (2010) PhiMSAMP: philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London
Mancosu P (ed) (2008) The philosophy of mathematical practice. Oxford University Press, Oxford
Pólya G (1945) How to solve it; a new aspect of mathematical method. Princeton University Press, Princeton
Rav Y (1999) Why do we prove theorems? Philos Math 7(1):5–4
Tymoczko T (1979) The four-color problem and its philosophical significance. J Philos 76(2):57–83

Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research

Valeria Giardino

Contents

1 Introduction 2046
2 The First Wave of Research on Diagrammatic Proofs: Looking for Case Studies 2050
2.1 The Philosophers' Darling: Greek Mathematics 2050
2.2 Contemporary Mathematics Enters the Scene 2053
3 The Second Wave: Calling for Big Pictures 2057
3.1 Pioneering Work on the Richness of the Practice of Mathematics 2057
3.2 Today's Views on Diagrammatic Proofs 2059
4 Conclusions 2064
5 Cross-References 2064
References 2065

Abstract

The objective of this chapter is to introduce some of the views that have been put forward in almost 20 years of research on diagrammatic proofs in the philosophy of mathematical practice. In Sect. 1, some contextual elements will be presented on the reasons why diagrammatic proofs have attracted so much philosophical attention in the past years. In Sect. 2, the "first wave" in the research on diagrammatic proofs based on the analysis of case studies will be described: this is how it started. In Sect. 3, the "second wave," which is calling for "big pictures" and divides into two different strategies, will be discussed: this is where it is going. In Sect. 4, some conclusions will be drawn.

Keywords

Diagrammatic proofs · Greek (Euclidean) geometry · Enhanced manipulative imagination · Fruitful representations · Informal and formal proofs

V. Giardino (✉)
Chargée de Recherche/Research Fellow (CRCN) Centre National de la Recherche Scientifique, Institut Jean Nicod, Paris, France
e-mail: valeria.giardino@ens.psl.eu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_46

1 Introduction

It is undeniable that diagrams are part and parcel of the practice of mathematics: mathematicians draw them on blackboards and occasionally on napkins at conference lunches; they use them to teach students and to show to colleagues that some relations obtain or more simply to share ideas with them; in some cases – but not as often compared to their heavy presence in more informal exchanges – diagrams end up being published in handbooks, encyclopedias, or research articles. How can the pervasiveness of diagrams in the practice of mathematics be accounted for? To clarify, this phenomenon can be investigated in two related but conceptually distinct ways.

First, very general questions can be asked. Why do diagrams seem so effective in mathematics (and beyond)? How are they used? Do they help understanding and explanation? If so, do they in all cases? What cognitive advantages do they offer? What are their cognitive costs? What are the cognitive foundations of the reasoning based on them? Are diagrams typical of Western mathematics or are they also found in other mathematical cultures? And how are they related to other forms of visualization? All these questions are of crucial interest not only for philosophy but also for other disciplines in its vicinity, such as history, sociology, anthropology, cognitive science and education.

Additionally, a second, more specifically philosophical, question can be raised about the role that diagrams might – or might not – play in mathematical *proofs*. Of course, answers to the first set of questions are relevant to discuss diagrammatic proofs; nonetheless, it is important to note that the second question is relatively independent from the first set of questions: it is indeed possible to accept that diagrams are omnipresent and useful tools in the practice of mathematics, that they are important heuristic tools used for discovering or even for proving some mathematical results, and yet to consider that they are not and cannot be parts of proofs or necessary for them.

In this chapter, I will focus on the second question, that is, on the possibility of finding proofs in the practice of mathematics that make use of diagrams. My objective will be to survey some of the views that have been put forward so far in the literature in the philosophy of mathematical practice. In order to do that, at times I will raise some of the more general questions; however, this will be done only to the extent that some of the replies can be used as evidence for the presence of diagrams in proofs. I will proceed by presenting some of the available proposals and by describing how the discussion in the philosophy and partially in the history of mathematical practice has taken shape in the last 20 years or so.

In the following, I will introduce some contextual elements that are essential to understand the reasons why diagrammatic proofs have attracted so much philosophical attention in the past years. As already said, it is undeniable that diagrams appear in some proofs; however, the nature of their contribution is controversial; that is, there is no agreement on whether diagrams are necessary for the proof to go through or whether they are just useful illustrations that can be in principle removed from the proof without epistemic loss. Before proceeding, a word of caution is necessary

regarding the meaning of the term "diagram." It is evident from the literature that there is no consensus on how to define diagrams. "Diagram" has indeed been used for a long time in a loose way, in order to refer to different objects, spanning from Euclidean diagrams (figures?), to diagrams (schemas?) in category theory. As I will show, some scholars consider two-dimensionality as a necessary feature for something to be considered as a diagram, while others focus on the manipulations that can be performed on the diagrams to be their characteristic feature, thus opening the way for linear notations to be included into the class of diagrams. In this chapter, I will keep using "diagram" in a general fashion, as a sort of umbrella term encompassing several kinds of representations, without taking any stance on the debate on how to better define them.

The relation between diagrams and proofs is as old as mathematics. Famously in his *Parallel Lives*, in the section entitled *Marcellus* (second century AD), Plutarch describes the death of Archimedes, which deeply afflicted Marcellus. In this account, Archimedes happened to be by himself, "working out some problem *with the aid of a diagram*, and having fixed his thoughts and his eyes as well upon the matter of his study, he was not aware of the incursion of the Romans or of the capture of the city" (*emphasis mine*). A soldier ordered him to go with him to see Marcello, but Archimedes refused because he wanted to first establish his proof. Hearing these words, the soldier got angry, drew his sword, and killed him. Archimedes was thus attempting to prove a result by fixing *not only his eyes but also his thoughts* upon the matter of his study – the diagram itself or what it represents? Whatever the reply to this latter question, Plutarch's story tells us that nothing at the time prevented a mathematician to use a diagram in a proof – at the exception of a Roman intruder.

As is well known, things had changed by the late nineteenth century, when the urge to make mathematics rigorous had grown stronger and stronger, to the point of pressing Pasch to claim that "the theorem is only truly demonstrated if the proof is completely independent of the figure" (see Pasch 1882). Dedekind was also very much dissatisfied with the appeal to diagrams that exploited geometric intuition in analysis – notwithstanding the fact that his manuscripts are replete of tables, doodles, and sketches (see Haffner forthcoming) – and Russell (1901) agreed that the best (mathematical) books do not contain figures. (For the details of this historical background and a general discussion of the epistemology of visual thinking in mathematics, see Giaquinto (2020).)

These attitudes contributed to the definition of what has been labeled the "received view" of mathematics (see van Bendegem 1993), according to which only reasoning that is expressed in formulas can bear epistemological weight, while the rest of the representations that are found in the practice of mathematics – for example, geometric figures – are just illustrations or at best a form of heuristics that would facilitate the grasp of the content of their linguistic counterparts, but nothing more. This received view, inherited from the nineteenth century, became the standard in the philosophy of mathematics of the twentieth century, as summed up in Tennant (1986, p. 304)'s claim: "[the diagram] has no proper place in the proof as such. For the proof is a syntactic object consisting only of sentences arranged in a finite and inspectable array." (This passage from Tennant (1986) is quoted in

Barwise and Etchemendy (1996, p. 3) as expressing the "dogma of logocentricity" that the authors want to challenge. Despite the fact that the context here is logic, the quote is still relevant for mathematics. I will briefly refer to the relation between the question about diagrammatic proofs in logic and in mathematics later in the article.) Despite the fact that diagrams can be important for understanding or explanation, they have no place in proofs.

The context changed again at the beginning of the present century (for a survey, see Giardino 2017a, and Carter 2019a): in 2005, Mancosu, Jørgensen, and Pedersen edited a volume of essays that were devoted to "new" – compared to the received view – research subjects: *Visualization, explanations and reasoning styles in mathematics* (Mancosu et al. 2005). In the introduction to this volume, the editors explain how the focus of the research in philosophy of mathematics should be on "what mathematicians are actually doing when they produce mathematics," thus posing the foundations for a philosophy of the practice of mathematics. Indeed, 3 years later, Mancosu edited a new collection of articles on *The Philosophy of the Mathematical Practice*, to show that several topics – concept-formation, understanding, heuristics, changes in style of reasoning, analogies, and diagrams – still belong in the domain of philosophy, despite the fact that philosophers had neglected them for a century or so (Mancosu 2008). The essays collected in these volumes constitute evidence for a restored interest in subjects that, according to him, were complementary to more traditional and "received" questions, for instance, the one about the access to mathematical objects. A challenge for research was to put these questions in relation with each other, for example, by exploring how mathematical concepts and objects are generated and how the process of justification works. In Mancosu's words, epistemology of mathematics could not be reduced to "a torso" (Mancosu 2008, p. 1), and it was time to return to issues like fruitfulness, evidence, visualization, diagrammatic reasoning, understanding, and explanation.

In this renewed philosophical panorama, studies on diagrams ended up occupying a preeminent position. The aim of this chapter is to trace back the literature about the role of diagrams in proofs starting from the publication of Mancosu's 2008 volume, that is, 15 years ago. However, given this long period of time, it will be impossible to cover all of the details – other people, including myself, have done it already (see for example Giardino 2017b). What I will try to do is rather to propose a narrative of how this research started and where it is going today. Given the distinction I made between two different sets of questions concerning diagrams, I will focus on the second one and therefore on a selection of studies devoted to giving an account of cases in which diagrams are part and parcel of mathematical proofs. In order to do that, it is necessary to specify which features they possess and how they are intended to be used and furthermore which relations they entertain with other elements of the practice of mathematics.

On this last point, let us recall an important methodological issue. To say that a proof is "diagrammatic" does not mean that it is "without words," as three celebrated volumes edited by Nelsen seem to suggest (Nelsen 1997, 2001, 2015). The proofs making use of diagrams that scholars are interested in are *always* proofs that also include some form of *text*, be it the labels accompanying the diagrams or the

formulas obtained by considering some of the diagram's features. The point about diagrammatic proofs is precisely to understand how diagrams and text – two crucial elements in the practice of mathematics – are related and how the contribution of the diagram to the information presented in the text can be (or not) essential for the proof to go through. In epistemological jargon, the interesting question is about how a mathematical belief that is obtained by looking at the spatial features of the diagram and by manipulating it can be raised to the status of knowledge in the absence of independent textual grounds, as well as be related to other pieces of knowledge expressed in the text.

The narrative that I propose is the following. In Sect. 2, I will quickly describe the "first wave" in the research on diagrammatic proofs. As I mentioned above, at the beginning of the 2000s, old topics were brought back to the front by raising questions that had been almost forgotten. This is how it started. Since then, many articles have been published that present case studies of diagrammatic proofs, with Greek mathematics taking center stage; as we will see, we have to wait until the beginning of the 2010s for case studies in contemporary mathematics. In Sect. 3, I will describe what I call the "second wave" in the research on diagrammatic proofs. This is where it is going. Now that the richness and the variety of diagrammatic proofs have been demonstrated, it is time to draw big pictures, which is what some scholars are presently trying to do. In order to introduce this second wave, I will first focus on pioneering work on the richness of mathematics and then on two strategies that are emerging today, aiming to specify what is meant by "proofs" (first strategy), and by "diagrammatic" (second strategy), in "diagrammatic proofs." In Sect. 4, I will briefly sum up the topics of the chapter.

Two last methodological notes. First, nothing of what I will write about in this chapter will refer to research on diagrammatic proofs *in logic*, despite the fact that some of the questions and resources in the literature are common to the two disciplines. The reason for this is that I consider that the practice of logic has its own specificities and thus deserves a separate treatment. (There are several studies that are relevant in this respect, for example, Shin (1995), Allwein and Barwise (1996), or Stenning (2002). More recently, Dutilh-Novaes (2012) discussed the role of formal language and notation in logic.) However, logic will be evoked at times in discussing rigor in contemporary mathematics. Second, I will mainly focus on studies in the *philosophy* of mathematical practice. I am aware that enormous work has been done in history – and I would say also archeology – of mathematics on diagrammatic proofs from the past, not only in the West but also in places such as old China and Mesopotamia (see for reference Chemla 2005, and Høyrup 2005). Many of the chapters of this handbook are devoted to the history of the practice of mathematics and will tackle the subject of diagrammatic proofs among others; therefore, for pragmatic reasons, I will not go into that research here. However, integration between the historical and philosophical points of view on diagrammatic proofs is desirable: as testified by the first of Mancosu's collections in 2005, historians and mathematicians started working together on these new-old topics. Since then, the two disciplines took their own research paths; it is time to foster again collaboration between the two to give an account of the practice of mathematics.

(Another subject I will not touch upon in this section is computational approaches to diagrammatic proofs. Some relevant references can be found in Giardino (2017b).)

2 The First Wave of Research on Diagrammatic Proofs: Looking for Case Studies

2.1 The Philosophers' Darling: Greek Mathematics

The Greek diagram is to some extent the paradigm of diagrammatic reasoning in mathematics, which explains why it has been a privileged target of research for philosophy for a long time. There is no need to be a philosopher or a mathematician to be familiar with Euclidean geometry, and in Euclidean geometry, the recourse to diagrams is natural and spontaneous: starting from a very early age, these geometrical figures become part of our visual repertoire. From this, can we take their presence and effectiveness in proofs for granted?

As Netz (1999) pointed out, diagrams are interdisciplinary research objects, because they are of interest both for the science of cognition that is looking at knowledge from the point of view of its practices and for history of science that studies cultural artifacts. For this reason, diagrams are the perfect subjects for what he calls a "cognitive history" and can be approached from many different points of view. I agree with this and I consider philosophy as the approach targeting more specifically their role in proofs. In his book on deduction in Greek geometry – going also beyond Euclid's *Elements* – Netz defends the idea that the "lettered diagram" is an effective geometric *tool* because of the richness of the different levels at which it can be accessed, which is what makes it appropriate to promote and justify deduction and, as a consequence, to demonstrate mathematical propositions.

The lettered diagram can in fact be considered from several points of view: logical, cognitive, semiotic, and historical. At all of these levels, it combines two elements that would normally be considered to be in contrast to each another, if one follows the standard opposition between image and language: lines and letters at the logical level, visual features and models at the cognitive level, icons and indices at the semiotic level, and a kind of know-how together with a very sophisticated reflexivity at the historical one. According to Netz, a proof in Greek geometry is based on the lettered diagram and, as a consequence, is an *event* that occurs on a papyrus or in a given oral communication. Despite its singularity, it is *felt* to be valid: the materiality of the proof is important as well as its immateriality, since the proof is considered as an *invariant* under the variability of the single action of drawing one diagram on the papyrus or of presenting the particular proof orally. The proof can be *repeated*, which is what matters for it to be recognized as a proof; contrary to what we are used to think today, the focus is not on the possibility of generalizing the result. For Netz, in order to give an account of how the lettered diagram makes the proof go through, it is necessary to consider its cognitive appeal. The diagram is in fact a static object that is nonetheless related to the text via the letters, which makes it *kinesthetic*: a diagram gets constructed and can be manipulated.

In a very renowned paper, which has been circulating since the 1990s but was published only in 2008, Manders (1995/2008) investigates Euclidean geometry specifically. He reminds us that up to the nineteenth century no one would have denied that the practice of Euclidean geometry was *rigorous*; rather, such a practice was considered to be the *most* rigorous practice among the various human ways of accessing knowledge. In Manders' view, such a practice is based on a division of labor between two artifact types – the diagram and the text sequence – that must be considered together. As humans, we have cognitive limitations, and therefore we cannot control the interpretation of a diagram so as to avoid any case that is alternative to the correct one; for this reason, the text is introduced. In the practice, diagram and text share the responsibility of allowing the practitioners to respond to physical artifacts in a "stable and stably shared fashion" (p. 83). In this framework, proofs have a verbal and a graphical part – the *diagram* and the *discursive text*: the diagram makes the very objects of traditional geometry appear, while the text ascribes particular features to the diagram, the so-called diagram *attributions*. Letters are introduced so as to facilitate cross-references between the text and the diagram. In this reconstruction of the Euclidean practice, Manders introduces a distinction between *co-exact* and *exact* features of a diagram that became very influential. (Mander's work was influential to the point of inspiring cognitive research. See Hamami and Mumma (2013) and the more recent Hamami et al. (2021).) A co-exact feature is a feature that can be directly attributed to the diagram, having certain perceptual cues that are fairly stable across a range of variations; such a feature cannot be readily eliminated, because of what Manders calls the diagram *discipline*, that is, the proper exercise of skill in producing diagrams that is required by the practice. An exact feature is instead in general affected by deformation, and it is licensed by the discursive text: for example, the fact that in a diagrams two segments have to be considered as – seen as – equal. Note that in both Netz' and Manders' proposals, the diagram is part and parcel of the proof *because* of the combination of diagram and text and thanks to its *dynamic* features: invariances or co-exact features – not varying under deformations.

Macbeth (2012, 2014) proposed another account for the use of diagrams in Euclidean proofs that emphasizes their nature of "signs." She considers Euclid's *Elements* as a system of *natural deduction*, where common notions, postulates, and definitions are not intended as premises, but as *rules* or *principles* according to which to reason. Along these lines, she suggests that in order to understand and use a diagram, one has to recognize the *intention* that is behind its construction. Moreover, in her view, a diagram is fruitful because it is an *icon*; that is, it "resembles" what it signifies. The Euclidean diagram has to be considered at different levels: first, it has primitive parts – points, lines, angles, and areas; then, thanks to the concepts of geometry, one gets the elements that someone doing geometry will be interested in – for example, an equilateral triangle is "seen as" composed of three equal lines; finally, one has to consider the diagrams as a *whole*, which shows at a glance the elements in the other two levels. In this sense, a Euclidean diagram is not different from other kinds of representations such as the famous Wittgensteinian duck-rabbit.

Consider, for example, the diagram of an equilateral triangle obtained from a segment AB by drawing two circles with centers A and B, respectively, intersecting in C. Draw the two segments from A to C and from B to C – this is Euclid I, 1: the "same" segments, AB, AC, and BC, are "seen as" the sides of the triangle *as well as* the radii of the same circle (see Fig. 1). Thanks to the combination of all these different elements in the same icon, one reasons *in* the diagram and obtains the proof. (Actually, there is a difference between the duck-rabbit and the triangle in Euclid, since the second is not an instable percept; on the contrary, someone who is competent in geometry will be able to see the two aspects and acknowledge that they are not incompatible. For a discussion of this point, see Giardino and Wöpking (2019).)

Another interesting notion to apply to Euclidean proofs was proposed by Azzouni (2005). In his view, traditional mathematical proofs imply the appeal to a kind of know-how and tacit knowledge involving "inference packages" applied to the proof as "black box units." The elements encapsulated by an inference package are in principle not visible by introspection as a list of assumptions, for the reason that these assumptions have been "knit together" in the representations of the objects. In traditional proofs – compared to derivations in algorithmic systems – it is not the case that each step follows from some sentences appearing earlier in the proof: traditional proofs work "semantically" rather than "syntactically"; that is, mathematicians recognize properties of the objects their inferences are *about* by operating with the available representations. The details of the inferences provided by inference packages are thus not only left out because everyone is familiar with them but because they are not introspectively accessible to practitioners. In this semantic view of proofs, representations – for example, some particular diagram – have precisely the role of allowing for a form of intuition to apply the inference package that comes with them, and for this reason, they should be at the core of the philosophical investigation (Azzouni 2005, pp. 23–24). (To learn more about Azzouni's view, see ► "The Algorithmic-Device View of Informal Rigorous Mathematical Proof," this volume.)

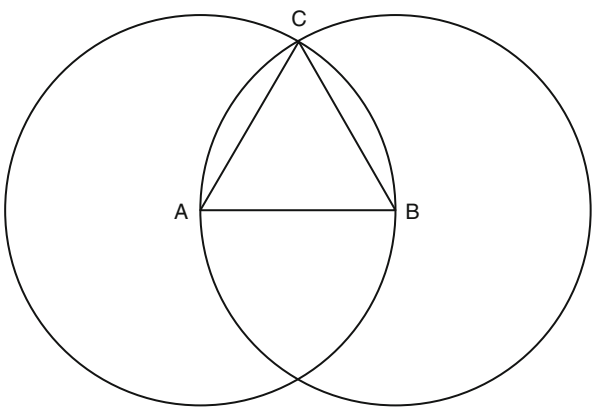


Fig. 1 *Euclid, Elements*, Book I, Proposition 1. *To construct an equilateral triangle on a given finite straight line*

I have discussed elsewhere these views and others focusing on the case of Greek diagrams and diagrammatic proofs (Giardino 2017b), and therefore I will not go here into further details. In the following, I will simply sum up some of the important points that have been proposed for the case of Euclidean (or Greek) diagrammatic proofs. It is crucial to note that all these studies show that in order to give an account of diagrammatic proofs, it is necessary to go beyond the insistence on the alleged opposition between a "visual" and a "linguistic" part of mathematical proofs, for two reasons. First, diagrams have to be considered together with the text, and therefore the focus of the research should be on the way in which the two relate. Second, diagrams can be interpreted as *dynamic*: for example, co-exact properties are not subject to modifications – which suggests that the diagram has to be manipulated in order to "test" our correct interpretation of it; the text clarifies what the intentions in the construction of the diagrams are in relation to the properties that may vary under the same manipulations. Such manipulations allow for particular inferences that are not accessible through the text only. Seeing the same configuration "differently," by identifying several levels that have to be considered separately but also as a whole, may be considered as a very basic way of making the diagram "dynamic" without modifying it. Moreover, the proof is an "event" that can be repeated, once the diagrams and the text are correctly put together.

In the following subsection, we will look at more recent case studies taken from contemporary mathematics and see how they relate to the features that emerged in the analyses of Greek mathematics.

2.2 Contemporary Mathematics Enters the Scene

If Greek diagrams have been at the center of philosophical discussion for a long time, case studies of diagrammatic proofs taken from contemporary mathematics have started becoming of interest to philosophers in most recent years, approximately after 2010. Why so late? Of course, a strong background in the practice of mathematics is necessary to get the gist of the examples taken from contemporary mathematics. Only recently, philosophers who were previously trained in mathematics started tackling specific questions about diagrammatic proof and collaborations between philosophers and mathematicians have developed – a good thing, given the interdisciplinary nature of the topic. It should also be noted that the support of interdisciplinarity in education today is encouraging more than before – when the training was more rigidly disciplinary – the creation of a new generation of interdisciplinary scholars. I have already presented elswhere some of these more recent case studies in detail (Giardino 2017b); in the rest of this section, I will sum up some of these results.

Carter (2010) was one of the first scholars to introduce case studies from today's mathematics – as well as one of the first who are now trying to think in terms of more comprehensive pictures, as we will see below. In her contributions, she focused on the role that diagrams have in a proof to define its *strategy*. She analyzes a case from free probability theory, where diagrams are introduced in informal discussions

among mathematicians to represent particular mathematical entities, for example, permutations (see Fig. 2). Carter argues that diagrams make it easier to evaluate relevant properties, e.g., whether the permutation is crossing or not or whether neighboring pairs can be canceled. Despite the fact that they are *not* published in the final article, it can be shown that they still suggest definitions and proof strategies, thus inspiring at least a part of the formal version of the proofs. Diagrams here function as a *framework* for the proof and help breaking it down into parts that are more manageable than the whole proof. We mentioned above the importance of the relationship between diagrams and text in Greek diagrams: here, some properties of the diagrams *correspond* to formal definitions, but *experiments* are performed on the diagrams. (On experiments and diagrams, see ▶ "Mathematical Experiments on Paper and Computer" in the chapter by Schlimm and Fernández González, this volume.) However, the relations used in the proof based on the diagrams refer to relations that *also* hold in the algebraic setting, which allows going back from the diagram to the algebraic setting. Moreover, some features of the diagrams – for example, crossings and neighboring pairs – are indeed co-exact properties, since they do not change under a series of variations. Carter's inspiration is C. S. Peirce's philosophy, according to which all mathematical reasoning is *diagrammatic*: in her case studies, diagrams are considered as "icons" in relation to the algebraic setting. However, one recognizes in them "indices" – letters – as an element that captures attention and, as Peirce writes, tells the user to look "There!" (*CP* 3.361).

Another case study Carter focuses on is directed graphs (Carter 2018). Also here diagrams are icons, that is, mediating objects, which link – metaphorically – two types of mathematical objects. This is possible because they constitute a "faithful" representation, which is able to put together the metaphorical aspects of using a mathematical diagram with the control that it is necessary to exercise on its manipulation. In fact, manipulations that are performed on the diagram respect the relations between the concepts that they represent, and this leads to the definition of new

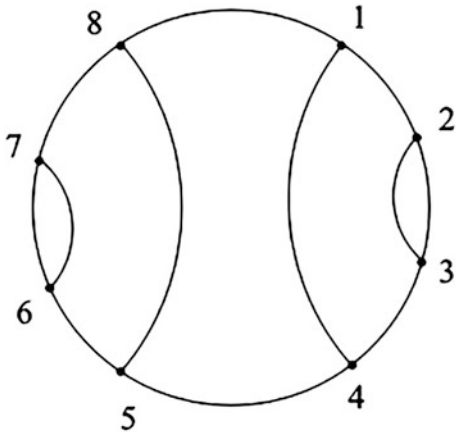


Fig. 2 An example of a diagram of a permutation, in Carter (2010)

relations between parts of those concepts that were not appearing at the beginning. (In a more recent paper, Carter (2021) reviews some of her previous examples by referring to the notion of "free ride" as proposed by Shimojima (2015). In the paper, she shows that "free rides" are not completely "for free": in most cases, some interpretation is required to "read off" some properties from a diagram, and therefore a free ride is not always a free sight.)

In Carter's view, diagrams are fruitful frameworks in parts of proofs because they trigger imagination. She refers to the notion of "manipulative imagination" as introduced by De Toffoli and Giardino (2014) in the context of an analysis of the role of knot diagrams in topology (see Fig. 3). In our reconstruction, experts in knot theory have developed a form of imagination that allows them to manipulate these diagrams – by redrawing them or by figuring out in their mind which kinds of actions can be performed on them. (In previous work, I defined diagrams as "props" – by referring to Walton's theory of imagination – for some mathematical object: working with them would correspond to test ways in which the mathematical objects they allude to might "behave" (Giardino 2018a). On the notion of manipulative imagination, see also Giardino (2018b).) Manipulative imagination is very effective in low-dimensional topology, where diagrams are used to explore the properties of mathematical objects and eventually to prove theorems about them, for example, in Rolfsen's demonstration of the equivalence of two presentations of the Poincaré homology sphere (see De Toffoli and Giardino 2015). In our view, the actual practice of proving in low-dimensional topology involves a kind of reasoning that requires intuition: the manipulations that are allowed for and on the representations as well as the representations themselves are *epistemologically relevant*, because they are integral parts both of the reasoning and of the justification provided by the proofs.

In a more recent paper, De Toffoli (2017) describes commutative diagrams and the "diagram-chasing" technique in homological algebra, where a new feature of diagrams in proofs emerges. Diagrams indeed can correspond to an abstract counterpart of geographic maps. (For case studies in category theory, see Halimi (2012)

Fig. 3 An example of a knot diagram. By applying the appropriate manipulations, one finds out that this is a knot diagram of the unknot

and Weber (2013).) As she explains, the *Five Lemma* and the *Fundamental Theorem of Homological Algebra* are both based on chasing an element in different places *in* the diagram, which in the second case is potentially infinite (see Fig. 4). The chase aims to identify paths in the diagram by exploiting its commutativity in order to verify some algebraic properties. The basic idea is to imagine moving an element around the diagram by following the relations that are there displayed (perhaps the analogs of the "co-exact" properties of Euclidean diagrams), according to some definitions (the "exact" attributions of the text). De Toffoli explains that in the case of homological algebra, commutative diagrams play the two roles already identified by Carter, that is, they inspire the proof strategy and they offer a framework that allows getting to it. However, they also have two additional roles: first, they display algebraic content in a way that is meaningful and *mathematically tractable*; second, they allow for *calculations* that can be integral parts of the proof. In De Toffoli (2023), she argues that commutative diagrams (as well as certain topological diagrams) not only can be used within rigorous proofs but are even essential to the proofs in which they figure. This would mean that non-diagrammatic conversions of such proofs would plausibly be classified not as different presentations of the same proof, but as different proofs.

To sum up, a look at case studies from contemporary mathematics shows that diagrams happen to be used effectively in proofs, for the reason that they can make new properties or relations appear. Of course, this is done in reference to some mathematical context. However, there are some caveats. First, not all diagrams are used specifically in proofs: some diagrams support intuition in the context of discovery. (For an interesting case on Caley graphs and their use for discovery, see Starikova (2010).) What emerges clearly from the literature is in fact the variety of diagrams: diagrams to display permutations have very different features compared to, say, commutative diagrams. This said, a common feature seems to be their way of triggering some form of intuition or imagination – whether they can be ultimately part and parcel of a proof depends on the control that may be applied to them by fixing their interpretation. To use an expression from Larvor (2012), within a certain diagrammatic practice, some actions on the diagrams are recognized by the experts as "permissible" or not. There is something more to note – to which I will return when I will discuss where we are now in the research on diagrammatic proofs: cases in contemporary mathematics show that diagrams and notations can be brought together. In her paper on homological algebra, for example, De Toffoli explicitly argues for diagrams as a subset of notations and gives three criteria to evaluate them as such: *expressiveness*, *calculability*, and *transparency*. For Carter as well, a piece

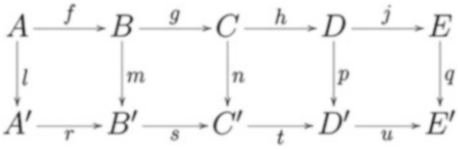


Fig. 4 The commutative diagram used in the proof of the *Five Lemma*

of notation such as $2\ k + 1$ for some natural number k is also a case of a faithful iconic representation of the concept of odd number. I will return to this issue below.

3 The Second Wave: Calling for Big Pictures

3.1 Pioneering Work on the Richness of the Practice of Mathematics

In the narrative that I propose in this chapter about these past 20 years or so of studies on diagrammatic proofs, I maintain that the urge for providing a big picture about their role in proofs has appeared only very recently. However, I will focus here on two scholars who were undoubtedly the first to try to offer a comprehensive account of the richness of the mathematical practice related to diagrammatic proofs. (Another important early publication on diagrams in mathematics is Brown (1999), where several cases of diagrammatic reasoning are discussed. Brown's view has strong Platonic assumptions: his "bold conjecture," as the author famously put it, is that (some) pictures *"are not really pictures, but rather are windows to Plato's heaven"* (p. 40, emphasis in the original).)

Giaquinto (2007/2011) was maybe the first to consider *visual* thought in mathematics in an interdisciplinary fashion, by relying on some explanations coming from cognitive science, which was on the rise at the time. The general aim of his book was to give an account of the complexity of mathematical thinking. To do so, he explored and presented some results in cognitive science about our capacity for visualizing in an attempt to integrate them with other more theoretical reconstructions of high-level processing of information. The book is interestingly devoted to discovery in mathematics, which, according to Giaquinto and in line with the revival for the attention to the practice of mathematics, is a subject that had been neglected by the most recent philosophy of mathematics as not legitimately philosophical. However, there are parts of the book where Giaquinto discusses the possible presence of diagrams in proofs. In his view, there is no need to give a uniform story about visual thinking in mathematics, which comes in a variety of ways: depending on the specific mathematical context, visual operations can be different, and consequently it may be possible to consider differently their epistemic value.

One point that seems important in the present context is Giaquinto's focus on *aspect shifting* in mathematics. Consider, for example, a possible proof of the Pythagorean theorem. In order to understand that it is valid, here is one possibility (see Fig. 5). First, one reasons geometrically and arranges a small square and four right triangles in a bigger square, so that the sum of the two sides of the triangle correspond to the side of the bigger square and their hypotenuses form the smaller square. Then, one reasons algebraically and calculates the area of the smaller square plus the areas of the four triangles and equates it to the area of the bigger square. Some more calculations and the claim is established. This procedure is accessible because we shift from one aspect, what we "see" in the diagram, to the other, what the rules on how to obtain the area of a square and of a triangle tell us. If this is the

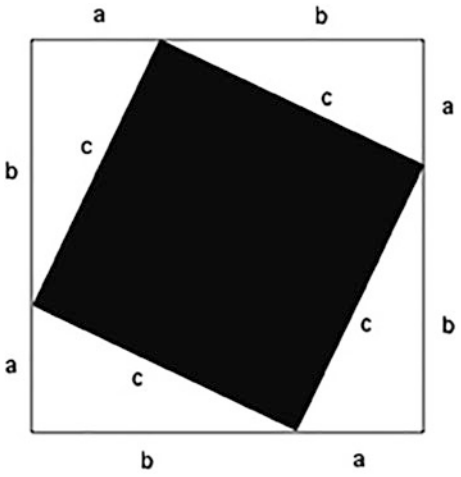


Fig. 5 One of the many so-called "visual proofs" of the Pythagorean theorem

case, would this reasoning be primary geometric or primary algebraic? In the last chapter of the book, Giaquinto suggests that neither of these two categories would be fully appropriate to capture it. Is this a general feature of mathematical reasoning?

Another interesting early work was Grosholz (2007, 2011), which discussed at depth the *heterogeneity* of mathematical reasoning as something that has been overlooked by what she identifies as the Russell-Carnap tradition – the received view? Also in this case, the practice of mathematics went back to the center of the philosophical investigation. In Grosholz' reconstruction, the Russell-Carnap tradition represents an epistemological ideal whose objective is to unify mathematics, or mathematics and the sciences, in one formalized theory. With this objective in the background, this tradition demands that logic is the homogeneous idiom constituting the only vehicle for deductive inference. Macbeth's view that was mentioned in the previous section on Euclidean diagrams recalled Grosholz' notion of "productive ambiguity" of diagrams, which is precisely what allows us to see different things in the same configuration – the sides of an equilateral triangle and the radii of the same circles – and to put the two together in order for the proof to go through. However, productive ambiguity and heterogeneous reasoning go beyond geometry: one of Grosholz' examples is Gödel's procedure in his incompleteness theorem. In her reconstruction, by his use of numbering as a strategic bridge between formalized arithmetic qua logical system and arithmetic, Gödel exploits a "carefully controlled and fruitful ambiguity" (Grosholz 2011, p. 421), which multiplies the information that is available to a mathematician. As she puts it, "it is hard for a logician to admit that he is *trafficking in heterogeneity*" (ibid., p. 425, *emphasis mine*). As we see, the richness and the variety of the representations involved in the reasoning process were extended to other parts of mathematics and even to logic. This opened up a new research horizon that will constitute a heritage for the views that I will present in the next subsection.

3.2 Today's Views on Diagrammatic Proofs

In Sect. 2, I summarized some of the case studies of diagrammatic proofs that were subject to philosophical scrutiny, starting from Greek mathematics and growing into cases from contemporary mathematics. This variety called for some kind of comprehensive framework that could account for them. However, how is it possible to talk in terms of a "big picture" for a phenomenon that is so diverse and that also depends – as Giaquinto suggests – on the specificity of the mathematical context considered in turn?

Let us focus on the term "diagrammatic proofs" and the way in which it is commonly intended. On the one hand, "proofs" is here meant narrowly, mostly in the sense of the received view: the question about diagrammatic proofs is whether forms of reasoning that make use of diagrams reach the standards of what is commonly considered as a proof – to the point that "diagrammatic proofs" may sound to some as an oxymoron. On the other hand, "diagrammatic" is interpreted as having a very wide scope so as to include many case studies and thus has the advantage of accounting for their richness. However, the fact that the term is used very loosely has also its disadvantages. What exactly is a diagram? Are geometrical figures diagrams? Are they analogous to a matrix? What is the relationship between diagrams and notations? Are they one and the same thing?

In my view, the strategies to give a comprehensive account of diagrammatic proofs can be divided into two groups. First, one might argue for a notion of proof that moves away from definitions *à la* Tennant and easily includes diagrammatic proofs as perfectly acceptable. Second, one might try to put forward some criteria for different categories of diagrams in mathematics, thus providing a taxonomy and more importantly drawing a line between diagrams that can be reliably used in a proof and others that play the role of useful illustrations or good heuristics but do not possess the necessary features to be considered part and parcel of a proof.

In the following sections, I will briefly present some of the accounts that correspond to these two strategies, which would give in my view a fair picture of where the research on diagrammatic proofs is going today. (Given that the studies I will refer to are newly published, I might happen not to know yet about other studies that could easily be included in this survey.)

3.2.1 First Strategy: What Is a "Proof"?

The first strategy that has recently emerged in the literature is the focus on mathematical proofs in the practice. Such an investigation does not target diagrams in particular, but aims to move away from the narrow definition of proof typical of the received view. I will present here some frameworks that would allow for diagrams to have a constitutive role in proofs: I will start from the recipe model of proofs introduced in Tanswell (forthcoming) and then briefly discuss recent work on informal proofs. It is important to note that the topic of informal proof in particular would demand a chapter of its own, and for reasons of space, I will not be able to go into the details; my discussion will be limited to the aspects of this debate that are relevant for diagrammatic proofs.

In Tanswell (forthcoming)'s "recipe model," proofs are to be intended as analogous to cooking recipes since, as cooking recipes, they give instructions on how to carry out actions – in the case of a proof, "mathematical" actions – in order to obtain the desired result. Of course, this does not conform with the received view, according to which the proof is a deductive syntactic argument that moves from some axioms to a conclusion by following some pre-established inference rules. In Tanswell's view, despite the success and fruitfulness of such a definition of proof, it is time to provide a new definition that would be more faithful to the practice of contemporary mathematics: written proofs are written records of how the theorem is obtained, and their main objective is to explain to the reader how to actually do the reasoning that is required in order for the proof of the theorem to go through. As in the case of recipes, proof has a communicative intention and there might be different styles of presenting the same proof.

The evidence that Tanswell brings about to argue for this analogy is first of all linguistic: the texts of proofs include instructions in the imperative tense (see "Imperatives in Mathematics" in the chapter by Tanswell and Inglis, this volume). However, such a view of proof allows including diagrams and figures interacting with the text as alternative means of giving instructions. The opposition between assertions on the one hand and pictures on the other is ill-posed: both of them can convey instructions, and for both of them, some competence for interpretation is required. However, in Tanswell's view, pictures give instructions more directly than text, as is the case for LEGO games or IKEA furniture: there is no reason to require that a picture corresponds to a particular sequence of assertions, since what matters are the actions that it guides the reader through (Tanswell 2017). It is important to note that such a view stresses the role of *action* in proving, which is in line with what we have seen so far in interpretations of diagrammatic reasoning as based on conceiving the legitimate operations and manipulations they make accessible. Moreover, such a conception is also related to wider debates; in the following, I will focus on three of them.

First, as Tanswell himself explains, the focus on action is inspired by other work on informal proofs, in particular Azzouni (2004) and Larvor (2012). The topic of informal proofs is very much discussed in the literature and clearly goes beyond the scope of this chapter. However, some of the elements of this debate are relevant for proofs using diagrams. I will refer to Burgess and De Toffoli (2022), who claim that the standard view, according to which a mathematical argument is a rigorous proof if and only if it can in principle be converted into a formal derivation, is a very general claim – a family of views rather than a single view – that can be specified in many different ways. However, the two authors themselves assume different positions in relation to the role of diagrams in proofs. According to Burgess (2015), there is a correspondence between the formalized version of the steps of an informal proof and the components of the formal proof. Formalization is thus a process of *filling in* the details, and the analysis of the role of diagrams is thus excluded from this account of rigor. However, De Toffoli (2021) argues that this is difficult to accept if one looks at the practice of mathematics: a definition of rigor such as this one would exclude a considerable number of proofs that mathematicians believe to be rigorous, which is not a desideratum if one moves from within a perspective that is interested in the

practice. (As I will explain in the next section, the way to go according to De Toffoli is to focus on what kinds of diagrams one refers to when asking questions about their being part and parcel of rigorous proofs.) As a consequence, diagrams, they claim, are relevant to the debate. This is because if the connection between a proof and one of its formal counterparts is simply the one of filling in, then diagrams cannot be part of rigorous proofs. However, the relation between informal and formal proofs might be more complicated and the existence of informal proofs still consistent with some versions of the standard view. (On the standard view, see also Hamami (2022).)

Second, the recipe model of proofs can be seen as an element of a larger picture where the focus is on the *activity* of proof (Hamami and Morris 2021). If a mathematical proof is to be intended as a sequence of deductive steps, then there is an isomorphism between such steps and the sequence of deductive *inferences* constituting the proof activity. The analogy here is between proof activity and other more ordinary activities such as travelling, where actions need to be planned in advance in order to be able to move from the current location to the desired one. (For a survey on the recent studies about agency in mathematical practice, see ▶ "Agency in Mathematical Practice" in the chapter by Hamami, this volume.)

Third, alternative models of proofs might have important consequences in other related research domains such as mathematics education. If students are taught the received view according to which proofs are sequences of assertions, they will miss important elements of the practice; for example, they will not be aware of the fact that there are imperatives in the written reports of a proof and that these are what would guide them through the steps of the proof (see Weber and Tanswell 2022). Why should students be taught definitions of proofs that do not correspond to what is really going on in the practice of mathematics and in the very practice of learning? More in general, the results of a philosophical approach to the practice of mathematics are opening up a potentially very fruitful space for discussion with education (see for reference Hamami and Morris 2020). Giaquinto, in his pioneering work, was one of the first to recognize the potential of such a collaboration and to foster it.

Once again, these topics go beyond diagrammatic proofs, and therefore for the details, I would encourage the reader to refer to the articles in the bibliography. However, from the narrative that I offer in the present chapter, it emerges that the questions that were raised in the context of diagrammatic proofs have at least in part contributed to the definition of these new research topics.

3.2.2 Second Strategy: What Is a "Diagram"?

The second strategy to formulate a big picture concerning diagrammatic proofs is based on a focus on the variety of diagrams and the different roles that they might play in the practice of mathematics, some of them being relevant for proofs.

One way to go is to "zoom out" and to consider diagrams as "representational cognitive tools," that is, as a kind of multi-recruiting systems constituting an "interface" on which information coming from perception or action, already at play in ordinary pragmatic contexts, and other more cognitive resources such as conceptual knowledge gets integrated (Giardino 2018b). However, this more cognitive approach will not suffice to respond to the specific question about the possibility of including

diagrams as part and parcel of proofs. In order to do that, one useful notion would be the one of "material anchor," as introduced in cognitive anthropology (Hutchins 2005). Mathematics would offer physical structures – diagrams and other representations – that result from a cultural process of "crystallization" of conceptual models into external and material objects. Such objects, when correctly interpreted, can be manipulated meaningfully; in some particular contexts, physical manipulations can correspond to mathematical operations so as to reduce memory and processing loads. In some cases, the particular physical structure would have an influence on the way in which the mathematical content is understood or conceptualized. Some other notions coming from cognitive science can in this respect be applied to such tools, so as, for example, the notion of *affordance*: diagrams afford some epistemic actions on them, given their physical features and more importantly their correct interpretation – the consideration of physical features, if not controlled, can lead us astray. As Rabouin (2018) has pointed out, the notion of material anchor might also be relevant to account for Netz' or Manders' studies of Greek mathematics, which consider how external resources are manipulated in order to operate on imagined structures. In his view, the historical evolution of mathematics has brought to a progressive "internalization" of these structures: representational tools such as figures still remained important, but new symbols were introduced in order to better control these operations on the imagined mathematical structures.

Another possibility to work on a big picture is to go in the opposite direction and "zoom in," by considering again the specifically mathematical case studies that have been produced so far in order to give a more precise characterization of what can be recognized in them as a diagram.

In a recent paper, Carter (2019b) goes back to her notion of *fruitful diagrams* to put together the several examples of diagrammatic proofs that were discussed in the literature. She refers to Macbeth's work on diagrammatic reasoning and her claim that diagrams and, more in general, other forms of notations are fruitful because they make mathematical objects "pop out," thanks to the way in which the representation is constructed. However, if this were the case, what would the peculiarity of diagrams be compared to other forms of representation? Carter's proposal is to define the appropriate class of objects to think in, which is the one of "faithful representations." Her background is once again Peirce's semiotics, according to which icons – a type of sign – divide into (i) *images*, resembling what they represent, (ii) *diagrams*, displaying relations, and (iii) *metaphors*, depending on conventional rules or structural similarities. Faithful representations are defined as icons, that is, as images, diagrams, or metaphors in Peirce's sense; moreover, manipulations made on them would respect manipulations made on "the object of study," which may be properties, relations, concepts, or more simply objects – we are not far from the idea that they are material anchors. These representations are fruitful precisely because manipulations performed on them can make new relations appear. It is important to note that the border between diagrams and notations in general gets blurred; Carter is aware of this and argues for two peculiarities of diagrams. First, diagrams show

that a relation holds between objects by joining their respective representations by a line; in some cases, the nature of the relation is given by the figure. Second, thanks to two-dimensionality, diagrams can be used to show that an object is related to multiple objects, as shown in Macbeth's example.

In some of her recent work, De Toffoli (2022) tries to give a more precise definition to the vague notion of diagrams that has been considered in the philosophy of mathematical practice. The taxonomy that she proposes is flexible enough to allow for borderline cases but has the advantage of offering a tool that is useful to navigate the literature by pointing to the relevant differences. In her reconstruction, and in line with her previous work, the starting point is notational systems in general, some of which are geometric-topological (GT), that is, they possess geometric-topological features that are relevant to enable the use of a form of intuition. GT figures can be diagrams or illustrations, but only the first are used in a systematic way, so as to allow them to eventually be part and parcel of a proof. Moreover, GT diagrams – it is the case of the knot diagrams briefly mentioned above – trigger enhanced manipulative imagination. However, other more precise distinctions have to be introduced. In fact, it is only thanks to further elements that the class of mathematical diagrams can be defined: mathematical diagrams are indeed systematic notational items that are either GT or two-dimensional or both.

There are two important points to note: first, the diagrams that constitute a challenge for the received view of proof – and the ones that we have mostly discussed so far in this chapter – are specifically GT. In fact, non-GT diagrams are generally accepted in proofs – this is, for example, the case of commutative diagrams. However, it is only after the middle of the twentieth century that non-GT mathematical diagrams entered the scene. (It would be interesting to discuss the relation between the process of internalization as mentioned above by Rabouin and the return to diagrams in a very different formal and formally controlled context.) To sum up, according to De Toffoli's big picture, notational systems can be distinguished across three orthogonal dimensions: (i) systematic vs. nonsystematic use, (ii) geometric-topological vs. non-geometric-topological, and (iii) two-dimensional vs. non-two-dimensional. Borderline cases are accepted. We now have a map to situate the different case studies, also in relation to the questions about which of these dimensions are decisive to allow for diagrams to be part and parcel of diagrammatic proofs.

This second strategy also opens up new research territories. On the one hand, diagrams can be considered as externalizations for thought, and therefore they can be inquired from the point of view of cognitive science – as Netz and Giaquinto had already been suggesting. Would that be some piece of evidence in favor of an "extension" of the mind? On the other hand, diagrams can be intended as signs and so they have to be studied in relation to other kinds of notations; as a consequence – as already proposed by scholars like Carter – diagrammatic proofs are a subject also for semiotics. (This latter research direction is developed in the section of this handbook on the "Semiology of Mathematics." See, for example, ▶ "Signs as a Theme in the Philosophy of Mathematical Practice" in the chapter by

Waszek, this volume that, as other papers in the same section, partially overlaps with some of the topics of the present chapter.)

4 Conclusions

In this chapter, I tried to build a narrative about how the views on diagrammatic proofs have developed in the philosophy of mathematical practice, starting from the first pioneering works that appeared at the beginning of the 2000s and getting to the most recent publications. I have shown that the interest in diagrammatic proofs was promoted by renewed attention to topics that were excluded from the philosophical research by the so-called received view of mathematics. The question about the role of diagrams in proofs is specifically philosophical despite the fact that it relates to other more general questions about the effectiveness of diagrams and their omnipresence in mathematics.

The research on diagrammatic proofs is characterized by two waves. The first wave focused on case studies, in particular from Greek geometry and more recently from contemporary mathematics. The second wave, to which we assist today, is instead aimed to build comprehensive frameworks on how to account for diagrammatic proofs. In my view, this wave is characterized by two strategies. First, it is possible to stretch the notion of proof as inherited from the received view and to consider alternative models for it, which would allow including diagrammatic representations in proofs. Second, it is desirable to focus on the term "diagram" and on its cognitive features or to distinguish several taxa for mathematical representations. For both strategies, it is evident that the research on diagrammatic reasoning has to widen its scope, either by relating to broader contexts – such as the reflection on rigor and formal and informal proofs in mathematics – or by opening up to new research domains, such as the consideration of the role of cognitive tools for thought and the philosophical study of notation in mathematics, which are both crucial issues for education and cognitive science. This sums up how the interest in diagrammatic proofs started – almost 20 years ago – and where it is going today.

5 Cross-References

- ▶ Agency in Mathematical Practice
- ▶ Mathematical Experiments on Paper and Computer
- ▶ Signs as a Theme in the Philosophy of Mathematical Practice
- ▶ The Algorithmic-Device View of Informal Rigorous Mathematical Proof

Acknowledgments I thank Joachim Frans, Bart van Kerkhove, and Bharath Sriraman for their editorial work and Silvia De Toffoli, Dirk Schlimm, and Fenner Tanswell for having commented on a previous version of this chapter.

References

Allwein G, Barwise J (eds) (1996) Logical reasoning with diagrams. Oxford University Press, Oxford

Azzouni J (2004) The derivation-indicator view of mathematical practice. Philosophia Mathematica (III) 12:81–105

Azzouni J (2005) Is there still a sense in which mathematics can have foundations? Essays on the Foundations of Mathematics and Logic 1:9–48

Barwise J, Etchemendy J, (1996) Heterogeneous logic. In Gerard Allwein & Jon Barwise (eds.) (1996), pp. 3–25

Brown JR (1999) Philosophy of mathematics: an introduction to the world of proofs and pictures. Routledge, New York

Burgess JP (2015) Rigor and structure. Oxford University Press

Burgess JP, De Toffoli S (2022) What is mathematical rigor? Aphex 25:1–17

Carter J (2010) Diagrams and proofs in analysis. Int Stud Philos Sci 24(1):1–14

Carter J (2018) Graph-algebras – faithful representations and mediating objects in mathematics. Endeavour

Carter J (2019a) Philosophy of mathematical practice – motivations, themes and prospects. Philos Math:1–32

Carter J (2019b) Exploring the fruitfulness of diagrams in mathematics. Synthese

Carter J (2021) "Free rides" in Mathematics. Synthese 199:10475–10498

Chemla K (2005) The interplay between proof and algorithm in 3rd century China: the operation as prescription of computation and the operation as argument. In Mancosu et al. (eds) (2005), pp. 123–145

De Toffoli S (2017) 'Chasing' the diagram – the use of visualizations in algebraic reasoning. The Review of Symbolic Logic 10(1):158–186

De Toffoli S (2021) Reconciling rigor and intuition. Erkenntnis 86:1783–1802

De Toffoli S (2022) What are mathematical diagrams? Synthese 200(86):1–29

De Toffoli S (2023) Who's afraid of mathematical diagrams? Philosophers' Imprint 23(1):1–20

De Toffoli S, Giardino V (2015) An inquiry into the practice of proving in low-dimensional topology. Boston Studies in the Philosophy and History of Science 308:315–336

De Toffoli S, Giardino V (2014) Forms and roles of diagrams in knot theory. Erkenntnis 79(3): 829–842

Dutilh Novaes C (2012) Formal languages in logic: a philosophical and cognitive analysis. Cambridge University Press

Giaquinto M (2007/2011) Visual thinking in mathematics: an epistemological study. Oxford University Press, Oxford

Giaquinto M (Spring 2020 Edition) The epistemology of visual thinking in mathematics. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. https://plato.stanford.edu/archives/spr2020/entries/epistemology-visual-thinking

Giardino V (2017a) The practical turn in philosophy of mathematics: a portrait of a young discipline. Phenomenology and Mind 12:18–28

Giardino V (2017b) Diagrammatic reasoning in mathematics. In: Magnani L, Bertolotti T (eds) Spinger handbook of model-based science, pp 499–522

Giardino V (2018a) Tools for thought: the case of mathematics. Endeavour 42:172–179

Giardino V (2018b) Manipulative imagination: how to move things around in mathematics. Theoria 33(2):345–360

Giardino V, Wöpking J (2019) Aspect seeing and mathematical representations. *Avant* X(2/2019)

Grosholz E (2007) Representation and productive ambiguity in mathematics and the sciences. Oxford University Press, Oxford

Grosholz E (2011) Logic, mathematics, heterogeneity. In: Ippoliti E, Cellucci C, Grosholz E (eds) Logic and knowledge. Cambridge Scholars Press, pp 291–300

Haffner E (forthcoming) Duality as a guiding light in the genesis of Dedekind's Dualgruppen. In: Krömer R, Haffner E (eds) Duality in 19th and 20th century mathematical thinking. Birkhäuser, Basel
Halimi B (2012) Diagrams as sketches. Synthese 186:387–409
Hamami Y (2022) Mathematical rigor and proof. The Review of Symbolic Logic 15(2):409–449
Hamami Y, Morris RL (2020) Philosophy of mathematical practice: a primer for mathematics educators. ZDM 52:1113–1126
Hamami Y, Morris RL (2021) Plans and planning in mathematical proofs. The Review of Symbolic Logic 14(4):1030–1065
Hamami Y, Mumma J (2013) Prolegomena to a cognitive investigation of Euclidean diagrammatic reasoning. Journal of Logic Language and Information 22(4):421–448
Hamami Y, Mumma J, Amalric M (2021) Counterexample search in diagram-based geometric reasoning. Cogn Sci 45(4):1–25
Høyrup J (2005) Tertium non Datur: on reasoning styles in early mathematics. In: Mancosu et al. (eds) (2005), pp. 91–121
Hutchins E (2005) Material anchors for conceptual blends. J Pragmat 37:1555–1577
Larvor B (2012) How to think about informal proofs. Synthese 187:715–730
Macbeth D (2012) Seeing how it goes: paper-and-pencil reasoning in mathematical practice. Philos Math 20(1):58–85
Macbeth D (2014) Realising reason. A narrative of truth & knowing. Oxford University Press, Oxford
Mancosu P (ed) (2008) The philosophy of mathematical practice. Oxford University Press, Oxford
Mancosu P, Jørgensen KF, Pedersen SA (eds) (2005) Visualization, explanation and reasoning styles in mathematics. Springer, Berlin
Manders K (1995/2008) The Euclidean diagram. In: Mancosu P (ed), pp. 80–133
Nelsen R (1997) Proofs without words: exercises in visual thinking, classroom resource materials. The Mathematical Association of America, Washington 1997
Nelsen R (2001) Proofs without words II: more exercises in visual thinking, classroom resource materials. The Mathematical Association of America, Washington
Nelsen R (2015) *Proofs without words III: further exercises in visual thinking*, classroom resource materials. The Mathematical Association of America, Washington
Netz R (1999) *The shaping of deduction in Greek mathematics: a study of cognitive history*. Cambridge University Press, Cambridge
Pasch M (1882) Vorlesungen über neuere Geometrie. Springer, Berlin. 1926, 1976 (with introduction by Max Dehn)
Rabouin D (2018) Logic of imagination. Echoes of Cartesian epistemology in contemporary philosophy of mathematics and beyond. Synthese 195(4):4751–4783
Russell B (1901) Recent work on the principles of mathematics. International Monthly 4:83–101. Reprinted as "mathematics and the metaphysicians" in mysticism and logic, George Allen and Unwin, London 1918
Shimojima A (2015) Semantic properties of diagrams and their cognitive potentials. CSLI Publications, Stanford
Shin S-J (1995) *The logical status of diagrams*. Cambridge University Press, Cambridge, NY/Oakleigh
Starikova I (2010) Why do mathematicians need different ways of presenting mathematical objects? The case of Cayley graphs. *Topoi* 29:41–51
Stenning K (2002) Seeing reason: image and language in learning to think. Oxford University Press, Oxford
Tanswell F (forthcoming) Go forth and multiply: on actions, instructions and imperatives in mathematical proofs. In: Bueno O, Brown J (eds) Essays on the work of Jody Azzouni
Tanswell F (2017) Playing with LEGO and proving theorems. In: Cook RT, Bacharach S (eds) LEGO and philosophy: constructing reality brick by brick. Wiley Blackwell, Oxford, pp 217–226

Tennant D (1986) The withering away of formal semantics? Mind and Language 1:302–318
van Bendegem JP (1993) Real-life mathematics versus ideal mathematics: the ugly truth. In: Krabbe ECW, Dalitz RJ, Smit PA (eds) Empirical logic and public debate, essays in honour of Else M. Barth. Rowman & Littlefield, Lanham, pp 263–272
Weber K, Tanswell F (2022) Instructions and recipes in mathematical proofs. Educ Stud Math 111: 73–87
Weber Z (2013) Figures, formulae, and functors. In: Moktefi A, Shin S-J (eds) Visual reasoning with diagrams, ed. Studies in universal logic. Springer, Basel, pp 153–170

The Social Epistemology of Mathematical Proof

Line Edslev Andersen

Contents

1 Introduction 2070
2 The Community Value of Justification as Deductive Mathematical Proof 2070
3 Challenging Cases I 2073
4 Proof Checking as a Social Process 2073
5 Challenging Cases II 2076
6 Conclusion 2077
7 Cross-References 2077
References 2078

Abstract

If we want to understand why mathematical knowledge is extraordinarily reliable, we need to consider both the nature of mathematical arguments and mathematical practice as a social practice. Mathematical knowledge is extraordinarily reliable because arguments in mathematics take the form of deductive mathematical ***proofs***. Deductive mathematical proofs are surveyable in the sense that they can be checked step by step by different experts, and a purported proof is only accepted as a proof by the mathematical community once a number of experts have checked the proof. Hence, the reliability of the body of mathematical knowledge is in part obtained through the surveyability of proofs and the ***social*** process of proof validation. This chapter reviews work relating to (1) the norm in the mathematical community of only counting as argument deductive mathematical proof, and (2) the norm of only counting as deductive mathematical proof an argument that has been confirmed to be a deductive mathematical proof by a number of experts. The chapter also presents cases of unusual mathematical proofs or arguments that challenge these norms.

L. E. Andersen (✉)
Aarhus University, Aarhus, Denmark
e-mail: line.edslev.andersen@gmail.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_33

Keywords

Social epistemology of mathematical proof · Norms in mathematics · Surveyability of proofs · Proof validation · Ghost theorems · Complex proofs

1 Introduction

The aim of this chapter is to provide an overview of work relating to what we may term the social epistemology of proof. Not much work has been done on the social epistemology of proof, but much work has been done that is highly relevant to the social epistemology of proof. The purpose of this chapter is to review this literature. The literature belongs to the philosophy of mathematical practice, and I hope to create connections between parts of the literature in the philosophy of mathematical practice that have not been connected before. Furthermore, I hope to promote social epistemology of mathematics as an interesting area of research.

If we want to understand why mathematical knowledge is extraordinarily reliable, we need to consider both the nature of mathematical arguments and mathematical practice as a social practice. Mathematical knowledge is extraordinarily reliable because arguments in mathematics take the form of deductive mathematical ***proofs***. Deductive mathematical proofs are surveyable in the sense that they can be checked step by step by different experts, and a purported proof is only accepted as a proof by the mathematical community once a number of experts have checked the proof. Hence, the reliability of the body of mathematical knowledge is in part obtained through the surveyability of proofs and the ***social*** process of proof validation.

This chapter has two main parts. The first part of the chapter reviews work relating to the norm in the mathematical community of only counting as argument deductive mathematical proof. The second part of the chapter reviews work relating to the norm of only counting as deductive mathematical proof an argument that has been confirmed to be a deductive mathematical proof by a number of experts. The chapter also presents cases of unusual mathematical proofs or arguments that either challenge the norm of only counting as argument deductive mathematical proof or the norm of only counting as deductive mathematical proof an argument that has been confirmed to be a proof by different experts.

2 The Community Value of Justification as Deductive Mathematical Proof

Many mathematicians have emphasized the major role of ***inductive*** reasoning in mathematics (see Borwein and Bailey 2003; Epstein and Levy 1995; Sørensen 2016). Mathematical results are typically discovered and tested in reliable ways before they are proved. To give an extreme example, the Riemann hypothesis has been verified in billions of instances. But there is really only one method of justification in mathematics, and this is deductive mathematical proof. Mathematical

results are established by proofs. And the Riemann hypothesis will remain a hypothesis until a deductive mathematical proof has been produced.

One way to identify the value of deductive proof is to ask why mathematicians are unwilling to accept other kinds of mathematical justification. We may wonder why the mathematical community is unwilling to accept proofs that make use of inductive evidence, even when the inductive evidence is extremely reliable, while the mathematical community does accept proofs that skip steps and are very long and complicated. I call the former randomized proofs in line with Jackson (2009). A proof that uses the Miller–Rabin primality test to establish that a specific number n is prime would be an example of a randomized proof that makes use of extremely reliable inductive evidence.

Philosophers have defended different views on whether the unwillingness of the mathematical community to accept randomized proofs is epistemically rational. Don Fallis (1997, 2000, 2002) argues that mathematicians' unwillingness to accept randomized proofs cannot be defended on epistemic grounds.

In response to Fallis, Kenny Easwaran (2009) offers an account of why mathematicians are unwilling to accept proofs that make use of inductive evidence. Easwaran claims that a randomized proof, as opposed to a traditional proof, is not transferable. By "transferable" he means that the author can give the proof to other mathematicians who can then verify that the claim established by the proof is true just by considering the steps in the proof. They do not have to rely on the testimony of the author of the proof. In the case of a proof that uses the Miller–Rabin primality test to establish that a specific number n is prime, by contrast, the reader would have to rely on the author to have performed the Miller–Rabin test correctly. Easwaran suggests that this explains why mathematicians are unwilling to accept randomized proofs. Without going into much detail, Easwaran (2009, 361) states that requiring proofs to be transferable is a way for the mathematical community to protect the quality of the body of mathematical knowledge. In this way, all the evidence for the knowledge is publicly available and can be independently checked by mathematicians with the relevant area of expertise.

Note that the fact that mathematical proofs can be independently checked by mathematicians separates mathematical proofs from scientific arguments. Scientific arguments cannot be independently checked by scientists. Scientists may have read and become convinced by a paper arguing that p from empirical data, but this does not exclude the possibility that the author of the paper has lied about having established p or has made a subtle mistake when collecting the data. Scientists can limit their dependence on the author by reproducing the experiments, but this is often not an option in practice. Reproducing experiments is expensive and time-consuming and generally does not give much credit.

Fallis (2011) agrees with Easwaran that the epistemic autonomy provided by transferable proofs can be epistemically valuable. However, Fallis argues that randomized proofs also allow mathematicians to be epistemically autonomous. Given the proof from the example above, mathematicians can just perform the Miller–Rabin test for themselves to check that n is indeed prime. It is unclear to Fallis, as it is to me, why it is important that mathematicians who independently

perform the Miller–Rabin test to establish that n is prime will not have the exact same evidence (2011, 172; Jackson 2009 gives a similar argument).

While Fallis and Easwaran disagree on what kinds of proof allow individual mathematicians to be epistemically autonomous, both Fallis and Easwaran emphasize that some level of autonomy of individual mathematicians is valuable to the mathematical ***community*** as a whole. Fallis (2011, 166–172) examines how some level of epistemic autonomy of individual mathematicians can be valuable to both the mathematical community as a whole and the individual mathematicians. It can be valuable to the mathematical community as a whole, he notes, because the community is more reliable when each proof is checked by more than one expert. Similarly, Easwaran calls transferability a "social epistemic virtue" presumably because protecting the quality of the body of mathematical knowledge is valuable to the mathematical community rather than any individual mathematician (Easwaran 2009, 343).

Don Berry (2016) also examines "mathematicians' strict insistence on proof." In line with Easwaran, Berry argues that there are good epistemic reasons for the mathematical community for only counting as argument deductive mathematical proof. Berry argues that there are some important virtues of mathematical research that are largely due to the norm in the mathematical community of only counting as argument deductive mathematical proof. And the success of mathematics, he further argues, is largely due to these virtues of mathematical research. Therefore, mathematicians are rational in their insistence on proof.

For example, Berry argues that the success of mathematics is partly due to the ***permanence*** of mathematical results. And the permanence of mathematical results, he argues, is a virtue that is due to the norm in the mathematical community of only counting as argument deductive mathematical proof. Like Easwaran and Fallis, Berry emphasizes as a virtue the possibility of ***autonomy*** for individual mathematicians when they read and assess a mathematical argument. For example, "When journal referees judge that a printed argument constitutes a sufficient basis for publication, they never rely upon trust in the testimony of the publishing mathematician" (Berry 2016, 117). This is due to the insistence on proof for "if a result were supported by inductive evidence alone, we would be required to accept the judgement of the author that this evidence is conclusive" (Berry 2016, 117). Berry emphasizes that the virtue of the possibility of autonomy for individual mathematicians helps maintain the virtue of permanence and other virtues.

To sum up, Easwaran, Fallis, and Berry agree that the norm of only counting as argument deductive mathematical proof allows individual experts to independently check mathematical arguments. And the fact that individual experts can independently check mathematical arguments is valuable to the mathematical community because it means that mathematical results can be extraordinarily reliable. This seems right. But the question remains as to how errors in a proof may be overlooked by individual mathematicians such that it makes a difference when more mathematicians check the same proof. Section "4" below addresses this question.

3 Challenging Cases I

It is worth briefly considering a phenomenon that seems to challenge the norm of only counting as argument deductive mathematical proof. I have in mind the phenomenon of ghost theorems. The notion of "ghost theorems" was introduced by Colin Rittberg, Fenner Tanswell, and Jean Paul Van Bendegem (2020, 3880) and refers to "mathematical results which are taken as accepted in a mathematical community, relied upon in talks, discussion, and proving further results, but which cannot be traced to a concrete proof in the literature." It remains to be studied how common ghost theorems are. Furthermore, it is unclear to what extent ghost theorems have unpublished proofs, but this is less important for our purposes because there is a substantial difference between unpublished proofs and published proofs. Rittberg et al. (2020, 3882–3884) emphasize that any unpublished proofs of ghost theorems are, all else being equal, less reliable than published proofs because they have not gone through the social process of proof checking that published proofs go through.

4 Proof Checking as a Social Process

A mathematical proof has traditionally been conceived as something very different from an ***argument*** as conceived in argumentation theory. In a key work in rhetoric, Chaïm Perelman and Lucie Olbrechts-Tyteca (1969) distinguish arguments from demonstrations and include mathematical proofs in the category of demonstrations. Demonstrations are impersonal as opposed to arguments which are relative to their intended audience. Perelman and Olbrechts-Tyteca (1969, 19) write that, "Since argumentation aims at securing the adherence of those to whom it is addressed, it is, in its entirety, relative to the audience to be influenced." In the same vein, Ralph Johnson writes that, "No mathematical proof had or needs to have a dialectical tier" (Johnson 2000, 270; quoted in Dufour 2013, 65). The dialectical tier of an argument is where the arguer anticipates and responds to objections to the argument.

While mathematical proofs have traditionally not been conceived as arguments, this is changing as philosophers of mathematics have started to pay more attention to ordinary mathematical proofs as they occur in actual mathematical practice. For example, Erik Krabbe (2008/2013, 193) holds a "dialectical and rhetorical view of proofs" on which "proofs may be more or less successful, depending upon context and audience." Krabbe notes that, "Some proofs are rather sketchy [relative to the intended audience], leaving much room for questioning (or work for the reader, if the author is not there to answer questions), whereas others, by answering these questions, are developed to great dialectical depth" (2008/2013, 193). Other studies similarly emphasize the audience-dependence of proofs and point out that the appropriate level of detail of proofs depends on their intended audience (see Andersen et al. 2019; Fallis 2003; Paseau 2016).

Ordinary mathematical proofs seem to be relative to their intended audience in at least some minimal sense. As mathematician William Thurston (1994, 175) has

pointed out, mathematicians "prove things in a certain context and address them to a certain audience." It seems fair to say that a proof aims to secure the adherence of the audience to whom it is addressed, since a purported proof is accepted as a proof by the mathematical community only when enough experts have read and assessed the purported proof and concluded that it, in fact, constitutes a proof. Hence, philosophers of mathematical practice have argued that for proofs to be thoroughly checked and accepted as proofs by mathematicians they need to go through a ***social*** process of verification in which the proofs are read by different experts and used in different contexts (see Rav 1999; Devlin 2003; Andersen et al. 2020).

Not until a number of experts have checked a proof does the proof count as a proof in mathematical practice. This is so because individual readers of proofs may quite easily overlook errors (see Geist et al. 2010; Müller-Hill 2011). (This may be one reason why it is becoming more common, although still rare, to formalize proofs as a way of verifying them (Avigad 2008; Avigad and Harrison 2014; Hales 2008). This kind of proof checking is not in general a social process, but the formalization of a proof will sometimes be the result of a collaborative effort.) Mathematicians themselves stress that errors in proofs may quite easily be overlooked. Philosopher Eva Müller-Hill has interviewed six mathematicians about their intuitions as to when a mathematician may be said to know that a mathematical claim is true. Five of them stressed that even an expert journal referee may overlook errors (Müller-Hill 2011, 305–310, 315–316, 319–322, 328, 342–343). In the words of one of the interviewees, prepublication peer review is just "one step in trying to make sure that nobody has overlooked a mistake in the proof" (Müller-Hill 2011, 309). That errors may quite easily be overlooked by any individual mathematician is also reflected in the fact that mathematicians seem to be hesitant to rely on a result without checking the proof of the result themselves unless a number of other experts have checked the proof (Andersen et al. 2020).

Because the vetting of a proof is a long and complex process and errors in proofs are sometimes only found after publication, some have argued that the mathematical community should have a better system for sharing publicly the errors found in published proofs (Grcar 2013; Frans and Kosolosky 2014).

I will now review a number of different accounts of proof since different accounts of proof come with somewhat different accounts of why it makes sense to have more experts check a proof. This is due to the fact that different accounts of proof come with different accounts of how individual mathematicians read and assess proofs. And, therefore, different accounts of proof come with different accounts of how errors in a proof may be overlooked by individual mathematicians.

A number of philosophers claim that ordinary proofs are what one could call ***sketches*** of formal proofs (see Avigad 2020; Azzouni 2004; Hamami 2019; MacLane 1986). A purported proof is a proof if it can be turned into a formal proof. If we conceive of proofs this way, then becoming convinced that a purported proof is a proof involves recognizing that the purported proof could be turned into a formal proof. Jeremy Avigad describes the process thus,

> Whether or not a mathematician reading a proof would characterize the state of affairs in these terms, a judgement as to correctness is tantamount to a judgment as to the existence of a formal derivation, and whatever psychological processes the mathematician brings to bear, they are reliable insofar as they track the correspondence. (Avigad 2020, 3)

On the proofs as sketches accounts, overlooking an error in a proof amounts to making a mistake in tracking the correspondence between the ordinary proof and the formal proof. Some philosophers have criticized the proof as sketch account (see Larvor 2012; Rav 1999; Tanswell 2015). The discussion is ongoing.

Others argue that we should conceive of proofs as ***narratives*** (Doxiadis 2012; Robinson 1991, 269; Thomas 2007). For example, we may conceive of a proof as a telling of how the author of the proof came from the premises to the conclusion, as a telling of a sequence of inferential actions performed by the author (Doxiadis 2012). This fits well with how mathematicians have occasionally described proving as taking a journey (as described in Lane et al. 2019, 199–200). If we conceive of proofs as narratives in this way, then coming to the conclusion that a purported proof is a proof involves an assessment of the actions performed by the author of the proof. Overlooking an error in a proof amounts to something like making a mistake when assessing whether a certain action is suitable for obtaining a certain goal.

Other accounts of proof are similar to the proof as narrative account in that they focus on actions, but they focus on the actions of the readers of proofs rather than the authors of proofs. Hence, some philosophers argue that we should conceive of proofs as ***recipes*** for how to prove propositions (Larvor 2012, 725–726; Sundholm 2012; Tanswell 2017). On these accounts, a proof as written is a recipe for the readers to follow to do an actual proof. If proofs are recipes, then reading and assessing a purported proof involve actually performing the steps in the recipe and seeing if they lead to the desired outcome. In this case, the reader is doing the proving herself under guidance, or taking the journey herself. If we conceive of proofs as recipes, then coming to the conclusion that a purported proof is a proof amounts to performing the inferential actions oneself. Overlooking an error in a proof amounts to making a mistake when doing the proof.

Finally, some philosophers have argued that we should conceive of proofs as ***dialogues*** (Dutilh Novaes 2016, 2018; Ernest 1994). Catarina Dutilh Novaes gives an account of a proof as a dialogue between two fictitious interlocutors whom she calls Prover and Skeptic. On her account,

> A deductive proof corresponds to a dialogue between the person wishing to establish the conclusion [. . .] and an interlocutor who will not be easily convinced and will bring up objections, counterexamples, and requests for further clarification. A good proof is one that convinces a fair but 'tough' opponent. (Dutilh Novaes 2016, 2617)

She adds: "Ultimately, most of the work is done by Prover, but Skeptic has an important role to play, namely to ensure that the proof is persuasive, perspicuous, and valid" (Dutilh Novaes 2016, 2618). If proofs are dialogues in this sense, then reading and assessing a proof involves assessing whether Skeptic is skeptic enough. Overlooking an error in a proof often amounts to overestimating the skepticism of

Skeptic. Note that the accounts of proofs as dialogues are perhaps especially interesting from the point of view of social epistemology because they include the social nature of dialogue in the very nature of proof.

Conceiving of proofs ***as*** dialogues should be distinguished from conceiving of proofs as the ***result of*** dialogues, as Lakatos (1976) famously did. In *Proofs and Refutations*, Lakatos illustrates how mathematical concepts can grow through a dialectic of proofs and counterexamples, using the historical example of the concept of polyhedron as used in proofs of Euler's conjecture. The concept of polyhedron becomes more clear and precise in the context of a long debate of the proofs, which are also refined in the process. The book is structured as a dialogue between a teacher and his students, with the students representing different methodological stances (see Larvor 1998, 10–11).

5 Challenging Cases II

The different accounts of proofs we have considered provide different answers to the question of why a purported proof is only accepted as a proof by the mathematical community when enough experts have read and assessed the purported proof and found no errors, that is, when the purported proof has gone through a social process of proof validation. But there are some unusual cases of proofs that challenge the validation of proofs as a social practice. More specifically, the cases challenge the norm of only counting as deductive mathematical proof an argument that has convinced a number of experts that it is indeed a deductive mathematical proof.

A number of prominent and highly praised proofs have received attention for being so long or complex that only a few mathematicians have read and checked the proofs. Famous examples of proofs that have received attention for being so long that they would be incredibly hard for any individual mathematician to read include the proof of the classification of finite simple groups (see Steingart 2012) and the proof of the Kepler Conjecture by Thomas Hales (see Devlin 2003). The famous proof of Fermat's Last Theorem by Andrew Wiles is an example of a proof that draws on different domains in such a way that only few individual experts have been able to read the proof (see Jones and Jones 1998, 217–237; Penrose 1997). These proofs are in practice not very open to checking by mathematicians. For this reason, they pose a challenge to the norm of only counting as deductive mathematical proof an argument that has been checked by a number of experts.

These are proofs of very high-profile results that ***challenge*** the norm. Very long or complex purported proofs of less important results presumably often ***succumb*** to the norm in the sense that they are not accepted as proofs by the mathematical community because they have not been vetted by enough experts. The norm of only counting as deductive mathematical proof an argument that has been checked by a number of experts limits the extent to which purported proofs can be long and complex and still be accepted as proofs by the mathematical community.

But even proofs of high-profile results by high-profile mathematicians may succumb to the norm of only counting as deductive mathematical proof an argument

that has been checked by a number of experts. The purported proof of the abc conjecture by the prominent mathematician Shinichi Mochizuki has been described as being extremely hard to even begin to comprehend (see Castelvecchi 2015; Hartnett 2014). The purported proof is from 2012 and has not yet been accepted as a proof by the mathematical community and perhaps will not be. Mochizuki had been working on the proof for many years on his own and had in that time "constructed his own mathematical universe and populated it with arcane terms like 'inter-universal Teichmüller theory' and 'alien arithmetic holomorphic structures'" (Hartnett 2014, 225). Hence, the purported proof is written in a new language and, in addition, the author is rather unwilling to explain his thinking. The behavior of the author can perhaps be described as going against the norm of the mathematical community of only accepting as justification in mathematics arguments that are open to independent checking by different mathematicians.

6 Conclusion

If we want to explain the extraordinary reliability of mathematical knowledge, we may point to the central role of deductive mathematical proofs in mathematics. But to understand the reliability of mathematical proofs, we also need to consider the relationship between mathematical proofs and mathematical practice as a social practice. We need to consider the social epistemology of mathematical proof. In this chapter, I have reviewed how philosophers have examined how mathematical proofs are valuable to the mathematical community because of their being open to independent checking by different mathematicians. I have also reviewed how philosophers have offered different accounts of how mathematicians read and assess proofs. These accounts offer different answers to the question of how individual mathematicians may overlook errors and thus of why we need a social process of proof checking.

7 Cross-References

- ▶ Agency in Mathematical Practice
- ▶ Argumentation in Mathematical Practice
- ▶ Formal Proofs in Mathematical Practice
- ▶ The Algorithmic-Device View of Informal Rigorous Mathematical Proof
- ▶ The Values of Mathematical Proofs

Acknowledgments I would like to thank an anonymous reviewer for helpful comments. The research for this chapter was funded by K. Brad Wray's grant from the Aarhus Universitets Forskningsfond—StartingGrant, AUFF-E-2017-FLS-7-3

References

Andersen LE, Johansen MW, Sørensen HK (2019) Mathematicians writing for mathematicians. Synthese (forthcoming)
Andersen LE, Andersen H, Sørensen HK (2020) The role of testimony in mathematics. Synthese (forthcoming)
Avigad J (2008) Computers in mathematical inquiry. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 302–316
Avigad J (2020) Reliability of mathematical inference. Synthese (forthcoming)
Avigad J, Harrison J (2014) Formally verified mathematics. Commun ACM 57:66–75
Azzouni J (2004) The derivation-indicator view of mathematical practice. Philos Math 12:81–105
Berry D (2016) Proof and the virtues of shared enquiry. Philos Math 26:112–130
Borwein J, Bailey D (2003) Mathematics by experiment: plausible reasoning in the 21st century. AK Peters, Wellesley
Castelvecchi D (2015) The impenetrable proof. Nature 526:178–181
Devlin K (2003) When is a proof? Column for Mathematical Association of America. https://www.maa.org/external_archive/devlin/devlin_06_03.html. Accessed 1 June 2020
Doxiadis A (2012) A streetcar named (among other things) proof. In: Doxiadis A, Mazur B (eds) Circles disturbed: the interplay of mathematics and narrative. Princeton University Press, Princeton, pp 281–388
Dufour M (2013) Arguing around mathematical proofs. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, Dordrecht, pp 61–76
Dutilh Novaes C (2016) Reductio ad absurdum from a dialogical perspective. Philos Stud 173:2605–2628
Dutilh Novaes C (2018) A dialogical conception of explanation in mathematical proofs. In: Ernest P (ed) The philosophy of mathematics education today: ICME-13 monographs. Springer, Cham, pp 81–98
Easwaran K (2009) Probabilistic proofs and transferability. Philos Math 17:341–362
Epstein D, Levy S (1995) Experimentation and proof in mathematics. Not Am Math Soc 42:670–674
Ernest P (1994) The dialogical nature of mathematics. In: Ernest P (ed) Mathematics, education and philosophy: an international perspective. Falmer Press, London, pp 33–48
Fallis D (1997) The epistemic status of probabilistic proof. J Philos 94:165–186
Fallis D (2000) The reliability of randomized algorithms. Br J Philos Sci 51:255–271
Fallis D (2002) What do mathematicians want? Probabilistic proofs and the epistemic goals of mathematicians. Log Anal 45:1–16
Fallis D (2003) Intentional gaps in mathematical proofs. Synthese 134:45–69
Fallis D (2011) Probabilistic proofs and the collective epistemic goals of mathematicians. In: Schmid HB, Weber M, Sirtes D (eds) Collective epistemology. Ontos Verlag, Frankfurt am Main, pp 157–175
Frans J, Kosolosky L (2014) Revisiting the reliability of published mathematical proofs: where do we go next? Theoria 29:345–360
Geist C, Löwe B, Van Kerkhove B (2010) Peer review and knowledge by testimony in mathematics. In: Löwe B, Müller T (eds) PhiMSAMP: philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London, pp 155–178
Grcar J (2013) Errors and corrections in mathematics literature. Not Am Math Soc 60:418–425
Hales TC (2008) Formal proof. Not Am Math Soc 55:1370–1380
Hamami Y (2019) Mathematical rigor and proof. Rev Symb Log (forthcoming)
Hartnett K (2014) An ABC proof too tough even for mathematicians. In: Pitici M (ed) The best writing on mathematics 2013. Princeton University Press, Princeton, pp 225–230
Jackson JC (2009) Randomized arguments are transferable. Philos Math 17:363–368
Johnson RH (2000) Manifest rationality: a pragmatic theory of argument. Lawrence Erlbaum Associates, Mahwah

Jones GA, Jones JM (1998) Elementary number theory. Springer, London
Krabbe ECW (2008/2013) Strategic maneuvering in mathematical proofs. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, Dordrecht, pp 181–193
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, New York
Lane L, Martin U, Murray-Rust D, Pease A, Tanswell F (2019) Journeys in mathematical landscapes: genius or craft? In: Hanna G, Reid D, de Villiers M (eds) Proof technology in mathematics research and teaching. Springer, Cham, pp 197–212
Larvor B (1998) Lakatos: an introduction. Routledge, London
Larvor B (2012) How to think about informal proofs. Synthese 187:715–730
MacLane S (1986) Mathematics: form and function. Springer, New York
Müller-Hill E (2011) Die Epistemische Rolle Formalisierbarer Mathematischer Beweise. Dissertation, Rheinische Friedrich-Wilhelms-Universität. http://hss.ulb.uni-bonn.de/2011/2526/2526.htm. Accessed 1 June 2020
Paseau AC (2016) What's the point of complete rigour? Mind 125:177–207
Penrose R (1997) Q.E.D. New York Times, 30 Nov 1997, sect 7, p 12
Perelman C, Olbrechts-Tyteca L (1969) The new rhetoric: a treatise on argumentation. The University of Notre Dame Press, Notre Dame
Rav Y (1999) Why do we prove theorems? Philos Math 7:5–41
Rittberg C, Tanswell FS, Van Bendegem JP (2020) Epistemic injustice in mathematics. Synthese 197:3875–3904
Robinson JA (1991) Formal and informal proofs. In: Boyer RS (ed) Automated reasoning. Kluwer Academic Publishers, Dordrecht, pp 267–282
Sørensen HK (2016) "The end of proof"? The integration of different mathematical cultures as experimental mathematics comes of age. In: Larvor B (ed) Mathematical cultures: the London meetings 2012–2014. Birkhäuser, Basel, pp 139–160
Steingart A (2012) A group theory of group theory: collaborative mathematics and the "uninvention" of a 1000-page proof. Soc Stud Sci 42:185–213
Sundholm G (2012) "Inference versus consequence" revisited: inference, consequence, conditional, implication. Synthese 187:943–956
Tanswell F (2015) A problem with the dependence of informal proofs on formal proofs. Philos Math 23:295–310
Tanswell F (2017) Proof, rigour & informality: a virtue account of mathematical knowledge. Dissertation, University of St. Andrews. https://research-repository.st-andrews.ac.uk/handle/10023/10249. Accessed 1 June 2020
Thomas RSD (2007) The comparison of mathematics with narrative. In: Van Kerkhove B, Van Bendegem JP (eds) Perspectives on mathematical practices: bringing together philosophy of mathematics, sociology of mathematics, and mathematics education. Springer, Dordrecht, pp 43–59
Thurston WP (1994) On proof and progress in mathematics. Bull Am Math Soc 30:161–177

The Values of Mathematical Proofs

Rebecca Lea Morris

Contents

1 Introduction 2082
2 Practical Values 2084
2.1 Individual Values 2084
2.2 Discussion of Practical Values 2086
2.3 Summary 2087
3 Abstract Values 2087
3.1 Explanation 2088
3.2 Beauty 2091
3.3 Fit 2094
3.4 Motivation 2096
4 Relationships Between Abstract and Practical Values 2099
4.1 Explanation 2099
4.2 Beauty 2100
4.3 Fit 2101
4.4 Motivation 2102
4.5 Summary 2103
5 Relationships Between Abstract Values 2103
5.1 Fit and Motivation 2103
5.2 Explanation and Beauty 2105
5.3 Explanation and Fit 2105
5.4 Explanation and Motivation 2106
5.5 Beauty and Fit 2107
5.6 Beauty and Motivation 2107
5.7 Summary 2108
6 Conclusion 2108
7 Cross-References 2109
References 2109

R. L. Morris (✉)
Independent Scholar, Minneapolis, MN, USA
e-mail: email@rebeccaleamorris.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_34

Abstract

Proofs are central, and unique, to mathematics. They establish the truth of theorems and provide us with the most secure knowledge we can possess. It is thus perhaps unsurprising that philosophers once thought that the *only* value proofs have lies in establishing the truth of theorems. However, such a view is inconsistent with mathematical practice. If a proof's only value is to show a theorem is true, then mathematicians would have no reason to reprove the same theorem in different ways, yet this is a practice they frequently engage in. This suggests that proofs can have a wide variety of values. My purpose in this chapter is to provide a survey of some of them. I will discuss how proofs can have *practical value*, for example, by helping mathematicians to make new discoveries or learn new mathematical tools, and how they can have *abstract value*, for example, by being beautiful or explanatory. I will also explore the relationships between different proof values.

Keywords

Mathematical practice · Proofs · Values

1 Introduction

Proofs are central, and unique, to mathematics. They establish the truth of theorems and provide us with the most secure knowledge we can possess. It is thus perhaps unsurprising that philosophers have thought their only job, and only value, was to establish the truth of theorems. For example, Rav describes the "standard view" in philosophy as follows: "proofs serve *only* to validate theorems on the basis of accepted axioms" (Rav 1999, 15, emphasis in original). However, as Rav also points out, the standard view is at odds with mathematical practice (Rav 1999, 15) though he has different aspects of mathematical practice in mind than what I focus on below.

If the standard view were true, then once a theorem has been proven, mathematicians would have no reason to prove it again. After all if the only value a proof has is establishing the truth of its theorem, then once mathematicians have one proof adding more would be entirely superfluous. Yet, as Avigad (2006) and Dawson (2006, 2015) emphasize, mathematicians *do* routinely reprove theorems, sometimes giving hundreds of different proofs of the same theorem. For example, 370 proofs of the Pythagorean theorem are presented in Loomis (1968), and as Avigad notes (Avigad 2006, note 2), Franz Lemmermeyer maintains a list of proofs of quadratic reciprocity (Lemmermeyer n.d.), which currently has over 240 entries. This indicates that, contra the standard view, proofs have value beyond establishing their theorems.

The purpose of this chapter is to survey some of the wide variety of values proofs can have, other than establishing the truth of theorems. As we will see, some of these values are possessed by all proofs, but the majority are only possessed by some proofs. I will here be concerned only with informal proofs and not with formal

derivations. That is to say, I will focus only on proofs as written by mathematicians in journal articles or textbooks, for example, and not on logical derivations made in formal deductive systems. This is not to deny that formal proofs are valuable. For example, they can be more valuable than informal proofs when it comes to ensuring the correctness of a result. Indeed significant mistakes have been made in published informal mathematical proofs. For example, Kempe's (1879) invalid proof of the four color theorem stood for 11 years before Heawood (1890) discovered the error. The prevalence of mistakes, as well as the ever increasing complexity of mathematics, has led some mathematicians to adopt the use of proof assistants (a type of computer software) in their work. Proof assistants help mathematicians to translate their informal proofs into formal ones where each step is made explicit and verified to be logically correct. Thus some mathematicians argue "that the use of a formal proof . . . will contribute to enhancing rigour and gaining confidence in the correctness of long and complex proofs that are not amenable to human verification" (Hanna and Larvor 2020, sec. 4.1). For an overview, see Avigad (2018).

In what follows I will organize the values proofs can have into two categories: practical and abstract. Proofs have practical value when they help mathematicians to achieve pragmatic goals, such as making new discoveries or learning new mathematical tools. Proofs have abstract value when they go beyond helping mathematicians to achieve pragmatic goals by having a quality that is either intrinsically valuable or that yields something that is. Examples of abstract values include explanatory power and beauty. As we will see, there are interconnections between these different values. For example, sometimes a particular abstract value is also often associated with specific practical values, and sometimes one particular abstract value is strongly associated with other abstract values.

Admittedly, the distinction between practical and abstract values is not always sharp. In particular, it can sometimes be difficult to determine whether a property that yields practical benefits is intrinsically valuable (or provides us with something that is). For example, intuitiveness is a valuable proof property and arguably has practical benefits like helping mathematicians to make new discoveries because intuitive proofs are easy to work with. But it seems hard to say whether intuitiveness is valuable for its own sake or whether it is valuable only for its practical consequences (Many thanks to an anonymous reviewer for this example). This does not mean, however, that the distinction between abstract and practical values is untenable. For while there will be borderline cases, there will be others that are more clear-cut. In Sect. 3, I identify four abstract values whose classification appears to be less controversial. This is not to say, however, that the classification of these values as abstract cannot be challenged. For example, further analysis of an ostensible abstract value could reveal that it is not an abstract value after all, perhaps by reducing it to something else that is more easily recognized as not having (or providing anything of) intrinsic value. At the beginning of the discussion of each value, however, I explain why I classify it as an abstract value. I hope that this will illuminate how the notion of abstract value is to be understood in practice.

The structure of this chapter is as follows. I first discuss practical values that proofs can have in Sect. 2 before considering abstract values in Sect. 3. Then, in Sect. 4,

I discuss the relationships between particular abstract and practical values. Next, in Sect. 5, I examine connections between the individual abstract values. Finally, in Sect. 6, I look at potentially fruitful avenues for future work on the values of proofs.

2 Practical Values

Philosophers, mathematicians, and mathematics educators have identified a variety of ways that proofs can have practical value. That is to say, they have identified ways that proofs help mathematicians to achieve important goals that ensure continued mathematical progress. As we will see, some of these practical values are common to all proofs, while others are only provided by some proofs. We will also see that there are close connections between some of these individual practical values. Below I will first briefly summarize the individual practical values before discussing (i) whether the individual values are provided by all or only some proofs and (ii) how they are related.

2.1 Individual Values

2.1.1 Making Mathematical Knowledge Accessible

Easwaran (2009) argues that traditional mathematical proofs have a property that distinguishes them from other arguments: transferability. To say an argument is transferable means that an expert in the field would already believe the premises of the argument and would come to accept any inferences made simply by considering them themself (Easwaran 2009, 354). So, for example, a mathematician reading a proof does not have to rely upon the testimony of the proof author or accept any appeals to authority to know the theorem. Proofs thus contrast with arguments in the physical sciences, which often do require readers to rely on testimony. For example, scientific arguments require readers to accept the authors' testimony about the experiments that figure in their arguments (Easwaran 2009, 359).

Easwaran suggests that transferability has an important practical benefit: it makes mathematical knowledge more accessible to the community. This is because transferability means that the mathematical community can be "self-monitoring" (Easwaran 2009, 356), i.e., members of the mathematical community can check through proofs and ensure any errors are corrected. If proofs were not transferable this would not be possible, since some steps could not be checked, relying, for example, on testimony or appeals to authority (Easwaran 2009, 356). As traditional mathematical proofs are transferable, one way they have practical value is thus by making mathematical knowledge accessible.

2.1.2 Power of Different Methodologies

Dawson notes that one reason mathematicians reprove theorems is to "demonstrate the power of different methodologies" (Dawson 2006, 277). For example, he

mentions Bishop's *Foundations of Constructive Analysis*, which contained proofs to show "how large a part of abstract analysis can be developed within a constructive framework" (Dawson 2006, 277). Avigad offers another example: "Dedekind often went out of his way to show how the new methods developed in his theory of ideals result in perspicuous proofs of established theorems of number theory, from Fermat to Kummer" (Avigad 2006, 107). A second way that proofs can thus have practical value is by showcasing what different methodologies can do.

2.1.3 New Ways of Thinking

Weber (2010) notes that mathematicians hope to learn new ways of thinking when they read proofs. He identifies two ways mathematicians learn to think differently: by "reconceiving mathematical domains and learning new proof techniques" (Weber 2010, 32). Proofs that reconceive mathematical domains help mathematicians to develop their intuitions and refine their mental models, giving them a new way of representing the mathematics in question (Weber 2010, 32–34). This in turn can help mathematicians to see why a theorem they initially found implausible turned out to be true (Weber 2010, 32).

Proofs that help mathematicians learn new techniques provide them with tools to add to their mathematical toolbox that can then be used and adapted to solve new problems (Weber 2010, 34–35). As a simple illustration, Weber borrows an example from Rav (1999, 21): Euclid's proof of the infinitude of the primes. This proof shows that given any finite collection of primes we can construct a new but distinct prime. The technique for constructing this new prime can be used to prove new theorems, for example, that there are infinitely many primes of the form $4n + 3$ (Rav 1999, 21).

Although Weber focuses on techniques, Rav (1999) points out that proofs can convey other tools, for example, concepts and ideas. As he puts it "Proofs are for the mathematician what experimental procedures are for the experimental scientist: in studying them one learns of new ideas, new concepts, new strategies—devices which can be assimilated for one's own research and be further developed" (Rav 1999, 20).

We have thus seen a third way that proofs can have practical value: by helping mathematicians to learn new ways of thinking, i.e., by developing new mental models, refining their intuitions, and learning new tools.

2.1.4 Systematization, Discovery, and Communication

De Villiers (1990) draws attention to a number of practical values of proof. First he notes that when mathematicians have a number of known results, they often want to better organize them in a more systematic way. They thus engage in "a posteriori axiomatization and defining" (De Villiers 1990, 20) to obtain "a deductive system of axioms, definitions and theorems" (De Villiers 1990, 20). As De Villiers points out, proofs play an important part in this systematization by relating different mathematical statements together as part of an organized whole.

Secondly, while proofs are traditionally thought of as belonging firmly to the context of justification, De Villiers points out that they can stimulate new discoveries. Using examples from geometry he shows that sometimes a proof of one theorem

can lead us to see that the result generalizes and thus to discover a new theorem. In his examples, it is the proofs themselves that highlight the crucial feature that allows the result to generalize, a feature that, De Villiers claims, would likely have been missed by other attempts to generalize the original theorem, for instance by experimenting with examples (De Villiers 1990, 21).

Finally De Villiers highlights the importance of proofs for mathematical communication in a way that bears some similarities to Easwaran's discussion of transferability. As De Villiers notes, mathematicians communicate with each other in proofs, and so proofs can be seen as "a form of social interaction" (De Villiers 1990, 22). This means that mathematicians have to agree on what counts as an acceptable argument and so proofs are scrutinized, refined, and rejected if a counterexample is found. This in turn has safeguarded mathematics, preventing mathematicians as a group from making or accepting big mistakes.

De Villiers thus identifies three ways that proofs can have practical value: (i) by helping to systematize a mathematical subfield; (ii) by stimulating new discoveries; and (iii) by safeguarding mathematics from mistakes.

2.2 Discussion of Practical Values

Some of the practical values of proofs discussed above are shared by all proofs in virtue of their being proofs. These include: making mathematical knowledge accessible, systematizing mathematical knowledge, and safeguarding mathematics from mistakes. As all traditional proofs are transferable, they all make knowledge of their corresponding theorem accessible to the community. As all proofs highlight the relationships between mathematical statements, at least to some degree, presumably they all help contribute to the systematization of mathematics. Finally, as all proofs serve as a means of communication between mathematicians and so will be refined if necessary, presumably all proofs help prevent the community from accepting mistakes.

Other practical values only seem to be provided by some proofs. These include: showcasing the power of different methodologies, providing mathematicians with new ways of thinking, and promoting discoveries. For example, a proof that uses standard techniques in a standard way is not going to illustrate the power of different methodologies; a proof that uses a variety of new methods, however, might. Similarly a proof that uses standard tools in a standard way is unlikely to teach mathematicians new ways of thinking. A proof that introduces a new technique or a new way of representing a mathematical concept, however, can. Finally a brute force proof using standard tools, for example, is unlikely to inspire mathematicians to make new discoveries. A proof that exploits a relationship in an unusual way or that uses new concepts or methods, however, may well do so.

Some of the individual practical values discussed above are interrelated. First, proofs that illustrate new methodologies seem likely to teach mathematicians new ways of thinking. After all, proofs that showcase different methodologies will use new ideas or tools or use existing ones in unusual or unexpected ways. This means

they will plausibly convey new representation systems or new methods and techniques to mathematicians who read them. Proofs that provide new ways of thinking, however, will not necessarily also illustrate new methodologies. For example, a proof that provides a new way of thinking could introduce a new method, but that method might not be part of a larger system that constitutes a new or different methodology.

Secondly, proofs that provide new ways of thinking also tend to stimulate discoveries. This is because proofs that yield new ways of thinking help provide mathematicians with new tools that they can use to work on new problems, increasing their chances of success. However, there may be proofs that stimulate discoveries without providing mathematicians with new ways of thinking. For example, it might be possible for a proof to promote further discoveries by drawing attention to a certain relationship between mathematical objects that can be proven to generalize using standard techniques. This would lead to new discoveries but would not necessarily provide new ways of thinking.

Finally, note that the six practical values of proofs discussed here may not be the only practical values proofs can have. Future work may reveal additional such values possessed by some or all proofs and should more deeply explore the relationships between them.

2.3 Summary

We have discussed six ways proofs can have practical value by: (i) making mathematical knowledge more accessible; (ii) showing what different methodologies can do; (iii) helping mathematicians learn new ways of thinking; (iv) systematizing mathematics; (v) stimulating new discoveries; and (vi) safeguarding mathematics against mistakes. Three of these practical values are possessed by all proofs: (i) making mathematical knowledge more accessible; (iv) systematizing mathematics; and (vi) safeguarding mathematics against mistakes. The remaining three are only possessed by some proofs. Further, as we saw, some of the individual practical values are connected. In particular, if a proof shows what a different methodology can do then it also likely helps mathematicians learn new ways of thinking and if a proof helps mathematicians to learn new ways of thinking it also likely stimulates new discoveries.

3 Abstract Values

Philosophers, mathematicians, and mathematics educators have identified and analyzed different ways in which proofs can have abstract value. That is to say, they have developed accounts of qualities that proofs can have that are (or provide something that is) intrinsically valuable. Due to space considerations I discuss below only four such abstract values: (i) explanation; (ii) beauty; (iii) fit; and (iv) motivation. Other important abstract values include, for example, depth

(see, e.g., Arana 2015; Gray 2015; Lange 2015; Stillwell 2015; Urquhart 2015) and purity (see, e.g., Detlefsen 2008; Detlefsen and Arana 2011). Unlike some of the practical values discussed in Sect. 2, none of the abstract values discussed here are possessed by all proofs. In other words, while some proofs will possess one (or more) of these abstract values, others will not.

Proofs with abstract value can also have practical value. We have already seen that all proofs have the practical values of making mathematical knowledge accessible, systematizing mathematics and safeguarding it from mistakes, so proofs with abstract value will have these practical values as well. Moreover, certain abstract values are closely associated with particular practical values, so that a proof with such an abstract value will likely have the associated practical values as well. I discuss these connections in Sect. 4. Then, in Sect. 5, I take up the issue of connections between the individual abstract values themselves.

3.1 Explanation

Mathematicians aim not just to prove new theorems, but to give mathematical explanations. Sandborg (1997) looks at reviews of mathematical papers from the 1980s and 1990s and identifies a wide variety of positive remarks about the explanations those papers provide. Below are just two examples:

> The paper under review makes a very nice contribution to the theory by giving a theoretical explanation of this bifurcation (Marsden 1985, quoted on pg 42 of Sandborg (1997)).

> Although [the authors] had previously proved (by a different method) that their algebras were Noetherian, the present paper gives a more fundamental explanation of this fact (Smith 1991, quoted on pg 35 of Sandborg (1997))

Perhaps the most obvious way mathematicians can provide explanations is by giving explanatory *proofs*. For example, Haigh, whose proof of Pick's theorem is discussed by Sandborg, remarks "I was surprised when I first met this theorem, and the proofs I read [. . .] did not allay my surprise, so I looked for a more explanatory proof" (Haigh 1980, 173).

This suggests that mathematical explanation, and explanatory proofs in particular, are highly valuable. Comparison with scientific explanation suggests that mathematical explanation is also an abstract value. For example, Strevens writes:

> If science provides anything of intrinsic value, it is explanation. Prediction and control are useful, and success in any endeavor is gratifying, but when science is pursued as an end rather than as a means, it is for the sake of understanding—the moment when a small, temporary being reaches out to touch the universe and makes contact (Strevens 2011, 3)

We could very easily imagine something similar being said about explanation in mathematics, and, in particular, explanatory proofs:

> If mathematics provides anything of intrinsic value, it is explanation. [. . .] When mathematics is pursued as an end rather than as a means, it is for the sake of understanding—the moment when a small, temporary being reaches out to touch the mathematical universe and makes contact.

Spelling this out more explicitly, explanation give us something that is intrinsically valuable: mathematical understanding. This means that explanation is an abstract value.

Although explanatory proofs have abstract value, we have not identified what, exactly, such proofs are. In fact, there is little consensus over the nature of explanatory proofs. For example, Lange describes the situation regarding the explanatory power of just one type of proof, proof by induction, in the following terms:

> Some philosophers appear quite confident that these arguments are generally explanatory, even in the face of other philosophers who appear equally confident of their intuition to the contrary (Lange 2009, 204).

Many philosophers have, nonetheless, developed accounts of explanatory proofs, including Steiner (1978), Kitcher (1989), Sandborg (1997), Cellucci (2008), Frans and Weber (2014), Lange (2014), Inglis and Meja-Ramos (2019), and Baron et al. (2019). Here I will focus on the accounts provided by Steiner, Kitcher, and Lange.

3.1.1 Steiner's Account

Central to Steiner's account of explanatory proofs is his notion of characterizing properties. A characterizing property is "a property unique to a given entity or structure within a *family* or domain of such entities or structures" (Steiner 1978, 143, emphasis in original). Note that a mathematical object can have different characterizing properties. Steiner gives the example of 18, which can be characterized as the successor of 17 or as $2 \cdot 3^2$ (Steiner 1978, 143). For Steiner, an explanatory proof is one which is clearly dependent on a characterizing property of something mentioned in the theorem. Moreover, an explanatory proof has to be clearly generalizable in the following sense: if we replace the original characterizing property with a characterizing property of a different, but related, object then we obtain a different theorem (Steiner 1978, 147).

An example of an explanatory proof on Steiner's account is the following proof of the theorem that $1 + 2 + \ldots + n = n(n + 1)/2$ (Steiner 1978, 136):

Proof Let $S = 1 + 2 + \ldots + n - 1 + n$. Then note:

$$\begin{array}{rcccccccccc} S & = & 1 & + & 2 & + & \ldots & + & n-1 & + & n \\ S & = & n & + & n-1 & + & \ldots & + & 2 & + & 1 \\ \hline 2S & = & (n+1) & + & (n+1) & + & \ldots & + & (n+1) & + & (n+1) \end{array}$$

Consequently $2S = n(n + 1)$ and so $S = n(n + 1)/2$. □

This is explanatory, according to Steiner, because it relies on a characterizing property of the sum *S*: its symmetry (Steiner 1978, 144–145). Moreover, as his account requires, we can deform it to obtain new results. For example, by omitting the even terms from the sums, we obtain a proof that $1 + 3 + 5 + \ldots + 2n - 1 = n^2$.

Steiner's account has been much discussed and criticized, with numerous counterexamples being proposed (see Resnik and Kushner 1987; Weber and Verhoeven 2002; Hafner and Mancosu 2005; Lange 2014; Pincock 2015). As just one example Hafner and Mancosu (2005) analyze Pringsheim's proof of Kummer's convergence test. This proof was designed to be explanatory, but Hafner and Mancosu argue that there is no appropriate characterizing property that can be appealed to in this case, and so the proof fails to be explanatory on Steiner's account.

3.1.2 Kitcher's Account

Kitcher (1989) develops an account of scientific explanation based on unification. He writes:

> Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again, and, in demonstrating this, it teaches us how to reduce the number of types of facts that we have to accept as ultimate (or brute) (Kitcher 1989, 432).

As Kitcher notes, his account can also be applied to mathematics, though he does not fully flesh out how to do this. A very rough adaptation is as follows: A proof is explanatory if and only if it instantiates argument patterns from the *explanatory store* of the field it is in, where the explanatory store is the collection of argument patterns that does the best job of unifying the field, i.e., maximizing the number of theorems that can be derived while minimizing the number of legitimate argument patterns needed for the derivations.

Like Steiner's account of explanation, the Kitcher-style account has faced considerable criticism. For example, Hafner and Mancosu (2008) present a detailed case study from real algebraic geometry and argue that the account does not judge the explanatoriness of proofs in a way that matches up with mathematicians' judgments. Further Lange criticizes a Kitcher-style account on the grounds that it would label as explanatory many proofs that do not appear to be genuinely so, such as proofs of geometric theorems that proceed by "algebraically grinding out the result" (Lange 2014, 523).

3.1.3 Lange's Account

More recently, Lange has developed an account of mathematical explanation. He argues that there is only a distinction between explanatory and nonexplanatory proofs when there is a striking or "salient" feature in the result itself, such as a prominent symmetry or unity (Lange 2014, 489). When there is such a feature, Lange suggests that an explanatory proof is one which traces that striking feature back to a similar type of striking feature involved in the "setup" or "problem" (Lange 2014, 525).

However, Lange's account is also subject to criticism. For example, Lange (2016a, 440) claims his account would judge Galois's proof of the unsolvability of

the quintic as explanatory but not Abel's, which means it would agree with judgments many mathematicians make about the explanatory power of these proofs. However, D'Alessandro (forthcoming) argues that Lange's account *does* judge Abel's proof to be explanatory. If he is right, then Lange's account diverges from intuitive judgments about the explanatory value of this proof.

3.1.4 Summary

Mathematicians value proofs that are explanatory. As explanation provides understanding, which is intrinsically valuable, it is an abstract value. There is, however, little consensus as to what an explanatory proof really is and so further work is needed to gain a deeper understanding of this value. For example, additional case studies could be used to assess existing accounts of explanation or to serve as a starting point from which to develop new accounts of explanatory proofs.

3.2 Beauty

Many mathematicians and philosophers value proofs that are beautiful. Below are some representative quotations:

> Mathematics, rightly viewed, possesses not only truth, but supreme beauty—a beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show (Russell 1907, 60).

> A mathematician, like a painter or a poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made with ideas.
>
> [. . .]
>
> The mathematician's patterns, like the painter's or the poet's must be beautiful; the ideas like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics (Hardy 1992, 13–14).

> Theorems, proofs, entire mathematical theories, a short step in the proof of some theorem, and definitions are at various times thought to be beautiful or ugly by mathematicians. Most frequently, the word "beautiful" is applied to theorems. In the second place we find proofs; usually, but not always, a proof that is deemed to be beautiful tends to be short (Rota 1997, 171).

Many mathematicians think beauty is intrinsically valuable. For example, Montaño reports "Mathematicians such as Bertrand Russell (1919), Hardy (1967), or Paul Erdös (Hoffman 1998) have argued that beauty is among the things that make mathematics valuable in itself" (Montaño 2012, 21). This suggests that beauty is an abstract proof value.

But what, exactly, does it mean for a proof to be beautiful? Perhaps surprisingly, there has been less interest in developing a philosophical account of beautiful proofs than explanatory ones. For example, Breitenbach and Rizza (2018) note that recently there have only been "a handful of attempts" to analyze mathematical beauty.

Nonetheless a number of philosophical accounts have been given, including those by Rota (1997), Cellucci (2015), Breitenbach (2015), Giaquinto (2016), and Dutilh Novaes (2019). Here I focus on Rota's well-known account and Cellucci's account that criticizes and builds on Rota's.

3.2.1 Rota's Account of Beautiful Proofs

While it may be natural to think of mathematical beauty as something that is immediately recognized in a flash, Rota argues that this is an error which he calls the "light bulb mistake" (Rota 1997, sec. 6). Instead, he claims that the process of appreciating a beautiful piece of mathematics is much more messy and requires considerable effort. For example, "[a] great many beautiful arguments require an extensive buildup, and are long-winded" (Rota 1997, 179). Moreover, we may only recognize a proof as beautiful once we have seen other ugly or cumbersome proofs.

Rota claims that understanding the "light bulb" mistake can help us to unravel the nature of mathematical beauty. To see why, we first need to understand three other claims that he makes. The first is that mathematical beauty does not come in degrees (Rota 1997, 180). In other words, a piece of mathematics is either beautiful or it is not—there is no middle ground. The second is that mathematicians criticize each other's work by asking what it is good for (Rota 1997, 178). When they ask this, mathematicians are not indicating that they cannot verify the correctness of a theorem or a proof. Rather, although they can verify it, they feel something important is still missing: enlightenment (Rota 1997, 180–181).

Rota does not offer a full analysis of enlightenment. However, he suggests that it involves, for example, grasping the role of a mathematical statement, seeing how different mathematical statements are connected to each other and recognizing how a statement is relevant (Rota 1997, 181). Rota emphasizes that this is not just psychological. Rather it has a logical component and is objective: "The property of being enlightening is objectively attributed to certain mathematical statements and denied to others" (Rota 1997, 181). Nonetheless enlightenment, unlike beauty, comes in degrees, as it allows for some statements to be more or less enlightening than others. This leads us to Rota's third important claim: mathematicians do not like the concept of enlightenment, in part because it is fuzzy, but also because it is not amenable to formalization (Rota 1997, 181).

With Rota's three core claims identified, we can now understand how grasping the "light bulb" mistake reveals the nature of mathematical beauty on his account. When mathematicians ask "What is this good for?" they are asking for enlightenment. However, they do not want to explicitly acknowledge enlightenment because it is fuzzy and not easily formalizable. The concept of mathematical beauty, however, is not fuzzy—it does not admit degrees. Moreover, in the grip of the "light bulb" mistake, mathematicians think that grasping beauty is instantaneous and forget how messy the process really is. They thus talk about mathematical beauty when they really mean enlightenment: "The comfortable one-shot but misleading idea of mathematical beauty saves us from having to deal with the messy situation of a concept having degrees. All talk of mathematical beauty is a copout from confronting the logic of enlightenment" (Rota 1997, 182). A beautiful proof, on Rota's account, is thus really an enlightening

one: "We say that a proof is beautiful when such a proof finally gives away the secret of the theorem, when it leads us to perceive the actual, not the logical inevitability of the statement that is being proved" (Rota 1997, 182).

3.2.2 Cellucci's Account of Beautiful Proofs

Cellucci (2015) criticizes Rota's account of mathematical beauty for not providing further details about the nature of enlightenment (Cellucci 2015, sec. 6). He further notes that Rota's use of the term "enlightenment" runs dangerously close to making the same light bulb mistake that Rota decried. More precisely, Cellucci notes that "enlighten" is often used to mean "to provide insight" which has connotations of grasping something instantly (Cellucci 2015, sec. 8).

Cellucci's own proposal is to analyze mathematical beauty in terms of understanding rather than enlightenment. On his account, a piece of mathematics, such as a proof, is beautiful when it provides understanding. Understanding itself is analyzed as "the recognition of the fitness of the parts to each other and to the whole" (Cellucci 2015, sec. 8). Because of this, understanding is taken to be a truly aesthetic property on Cellucci's account (Cellucci 2015, sec. 8). Applying his account to proofs, Cellucci says "There is a fitness of the parts of a demonstration to each other and to the whole when it is clear what the whole idea of the demonstration is, what the contribution of each part of the demonstration to the whole idea is, and why such contribution is essential" (Cellucci 2015, sec. 10). A proof that exhibits such fit yields understanding and is thus beautiful.

Cellucci (2015, sec. 11) presents the solution to the problem of constructing a square with half the area of a given square as an example of a beautiful proof. The diagram in Fig. 1 shows the large, given square with solid lines which has an area equal to eight right-angled triangles. The area of the square with dotted lines, on the other hand, is easily seen to be equal to four right-angled triangles. This square therefore has half of the area of the given square. As Cellucci simply remarks "The demonstration is beautiful because it provides understanding" (Cellucci 2015, sec. 11).

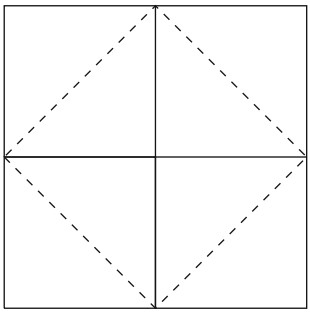


Fig. 1 The dotted square has an area half that of the large given square

3.2.3 Summary

Beautiful proofs are highly valued by mathematicians. Moreover, as mathematical beauty is intrinsically valuable, it counts as an abstract proof value. However, the topic of beautiful proofs has been less thoroughly investigated than explanatory proofs. Further work to analyze beautiful proofs and to evaluate existing accounts of such proofs would therefore help us to gain a deeper understanding of mathematical beauty.

3.3 Fit

Closely related to mathematical beauty is the notion of *fit*. Proofs that are particularly fitting are thus unsurprisingly judged to be valuable:

> [O]ne can often experience a feeling in mathematics that a claim is *right*, that a certain proof *fits* a theorem, or that a particular argument is *exactly* the one needed. Similarly one can feel that one proof has a better fit for a certain theorem than another, or that another proof, while being technically correct, jars with the statement of a theorem, being somehow the wrong kind of argument for the given claim (Raman-Sundström and Öhman 2018, 184–185, emphasis in the original).

> Scientists talking about their own work and that of other scientists use the terms "beauty," "elegance," and "economy" with the euphoria of praise more characteristically applied to painting, music, and poetry. Or there is the exclamation of recognition—the "Aha" that accompanies the discovery of a connection or an unexpected but utterly right realization in art and science. These are epithets of the sense of "*fit*"—of finding the most appropriate, evocative and correspondent expression for a reality heretofore unarticulated and unperceived, but strongly sensed and actively probed (Wechsler 1988, 1, emphasis added).

Quotes such as these already suggest that fit is an abstract value. For further evidence of this, note that elsewhere Raman-Sundström argues that mathematics "involves family relationships, and inner coherences that are deeply, and objectively, satisfying" (Raman-Sundström 2016, 284) and that "the better the fit between ideas, and the better we understand mathematics, the better it coheres" (Raman-Sundström 2016, 283). Given that fit is a property involving "deeply, and objectively, satisfying" (Raman-Sundström 2016, 284) coherences, it appears to be intrinsically valuable (Raman-Sundström is, however, somewhat hesitant about calling fit a value of proofs in the sense of being "something valued on a biological level, and/or valued as part of a good mathematical education" (Raman-Sundström 2016, 283)). Further, by providing us with understanding it gives us something else that is also intrinsically valuable.

But what, exactly, is a fitting proof? This question has been explored by Raman-Sundström (2016) and Raman-Sundström and Öhman (2018). Below we will consider the account developed in Raman-Sundström and Öhman (2018).

3.3.1 Raman-Sundström and Öhman's Account of Fitting Proofs

Raman-Sundström and Öhman (2018) offer an analysis of a variety of different ways in which a proof can be fitting. They identify three general types of fit—direct, presentational, and familial—which they then break down further, giving a total of six different criteria. Direct fit concerns the way the proof connects to the theorem,

presentational fit focuses on how the proof fits with the reader, and familial fit is about how the proof fits with mathematics more generally.

Direct fit has two subtypes: coherence and specificity (Raman-Sundström and Öhman (2018), 3–4). A proof meets the criterion of *coherence* when it uses the same language and conceptual resources as its theorem. A proof meets the criterion of *specificity* when the tools used by the proof are of the appropriate power. In other words, the proof neither resorts to brute force methods nor high-powered methods that deliver more than is needed.

Presentational fit has two flavors: level of detail and transparency (Raman-Sundström and Öhman 2018, 4–5). A proof meets the criterion of *level of detail* when it provides appropriate details, given the background of the intended reader. For example, a proof that gives every single step in a computation is likely to fail the level of detail criterion, because a reader could easily do some of these steps herself. A proof that skips too many steps is also likely to fail this criterion, because a reader cannot figure out what these steps are by himself/herself. A proof meets the criterion of *transparency* when its argument is clear, without any deus ex machina steps, so that a reader can easily understand the idea of the proof. Proofs that are highly transparent make it easy for the reader to see "what is going on" (Raman-Sundström and Öhman 2018, 188) in the argument.

Familial fit has two subtypes: generality and connectedness (Raman-Sundström and Öhman 2018, 5). A proof meets the criterion of *generality* when it can be generalized to prove other theorems. A proof meets the criterion of *connectedness* when the ideas or tools used in the proof connect to the ideas or tools used in other proofs, though not via generalization.

Raman-Sundström and Öhman illustrate how these criteria are to be applied in practice by analyzing eight proofs of four different theorems (Raman-Sundström and Öhman 2018, §3). Below we will consider one of these proofs: a trigonometric proof of the Pythagorean theorem. Recall that the Pythagorean theorem states that for a right-angled triangle with hypotenuse c and other sides a and b, the sum of the squares on the sides a and b is equal to the square on side c. A trigonometric proof, found in 2009, is given below (Raman-Sundström and Öhman 2018, 195–197). The proof relies on the subtraction formulas: $\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$ and $\sin(\alpha - \beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta$.

Proof Suppose that c is the hypotenuse of the right-angled triangle and the other sides are a and b. Let α be the angle opposite a and β be the angle opposite side b. Assume WLOG that $0 < \beta \leq \alpha < 90^{\circ}$. Then we have

$$\begin{aligned}\cos\beta &= \cos(\alpha - (\alpha - \beta)) = \cos\alpha\cos(\alpha - \beta) + \sin\alpha\sin(\alpha - \beta)\\ &= \cos\alpha(\cos\alpha\cos\beta + \sin\alpha\sin\beta) + \sin\alpha(\sin\alpha\cos\beta - \cos\alpha\sin\beta)\\ &= \left(\cos^2\alpha + \sin^2\alpha\right)\cos\beta.\end{aligned}$$

Hence $\cos^2\alpha + \sin^2\alpha = 1$ and thus, from the definitions of sine and cosine and scaling, the Pythagorean theorem is proven. □

According to Raman-Sundström and Öhman, this proof is fitting in only two of the senses they identify: specificity and level of detail (Raman-Sundström and Öhman 2018, 196–197). The proof meets the criterion for specificity because it uses appropriately powerful resources: the subtraction formulas. The proof meets the criterion for level of detail because it is concise but still provides all of the steps needed to follow the proof easily.

The proof fails to meet the remaining four criteria of fit: coherence, transparency, generality, and connectedness (Raman-Sundström and Öhman 2018, 196–197). The proof fails the criterion of coherence because the Pythagorean theorem, stated above, is given in terms of areas, while the proof makes no mention of areas and is instead entirely trigonometric. The proof fails to be transparent because it is not a very natural approach to take. For example, the reader may wonder why the proof author decided to rewrite $\cos\beta$ as $\cos(\alpha - (\alpha - \beta))$ (Raman-Sundström and Öhman 2018, 196). The proof fails to be general because the core idea of the proof does not easily generalize to other theorems. Finally, the proof is not connected because the only family of proofs it could be said to belong to are those that use trigonometric identities. However, this does not really seem to be a meaningful family, given that it is so broad (Raman-Sundström and Öhman 2018, 196).

3.3.2 Summary

Mathematicians value proofs that are fitting, where a proof can be fitting in a variety of different ways. As fit involves deeply pleasing coherences, it appears to be intrinsically valuable, but it also provides something else of intrinsic value: understanding. It is therefore an abstract proof value. However, the topic of fitting proofs has only recently attracted attention. It would thus benefit from further investigation to evaluate the criteria of fit already identified by Raman-Sundström and Öhman and to look for additional ways in which a proof can be fitting.

3.4 Motivation

The final abstract value of proofs that we will consider is motivation. Mathematicians care about motivation, as evidenced by the fact that they often make explicit remarks about the motivational power of a proof. For example:

> Here we offer a new and motivated proof of Eisenstein's theorem, relying on the basic theory of differential operators for Jacobi forms together with a classical identity for the first quasi-period of a lattice (Rolen 2015, abstract).

> We present a short and motivated proof of Riemann's functional equation for Riemann's zeta function (Knopp and Robins 2001, 1915).

> We give a generalization for Hajós' theorem. The proof of this generalization is simpler and the steps are better motivated than the proof of the original result (Corrádi and Szadó 1993, abstract).

Motivated proofs, like explanatory ones, provide understanding. For example, Pólya remarks that when an unmotivated proof of Carleman's inequality was supplemented with discussion designed to make it more motivated, "The derivation [. . .] became also more understandable" (Pólya 1949, 690). As motivated proofs provide us with understanding, which is intrinsically valuable, motivational power is an abstract value.

But what makes a proof motivated? Pólya (1949), Sandborg (1997), and Morris (2020) have all tackled this question. I will present the accounts of Pólya and Morris below.

3.4.1 Pólya's Account

As mentioned above, Pólya discusses motivation in the context of a proof of Carleman's inequality. Carleman's inequality states the following: Let a_1, a_2, a_3, . . . be a sequence of nonnegative real numbers, not all equal to 0. Then $\sum_{n=1}^{\infty}(a_1a_2a_3\ldots a_n)^{\frac{1}{n}} < e\sum_{n=1}^{\infty}a_n$. The proof, presented below, is taken verbatium from Pólya (1949, 684–685):

Proof Define the numbers $c_1, c_2, c_3, \cdots$ by

$$c_1c_2c_3\cdots c_n \;=\; (n+1)^n$$

for $n = 1, 2, 3, \cdots$. We use this definition, then the inequality between the arithmetic and the geometric means, and finally the fact that the sequence defining e, the general term of which is $[(k+1)/k]^k$, is increasing. We obtain

$$\begin{aligned}
\sum_{n=1}^{\infty}(a_1a_2\cdots a_n)^{1/n} &= \sum_{n=1}^{\infty}\frac{(a_1c_1a_2c_2\cdots a_nc_n)^{1/n}}{n+1}\\
&\le \sum_{n=1}^{\infty}\frac{a_1c_1+a_2c_2+\cdots+a_nc_n}{n(n+1)}\\
&= \sum_{k=1}^{\infty}a_kc_k\sum_{n\ge k}\frac{1}{n(n+1)}\\
&= \sum_{k=1}^{\infty}a_kc_k\sum_{n=k}^{\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right)\\
&= \sum_{k=1}^{\infty}a_k\frac{(k+1)^k}{k^{k-1}}\frac{1}{k}\\
&< e\sum_{k=1}^{\infty}a_k.
\end{aligned}$$

□

Although this proof is perfectly correct, Pólya notes that readers will be puzzled by the very first step: the introduction of the auxiliary c_i sequence. He imagines that complaints could reasonably be made, for example:

> It pops up from nowhere. It looks so arbitrary. It has no visible motive or purpose (Pólya 1949, 685).

> Look here, I am not here just to admire you. I wish to learn how to do problems by myself. Yet I cannot see how it was humanly possible to hit upon your . . . definition. So what can I learn here? How could I find such a . . . definition by myself? (Pólya 1949, 685)

Pólya suggests that reactions like these indicate that a reader wants the steps in a motivated proof to be not just correct, but also *recognizably appropriate* (Pólya did not introduce the specific term "recognizably appropriate"—I am introducing it here for clarity and convenience. Instead, Pólya notes that a step must be appropriate and should also "*appear so* to the reader" (Pólya 1949, 685, emphasis in the original)). For Pólya, a proof step is appropriate "if it is essentially connected with the purpose, if it brings us nearer to the goal" (Pólya 1949, 687). Note, however, that it is crucial that the steps in a proof are *recognizably* appropriate if the proof is to be well motivated. Indeed, the step that introduces the c_i sequence in the unmotivated proof of Carleman's inequality *is* appropriate, though it fails to be *recognizably* so. In Pólya's words: "this step [that introduces the c_i sequence] is certainly appropriate. In fact, the argument based on this definition proves the proposed theorem, and proves it rather quickly and clearly. The trouble is that the step in question, although vindicated in the end, does not appear as justified from the start" (Pólya 1949, 686). Unfortunately, Pólya does not provide further analysis of what it means for a step to be recognizably appropriate.

3.4.2 Morris's Account

Morris (2020) suggests that there are two related, but distinct, reasons why the introduction of the c_i sequence in Pólya's proof is problematic. The first reason is that the role of the step, that is the task it is meant to perform in the overall argument, is unclear (Morris 2020, 29). In other words, it is hard for a reader to identify how exactly the c_i sequence interacts with other parts of the proof to ensure the result follows. The second reason is that the insight that led to the introduction of the c_i sequence, i.e., where it came from, is unclear (Morris 2020, 30). In other words, it is hard for a reader to figure out what led the proof author to come up with and introduce that sequence.

This leads Morris to suggest that a proof is motivated if and only if the role of each step and the insight behind each step are clear, i.e., the reader can identify the tasks each step is meant to carry out and where each step could have come from (Morris 2020, 31). Whether the role of and insight behind each step are clear depends on the audience and so the notion of motivated proofs she develops is a relative one. She takes it to be relative to a given mathematical *context*, that is, a specific

mathematical background that readers of the proof are assumed to have. So while the proof of Carleman's inequality is unmotivated for students with only a limited mathematical background or context, it may be motivated for experts with an extensive background in inequalities.

3.4.3 Summary

Mathematicians value proofs that are well motivated. Like explanation, motivation provides understanding, which is intrinsically valuable, and so motivation is an abstract proof value. To my knowledge, however, motivation has only been investigated by Pólya, Sandborg, and Morris and would therefore benefit from further work. For example, Morris notes that there may be other senses in which a proof can be motivated which "focus on . . . a proof's interestingness or its connections to other results" (Morris 2020, 28). Developing accounts of such alternative forms of motivation would therefore deepen our understanding of this abstract proof value.

4 Relationships Between Abstract and Practical Values

A proof can have both practical and abstract value. Recall from the discussion of practical values in Sect. 2.2 that *all* proofs have the practical values of making mathematical knowledge accessible, systematizing mathematics and safeguarding mathematics against mistakes. Proofs with abstract value thus always have these practical values as well. However, they are often associated with additional practical values, as I will discuss below.

4.1 Explanation

It seems plausible to think that an explanatory proof may explain its theorem by introducing new concepts and methods and thus have the practical value of helping the reader to learn new ways of thinking. Further, it seems reasonable to think, as De Villiers does, that since an explanatory proof "reveals the underlying properties and reasons for the validity of a result, [it] may upon further reflection, sometimes lead to the discovery of immediate generalizations, specializations or other variations" (de Villiers 2017, 2). In other words, it seems plausible to think that explanatory proofs also have the practical value of promoting new discoveries. However, Morris's (forthcoming) arguments suggest that the situation is not quite so straightforward.

First consider the practical value of helping mathematicians learn new ways of thinking. Morris (forthcoming) argues that to help mathematicians learn tools like new concepts and methods from a proof, the proof must be well designed and presented so that it manages information well. She claims that, in general, current accounts of explanatory proofs are *not* sensitive to such issues, meaning that proofs deemed explanatory by such accounts may manage information poorly and make it

difficult for readers to learn new tools from them. She provides a concrete example in the form of a proof of Fermat's little theorem that, she argues, is explanatory according to Lange's account but which does a bad job of managing information, making it difficult for readers to learn the tools it uses. If Morris is right, then while some explanatory proofs may teach mathematicians new tools, others will be poorly designed and presented and so will not. We thus cannot say that explanatory proofs in general have this practical value.

Next consider the benefit of stimulating discovery. Recall that, in the quote from De Villiers above, explanatory proofs were *sometimes* said to lead to new discoveries. This qualification is important—we should not expect explanatory proofs to have this benefit generally (Although see Cellucci (2008, 2014) for an account of mathematical explanation that connects explanation directly to discovery). Indeed Lange is quite explicit that something can be explanatory without leading to further discoveries: "A new proof technique can explain why some theorem holds even if that technique allows no new theorems to be proved" (Lange 2014, 523). Further, Morris's argument about the practical value of helping mathematicians to learn new tools seems to apply here as well. If accounts of explanatory proofs are not sensitive to issues of design and presentation, then, at least according to these accounts, there can be explanatory proofs that manage information poorly. Although such a proof will still identify the reason why the theorem is true, it will not make it easy for the reader to grasp that reason. If the reader does not grasp this, however, it is difficult to see how the proof can help her to make further discoveries.

4.2 Beauty

There is reason to think that beautiful proofs often have the practical values of stimulating new discoveries and helping mathematicians to learn new ways of thinking.

First consider the practical value of stimulating new discoveries. Cellucci (2015, sec. 15) argues that there is a relationship between beautiful *hypotheses* and discovery, but there is reason to think that there is also a connection between beautiful *proofs* and discovery. On Cellucci's account, beautiful proofs provide understanding by making clear what the idea of the argument is, how each part of the proof contributes to it, and why each such contribution is necessary. Plausibly this may suggest ways to adapt or generalize the argument and thus put mathematicians in a good position to discover and prove new theorems.

Moreover, because of their connection to understanding, beautiful proofs may plausibly introduce mathematicians to new ways of thinking. For example, by making it clear how its different parts fit together, a beautiful proof may help mathematicians to refine intuitions or mental models of the mathematics in question. Beautiful proofs may also promote understanding of how their parts fit together by introducing new methods or concepts, thus providing those who read them with new mathematical tools.

4.3 Fit

There are potential connections between the criteria of fit identified by Raman-Sundström and Öhman and the practical values of helping mathematicians to learn new ways of thinking and making new discoveries.

Consider first fitting proofs that meet the criterion of coherence. Such proofs use the same language and conceptual resources as the theorem and so are very similar to proofs that are *pure*. For example, Detlefsen roughly characterizes purity as "enforc [ing] a certain symmetry between the conceptual resources used to prove a theorem and those needed for the clarification of its *content*. The basic idea is that the resources of proof ought ideally to be restricted to those which determine its content" (Detlefsen 2008, 192–193, emphasis in original). It has been claimed that pure proofs can help stimulate new discoveries since they serve as "a more effective instrument for gaining further knowledge" (Detlefsen 2008, 191). Detlefsen unpacks the reasoning behind such a claim as follows:

> The idea seems to be that (i) a specialist is most likely to do something significant with any solution to a problem in her speciality, [. . .] (ii) this advantage is enhanced if a solution is given in terms of concepts most familiar to her, and [. . .] (iii) the purer the proof the more it conforms to this type of solution (Detlefsen 2008, 190–191)

Given the similarity between pure and coherent proofs, if the former stimulate new discoveries, we should expect the latter to as well. However, there is reason to be skeptical of the claim that pure proofs stimulate new discoveries, since they may be harder to understand than impure proofs. As an example of a pure proof that may be more difficult to understand than impure proofs of the same theorem, we can consider Selberg's (1949) proof of the prime number theorem. The prime number theorem is a result in number theory describing the distribution of the primes. The original proofs by Hadamard (1896) and de la Vallée Poussin (1896) made heavy use of analysis. This means they are impure, since analysis involves conceptual resources that go beyond the statement of the theorem. Selberg's proof did not use analysis and was thus much purer than previous proofs. However it has also been described as "more difficult to understand" (Hua 2012, 217) than more modern proofs that do use analysis and are thus impure.

Presumably, if a proof is harder to understand, regardless of whether it involves "concepts most familiar to" (Detlefsen 2008, 191) a specialist, it will make it more difficult to adapt or generalize to make new discoveries than an impure proof. The argument that coherent proofs promote new discoveries because they are like pure proofs is thus unconvincing.

Fitting proofs that are specific may have the practical benefit of helping mathematicians learn new ways of thinking. For example, Raman-Sundström and Öhman note that proofs that score high on the criterion of specificity use tools that go "above and beyond the standard brute-force methods for approaching a proof of a theorem at hand" (Raman-Sundström and Öhman 2018, 187). Specific proofs may thus help the reader to learn new mathematical tools, at least if they are appropriately designed and

presented. As we have seen in Sect. 2.2, introducing mathematicians to new mathematical tools often leads to new discoveries. Proofs that satisfy the criterion of specificity may therefore stimulate new discoveries as well.

Fitting proofs that have an appropriate level of detail may not, in virtue of this alone, teach readers new ways of thinking or prompt them to make new discoveries. However, the criterion of level of detail is still important to the attainment of these practical benefits. Suppose, for example, a proof has too much or too little detail. Then the proof manages information poorly. A proof that manages information poorly can be confusing to the reader and thus prevent him/her from learning new tools or making new discoveries from it. So while level of detail may not by itself bring practical benefits, it contributes to their attainment.

Transparent proofs seem likely to introduce a reader to new ways of thinking. Raman-Sundström and Öhman remark that "if a proof is transparent, a reader with the appropriate background should be in *an ideal position to grasp the ideas of the proof*" (Raman-Sundström and Öhman 2018, 188, emphasis added). Ensuring that readers can grasp the ideas of the proof helps them to refine their intuitions and mental models about relevant mathematical concepts and potentially to understand new methods and techniques used within the proof as well. Given the connection between helping mathematicians learn new ways of thinking and making new discoveries, we can expect transparent proofs to stimulate new discoveries as well.

Fitting proofs that score highly on the criterion of generality are likely to have the practical value of promoting new discoveries, at least if they are appropriately designed and presented. After all, the definition of proofs that score highly on this criterion is that they can be generalized to prove new theorems!

Finally, the criterion of connectedness can plausibly help readers to develop new ways of thinking, so long as mathematicians reading the proof are in a position to recognize a fitting proof's connectedness. If a mathematician is able to recognize a proof's connectedness, then by seeing that the ideas or tools it uses are shared by a family of other proofs, the reader can refine their mental models or intuitions. Further, if a reader comes to see that a technique or method can be used in a variety of different circumstances, their may learn that the scope of that tool is more extensive than they first thought. While this is not exactly the same as learning a completely new method, it is very similar to it. Given that proofs that help mathematicians learn new ways of thinking often stimulate new discoveries, connected proofs should also lead mathematicians to new discoveries, at least if their connectedness is recognizable.

4.4 Motivation

There are strong connections between motivated proofs and the practical values of helping mathematicians to learn new ways of thinking and stimulating new discoveries.

Morris claims that motivated proofs help their readers to understand how to use the tools they contain as well as the circumstances under which it is useful to apply

them, both of which are important to successfully learning new mathematical tools. In particular, the reader of a motivated proof can identify the tasks each step performs which helps them learn how to use new resources the proof contains (Morris 2020, 39). The reader can also identify where each step could have reasonably come from, which helps them to get an idea of the circumstances under which it might be helpful to apply new resources the proof uses (Morris 2020, 39). Motivated proofs are thus likely to help mathematicians learn new ways of thinking.

As we have seen in Sect. 2.2, proofs that help mathematicians learn new ways of thinking often also stimulate new discoveries. Further, Morris argues explicitly that identifying the tasks a step is meant to perform and where it could have reasonably come from can stimulate new discoveries (Morris 2020, 39).

4.5 Summary

We have seen that proofs with abstract values are connected, to varying degrees, to the practical values of learning new tools and making new discoveries. We also noted that, just by being proofs, proofs with abstract values have the practical values of helping to make mathematical knowledge accessible, systematizing mathematics and safeguarding mathematics from mistakes. This leaves only one practical value unmentioned: illustrating the power of different methodologies. There is, in fact, little reason to think that proofs that have the abstract values discussed here will also tend to have this practical value. Indeed, there does not seem to be anything in the nature of explanation, beauty, fit, or motivation that suggests that proofs with these values often introduce or utilize an entirely different methodology.

5 Relationships Between Abstract Values

Given the discussion of the relationships between practical and abstract values of proofs, we may wonder whether, and to what extent, the abstract values themselves are related (Important empirical work on relationships between proof values has been conducted by Inglis and Aberdein (2015) and Inglis and Aberdein (2016)). In particular, are proofs with a given abstract value likely to also have a different abstract value? For example, are explanatory proofs also likely to be beautiful? Below I consider the relationship between each of the abstract values discussed in Sect. 3. I start with, and provide the most detailed discussion of, the relationship between fit and motivation and provide shorter sketches of the relationships between the other abstract values.

5.1 Fit and Motivation

Raman-Sundström and Öhman's criteria of presentational fit—transparency and level of detail—have strong connections to Morris's account of well-motivated

proofs. First consider transparency. Morris (2020) argues that a well-motivated proof is likely to be transparent. If a proof is well motivated, relative to a given context, then the reader can identify the tasks each step is intended to perform as well as where they came from and this means they should also be able to see "what is going on" in the proof, thus making it transparent (Morris 2020, 43). However, Morris suggests that there may be transparent proofs that fail to be well motivated: "perhaps it is possible to understand the core ideas of a proof and have a high level grasp of 'what is going on' within it while still finding a few steps involving the technical details to be puzzling" (Morris 2020, 43).

Next consider the level of detail. Given that a well-motivated proof, on Morris's account, must manage information well, we can also expect well-motivated proofs to be fitting in the sense of having an appropriate level of detail. However, having an appropriate level of detail is not, by itself, sufficient for a proof to be well motivated. Presumably the trigonometric proof of the Pythagorean theorem given in Sect. 3.3 will fail to be motivated relative to a standard geometric context, but, as was noted there, it does satisfy the level of detail criterion.

There are also strong connections between the two types of familial fit—generality and connectedness—and Morris's account of motivated proofs. In particular, it seems plausible to suggest that motivated proofs are likely to score highly on these two criteria. As we have seen, Morris argues that motivated proofs are likely to have the practical values of helping mathematicians to learn new tools and promoting new discoveries. If a motivated proof promotes discoveries because the underlying idea generalizes, then it will score highly on generality. And if a motivated proof promotes discoveries because it helps the reader to learn and subsequently reuse new tools that it contains, then it will likely score highly on connectedness. However proofs that score highly on generality or connectedness will not necessarily be motivated, at least on Morris' account. This is because a proof can presumably be generalizable or connected and yet fail to be well designed and presented, making it difficult for the reader to grasp the role of and insight behind each step.

There are fewer connections between the criteria of direct fit—coherence and specificity—and well-motivated proofs. First consider coherence. As we have seen in Sect. 3.3.2, coherence is similar to purity, in that a coherent proof uses the same conceptual resources and language as the theorem. However it seems plausible to imagine a coherent proof that nonetheless fails to be well motivated on Morris's account because it is too difficult for a reader to identify the role of and insight behind its steps. Selberg's proof of the prime number theorem, already discussed in Sect. 3.3.2, may be such an example. This proof scores high on coherence because it does not use analysis, but it has been described as "difficult to understand" (Hua 2012, 217). It is thus likely that it will be difficult for a reader to identify the role of each step as well as the insight behind it relative to certain contexts, meaning that it is not well motivated. In fact, the mathematician Levinson wrote an article which aimed "to give a self-contained and motivated account of an elementary proof of the prime number theorem" (Levinson 1969, 225). Presumably he would have no reason to do this if Selberg's original proof had been motivated. Similarly, a

motivated proof does not need to be coherent. As mentioned earlier, there are proofs of the prime number theorem that use analysis, and so fail to be coherent, but are easier to understand, and so may be motivated relative to an appropriate context.

Consider now specificity. Again there is little reason to expect a connection between specificity and well-motivated proofs. The trigonometric proof of the Pythagorean theorem, given in Sect. 3.3, for example, scored highly on specificity, but would likely be unmotivated relative to a standard geometric context. Similarly, proofs that are motivated may fail to be specific. For example, the well-known proof of the irrationality of $\sqrt{2}$ that uses the unique factorization theorem is plausibly well motivated relative to appropriate contexts. However, Raman-Sundström and Öhman argue that this proof is not specific: "[The unique factorization theorem] is adequate to prove the theorem, but is overly powerful. In fact, only the number of factors 2 in [the prime decomposition of certain numbers] is needed for the conclusion to be drawn" (Raman-Sundström and Öhman 2018, 191).

5.2 Explanation and Beauty

It seems plausible to think that there is a strong connection between proofs that are beautiful and proofs that are explanatory. Lange (2016b), for example, argues that explanatory proofs, on his account, are also beautiful. He writes that explanatory proofs "derive their beauty from exactly what gives them their explanatory power, namely, from their exploiting precisely the feature in the setup that is salient in the theorem (such as its symmetry, unity, or simplicity)" (Lange 2016b, 45). However, Lange remains agnostic about whether beautiful proofs are also explanatory, noting that his account "does not rule out beautiful, nonexplanatory proofs of theorems that have explanations" (Lange 2016b, 46).

On the other hand, Giaquinto (2016) suggests that there are reasons to be skeptical of the claim that explanatory proofs are beautiful. In particular, he argues that many explanatory proofs are also "facile" and that, for this reason, they fail to be beautiful (Giaquinto 2016, 67–68). Giaquinto also suggest we should be skeptical of the claim that beautiful proofs are explanatory, at least until we have more data. For while it may initially seem plausible to suggest that beautiful proofs tend to be explanatory, he presents a proof of the infinitude of the primes due to Erdös as a potential counterexample. This proof has been described as beautiful, but Giaquinto argues that it fails to be explanatory (Giaquinto 2016, 68–70).

5.3 Explanation and Fit

Raman-Sundström and Öhman (2018) suggest that there are relationships between explanatory proofs and their criteria of coherence, transparency, and connectedness. Recall that, on Steiner's account, an explanatory proof clearly depends on a characterizing property of something mentioned in the theorem. Raman-Sundström and

Öhman then note that this relates to coherence: "A proof that has the same terms as a theorem, which is how we have characterized coherence, seems similar to a proof that evidently gives rise to a particular result" (Raman-Sundström and Öhman 2018, 205). They further suggest that connectedness is related to both Steiner's and Kitcher's account of explanatory proofs, since familial relationships enter into Steiner's account via characterizing properties and into Kitcher's via the explanatory store (Raman-Sundström and Öhman 2018, 206).

Raman-Sundström and Öhman also suggest that their notions of coherence and transparency have a place in Lange's account. Recall that, for Lange, an explanatory proof traces a salient feature, like symmetry or unity, in the result back to its setup. Raman-Sundström and Öhman argue that coherence is similar to how Lange understands unity and that transparency relates to his notion of salience. Lange describes a unified proof as one which "exploits a property that all of the cases covered by the theorem have in common and treats all of those cases in the same way" (Raman-Sundström and Öhman 2018, Lange quoted on pg 206) and Raman-Sundström and Öhman take this to be connected to coherence in a "fairly straightforward" (Raman-Sundström and Öhman 2018, 206) way. They further suggest that coherence can justify the salience of a feature of the result and that transparency can help us to access "the underlying ideas" (Raman-Sundström and Öhman 2018, 206).

5.4 Explanation and Motivation

While we might expect there to be a close relationship between explanatory proofs and well-motivated ones given that both types of proof are connected to understanding, Morris (2015, 2020) argues that there are significant differences between them. In particular, she notes that accounts of explanatory proofs do not usually take into consideration whether a reader will *recognize* the explanation. However, for a proof to be well-motivated with respect to a given context, the reader must identify and thus recognize the tasks each step performs as well as where each step came from (Morris 2020, 42). She also points out that explanatory and well-motivated proofs differ in their practical value, as motivated proofs are likely to successfully convey new ways of thinking to their readers but the same cannot be said for explanatory proofs (Morris 2020, 42).

Morris further provides examples to illustrate that there can be explanatory proofs which are not well motivated and vice versa. For example, Steiner remarks that many inductive proofs will fail to be explanatory on his account but Morris suggests that inductive proofs can be well motivated relative to a suitable context (Morris 2015, 108–110). Hafner and Mancosu (2008) argue that proofs using a decision procedure in the theory of real-closed fields are explanatory according to Kitcher's account (and use this to criticize his account) but Morris argues that, because these proofs can be inordinately long and tedious, they are not well designed and so may fail to be motivated (Morris 2015, 106–108).

5.5 Beauty and Fit

It seems that there are strong connections between beautiful proofs and fitting proofs, especially given that Cellucci's account of beauty was analyzed in terms of fit. As we saw in Sect. 3.3, a proof is beautiful on Cellucci's account when the parts of a proof fit together, i.e., "when it is clear what the whole idea of the demonstration is, what the contribution of each part of the demonstration to the whole idea is, and why such contribution is essential" (Cellucci 2015, sec. 10). Proofs that are beautiful, on Cellucci's account, thus also seem to be what Raman-Sundström and Öhman call transparent, since they will help the reader to grasp "what is going on" (Raman-Sundström and Öhman 2018, 188) in the proof.

It is less clear, however, whether proofs that are transparent will also always be beautiful. For example, perhaps it is possible to grasp "what is going on" in the proof at a high level without fully grasping the contribution made by some of the more technical parts of the proof, or without fully grasping why they are essential. Such a proof would be transparent, at least at a high level, but would not be fully beautiful on Cellucci's account.

In addition to a connection between beauty and transparency, Raman-Sundström and Öhman note that there may be connections between beauty and level of detail and connectedness. As level of detail connects to the length of proofs, they suggest that it may also be related to beauty, given the common intuition that short proofs are more beautiful (Raman-Sundström and Öhman 2018, 207). Regarding connectedness, Raman-Sundström and Öhman suggest that a proof that scores highly on this criterion lets you see it from a different perspective, and this, in turn, may make it beautiful (Raman-Sundström and Öhman 2018, 207).

5.6 Beauty and Motivation

Morris (2020) argues that there is a connection between beautiful proofs and well-motivated proofs in the following sense: well-motivated proofs will tend to be beautiful, at least on Cellucci's account. As a motivated proof, on Morris's account, will help the readers to identify the tasks each step is intended to perform as well as where each step could have reasonably come from, they should thus be able to grasp the idea of the proof, the contribution of each of its parts, and why each part was required (Morris 2020, 43). Thus motivated proofs will be beautiful.

Perhaps it is possible, however, for there to be proofs that are beautiful on Cellucci's account, but not well motivated on Morris' account. After all, Cellucci's criteria for a beautiful proof do not explicitly include making clear where each step in the argument could have come from. Perhaps, then, there could be a proof whose overall idea is clear, whose steps have a clear contribution and are clearly needed, but which contains a step that does not make it easy to see how the author could have come up with it.

5.7 Summary

We have seen that there appear to be significant connections between many of the abstract values of explanation, beauty, fit, and motivation. However, deeper investigations are required to fully understand the nature of the relationships between them. For example, as we saw above, Giaquinto urges caution in accepting the claim that beautiful proofs are explanatory until we have more data and case studies. Further, although we saw Morris highlights important differences between motivated and explanatory proofs, she also remarks that the relationship between them is "both interesting and complicated and should be further explored" (Morris 2020, 43). Finally, Raman-Sundström and Öhman invite additional analysis on the relationships between fit and other values when they describe their own discussions on the topic as "more stubs than full articulations" (Raman-Sundström and Öhman 2018, 205).

6 Conclusion

In this chapter I have surveyed some of the wide variety of ways in which proofs can have value. I divided these values into two categories: practical and abstract. Proofs have practical value when they help mathematicians to achieve their goals, such as making new discoveries or learning new tools. Proofs have abstract value when they have qualities that are (or provide something that is) intrinsically valuable, such as beauty or explanatory power.

As we have seen, there are important interconnections between the values proofs can have. Some practical values are closely associated with others, so that proofs with one value often also have the other. For example, proofs that have the practical value of illustrating new methodologies will likely also have the practical value of helping mathematicians to learn new ways of thinking. Abstract values are also closely associated with practical values. For example, motivated proofs are likely to also have the practical values of helping mathematicians learn new mathematical tools and stimulating new discoveries. Finally some of the abstract values themselves are closely related. For example, the abstract value of fit has connections to beauty, explanation, and motivation.

An important future task is to investigate more deeply these individual proof values as well as the relationships between them. For example, we have seen that there is little consensus over the nature of explanatory proofs, while beautiful, fitting, and motivated proofs have received correspondingly little attention. These values would thus all benefit from further analysis. As another example, while intuitively there seems to be a close relationship between beautiful proofs and explanatory ones, as Giaquinto pointed out, we need a more thorough investigation to be able to make any definitive conclusions.

While a wide variety of proof values have been identified and analyzed, there are still more to be investigated. As just one example, mathematicians often indicate,

even in their research papers, that they value proofs that are "intuitive." Below are some representative quotes:

> We give a simple, intuitive proof of a known theorem (Scovel 1985, 751).

> Our proof here is, hopefully, more intuitive and easier to motivate than the existing proofs (Granville et al. 2018, sec. 1).

> The purpose of this paper is to provide a simple and intuitive proof of Weber's result (Derks 2005, sec. 1).

A second future task is thus to identify additional proof values and to develop precise accounts of them.

Continuing to investigate the value of proofs has three main benefits. The first two are practical: (i) further analysis may yield advice about how to *write* proofs with particular values and (ii) understanding the value of proofs may be of use in mathematics education. As an example of (i), Morris's (2020) analysis of motivated proofs includes discussion of how to write them, drawing on work by Avigad (2020) and Sieg (2010) on the design of mathematics. Regarding (ii), mathematics teachers may wish to adopt instructional strategies that emphasize the value of proofs beyond simply establishing that their theorems are true (see, e.g., Hanna (1995)). Having more precise accounts of the values of proofs can be helpful in this regard (see e.g. Hanna (2018)). The third benefit of continued investigation into the value of proofs is more abstract: proofs are central and unique to mathematics and so by deepening our understanding of their value, we similarly deepen our understanding of mathematics itself.

7 Cross-References

▶ Formal Proofs in Mathematical Practice
▶ The Social Epistemology of Mathematical Proof

References

Arana A (2015) On the depth of Szemerédi's theorem. Philos Math III 23(2):163–176
Avigad J (2006) Mathematical method and proof. Synthese 153(1):105–159
Avigad J (2018) Opinion: the mechanization of mathematics. Not Am Math Soc 65(06):1
Avigad J (2020) Modularity in mathematics. Rev Symb Logic 13(0):47–79
Baron S, Colyvan M, Ripley D (2019) A counterfactual approach to explanation in mathematics. Philos Math III, December
Breitenbach A (2015) Beauty in proofs: Kant on aesthetics in mathematics. Eur J Phil Sci 23(4): 955–977
Breitenbach A, Rizza D (2018) Introduction to special issue: aesthetics in mathematics. Philos Math III 26(2):153–160
Cellucci C (2008) The nature of mathematical explanation. Stud Hist Philos Sci B 39(2):202–210

Cellucci C (2014) Explanatory and non-explanatory demonstrations. In: Logic, methodology and philosophy of science. Proceedings of the Fourteenth International Congress. College Publications, pp 201–218
Cellucci C (2015) Mathematical beauty, understanding, and discovery. Found Sci 20(4):339–355
Corrádi K, Szadó S (1993) A generalized form of Hajós' theorem. Commun Algebra 21(11):4119–4125
D'Alessandro W (forthcoming) Mathematical explanation beyond explanatory proof. Br J Philos Sci
Dawson JW (2006) Why do mathematicians re-prove theorems? Philos Math III 14(3):269–286
Dawson JW Jr (2015) Why prove it again? Alternative proofs in mathematical practice. Birkhäuser, Cham
de La Vallée Poussin CJ (1896) Recherches analytiques de la théorie des nombres premiers. Ann Soc Sci Bruxelles 20(0):183–256
de Villiers M (2017) An explanatory, transformation geometry proof of a classic treasure-hunt problem and its generalization. Int J Math Educ Sci Technol 48(2):260–267
Derks J (2005) A new proof for weber's characterization of the random order values. Math Soc Sci 49(3):327–334
Detlefsen M (2008) Purity as an ideal of proof. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 179–197
Detlefsen M, Arana A (2011) Purity of methods. Philosophers' Imprint 11
Dutilh Novaes C (2019) The beauty (?) of mathematical proofs. In: Aberdein A, Inglis M (eds) Advances in experimental philosophy of logic and mathematics. Bloomsbury Academic, London, pp 63–93
Easwaran K (2009) Probabilistic proofs and transferability. Philos Math 17(3):341–362
Frans J, Weber E (2014) Mechanistic explanation and explanatory proofs in mathematics. Philos Math III 22(2):231–248
Giaquinto M (2016) Mathematical proofs: the beautiful and the explanatory. J Humanist Math 6(1): 52–72
Granville A, Harper AJ, Soundararajan K (2018) A more intuitive proof of a sharp version of halász's theorem. Proc Am Math Soc 146(10):4099–4104
Gray J (2015) Depth – a Gaussian tradition in mathematics. Philos Math III 23(2):177–195
Hadamard J (1896) Sur la distribution des zéros de la fonction ζ(s) et ses conséquences arithmétiques. Bulletin Soc Math France 24(0):199–220
Hafner J, Mancosu P (2005) The varieties of mathematical explanation. In: Mancosu P, Jørgensen KF, Pedersen SA (eds) Visualization, explanation and reasoning styles in mathematics. Springer, Dordrecht, pp 215–250
Hafner J, Mancosu P (2008) Beyond unification. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 151–178
Haigh G (1980) A 'Natural' approach to pick's theorem. Math Gaz 64(429):173–177
Hanna G (1995) Challenges to the importance of proof. Learn Math 15(3):42–49
Hanna G (2018) Reflections on proof as explanation. In: Stylianides AJ, Harel G (eds) Advances in mathematics education research on proof and proving: an international perspective. Springer International Publishing, Cham, pp 3–18
Hanna G, Larvor B (2020) As thurston says? On using quotations from famous mathematicians to make points about philosophy and education. ZDM, March
Hardy GH (1967) A Mathematician's apology. Cambridge University Press, reissue edition edition
Hardy GH (1992) A Mathematician's apology. Cambridge University Press, January
Heawood PJ (1890) Map-Color theorem. Quar J Math 24(0):332–338
Hoffman P (1998) The man who loved only numbers: the story of Paul Erdös and the search for mathematical truth. Hyperion, 1st edn, January
Hua LK (2012) Introduction to number theory. Springer Science & Business Media
Inglis M, Aberdein A (2015) Beauty is not simplicity: an analysis of mathematicians' proof appraisals. Philos Math III 23(1):87–109

Inglis M, Aberdein A (2016) Diversity in proof appraisal. In: Mathematical cultures. Springer International Publishing, pp 163–179
Inglis M, Meja-Ramos JP (2019) Functional explanation in mathematics. Synthese, May
Kempe AB (1879) On the geographical problem of the four colours. Am J Math 2(3):193–200
Kitcher P (1989) Explanatory unification and the causal structure of the world. In: Kitcher P, Salmon W (eds) Scientific explanation. University of Minnesota Press, Minneapolis, pp 410–505
Knopp M, Robins S (2001) Easy proofs of Riemann's functional equation for $\zeta(s)$ and of lipschitz summation. Proc Am Math Soc
Lange M (2009) Why proofs by mathematical induction are generally not explanatory. Analysis 69(2):203–211
Lange M (2014) Aspects of mathematical explanation: symmetry, unity, and salience. Philos Rev 123(4):485–531
Lange M (2015) Depth and explanation in mathematics. Philos Math III 23(2):196–214
Lange M (2016a) Because without cause: non-causal explanations in science and mathematics. Oxford University Press, New York
Lange M (2016b) Explanatory proofs and beautiful proofs. J Humanist Math 6(1):8–51
Lemmermeyer F (n.d.) Proofs of the quadratic reciprocity law. https://www.rzuser.uni-heidelberg.de/hb3/rchrono.html. Accessed 30 Mar 2020
Levinson N (1969) A motivated account of an elementary proof of the prime number theorem. Am Math Monthly 76(3):225–245
Loomis ES (1968) The Pythagorean proposition. National Council of Teachers of Mathematics
Marsden JE (1985) Review of Mill, F., "An effective potential for classical Yang-Mills fields as outline for bifurcation on gauge orbit space". Ann Phys Rehabil Med 149:179–202. (MR# 85c:58080)
Montaño U (2012) Ugly mathematics: why do mathematicians dislike computer-assisted proofs? Math Intell 34(4):21–28
Morris RL (2015) Appropriate steps: a theory of motivated proofs. PhD thesis, Carnegie Mellon University
Morris RL (2020) Motivated proofs: what they are, why they matter and how to write them. Rev Symb Logic 13(0):23–46
Morris RL (forthcoming) Do mathematical explanations have instrumental value? Synthese, February
Pincock C (2015) The unsolvability of the quintic: a case study in abstract mathematical explanation. Philos' Imp 15(3):1–19
Pólya G (1949) With, or without, motivation? Am Math Monthly 56(10):684–691
Raman-Sundström M (2016) The notion of fit as a mathematical value. In: Mathematical cultures. Springer International Publishing, Berlin, pp 271–285
Raman-Sundström M, Öhman L-D (2018) Mathematical fit: a case study. Philos Math 26(2):184–210
Rav Y (1999) Why do we prove theorems? Philos Math 7(1):5–41
Resnik MD, Kushner D (1987) Explanation, independence and realism in mathematics. Br J Philos Sci 38(2):141–158
Rolen L (2015) A new construction of eisenstein's completion of the weierstrass zeta function. Proc Am Math Soc 144(4):1453–1456
Rota G-C (1997) The phenomenology of mathematical beauty. Synthese 111(2):171–182
Russell B 1907 The study of mathematics. The New Quarterly 1
Russell B (1919) The study of mathematics. In: Mysticism and logic. Longman, London
Sandborg D (1997) Explanation in mathematical practice. PhD thesis, University of Pittsburgh
Scovel JC (1985) A simple intuitive proof of a theorem in degree theory for gradient mappings. Proc Am Math Soc 93(4)
Selberg A (1949) An elementary proof of the prime-number theorem. Ann Math 50(2):305–313
Sieg W (2010) Searching for proofs (and uncovering capacities of the mathematical mind). In: Feferman S, Sieg W (eds) Proofs, categories and computations: essays in honor of Grigori Mints. College Publications, London, pp 189–215

Smith SP (1991) Review of Artin, M., and Van den Bergh, M., "Twisted homogeneous coordinate rings". J Algebra 133 (1990), 249–271. (MR# 91k:14003)
Steiner M (1978) Mathematical explanation. Philos Stud 34(2):135–151
Stillwell J (2015) What does 'depth' mean in mathematics? Philos Math III 23(2):215–232
Strevens M (2011) Depth: an account of scientific explanation. Harvard University Press, Cambridge, MA
Urquhart A (2015) Mathematical depth. Philos Math III 23(2):233–241
Michael De Villiers. The role and function of proof in mathematics. Pythagoras, 24(24):17–24, 1990
Weber K (2010) Proofs that develop insight. Learn Math 30(1):32–36
Weber E, Verhoeven L (2002) Explanatory proofs in mathematics. Log Anal 45(179–180):9–180
Wechsler J (1988) On aesthetics in science (Design science collection). Birkhäuser, 1988 edition, February 1988

Formal Proofs in Mathematical Practice

Danielle Macbeth

Contents

1 Introduction 2114
2 Some Early History 2115
3 The Contemporary Debate 2118
4 More History 2123
5 What Is a Mathematical Proof? 2124
6 Resolving the Antinomy 2126
7 Conclusion 2132
References 2133

Abstract

Over the past half-century, formal, machine-executable proofs have been developed for an impressive range of mathematical theorems. Formalists argue that such proofs should be seen as providing the fully worked out proofs of which mathematicians' proofs are sketches. Nonformalists argue that this conception of the relationship of formal to informal proofs cannot explain the fact that formal proofs lack essential virtues enjoyed by mathematicians' proofs, the fact, for example, that formal proofs are not convincing and lack the explanatory power of their informal counterparts. Formal proofs do not yield mathematical insight or understanding, and they are not fruitful in the way that informal proofs can be, for instance, in suggesting novel approaches to old problems. And yet, there is a clear sense in which the formalist is right: Formal proofs do seem to make explicit what is otherwise taken for granted in the mathematician's reasoning. Very recent work uncovers the source of the dilemma by showing that rigor does not entail formalism insofar as rigor, unlike formalism, is compatible with contentfulness. Mathematical proofs, were they to be recast in a Leibnizian language, at once a characteristica and a calculus, would be, or could be made to be, completely gap-

D. Macbeth (✉)
Haverford College, Haverford, PA, USA
e-mail: dmacbeth@haverford.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_35

free and hence machine checkable. Because as so recast they would be grounded in mathematical ideas, they would not be machine executable. It is on just this basis that a compelling account can be given of the fact that formal proofs are mostly irrelevant to mathematicians in their practice.

Keywords

Automated deduction · *Begriffsschrift* · Computer proof · Formal proof · Formalization · Informal proof · Leibnizian universal language · Machine reasoning · Mathematical proof

1 Introduction

Formal, machine-executable proofs, envisaged already by Leibniz, are a reality. And at least some mathematicians see in such proofs the future of mathematics. Most mathematicians hold that automated deduction belongs, and always will belong, only at the fringes of mathematics. For philosophers, formal proofs of mathematical theorems constitute a problem. Such proofs are not compelling to the practicing mathematician. They cannot serve as vehicles of mathematical understanding. And they are of no use in teaching mathematics to students. The philosopher aims to know why it is given that formal proofs appear, at least, to be nothing more than fully rigorous versions of the everyday proofs that mathematicians esteem. The problem, in short, is to understand the nature of the relationship between a mathematician's proof and a formalization of that same proof. Is, for example, the formalization even correctly described as a mathematical counterpart of the proof it purports to formalize? And if it is not, how is it that something of intrinsic mathematical interest and value is transformed in the process of formalization into something that is not, apparently, mathematical at all?

According to formalists, although mathematicians do not in their practice produce properly formal, machine-executable proofs, mathematicians' proofs do serve as sketches of real, that is to say, formal, rigorous proofs. Nonformalists reject the account, and in recent decades, philosophers of mathematical practice have shown what mathematicians in their practice already intimate, that formal proofs do not play any of the roles mathematicians' proofs play in mathematical practice. Not only do they not convince or verify, but they also do not, for instance, deepen mathematical understanding, provide new resources for tackling outstanding problems, or reveal any interesting relations among apparently disparate sorts of mathematical issues and ideas. And there is a deeper problem as well, one that is of real philosophical significance outside the domain of the philosophy of mathematics. Mathematical reasoning, at least in the cases of concern here, is deductive, and the logic on the basis of which formal proofs (and computing generally) are conducted provides a complete theory of deduction. It would seem, then, that formal proofs of mathematicians' theorems must reveal the essence of mathematical reasoning, or at least something of fundamental significance to such reasoning, and thereby something of

significance to reason itself. And yet, again, formal proofs seem, on the whole, to be utterly irrelevant to mathematicians' practice. *How can this be?* Insofar as they are strictly and manifestly logical, formal proofs must reveal something fundamental about mathematical reasoning, and thereby about rationality generally. And yet, because they are strictly logical, purged of all mathematical content, mathematical concepts, ideas, and methods, formal proofs seem not to reveal anything fundamental about mathematical reasoning in practice nor, presumably, about what it is to be rational. What, then, *is* the relationship between a mathematician's proof and its fully formalized counterpart?

2 Some Early History

Although it was only developments in the nineteenth and twentieth centuries that would reveal the full nature and extent of the problem of formal proof in mathematical practice, the roots of the difficulty stretch back to the seventeenth century, to the first full flowering of the symbolic language of algebra in Descartes' *La Géométrie* (1637). Indeed, its outlines are discernable already in the ancient Greeks, in the dispute between Plato, for whom mathematical knowledge is a matter of grasping, seeing with the mind's eye the forms of things mathematical, and Aristotle, the founder of the science of logic, who looks instead to mathematical reasoning, the syllogism, to understand the nature and peculiar value of mathematical knowledge. In the seventeenth century, Descartes will take the part of Plato, arguing that it is not by reasoning but instead by intuition that one truly knows in mathematics. Leibniz, taking the part of Aristotle, holds that it is proof rather than insight and intuition that establishes truth in mathematics. In the background of both is a conception of language that had by then already been realized, at least in part, in Descartes' symbolic language.

Leibniz first explores the idea of what he thinks of as a universal language (which we here consider as a language for mathematics in particular) in his "Dissertation on the Art of Combinations" (1666). It is the idea of a language that is at once a *characteristica,* a language within which to exhibit the contents of concepts, and a *calculus,* a system of rules governing reasoning on the basis of those contents. Descartes mentions the idea of such a language in a letter written to Mersenne on November 20, 1629. To learn the positional system of Arabic numeration is to learn to write and to compute with numbers in that written system of signs no matter what one's native tongue. Perhaps, then, the same could be done for all the words of natural language – or at least, for such words that pertain to mathematics and mathematical reasoning.

The ways words of natural language are formed from sounds and other words clearly bear no internal, systematic relationship to the ways the concepts signified by the words are formed from simpler concepts. But perhaps a system of written signs could be devised in which there was such a relationship. Such a system would involve simple signs for primitive concepts and complex signs for other concepts that would show in their composition how the contents of those concepts depend on

and are built up out of the contents of the primitive concepts. Even more than the positional system of Arabic numeration that Descartes refers to in his letter to Mersenne, Descartes' symbolic algebra provides the model. In that language, one displays what it is to be, for example, an even number, $2n$ for natural number n, an odd number, $2n - 1$, the length of the hypotenuse of a right triangle the other two sides of which have lengths a and b, $\sqrt{(a^2 + b^2)}$, and so on. This expressive project is, however, only half the story. A Leibnizian universal language also provides rules governing the manipulation of those signs in chains of reasoning; it is not only a characteristica but a calculus. Again, Descartes' symbolic language provides the model. We begin, for example, with the idea of a product of two sums of integer squares. What is to be shown is that that product is itself a sum of integer squares. The first step is to formulate the concept of a product of two sums of integer squares in the language, to show in a complex of primitive signs of the language what is it to be such a product: $(a^2 + b^2)(c^2 + d^2)$. Familiar rules license successive rewritings of that collection of signs as follows.

$$\begin{aligned}
&\left(a^2 + b^2\right)\left(c^2 + d^2\right) \\
&= a^2c^2 + b^2c^2 + a^2d^2 + b^2d^2 \\
&= a^2c^2 + b^2d^2 + b^2c^2 + a^2d^2 \\
&= a^2c^2 + 2acbd + b^2d^2 + b^2c^2 - 2bcad + a^2d^2 \\
&= (ac + bd)^2 + (bc - ad)^2
\end{aligned}$$

Because this last complex of signs sets out the content of the idea of a (particular sort of) sum of two integer squares, we are done. We have demonstrated in a way that is quite rigorous and largely gap-free that a product of two sums of integer squares is itself a sum of integer squares. For Leibniz, such a chain of reasoning is a paradigm of proof, and thereby of knowledge, insofar as it "renders truth stable, visible, and irresistible, so to speak, as on a mechanical basis" (Leibniz 1675, p. 166). The symbolic and rule-governed reasoning of the new algebra provided in this way a model for a kind of mechanical reasoning, one that would, Leibniz thought, put an end to all controversies. Where there are disagreements, disputants need only take up their pens and calculate; the result of the calculation must be agreed to by all, just as it is in the case of an arithmetical calculation.

For Leibniz, this power of the calculus to "[render] truth stable, visible, and irresistible . . . as on a mechanical basis" was sufficient reason to esteem formal proof above all in mathematics. Descartes disagreed because he was less concerned with disputes and compelling agreement than he was with insight and understanding. Already in Rule Three of the *Regulae,* Descartes suggests that intuition is simpler and more certain than deduction. It is in Descartes' Second Replies (to Objections to his *Meditations),* in a discussion of the analytic and synthetic methods of demonstration in geometry, that Descartes explains the distinctive value of intuition relative to deduction.

As Descartes explains, whether the method is that of analysis or that of synthesis, the order of argumentation is the same: "the items which are put forward first must be

known entirely without the aid of what comes later; and the remaining items must be arranged in such a way that their demonstration depends solely on what has gone before" (Descartes 1641, p. 110). The reasoning in both cases is deductive, at least in a broad sense ("what comes later depends solely on which has gone before"), and it is noncircular (what is first put forward "must be known entirely without the aid of what comes later"). Nevertheless, Descartes thinks, the two methods are very different.

> Analysis shows the true way by means of which the thing in question was discovered methodically and as it were a priori, so that if the reader is willing to follow it and give sufficient attention to all points he will make the thing his own and understand just as perfectly as if he had discovered it for himself. But this method contains nothing to compel belief in an argumentative or inattentive reader, for if he fails to attend even to the smallest point, he will not see the necessity of the conclusion. Moreover, there are many truths which – although it is vital to be aware of them – the method often scarcely mentions, since they are transparently clear to anyone who gives them his attention (Descartes 1641, p. 110).

This method is, furthermore, "a priori" because it is prior to the method of synthesis.

> Synthesis . . . employs a directly opposite method where the search is, as it were, a posteriori (though the proof is often more a priori than it is in the analytic method). It demonstrates the conclusion clearly and employs a long series of definitions, postulates, axioms, theorems, and problems, so that if anyone denies one of the conclusions it can be shown at once that it is contained in what has gone before, and hence the reader, however argumentative or stubborn he may be, is compelled to give his assent (Descartes 1641, pp. 110–111).

In a synthetic demonstration, everything on which the proof depends is explicitly stated in advance in axioms, definitions, postulates, (previously demonstrated) theorems, and (previously solved) problems. Because it is, the proof compels assent by anyone, no matter how argumentative or stubborn, who accepts those axioms, definitions, postulates, theorems, and problems. But of course, one can set out everything on which a proof depends only after one has the proof itself, only after one has achieved, analytically, knowledge of the "primary notions" and has drawn (materially valid) conclusions from them. It is for just this reason that the analytic method, unlike the synthetic method, requires a reader to "make the thing has own and understand it as perfectly as if he had discovered it for himself." The analytic method employs unstated rules of material inference to draw conclusions from evidently true premises. In this case, one is to *see* that the truth of the premises entails the truth of the conclusion, but in order to see that, one must actually think about what the premises mean, and what they entail. The synthetic method, because and insofar as it relies only on formally valid rules of inference, does compel assent. But because it does not require that one actually thinks about what is being claimed in the premises, it does not ensure understanding. It follows, Descartes thinks, that the synthetic method "is not as satisfying as the method of analysis," that the analytic method is "the best and truest method of instruction" (Descartes 1641, p. 111).

What, then, is wanted in a mathematical demonstration? Does one want to be convinced that the conclusion follows from the premises by seeing precisely on what it depends, all assumptions, rules, and definitions having been laid out in advance? Or does one instead seek to understand, to grasp the course of reasoning in a way that provides mastery of that reasoning, the sort of mastery that is displayed in, for example, applying the ideas of the proof to new and different mathematical problems? Clearly, one who understands a proof will be convinced by it. The problem for Descartes is that one who is properly convinced by a proof may nonetheless not understand it, and will not understand it if all they focus on is the necessity of the steps of reasoning in light of what has gone before. For Leibniz, the principal issue is conviction, and thereby agreement: what matters to him is that a proof be compelling independent of understanding, that it be (capable of being) seen to be nothing more than the mechanical manipulation of meaningless signs according to rules. And we can understand how this is to work if we consider again our little demonstration that the product of two sums of integer squares is itself a sum of integer squares. Although as presented that demonstration was not fully rigorous, gap-free, clearly it could have been made to be; we could have stated in advance all the rules and assumptions on which that course of reasoning depends so that even someone knowing nothing of what the collections of signs mean, what concepts they articulate, could see that the end point is derivable from the starting point by applying rules to what is given in the order and manner displayed. But of course, the relevant ideas, that of a sum of integer squares, of a product of such sums, and so on, would still be there. One could still follow the reasoning as a course of reasoning on the basis of mathematical ideas. What distinguishes the current debate between formalists and nonformalists, we will see, is that this seems no longer to be true. As currently understood, and as Rav (1999) already emphasizes, the problem of formal proof in mathematical practice is that *either* the proof is grounded in mathematical ideas, *or* it is merely a matter of the mechanical manipulation of meaningless signs according to stipulated rules. It cannot be both.

3 The Contemporary Debate

The paradigm of mathematical reasoning for both Descartes and Leibniz is the sort of symbolic, algebraic construction first introduced in Descartes' *La Géométrie* and developed further by Leibniz, Newton, Bernoulli, Euler, and others. By the dawn of the twentieth century, it had come to seem that mathematical reasoning as it ought to be conducted is not a matter of constructive algebraic problem solving in the symbolic language of arithmetic and algebra but instead deductive from concepts. It involves, to put it as Kant had by then taught us to do, not an intuitive use of reason, the construction of concepts in pure intuition, but instead a discursive use of reason, strictly logical, deductive reasoning on the basis of defined concepts. But how can one learn anything new by deductions from definitions of concepts? How can a deductive proof extend our knowledge? If Kant is right, logic is merely formal, devoid of all content, and deductive reasoning is merely analytic, explicative.

Though one might through such reasoning make something hitherto implicit explicit, nothing truly new can emerge from such reasoning. There appeared, then, to be only two options: either to argue that mathematical reasoning is deductive, and hence merely formal and explicative rather than ampliative, or to argue contrapositively, that mathematical reasoning is ampliative, a source of new knowledge and insight, and so cannot be merely formal, merely deductive.

Poincaré (1920, pp. 133–134) argued contrapositively. Although one *can* break a strictly mathematical mode of inference down into a series of little, strictly logical steps, to break it down is to destroy the mathematics of the reasoning. One can no more learn about mathematics, or indeed learn *mathematics,* by focusing on the strictly logical inferences of which its modes of inference are in some sense composed than one can learn about, say, an elephant (Poincaré's example) by focusing on the cells of which it is in some sense composed.

Suppose that the mathematician infers some conclusion from a set of premises by means of a form of inference that the mathematician claims is mathematically valid, and that our mathematician can convince other mathematicians that the conclusion does indeed follow, can get them to see, to grasp the validity of the inference. Others, those who are not mathematicians, are utterly baffled; they simply cannot see how the conclusion follows. So, the logician breaks the inference down into a series of little, strictly logical steps, and now even a nonmathematician can see that the conclusion can be derived from the premises. The problem, as Poincaré sees it, is that what the nonmathematician is now seeing is not mathematics but only logic. The mathematics has been lost in the course of the dissection of its modes of inference into their logical parts. What is distinctive of mathematical insight, on this account, is that it enables one to recognize irreducible mathematical wholes; valid mathematical modes of inference that *can* be broken down into strictly logical steps but *qua* mathematical are essential unities and cannot be so reduced to logic. On this view, what the nonmathematician lacks is the ability to discern those unities in addition to the logical steps into which they can be broken. The nonmathematician is like one who can see only the cells of the elephant, the elephant as a collection of cells but not also the elephant, that is, the essential unity in virtue of which it is an elephant, an animal of a certain sort.

But there is also another way in which there might be distinctively mathematical modes of reasoning in addition to logical ones, namely, if there were primitive modes of material inference in mathematics, ineliminable relations of entailment among the concepts of mathematics.

Suppose that being F entails being G for mathematical concepts F and G, not logically, as would be the case if, say, being F were a matter of being G and H, but in virtue of what it means to be F and G. A mathematical inference would then be an inference that depended in this way on the meanings of the mathematical ideas involved. Understanding those ideas would enable one to recognize the validity of the inferences, and although one could always express those relations of entailment in generalized conditionals and treat them as premises, thereby turning one's materially valid inference into a formally valid inference, to do so would again be to lose sight of the mathematics involved.

> Even though the warrant for "if **p**, then **c**" may be based on a grasp of the mathematical architecture linking **c** with **p**, the modus ponens counterpart nonetheless *abstracts away from* this grasp of architecture itself and focuses instead on its net classical effect, that is, on the classical semantico-epistemic status (e.g., truth, certitude of truth, a priori certitude of truth, etc.) and logical form of the belief (viz., "if **p**, then **c**") which it warrants … It treats grasp of an architecture as merely the means by which the classical semantico-epistemic status of "if **p**, then **c**" is established while viewing the semantico-epistemic status itself, rather than the means of establishing it, as the matter of primary epistemic importance (Detlefsen 1992, p. 369).

Although the (materially valid) rule that enables one to infer **c** on the basis of **p** shows that knowing that **p** provides a good reason for inferring that **c**, if the inference the rule governs is replaced by another, strictly logical inference involving also the conditional "if **p**, then **c**," then **p** no longer serves as the ground for **c**. What one now has is a logically valid inference, and what one now knows is only that, given that the premises are true, the conclusion is also true. One has lost sight of why the conclusion is true: that **p** no longer can be seen to serve as the reason for inferring that **c**. And this happens because of what had been functioning as a rule of inference regarding what is a reason for what has been reduced to a merely truth-functional conditional, that this is true if that is true, thereby robbing the original premise of its status as a reason. What had been a good mathematical ground for inference is reduced to a merely logical ground. And now, just as Descartes says, one need not think at all to see that the conclusion follows but only (mindlessly) apply rules of logic to given claims among which are the conditionals codifying primitive entailment relations among mathematical concepts.

Poincaré, like Kant, is concerned to understand how mathematical proofs can constitute real extensions of our knowledge, but whereas Kant focuses on mathematical judgment, the possibility of synthetic a priori knowledge in mathematics; Poincaré is concerned instead with mathematical reasoning, the possibility, as we might say, of synthetic a priori inference. Much as a synthetic a priori judgment is necessary without being logically necessary, so the reasoning of the mathematician seems to be necessary, that is, truth preserving, without being logically necessary, merely a matter of making explicit something contained already implicitly in one's starting point. Poincaré does not deny that the mathematician's proof can be formalized, rendered logically valid, merely explicative. What he denies is that such a formalized proof is mathematics.

The idea of fully formalized, machine-executable mathematical proofs has now been realized in the formalization of many interesting and significant proofs (among them, Gödel's incompleteness theorem, the fundamental theorem of algebra, the prime number theorem, and the four-color theorem). And one important consequence of such work has been the development of extensive libraries of mathematical theories (for example, of elementary number theory, linear algebra, and finite group theory) to be drawn on in formalizations. Given these advances in automated deduction, it can seem not at all unreasonable to think – following Mac Lane and Bourbaki (see Hamami 2019), and more recently, Azzouni (2004, but see also Azzouni 2013), Avigad (2006), and Weir (2016) – that although mathematicians' proofs provide sketches of real, or complete proofs, it is the formalization that

actualizes the potential of such sketches to prove this or that theorem. One might well think, then, that "over the next hundred years or so, computers will gradually be able to do more and more of what mathematicians do . . . eventually supplanting us entirely" (Gowers 2002, p. 134). Nonetheless, as Gowers also notes, his is a minority view: "most mathematicians are far more pessimistic (or should that be optimistic?) about how good computers will ever be at mathematics." After all, it is one thing to come up with a conjecture, that is, to postulate some mathematically interesting connection among mathematical ideas, and reflecting on those ideas, to find some way of showing that the postulated connection is real, both of which mathematicians are very good at, and quite another to have a formalized, machine-executable proof. Indeed, as Robinson (1997, p. 54) notes, "the translation from informal to formal is by no means a mere matter of routine. In most cases it requires considerable ingenuity, and has the feel of a fresh and separate mathematical problem in itself. In some cases the formalization is so elusive as to seem impossible." More than two decades on, although much has changed in the domain of formalization, the problem of formalizing a given mathematical proof remains, in general, "painstakingly difficult" (Avigad 2018, p. 687). And yet, Avigad is surely also right to suggest that "we should be open to the possibility that new technologies can open new mathematical vistas and afford new types of mathematical understanding" (Avigad 2018, p. 688).

Formalized proofs, according to many, are not more rigorous by mathematicians' standards of rigor than informal proofs (Kreisel 1970; Thurston 1994; Robinson 1997; Marfori 2010). Nor, others have argued, do such proofs explain or ground such rigor as mathematicians' proofs enjoy (Tanswell 2015; Paseau 2016). Formal proofs do not seem to advance mathematical understanding (Rav 1999; Detlefsen 2008). They generally lack the explanatory power of ordinary proofs (Robinson 2000). And sometimes, as Manders (1987) argues, they are simply unintelligible. In short, "a formal, symbolic proof is for the most part humanly incomprehensible and does nothing to augment our understanding" (Avigad and Harrison 2014, p. 73). But this, as Avigad and Harrison go on to suggest, is unproblematic: "formal verification is not supposed to replace human understanding or the development of powerful mathematical theories and concepts . . . Rather, they are intended to supplement the mathematics we do with precise formulations of our definitions and theorems and assurances that our theorems are correct."

We have seen that according to some, for instance, Plato, Descartes, and non-formalists, the virtue of a (successful) train of mathematical thought lies in the insight and understanding it affords. For others, among them, Aristotle, Leibniz, and the formalist, it lies instead in the power of the train of thought to show indubitably that some theorem is true and thereby to end once and for all disagreements as to its truth. The two desiderata, insight and understanding, on the one hand, and conviction and truth, on the other, are not incompatible. And the Leibnizian idea of a universal language, at once a characteristica and a calculus, shows that they are not insofar as such a language enables one to express in a specially devised system of written signs the contents of concepts on the basis of which to reason. The manipulation of signs in such a language is everywhere governed by antecedently stated

rules, and because it is, the reasoning can be mimicked by a machine. But the mathematician's reasoning in such a language is not, or at least need not be, merely mechanical; a Leibnizian language is not only a calculus, but also a characteristica within which to display the contents of mathematical ideas, what it is to be this or that mathematical object, or property, or relation, and to do so in a way that renders the reasoning tractable. In a Leibnizian language, ideas are expressed in a way enabling rigorous rule-governed reasoning *in* the system of signs. Thus, while it is possible in a case of such reasoning to ignore meanings and mathematical ideas, it is equally possible to attend to them, to ignore trivial and routine steps so as to focus on the *ideas* of the proof. A demonstration in a Leibnizian universal language in this way provides, or at least can provide, *both* insight and understanding *and* conviction and truth.

The idea of a Leibnizian language for mathematics is realized in part in the symbolic language of arithmetic and algebra within which to exhibit in written signs the contents of concepts and to reason rigorously on the basis of those contents to new and significant mathematical truths. The reasoning is at once compelling and a vehicle of understanding. The same cannot be said for reasoning in a formalized proof of contemporary mathematics. Although a mathematician's informal proof of, say, the prime number theorem enables insight and understanding, and thereby conviction, at least for the mathematician, the corresponding formalization, while it may compel assent (even in the most stubborn and argumentative reader), does not seem to enable insight and understanding. Formalization thus seems, as Poincaré already suggested, somehow to undermine the unity and integrity of the mathematician's course of reasoning. Breaking down a chain of mathematical reasoning into a series of little, strictly logical steps transforms a rich and meaningful mathematical demonstration into something lifeless, mathematically inert. And in a way, this is not surprising given that, in logic, validity is, we think, a matter of form rather than of content. Indeed, that is precisely why such reasoning can be mechanized, executed in a computer program run on a machine. But if reasoning is merely mechanical, insofar as it is reasoning, then as the formalist tends to think, the solution of our difficulty must lie elsewhere, perhaps outside of logic and the philosophy of mathematics altogether, in, say, human psychology (Goldfarb 1988), or in some other not strictly rational dimension of mathematical practice (Azzouni 2004).

If deductive rationality is, in essence, a matter of the manipulation of meaningless signs according to rules, then the problem of formal proof in mathematical practice is not a problem about the nature and function of mathematical reasoning but instead an altogether different sort of problem, a problem in social epistemology, say, or in psychology or sociology. But can this be right? The task of logic is to provide an adequate account of the nature and form of deductive reasoning, a paradigm of which is reasoning in contemporary mathematical practice. Indeed, as Burgess (1992, p. 9) points out, mathematical logic "was developed ... as an extension of traditional logic mainly, if not solely, about proof procedures in mathematics." And certainly, this logic has proved extraordinarily fruitful in the domain of machine reasoning, in computing: "the clear recognition of just how much reasoning – that is, as far as results are concerned, never mind the processes – can be mechanized is

surely the most outstanding contribution of 20th century logic *sub species aeternitatis*" (Kreisel 1985, pp. 143–144). And yet, as developments over the past half-century seem clearly to show, formal proofs are irrelevant to the practice of most mathematicians. Can it really be true, as some philosophers of mathematical practice have argued, that "mathematical logic cannot provide the tools for an adequate analysis of mathematics and its development"? (Mancosu 2008, 5)

4 More History

The symbolic language of algebra that Descartes introduced in 1637 proved over the course of the seventeenth and eighteenth centuries to be "the greatest instrument of discovery in the history of mathematics" (Grabiner 1974, p. 357), so much so that to an eighteenth-century mathematician such as Euler it could seem that his pencil surpassed him in intelligence. Grabiner (1974, p. 355) provides a nice example of just this intelligence.

> He [Euler] began with the identity
>
> $$\cos nz = (\cos z)^n - n(n-1)/2!(\cos z)^{n-1}(\sin z)^2 + [n(n-1)(n-2)(n-3)]/4!(\cos z)^{n-4}(\sin z)^4 - \ldots$$
>
> Let z be an infinitely small arc, and let n be infinitely large. Then:
>
> $$\cos nz = 1,\ \sin z = z, n(n-1) = n^2, n(n-1)(n-2)(n-3) = n^4, \text{etc.}$$
>
> The equation now becomes recognizable:
>
> $$\cos nz = 1 - n^2z^2/2! + n^4z^4/4! - \ldots$$
>
> But since z is infinitely small and n infinitely large, Euler concludes the nz is a finite quantity. So let $nz = v$.
>
> The modern reader may be left slightly breathless; still we have
>
> $$\cos v = 1 - v^2/2! + v^4/4! - \ldots$$

Although in no way a rigorous course of reasoning, this demonstration is nonetheless compelling, and, as was invariably true of Euler's work, it is perfectly correct. It is a paradigm of the eighteenth-century practice of mathematics, a practice that would yield an astonishing array of new results, however informal, lax even, the method of demonstration.

Eventually, the requirement of rigor would reassert itself – though not in order to avoid errors. Like mathematicians of any era, "eighteenth-century mathematicians had an almost unerring intuition" (Grabiner 1974, p. 358). The motivation to be more rigorous in their definitions and reasoning lay instead in "the mathematician's

need to teach": "A mathematician could understand well enough about a concept to use it, and could rely on the insight he had gained through his experience. But this does not work with freshmen, even in the eighteenth century" (Grabiner 1974, p. 360; see also Kitcher 1981). In mathematics, the problem is not, and may never have been, argumentative and stubborn interlocutors. The problem is nonmathematicians, that is, most immediately, the need to teach those who are not yet and may never become mathematicians. The foundational work of the eighteenth and nineteenth centuries, the paradigm of which was the rigorization of the calculus, was first done in, or for, the classroom.

This foundational work was, furthermore, just one manifestation of the revolutionary transformations in mathematical practice that occurred over the courses of the nineteenth century, one aspect of what Stein (1988, p. 238) describes as "a second birth of the subject—its first birth having occurred among the ancient Greeks" (see also Macbeth 2014). Over the course of the nineteenth century, most markedly in the work of Riemann and Dedekind following him, mathematicians came to eschew the symbolic language of arithmetic and algebra that had hitherto proven so fruitful, to reflect directly on concepts and to reason deductively on the basis of their contents, "to conquer problems," as it was sometimes put, "with a minimum of blind calculation, a maximum of clear seeing thought" (Stein 1988, p. 241). The algebraic, computational, constructive practice of the seventeenth and eighteenth centuries was in the nineteenth century to be replaced by a new, wholly conceptual approach, *Denken in Begriffen* (Laugwitz 1999). And at least at first, there was no Leibnizian symbolic language within which to do this work, nothing that could play the role the symbolic language of arithmetic and algebra had played, nothing to serve at once as a characteristic and a calculus for this new practice of thinking in concepts. All the great works of Riemann and Dedekind, of Weierstrass and Bolzano, could only be reported (in natural language); it could not be displayed. And this remains true even today. As Avigad (2006) notes, even a canonical text such as Birkhoff and Mac Lane's *A Survey of Modern Algebra* does not show chains of mathematical reasoning, whether or not gap-free, but only describes how the reasoning goes, or would go were one actually to reason from the given concepts to the derived theorem.

5 What Is a Mathematical Proof?

A mathematical proof, as described by Hanna (1995, p. 46), "is a transparent argument, in which all the information used and all the rules of reasoning are clearly displayed and open to criticism." And this is the right way to think about a mathematician's proof insofar as "it is in the very nature of proof that the validity of the conclusion flows from the proof itself, not from any external authority." Of course, what counts as a transparent argument will depend on the context. What counts as transparent to a mathematician working in the same subfield will not so count for a mathematician working in a different field; what counts as transparent for mathematicians generally will not so count for graduate students; what counts as

transparent for graduate students will not so count for undergraduates, and so on – though not, as Rav (1999) suggests, infinitely. What may be taken for granted as known and beyond criticism depends on one's audience. Nevertheless, what Hanna describes as transparency is crucially important, something that is "in the very nature of proof." A proof is self-standing. That is, unlike a report of empirical findings, a mathematical proof itself carries its evidential force; the intentions, reliability, and trustworthiness of its author are irrelevant to its cogency. One can, at least in principle, see for oneself that things are as the author of the proof claims. The proof is, as it is sometimes put, "transferable" (Easwaren 2009). Indeed, it is in just this sense that mathematical knowledge is a priori: It is a priori not because it is infallible but because and insofar as it does not rely on empirical evidence, whether that of one's own senses or that given on the testimony of another. A mathematical proof is a priori insofar as "the validity of the conclusion flows from the proof itself."

What Hanna describes as transparency is obviously very closely related to the idea of rigor insofar as a rigorous proof is one in which everything on which the chain of reasoning depends is stated explicitly and in advance, all definitions, all assumptions, and all rules of inference. In a rigorous proof, the reasoning is, as Frege thinks of it, gap-free (see Fallis 2003; Hamami 2014; Anderson 2020 for discussion of the idea of a proof gap). But as we have already seen, and contrary to what is often assumed, rigor does not require formalization. Being mechanically checkable is a good test of rigor: If all the steps of reasoning can be made merely mechanically, without appeal to meanings, solely by applying given rules to the given collections of signs in virtue of their shapes, then it is evident that everything on which the proof depends has been set out in advance. The proof is rigorous. But one need not purge a proof of meaning in order to state in advance everything on which it depends. To rigorize a proof is not thereby to formalize it, to replace all its meaningful signs with meaningless marks; it is perfectly possible to render a proof rigorous and gap-free without formalization. And that is to say that although a fully rigorous proof must be machine *checkable,* it does not follow that such a proof is machine *executable,* a chain of reasoning that could (in principle) be carried out by a machine just as well as, and in the same sense as, it is carried out by the mathematician. The fact that a mathematician's proof can be made rigorous and gap-free does not show, as many formalists and nonformalists assume, that the reasoning of the mathematician is merely mechanical, nothing more than the manipulation of meaningless marks according to stipulated rules.

A mathematician's proof can be at once rigorous and fully meaningful. And yet, it is often assumed (falsely) that to render a proof fully rigorous, wholly gap-free, just is to just formalize it, to make it into something that a machine can do as well as any human being. In order to avoid unthinkingly making that assumption, we need explicitly to distinguish between the language of logic that is used in formalizations, on the one hand, and a Leibnizian language, on the other. In particular, we need to remember that our logic is a calculus, that it is not also a characteristica. Indeed, the whole point of logic is to focus on form rather than content and meaning. The aim is not to express content but to abstract from it, leaving only (logical) form. And so it is with a formalization: "the very purpose of formalization is to squeeze out the sap of meanings in order not to blur focusing only on the logico-structural properties of

proof" (Rav 1999, p. 12). By contrast with the strengths of the symbolic language of arithmetic and algebra within which to express mathematical ideas in ways that facilitate rigorous reasoning, the great strength of logic lies in its having nothing to do with meaning, in its being a matter of the merely mechanical manipulation of meaningless signs according to rules. And thus, it can seem to follow that to reason logically is to reason as a machine reasons. Should we be surprised, then, to find that the formalizations of mathematicians' proofs according to the standards of logic are at once a *tour de force* in the science of computing and seemingly irrelevant to the practice of mathematics?

6 Resolving the Antinomy

Formal proofs, unlike the proofs mathematicians develop and esteem, are completely rigorous and gap-free; all definitions and assumptions are set out in advance, and all steps of reasoning conform to some small set of basic rules of logic. Thus, it comes to seem to the formalist to be obvious that the relationship of the informal proofs of mathematicians to formal, machine-executed proofs is that of a sketch to a fully worked out chain of reasoning. The nonformalist sees that this cannot be right but seems to have no way of explaining to the formalist just what has gone wrong. Again, the reasoning of the mathematician is deductive, and logic provides a complete theory of deduction. It must then be the case that a formal proof is, in essence, a fully worked out version of the informal proof of the mathematician. Only, it seems not to be. The problem is somehow to move beyond this stalemate.

Over the course of the nineteenth century, we have seen the practice of mathematics was radically transformed, and with that transformation came a renewed call for rigor in the practice.

> After deserting for a time the old Euclidean standards of rigor, mathematics is now returning to them, and even making efforts to go beyond them . . . Proof is now demanded of many things that formerly passed as self-evident. Again and again the limits to the validity of a proposition have been in this way established for the first time. The concepts of function, of continuity, of limit, and of infinity have been shown to stand in need of sharper definition. Negative and irrational numbers, which have long since been admitted into science, have had to submit to a closer scrutiny of their credentials.
>
> In all directions, these same ideals can be seen at work – rigor of proof, precise delimitation of extent of validity, and as a means to this, sharp definition of concepts (Frege 1884, §1).

But again, this demand was not motivated by any perceived need to render the practice less error-prone or to resolve disputes regarding the validity of results. Mathematicians are not infallible; they do make mistakes. But they do not make mistakes very often, and generally such mistakes as they do make are mathematically significant insofar as they serve ultimately to improve mathematicians' grasp of mathematical concepts they have been working with all along. What enables mathematicians to discover and avoid errors is not formalization but understanding.

> Mathematicians can and do fill in gaps, correct errors, and supply more detail and more careful scholarship when they are called on or motivated to do so. Our system is quite good at producing reliable theorems that can be solidly backed up. It is just that the reliability does not primarily come from mathematicians formally checking formal arguments; it comes from mathematicians thinking carefully and critically about mathematical ideas (Thurston 1994, p. 170).

Thus,

> We are most certain to avoid gaps in reasoning when premises *explain* conclusions . . . The greater such explanatory transparency, the more confident we can be that unrecognized information has not been used to connect a conclusion to premises in ways that matter. To the extent, then, that formalization decreases explanatory transparency, it also decreases rigor (Detlefsen 2008, p. 19).

And as we have seen, formalizing a mathematician's proof does decrease transparency, and so also rigor, and it does so precisely because and insofar as it purges a proof of all content and meaning, all mathematical ideas. But again, we must not confuse the logic of formalizations with a Leibnizian language for mathematics, a language within which to express mathematical ideas and to reason from them.

Philosophers sometimes claim that much as the signs of standard logic are abbreviations for words so the signs that comprise the writing systems that have been devised for reasoning in mathematics, paradigmatically, the symbolic language of algebra, are nothing more than abbreviations for words. But this does not seem to be right. The systems of signs that have been devised (and universally used) for mathematics are instead the signs of a Leibnizian language. That is, as Jourdain notes, echoing again Euler's sentiment,

> it is important to realize that the long and strenuous work of the most gifted minds was necessary to provide us with simple and expressive notation which, in nearly all parts of mathematics, enables even the less gifted of us to reproduce theorems which needed the greatest genius to discover. Each improvement in notation seems, to the uninitiated, but a small thing; and yet, the pen sometimes seems to be more intelligent than the user. (Jourdain 1912, p. 16)

The formula language of arithmetic and algebra is one such language. But by the nineteenth century, mathematics had grown out of that language. Frege took it upon himself to provide mathematics with another; he devised what he called a *Begriffsschrift,* a concept-script, one that (although it would come to be read as a mere calculus) was self-consciously designed to serve not only as a calculus but also, and equally, as a characteristica, as a mathematical language within which to exhibit the contents of mathematical concepts and to reason deductively on the basis of those contents. Macbeth (2014; see also Macbeth 2005) argues that he succeeded, and that in so doing Frege shows how mathematical reasoning can be at once deductive and ampliative. What matters for purposes here is how exactly a Leibnizian language for contemporary mathematics would (and if *Begriffsschrift* is such a language, does) differ from the language of logic, what it is that enables such a

system of signs to function as a Leibnizian language, at once a calculus and a characteristica.

We have seen that for the philosopher the problem of formal proof in mathematical practice is the problem that, so it seems, a formal proof both must and cannot clarify the nature of mathematical practice. A formal proof must clarify the nature of mathematical reasoning because mathematical reasoning is deductive and the rules governing a formal proof are those of deductive reasoning. And formal proof cannot clarify the nature of mathematical reasoning because the reasoning in a formal proof is merely mechanical, without content. What we are now in a position to see (following Macbeth 2015, 2018) is that the move from an informal to a formal proof involves not one but two conceptually distinct steps. One step, the step of rigorization, is to fill in all the gaps, include all the steps of reasoning, even the most trivial and obvious, and to state in advance everything on which the reasoning depends, all definitions, all assumptions, and all rules of reasoning. The second, conceptually distinct step is to replace all meaningful expressions with meaningless signs.

Suppose, for example, that the mathematician infers from the fact that some number *n* is a square root of 4 that *n* is a fourth root of 16. Our mathematician can make the inference rigorous by explicitly adding the rule on which it depends, that being a square root of 4 entails being a fourth root of 16. Our mathematician can also go on to show how this entailment is derivable from basic truths of arithmetic. The reasoning, that is to say, can be rendered even more rigorous by proving everything that can be proved, by isolating a small set of basic truths of arithmetic and deriving from those basic truths all other entailments on the basis of which to make the original step of reasoning. It is in just this way that the original step of inference is rendered fully rigorous, gap-free.

But our mathematician can also do something quite different with the original inference. First, the relevant entailment is again made explicit in a premise so that the inference now goes like this: being a square root of 4 entails being a fourth root of 16; *n* is a square root of 4; therefore, *n* is a fourth root of 16. This inference is deductively, that is, logically, valid. It is valid in virtue of its form, and what that means, we think, is that the fact that it is about the number *n,* and square and fourth roots, is irrelevant to its validity. The inference is valid in virtue of the fact that it has this form: being *F* entails being *G, o* is *F*; therefore, *o* is *G*. To formalize the step of inference is to replace in this way all meaningful, nonlogical expressions with meaningless marks, here "*F*," "*G*," and "*o*." And now, we have something that can be implemented in a machine. We have reduced the mathematician's reasoning to the manipulation of meaningless marks according to stipulated rules.

In standard logic, to say that an inference is valid in virtue of its form is to say that its content, in our example, its being about numbers, squares, and fourth roots, is utterly irrelevant to its validity. There is, however, another way to understand the notion of validity in virtue of form. Following Peirce (1992, p. 131) and Ryle (1950), we can say that an inference is valid in virtue of its form just in case it is an instance of something more general, a rule that applies also in other cases. We can, that is, think of the rule "if being *F* entails being *G* and *o* is F, then *o* is *G*," as perfectly

meaningful despite being maximally general, as being meaningful in essentially the way that the rule "if a number is a square root of four then it is a fourth root of sixteen" is meaningful. So understood, the difference between the two cases is only that the law of logic is more general, a rule of logic rather than of mathematics. Logic, on such a view, is not the articulation of mere form, of form as it contrasts with content and truth. Logic (on such a view) is a science, the science, as Frege once put it, of the most general laws of truth.

As we are taught to understand it, the validity of an inference concerns form as it contrasts with content and truth. The validity of inference is thereby conceived not only formally, however, but also on the model of containment: An inference is valid just if the information in the conclusion is contained already in the premises, even if only implicitly. Deduction on this model is not only truth preserving, in that it is impossible to derive a falsehood from true premises; it is also merely explicative. Nothing can be inferred that was not contained already in one's starting points, one's premises (We have seen already that this presents a problem for any adequate understanding of reasoning in mathematics insofar as the mathematician's deductive proofs seem often to extend our knowledge.). Here again, there is an alternative, and our little example involving square and fourth roots can help to remind us what it is. We have an inference that we expressed thus:

> Being a square root of four entails being a fourth root of sixteen.
> The number *n* is a square root of four.
> Therefore, *n* is a fourth root of sixteen.

What is at issue now is the status of the first premise, what Aristotle calls the major. There are two options. Either we treat premise as a rule governing (material) inferences, or we treat it as a claim about how things actually are, whether or not they could be otherwise (Sundholm 2012 also makes this point.). On the latter option, we read the major as setting out a piece of information, the fact (possibly necessary) that all numbers that are square roots of 4 are fourth roots of 16. Given this piece of information, together with the information in the minor premise, that the number *n* is a square root of four, it clearly must be true, indeed, it has already been implicitly stated that *n* is a fourth root of 16. The inference so understood does not merely state in the premises everything that is needed to infer that *n* is a fourth root of 16. The premises already state that *n* is a fourth root of 16, though not in so many words. The information that the number *n* is a fourth root of 16 is contained already in the information provided by the premises: *If* all square roots of 4 are fourth roots of 16 *and n* is a square root of 4, then, whether or not anyone knows it, it is settled that *n* is a fourth root of 16. Such an inference can extend our knowledge only in the sense of revealing, or at least reminding us of, something to which (on pain of irrationality) we are already committed, something that we already believe – if only implicitly.

The other option is to treat the first premise, the major, as the expression of a rule of inference, as licensing inferential steps, acts of reasoning, from appropriate premises to the relevant conclusion. The minor then gives us a premise to the effect that the number *n* is a number to which the rule given in the major applies, and so,

applying the rule given in the major to the instance given in the minor, we make our inference, infer that *n* is a fourth root of 16. In this case, although our premises give us everything we need in order to draw the relevant inference, give us, that is, both the rule licensing the inference and an instance to which the rule applies, the premises do not contain already the information contained in the conclusion. Everything needed to *get* that information is supplied, but the inference must still be made; one must actually apply the rule to the instance in order to derive the conclusion. The step of reasoning from the premises to the conclusion must be taken, and it must be taken because, on this reading, the conclusion is contained in the premises not implicitly, needing only to be made explicit, but instead potentially: Actually to have the conclusion, realizing thereby the potential of the premises to yield a conclusion, something must happen. The inference must be drawn. In that case, to borrow Frege's (1884, §88) imagery, the conclusion is contained in the premises not as beams are contained in a house, that is, implicitly, but as the plant is contained in the seed, that is, potentially.

To say that a proposition has the form of a generalized conditional is understood in standard logic as saying that it is a claim about all objects of a certain sort. But the expression of a rule also has the form of a generalized conditional insofar as it includes, in the antecedent, a specification of the circumstances of application of the rule, and is something that applies, in principle, in a variety of cases. A rule is constitutively conditional and inherently general. In some contexts, then, something of the form of an English sentence, "all S is P," may mean not that every object that is S is also P but instead that it can be inferred from the fact that a thing is S that it is P, that being S entails being P – even if in other contexts, a sentence of just that same form means instead that as a matter of (contingent or necessary) fact, all the Ss there are also P. Recall, now, that in a Leibnizian language all rules governing inferences are made explicit as claims, that is, as generalized conditionals; however, they are to be read. The language is a calculus. But we also know that a Leibnizian language is equally a characteristica, a language within which to display the contents of concepts on the basis of which to reason. What we need to see (if only in outline here) is how this is to work, how it is that the contents of concepts can be displayed in complexes of signs that, in combination with the generalized conditionals of the system, enable reasoning *in* the system of signs.

We want to understand how a system of written marks might enable the mathematician to display the contents of mathematical concepts in a way that enables rigorous, gap-free reasoning in the system of signs. Following Macbeth (2015, 2018), we use the following simple example to illustrate the essential point. Imagine, to begin with, a person able to count but who does not know even the simplest facts of arithmetic, and suppose further that this person wishes to know what is the sum of seven and five. A good, if merely mechanical, way of finding that out would be to count out seven things, then five more, and then to count the whole collection that results. We can do this with strokes on the page (or screen). Here are seven things: ///////. Now we add five more to that given array: /////// (already given) plus /////. Counting the number of strokes in all gives the answer that is wanted, namely, 12. Clearly, this is not yet mathematics. It is furthermore clear that in such a case the

answer, 12, is contained already in the given parameters of the problem as those parameters have been set out: if one has 7 things (strokes, say) and adds 5 more, then one has 12 things (strokes), whether or not one knows that one does. That 7 and 5 is 12 is, to this way of thinking, analytic: Being 12 is contained already in being $7_{+}5$.

If one thinks of 7 + 5 merely mechanically, as a matter of adding 5 things to 7 things already given, then the fact that 7 + 5 equals 12 would seem to be analytic in Kant's sense. And yet, Kant claims in the B Introduction to the first *Critique* that the arithmetical truth that 7 + 5 equals 12 is not analytic but instead synthetic, that being 12 is not contained already in the concept of the sum of 7 and 5, despite its being a priori that the sum of 7 and 5 is 12. And as Macbeth shows, we can understand how this works if we conceive the strokes not as mere things but instead as tokenings of the primitive sign type of an exceedingly simple Leibnizian language. The language has one primitive sign, the stroke, and we are to display the arithmetical contents of numbers, in particular, their various multiplicities, in this language by combining strokes in complex signs. In this language, a collection of seven strokes is not merely that, a collection of seven things; it is instead a meaningful complex sign that designates a number, *one* number, namely, the number seven, and it does so through a display of what it is to be that number, namely, a particular multiplicity (Obviously for other mathematical purposes, we would want to think differently about what it is to be the number seven, for example, as the successor of the number six, and in that case would display a different content, which would require as well devising a different Leibnizian language.).

We start, then, with complex signs for our two numbers, the number seven, the content of which is exhibited in the language in a complex of seven strokes, and the number five, the content of which is exhibited in a complex of five strokes. So far, then, we have a display of two numbers in an array such as this: /////// /////. Each of the two collections of strokes designates a number, and it does so through what we can think of as a Fregean sense, something that displays what it is to be those numbers, namely, again, certain multiplicities. So far, then, there is nothing that can be taken to designate the number 12. If, however, we have a rule licensing a certain rewriting of signs, a rule to the effect that if one has a number designated by m strokes and a number designated by n strokes, then one may rewrite to give two numbers that are designated by $m + 1$ strokes and $n - 1$ strokes respectively. Applying the rule to our original numbers /////// and /////, and to the results of applications of that rule, we eventually arrive at the number ////////////, that is, a complex sign that designates the number 12. We have shown, through a chain of (excessively simple) mathematical reasoning, that the sum of the numbers 7 and 5 is the number 12. The result is manifestly synthetic a priori in Kant's sense. Although (a representation of) the number 12 is not contained already in (the representations of) the numbers 7 and 5, (a representation of) 12 can be constructed from (representations of) those numbers by successive applications of an explicitly stated rule. The rule and the starting point give us everything we need to construct a sign designating the number 12, but that sign does, nonetheless, need to be constructed. It is contained in the starting point not implicitly but only potentially. It is contained in the starting point as the plant is contained in the seed.

7 Conclusion

The philosophical problem of formal proof in mathematical practice is the problem of the relationship between a mathematician's proof and its fully formalized counterpart. It can seem that this problem is merely one of emphasis, of the relative value of, on the one hand, mathematical insight and understanding, and on the other, indubitability and persuasiveness. Mathematicians' proofs, so it can seem, focus on the former, formal proofs on the latter. Closer inspection showed that things are more complicated, that we need to distinguish among many, quite different parameters along which the two types of proof differ.

First, and most obviously, there was the issue of rigor, of gap-free proof. Although the mathematician's proof is transparent insofar as the validity of the conclusion "flows from the proof itself," as mathematicians are in a position to see, a mathematician's proof is almost never rigorous in the sense of being wholly gap-free. But there is no reason why it could not be made to be rigorous in this sense. Mathematicians' proofs could, with effort and application, be written so that everything on which the reasoning depends was set out in advance. Formal proofs are also so written, as of course they must be if they are to be machine executable.

A related but more substantial difference between a mathematician's proof and a formal proof is, we saw, that the mathematician's proof is meaningful and the formal proof is not. Even were it to be rendered altogether gap-free, the mathematician's proof would still be a chain of reasoning about mathematical concepts and how they are related; it would still contain whatever proof ideas the transparent but not rigorous proof contains – though it would likely take a trained eye to discern those ideas amid the welter of (in most contexts, needless) detail. The formal proof, by contrast, is nothing more than a series of moves in an automated system designed to transform arrays of marks into other arrays of marks.

But we have seen that even this does not exhaust the differences between the two sorts of proofs. After all, were the only difference between the two sorts of proofs to be that one is gap-free and purged of all meaning, while the other is neither, it would be inexplicable why mathematicians cannot simply reinject the meaning into a formal proof and recover thereby (a fully rigorous version of) the original mathematician's proof. To understand why this is not possible, we needed to distinguish between three different sorts of languages: (1) everyday, natural language, the language that is currently used in mathematics textbooks to describe the sort of reasoning from concepts that is involved in proving theorems in contemporary mathematical practice, (2) the language of standard logic, the signs of which are abbreviations of words of natural language and enable the perspicuous display of (what is assumed to be) the logical forms of sentences of natural language, and (3) a Leibnizian language, a specially devised written mathematical language within which to display the contents of mathematical concepts in a way enabling rigorous reasoning in the system of signs. We saw that the symbolic language of arithmetic and algebra is such a Leibnizian language, for at least a part of mathematics, and noted as well that Frege took his *Begriffsschrift* to provide such a language for the mathematical practice of deductive reasoning from concepts that came to dominate

mathematical practice over the course of the nineteenth century. As we saw, a Leibnizian language can replace natural language for the purposes of reasoning in mathematical practice precisely because and insofar as it is not merely a calculus but also a characteristica enabling one to reason in the language on the basis of the contents of concepts.

Were mathematicians' proofs to be formulated in a Leibnizian language, and made fully rigorous, gap-free, then they would be machine checkable. But as we have seen, they would still be different from formal proofs insofar as they would not also be machine executable. The mathematician's reasoning, even that conducted in a Leibnizian language, is fully contentful, a form of mathematical reasoning on the basis of mathematical ideas (as we see already in seventeenth- and eighteenth-century mathematics conducted in the symbolic language of algebra). A formal proof is not contentful. Indeed, it is only because it is not contentful that a formal proof is not only machine checkable but also properly machine executable. Of course, we can read expressions in a Leibnizian language mechanically, from which it follows, that were the mathematician's proof to be written in a Leibnizian language, and then automated, it would be possible for the mathematician to recover the ideas of the proof from the formalization. That this is not possible in the case of formalizations currently available is, we saw, because such formalizations begin not with a Leibnizian language within which to formulate mathematical ideas but instead with the relatively impoverished resources of standard logic. Indeed, it is that poverty of resources that may explain the fact that formalizing a proof generally amounts to a wholly new sort of mathematical problem. Is it any wonder, then, that formalized proof (as currently understood) has so little to offer to the practicing mathematician?

References

Anderson L (2020) Acceptable gaps in mathematical proofs. Synthese 197:233–247. https://doi.org/10.1007/s11229-018-1778-8

Avigad J (2006) Mathematical method and proof. Synthese 153:105–159. https://doi.org/10.1007/s11229-005-4064-5

Avigad J (2018) The mechanization of mathematics. Not Am Math Soc 65:681–690

Azzouni J (2004) The derivation-indicator view of mathematical practice. Philos Math 12:81–105. https://doi.org/10.1093/philmat/12.2.81

Azzouni J (2013) The relationship of derivations in artificial languages to ordinary rigorous mathematical proofs. Philos Math 21:247–254. https://doi.org/10.1093/philmat/nkt007

Avigad J, Harrison, J (2014) Formally verified mathematics. Comm ACM 57:66–75. https://doi.org/10.1145/2591012

Burgess J (1992) Proofs about proofs: a defense of classical logic. Part I: the aims of classical logic. In: Detlefsen M (ed) Proof, logic, and formalization. Routledge, London/New York, pp 8–23

Descartes R (1637) The geometry of René Descartes. Dover, New York

Descartes R (1641) Second set of replies. In: The philosophical writings of Descartes, vol II. Cambridge University Press, Cambridge, pp 93–120

Detlefsen M (1992) Poincaré against the logicians. Synthese 90:349–378. https://doi.org/10.1007/BF00500033

Detlefsen M (2008) Proof: its nature and significance. In: Gold B, Simons R (eds) Proof and other dilemmas: mathematics and philosophy. Mathematical Association of America, Washington, DC, pp 3–32
Easwaren K (2009) Probabilistic proofs and transferability. Philos Math 17:341–362. https://doi.org/10.1093/philmat/nkn032
Fallis D (2003) Intentional gaps in mathematical proofs. Synthese 134:45–69. https://doi.org/10.1023/A:1022131513275
Frege G (1884) Foundations of arithmetic. Northwestern University Press, Evanston, 1980
Goldfarb W (1988) Poincaré against the logicists. In: Aspray W, Kitcher W (eds) Minnesota studies in the philosophy of science, vol XI. University of Minnesota Press, Minneapolis, pp 61–81
Gowers T (2002) Mathematics: a very short introduction. Oxford University Press, Oxford
Grabiner J (1974) Is mathematical truth time-dependent? Am Math Mon 81:354–365
Hamami Y (2014) Mathematical rigor, proof gap and the validity of mathematical inference. Philos Sci 18:7–26
Hamami Y (2019) Mathematical rigor and proof. Rev Sym Logic. https://doi.org/10.1017/S1755020319000443
Hanna G (1995) Challenges to the importance of proof. Learn Math 15:42–49
Jourdain P (1912) The nature of mathematics. Dodge Publishing, New York
Kitcher P (1981) Mathematical rigor – who needs it? Noûs 15:469–493
Kreisel G (1970) The formalist-positivist doctrine of mathematical precision in the light of experience. L'age Sci 3:17–46
Kreisel G (1985) Mathematical logic: tool and object lesson for science. Synthese 62:139–151. https://doi.org/10.1007/BF00486043
Laugwitz D (1999) Bernhard Riemann 1826–1866: turning points in the conception of mathematics. Birkhaüser, Basil/Berlin/Boston
Leibniz G (1666) Dissertation on the art of combinations. In: Loemker, L (ed) Philosophical papers and letters. D. Reidel, Dordrecht/Boston
Leibniz G (1675) Letter to Henry Oldenberg. In: Loemker L (ed) Philosophical papers and letters, 2nd edn. D Reidel, Boston, pp 165–166
Macbeth D (2005) Frege's logic. Harvard University Press, Cambridge, MA
Macbeth D (2014) Realizing reason: a narrative of truth and knowing. Oxford University Press, Oxford
Macbeth D (2015) Reasoning in mathematics and machines: the place of mathematical logic in mathematical understanding. Proceedings AISB 2015 https://www.cs.kent.ac.uk/events/2015/AISB2015/proceedings/socialComp/papers/AISBIn vitedmacbeth.pdf. Accessed 20 May 2020
Macbeth D (2018) Proof in mathematics and in logic. In: Arazim P, Lávička T (eds) The logica yearbook 2017. College Publications, Rickmansworth, UK, pp 197–208
Mancosu P (2008) Introduction. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 1–21
Manders K (1987) Logic and conceptual relations in mathematics. In: The Paris Logic Group (ed) Logic colloquium '85. Elsevier Science, Amsterdam, pp 193–211
Marfori M (2010) Informal proofs and mathematical rigour. Stud Logica 96:261–272. https://doi.org/10.1007/s11225-010-9280-4
Paseau A (2016) What's the point of complete rigour? Mind 125:177–207. https://doi.org/10.1093/mind/fzv140
Peirce C (1992) Reasoning and the logic of things. Humanities Press, Atlantic Highlands
Poincaré H (1920) Science et Methode. Ernest Flammerion, Paris
Rav Y (1999) Why do we prove theorems? Philos Math 15:5–41. https://doi.org/10.1093/philmat/7.1.5
Robinson J (1997) Informal rigor and mathematical understanding. In: Gottlob G, Leitsch A, Mundici D (eds) Computational logic and proof theory. Springer, Berlin/Heidelberg, pp 54–64
Robinson J (2000) Proof = Guarantee + Explanation. In: Hölldobler S (ed) Intellects and computational logic. Kluwer Academic, Dordrecht/Boston/London, pp 277–294

Ryle G (1950) 'If,' 'so,' and 'because'. In: Black M (ed) Philosophical analysis: a collection of essays. Cornell University Press, Ithaca, pp 323–340
Stein H (1988) Logos, logic, and logistiké: some philosophical remarks on nineteenth-century transformation of mathematics. In: Aspray W, Kitcher P (eds) Minnesota studies in the philosophy of science, vol XI. University of Minnesota Press, Minneapolis, pp 238–259
Sundholm G (2012) 'Inference versus consequence' revisited: inference, consequence, conditional, implication. Synthese 187:943–956
Tanswell F (2015) A problem with the dependence of informal proofs on formal proofs. Philos Math 23:295–310. https://doi.org/10.1093/philmat/nkv008
Thurston W (1994) On proof and progress in mathematics. Bull Am Math Soc 30:161–177
Weir A (2016) Informal proof, formal proof, formalism. Rev Sym Logic 9:23–43. https://doi.org/10.1017/S1755020315000234

Further Reading

Aberdein A (2006) The informal logic of mathematical proof. In: Hersh R (ed) 18 unconventional essays on the nature of mathematics. Springer, New York, pp 56–70
Avigad J (2020) Reliability of mathematical inference. Synthese. https://doi.org/10.1007/s11229-02524-y
Burgess J (2015) Rigor and structure. Oxford University Press, Oxford
Cellucci C (2008) Why proof? What is a proof? In: Lupacchini R, Corsi G (eds) Deduction, computation, experiment: exploring the effectiveness of proof. Springer, Milan, pp 1–27
Corcoran J (1973) Gaps between logical theory and mathematical practice. In: Bunge M (ed) The methodological unity of science. D Reidel, Dordrecht, pp 23–50
De Toffoli S (2020) Reconciling rigor and intuition. Erkenntnis. https://doi.org/10.1007/s10670-00280-x
Dove I (2009) Towards a theory of mathematical argument. Found Sci 14:137–152. https://doi.org/10.1007/s10699-088-9156-5
Feferman S (1979) What does logic have to tell us about mathematical proofs? Math Intell 2:20–24. https://doi.org/10.1007/BF03024381
Goethe N, Friend M (2010) Confronting ideals of proof with ways of proving of the research mathematician. Stud Logica 96:277–292. https://doi.org/10.1007/s11225-010-9284-0
Gonthier G (2008) Formal proof – the four-color theorem. Not Am Math Soc 55:1382–1393
Hales T (2008) Formal proof. Not Am Math Soc 55:1370–1380
Hamami Y (2018) Mathematical inference and logical inference. Rev Sym Logic 11:665–704. https://doi.org/10.1017/S1755020317000326
Harrison J (2008) Formal proof – theory and practice. Not Am Math Soc 55:1395–1406
Larvor B (2012) How to think about informal proofs. Synthese 187:715–730. https://doi.org/10.1007/s11229-011-0007-5
Leitgeb H (2009) On formal and informal provability. In: Linnebo Ø, Bueno O (eds) New waves in philosophy of mathematics. Palgrave Macmillan, New York, pp 263–299
Myhill J (1960) Some remarks on the notion of proof. J Phil 57:461–471
Pelc A (2009) Why do we believe theorems? Philos Math 17:84–94. https://doi.org/10.1093/philmat/nkn030
Rav Y (2007) A critique of a formalist-mechanist version of the justification of arguments in mathematicians' proof practices. Philos Math 15:291–320. https://doi.org/10.1093/philmat/nkm023
Sjögren J (2010) A note on the relation between formal and informal proof. Acta Anal 25:447–458. https://doi.org/10.1007/s12136-009-0084-y
Wiedijk F (2008) Formal proof – getting started. Not Am Math Soc 55:1408–1414

The Dark Side of the History of Proof

Proving the Correctness of Algorithms in the Ancient World

Karine Chemla

Contents

1 An Introduction to Proof and Computation: Some Historiographic Remarks 2138
2 Proof and Computation in the Context of the History of Proving the Correctness of Algorithms 2142
2.1 Why There Is No Feeling of Compulsion in Calculation 2142
2.2 The Various Practices Attached to Checking 2144
2.3 Practices of Proving the Correctness of an Algorithm and the Issue of Understanding 2146
2.4 Understanding and Proving as Actors' Categories 2149
2.5 Proof as Calculation and Why Not All the Algorithms Are Transparent 2153
3 Conclusion 2160
References 2160

Abstract

In the historiography of mathematics, proof and computation have regularly been perceived as opposed to each other in ways that have significantly obscured the discussion of the relationships between these two mathematical practices. This essay intends to draw on recent work in the history and the philosophy of mathematics to offer some reflections on these relationships.

Keywords

Proof · Computation · Algorithms · Cuneiform texts · The nine chapters on mathematical procedures · Liu Hui · Algebraic proof · Formal operation · Understanding · Practical certainty

K. Chemla (✉)
SPHERE (UMR7219, CNRS & Université Paris Cité), Paris, France
e-mail: chemla@univ-paris-diderot.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_81

For the centenary of the birth of Robert S. Cohen (1923–2017),
In memory of his passion for a global history of proof and, more generally, of science

1 An Introduction to Proof and Computation: Some Historiographic Remarks

In the depictions of an opposition between proof and computation, there are recurring themes that need to be highlighted and analyzed. With a view to pointing out problems in these commonplaces, I will start by quoting a few statements that illustrate widespread representations of proof and computation and in which these mathematical activities appear opposed to each other.

Let me begin with a statement by Jean-Baptiste Biot (1774–1862) that I have often quoted, but which remains useful inasmuch as it presents a concentration of many of the usual *clichés* on this issue. Indeed, in an 1841 review of a book by Jean-Jacques Sédillot, published in 1834 under the title *Traité des instruments astronomiques des Arabes*, Biot pointed out what, for him, was:[1]

> this peculiar habit of mind, following which the Arabs, as the Chinese and Hindus, *limited* their scientific writings to the statement of a *series of rules*, which, once given, *ought only* to be *verified* by their *applications*, *without* requiring *any logical demonstration* or *connections* between them: this gives those *Oriental nations* a remarkable character of *dissimilarity*, I would even add of *intellectual inferiority*, *comparatively* to the *Greeks*, with whom *any proposition is established by reasoning, and generates logically deduced consequences.* This *fixed* writing of scientific methods, in the form of *prescriptions*, must have represented an important *hindrance* for the development of *new* ideas for peoples for which it was in use, and it is in sharp *contrast* with our *European* motto: *nullius in verba.*

Here, Biot outlines an opposition between two types of mathematical practice. The practice with which he begins is clearly associated with computation, the term "rules" referring to the lists of operations – which we now call "algorithm" – that need to be "applied" to solve mathematical problems (to which Biot refers using the term "applications"). In Biot's view, works that contain such algorithms are characterized by several features. First, the rules are given as a series of disconnected statements. In particular, they present no logical connection with each other. Moreover, practitioners did not establish that "rules" were correct, but they simply "verified," that is, checked them: For Biot, the "applications" (the mathematical problems) were used precisely to this effect. Last, but not least, as the end of the declaration makes clear, Biot considered that the operations composing the text of a

[1] (Biot 1841, 674). Except for the Latin quotation (*nullius in verba*, "on the word of no one"), my emphasis. The Latin sentence is the motto of the Royal Society, and can be interpreted as referring to modern science. The first part of this excerpt is quoted in Charette 2012: 274. It is a great pleasure to thank the editors, Joachim Frans and Bart van Kerkhove, as well as Richard Kennedy, for their close reading of this article.

rule were "prescriptions," which simply required obedience and thus impeded the development of the critical mind that was essential for modern science.

The mathematical practice that, in the contrast Biot draws, he associates with proof is opposed point by point to the previous one: in this context, propositions are proved to be true, and by the effect of proof, they are connected with each other. What is more, for Biot, this practice invites practitioners to exert a critical mind towards authorities, which enables "the development of new ideas."

Several dimensions of the declaration look outdated and shocking to the modern reader: The fact that Biot's account associates each of the two practices with "peoples," and notably non-Western "peoples" lumped together, on one side, and "the Greeks" as well as the "Europeans," on the other; the fact that he gives the distinct mathematical practices identified as deriving from characteristics of the "peoples," which the practices in turn illustrate; and the fact that he takes the opposition that he asserts as a sign of the "intellectual inferiority" of the former peoples in comparison to the latter. However, these dimensions of Biot's assertions should not hide the elements of the representation of an opposition between computation and proof that he puts forward and that are still quite widespread.

A similar opposition using elements of the same kind is encountered in many coeval sources. Suffice it to mention the assessment that we read in William Whewell's *History of the Inductive Sciences*, first published in 1837, which reads as follows:

> The Arabs are in the habit of giving *conclusions without demonstrations*, *precepts without the investigations* by which they are *obtained*; as if their *main object* were *practical* rather than *speculative*,—the *calculation* of *results* rather than the *exposition* of *theory*. (Whewell 1837: 249, my emphasis)

Here, the term "precepts" replaces that of "rules," the emphasis being placed on the intention of obtaining "results" instead of providing the reasoning that had led to the "conclusions," or of presenting theory. For Whewell, this difference derives from a contrast between two main aims of the different practices: he associates with calculation a "main object" that is "practical," whereas, when reasoning and theory are emphasized, the "main object" is "speculative." The context makes clear that, in Whewell's view, just as in Biot's, the latter aim and practice characterize "the Greeks," while the former goal and the related practice is the proper of "the Arabs." However, what matters more for us here is the representation of computation that emerges from this assessment.

In fact, to this day, we still encounter echoes of views of this kind. This can be seen, for example, in the comments on proof and calculation that Ian Hacking puts forward in his 2000 essay: "What Mathematics Has Done to Some and Only Some Philosophers" and notably in the section titled "The Unpuzzling Character of Calculation."[2] I will focus on this facet of his essay, without discussing the broader

[2] (Hacking 2000), in particular, pp. 101–103. Note that similar ideas are put forward in (Hacking 2009: 72–87) and in (Hacking 2014: 99–100, 138–139, 145).

issues in the context of which Hacking deploys his representation of the opposition.

For Hacking, "a great many of the philosophers whom we still read have been deeply impressed by mathematics, and have gone so far as to tailor much of their philosophy to their vision of mathematical knowledge, mathematical reality, or, what [he] *think*[s] is *crucial*, *mathematical proof*" (p. 83, my emphasis). In this context, Hacking discusses features of mathematical proof that appear as essential to him in this respect, relying to this effect on the dialogue that Socrates has with the slave in the context of Plato's dialogue *Meno*. Indeed, in this exchange, a proof is developed that establishes that the area of a square whose side is the diagonal of a given square has an area double of that of the original square (Hacking 2000: 92–99).

Hacking emphasizes features of the proof that he will subsequently use to outline a clear-cut contrast with calculation. First, the proof is "compelling" (p. 98), establishing that the theorem is "necessarily true" (p. 97). Moreover, through the proof, we "can see" "*why* what we have found out is true" – a "fact" that "is one of the core phenomena of *some* proofs (. . .)" (p. 94, my emphasis). In addition, he argues, "the proof in *Meno* is perspicuous" or "surveyable" (p. 95). Following Russell, Hacking stresses that the proof gives us "the power of *anticipating* facts about things of which we have *no experience*" (p. 96, my emphasis).

Each of these remarks has a counterpart in Hacking's comments on calculation, which are grouped together in a passage where he gives an example of calculation that illustrates his view of it. The passage reads as follows:

> So let us turn to calculation, ordinary calculation. There is *nothing perspicuous about calculation*. There is *no feeling of compulsion*. There is *no sense of understanding*.
> (. . .) Divide 471 by 9. I can do this in my head. Nine into 47 goes 5 times, 2 to carry, and into 21 it goes twice, 3 to carry, so 471 divided by 9 is 52 and 3/9, or 52 1/3. I *quickly check to see that is correct*, by multiplying 52 1/3 by 9 in my head; yes, that's right, 471."
> "That is the result: 471 divided by 9 equals 52 1/3. *Must* it be the result? Do I see why this is the result? No, it simply is what happens *when I follow rules that I had learned before I was six years old*, and have used ever since. (Hacking 2000: 101, my emphasis)

The assertions about calculation are manifestly correlated point by point with the previous remarks about proof, resulting in a description of a clear-cut contrast between proof and computation. This representation evokes the way the contrast was portrayed by nineteenth century authors that we have evoked above. What is more, for Hacking too, the opposition is correlated with an opposition between peoples, as is made clear by his ensuing remarks:

> Most mathematics is not like the proof in *Meno* (. . .). Kant distinguished arithmetic from geometry. (. . .) Kant's distinction (. . .) can be paired, very loosely, *with two distinguishable modes of human cognition* (. . .), the *visual* and the *combinatorial*. It is a *happy accident that geometry and algebra took root in different places in different ways, among different peoples.* Of course there was a Greek study of numbers, but *geometry was the model of the mathematical*, and numbers were to a large extent understood as measures, as lengths. *Arithmetic and algebra are of Indian and Arabic devising.* That most combinatorial of

> concepts, the *algorithm*, is named after a 10th century mathematician, whose name, when Europeanized, is Al Gorismi.[3]

The correlation asserted between, on the one hand, "the Greeks" and geometry and, on the other, some Orientals and computation cogently evokes the nineteenth century statements quoted above. That said, in fact Hacking opposes the knowledge that can be acquired by a mathematical proof – such as the one presented in the *Meno* – and the knowledge available to an "Athenian land surveyor," who, in Hacking's imagination, operated with measurement and calculation.

Two remarks highlight how Hacking represents the knowledge of the latter as inferior to that of the former. First, the surveyor can only "*check* the 'anticipation'" provided by the proof in the *Meno* (p. 96, my emphasis), and, if he does not, "he *must* have made an error, or perhaps the land has shifted" (p. 96, emphasis in the original). Second, for other theorems, such as the one stating the incommensurability of the diagonal of a square and its side, the surveyor does not "anticipate nature," by which Hacking means that the experience of the measurement does not enable us to "check" the statement. However, he continues, the proof enables us to assert that "our measurement is inexact." (p. 96).

Only for one aspect does Hacking point out an interaction between the two practices of proof and computation. This interaction derives from a virtue that Hacking attaches to "calculation," as follows:

> Calculation has, however, a *virtue* altogether different from the insight of the theorem imparted by Socrates' talk and his squiggles on the sand. In the case of a calculation *you get practical certainty*, it seems, because each individual step can be *checked*. (Hacking 2000: 103, my emphasis)

In Hacking's view, this virtue inspired Leibniz's conception of proof as calculation:

> Leibniz was a member of the first European generation that could imagine that proof was essentially combinatorial, a matter of *checking the manipulation of signs*. (. . .) The starting points for proofs would be statements of identity, not real statements at all, not worth calling axioms (which have some content), but just identities. *Proof would proceed by rules of inference for manipulating signs*." (. . .) Leibniz "was overwhelmed by his sense of *proof as a sequence of steps* each connected to its predecessor by a single combinatorial *transformation of signs.*[4]

[3] (Hacking 2000: 99), my emphasis. Al Gorismi probably refers to al-Khwarizmi (ninth century). The assertion about the nature of number in Greek texts such as Euclid's *Elements* require qualification. Interestingly, in this context, the assertion is correlated with the subsequent assertion of a social distinction between the philosopher and the mathematician, on the one hand, and the "land surveyor" on the other. See below.

[4] (Hacking 2000: 103, 105, respectively, my emphasis). On this point, see (Hacking 1973).

The foregoing analysis has shed light on several recurring features that have been used to oppose proof and computation. Is it the case that there is "nothing perspicuous about calculation"? Can we accept the idea that, in computation, there is "no feeling of compulsion," "no sense of understanding"? Is it the case that calculation amounts to following rules and can only be checked – and quickly checked? What exactly does "checking" mean in this context? And what is exactly the concept of "calculation" that presides over these representations? I will now turn to a discussion of these issues in the context of a completely different analysis of the relationship between proof and computation.

2 Proof and Computation in the Context of the History of Proving the Correctness of Algorithms

Just as Biot indicated, Chinese, Sanskrit, and Arabic sources from the ancient world abound in mathematical texts dealing in one way or another with computation. As we will see, *pace* Biot, this does not mean that these sources limit themselves to prescribing orders. Moreover, we should not forget that, if we set aside Euclid's *Elements* (and, in this case, only Greek sources) and Apollonius's *Conica*, Greek and Latin sources also as a rule refer to computations. However, Chinese, Sanskrit, and Arabic sources, to which we will add hieratic and cuneiform sources that were not known in Biot's time, will suffice here to highlight problems in the understanding of computation that we have outlined above and to develop our discussion about proof and computation.

2.1 Why There Is No Feeling of Compulsion in Calculation

As we have seen, Hacking opposed proof and computation from the viewpoint that the former elicited a feeling of compulsion, and not the latter. However, is this comparison justified?

In Hacking's example quoted above, to divide 471 by 9, a sequence of operations is followed, and it yields $52 + 1/3$. To this, Hacking immediately appends another procedure, which he concludes by "that's right." In other words, the result is correct. This second part highlights that the computation is in fact a statement: it asserts that, if I apply a given sequence of operations to two numbers, I obtain the quotient of one by the other. Like any assertion, it can become the object of a proof. More generally, any computation is an assertion claiming that the application of a list of operations to some given numbers yields another number sought-for. Its correctness requires a proof.

If we observe the text of Plato's *Meno* more closely, we see that the requirement to construct a square whose area is the double of that of a given square is followed by a complex exchange between Socrates and the slave. Indeed, the dialogue entwines two types of statements. On the one hand, Socrates performs the construction of the

square sought-for step by step, much in the same way as a computation is carried out. On the other, Socrates regularly adds to these steps assertions regarding the successive states of the diagram constructed. The sequence of the latter assertions, in the end, forms a proof establishing that the area of the square constructed is the double of that of the original square. What communicates a feeling of compulsion is not the construction part, but the proof part of the dialogue.

Similarly, from a logical viewpoint, what might communicate a "feeling of compulsion" is the proof of the correctness of the computation, not the statement of the computation. Socrates' geometrical construction is not compelling – in any event, as little compelling as a computation: only the proof of the statement might be. The opposition of proof and computation from this perspective is thus misleading (Chemla 2012b).

Concerning the *Meno*, Hacking writes: "The fact that we can see not only that the theorem is true, but also why it must be true, is one of the core phenomena of *some* proofs, the sheer feeling of having 'got it'" (Hacking 2000: 94, emphasis in the original). The same assertion could be put forward with respect to the proof of the correctness of a computation. However, to examine this issue, we need first to address a more basic question: how can a proof of this kind be conducted? In what follows, we will rely on ancient sources to highlight key operations in this type of proof and thereby point out a dimension of the history of proof that has been overlooked in a historiography that has mainly focused on proving the truth of theorems (Chemla 2012a).

To begin with, a remark about computation is necessary to clarify the discussion. Let us return to the text of Hacking's calculation to formulate it. The text in question combines a list of operations with the sequence of related results. The first operation can be formulated as follows: the divisor "into" the first sequence of digits to the left of the decimal expansion representing the number that is greater than the divisor "goes" so many "times." Its results are 5, on the one hand, and 2, on the other. Proving the correctness of the algorithm involves proving that the sequence of operations actually yields the quotient. It also involves proving that the execution of each of these operations yields correct numerical results.

The significance of distinguishing between these two facets appears clearly with the example of an algorithm given to compute the area of a circle, which I will take from *The Nine Chapters on Mathematical Procedures* (*Jiu zhang suanshu* 九章算術) (Chemla and Guo 2004). This Chinese mathematical classic, which, in my view, was completed in the first century CE, is composed of problems and related algorithms, and it illustrates several facets of computation in which I am interested here. In what follows, I will thus regularly return to it. This classic is all the more interesting because it was handed down with a commentary that Liu Hui 劉徽 completed in 263 and a subcommentary that Li Chunfeng 李淳風 presented to the throne in 656, these two layers of exegesis offering systematically proofs of the correctness of the algorithms.

The first algorithm that *The Nine Chapters* gives to compute the area of a circle reads as follows:

PROCEDURE: HALF OF THE CIRCUMFERENCE AND HALF OF THE DIAMETER BEING MULTIPLIED ONE BY THE OTHER YIELDS THE *BU* OF THE PRODUCT/AREA.
術曰: 半周半徑相乘得積步。[5]

The commentary ascribed to Liu Hui that features immediately after the text of the algorithm proves geometrically that the algorithm is correct. However, it points out that it cannot lead to any numerical computation that would be exact. The example thus highlights the distinction that we must make between the correctness of the list of operations and that of its execution. In this case, no exact value can be given simultaneously for the operands of the algorithm.

Further, the comparison between this algorithm text and the text given by Hacking to illustrate what he means by computation illustrates several ways in which a computation can be formulated.

The text of an algorithm can bring into play ordinary operations ("halving," "multiplying") or other operations ("to carry"). It can also refer to the operations in various ways. In the case of the algorithm in *The Nine Chapters*, the key operation of multiplication takes as operands the results of two operations of halving, which must be performed first. In Hacking's text, the computation unfolds in time linearly, following the sequence of operations given by the text.

In fact, this latter feature is correlated with another characteristic trait that opposes both texts. Hacking's computation refers to an algorithm through its execution on a paradigm, that is, using the case of the division of 471 by 9. One might consider this text as reflecting the treatment of a particular case. However, one might also consider that the same text outlines the general algorithm in the same way as, in a grammar book on the French language, the conjugation of all the verbs of the first class is given using that of the verb *chanter*. Indeed, in some sources, general procedures are expressed in this way. Such a formulation for an algorithm is complete only if the execution for the paradigm illustrates all the situations to which, for other numbers, the division might confront the practitioner. In the case of the area of the circle, the algorithm in *The Nine Chapters*, in contrast, is given without reference to any paradigm. In fact, it could be applied to numbers expressed with respect to any numeration system and to any way of carrying out the operations mentioned. With these remarks in mind, let us turn to some key operations involved in proving the correctness of algorithms.

2.2 The Various Practices Attached to Checking

For Hacking, the proof in the *Meno* establishes the fact stated by the theorem in such a manner that knowledge of the theorem enables someone possessing it to "check"

[5] (Chemla and Guo 2004). To distinguish the text of *The Nine Chapters* from that of the commentaries, I put the quotations of *The Nine Chapters* in capital letters and those of the commentaries in lower-case. *Bu* refers to a measurement unit of distance as well as a measurement unit of surface (in the latter case, the number of *bu* refers to the area of a rectangle whose length has this measurement in *bu* and whose width is 1 *bu*.)

the correctness of the surveyor's results. In contrast, in his view, checking is what would establish the correctness of the execution of division he presents. Checking is indeed one way of approaching the correctness of an algorithm. Let us examine the various meanings that checking could take in this context as well as some conditions of validity for proofs of this kind.

Hacking's way of checking his calculation proceeds by applying the inverse computation to the result (i.e., multiplying the result by the divisor). Restoring the dividend is what leads Hacking to assert: "that's right." In this case, checking is carried out by applying a *procedure*, and its outcome in fact *might* establish the correctness of the result. The same conclusion would apply to the common way of checking the result of any arithmetical operation by "casting out nines": a positive outcome simply enables us to assert that the result *might* be correct. For a checking like Hacking's to enable us to actually draw the conclusion that the result is correct, we would indeed need to establish the correctness of the procedure used for checking. Moreover, let us insist: what is checked in this case is *merely* the fact that the result of the execution of *this* division is correct. However, it says *nothing* about either the correctness of the procedure or that of its execution.

Checking procedures of the same kind as Hacking's are regularly encountered in the Rhind mathematical papyrus. This document, which was the result of illegal excavations in Egypt, is written in hieratic and is believed to contain mathematical problems and procedures dating to the first centuries of the second millennium BCE. T. Eric Peet translates a term introducing this type of checking as "proof" (Peet 1923: 22, 70–76), a translation that was used to the same effect and further discussed by Gillings (1972: 232–234).

However, there is more: if we grant the correctness of the procedure for multiplication used to carry out the inverse operation, Hacking's checking procedure can work to check the result of his division of 471 by 9 *only because* the result of the division is given as exact. If the result had been given as the integral part of the quotient (52), this practice of checking would fail. If, in addition, the result of the division of 10 by 3 was given as 3.33, the same checking procedure would not succeed either.

These last remarks hold true when checking the result of an *execution*. They also hold true when an algorithm *qua* algorithm is proved to be correct through showing that the inverse algorithm (whose correctness has already been established), applied to the magnitude and numerical value yielded by the former algorithm, restores a magnitude and its numerical value that this algorithm had taken as its starting point. We find a case of this kind in *The Nine Chapters* (Chemla 1997/1998). The commentary ascribed to Liu Hui that establishes the correctness of the procedure in this way links the validity of the cancellation of the division and the square root extraction put into play in the procedure, by application of the inverse operation, to the fact that the algorithm for root extraction given in *The Nine Chapters*, like that for division, always yields exact results. In this way, as long as the correctness of the algorithms for multiplication, division, and root extraction is established, the checking procedure that consists in applying the inverse algorithm establishes the correctness of the original algorithm. As we show in Sect. 2.5, in fact, this practice of

proving the correctness of algorithms is merely one type among others, which all fall under a more general practice of proof.

These practices of checking, whether they establish the correctness of the result of the execution or that of the algorithm *qua* algorithm, can be characterized by the fact that they start from the result of a procedure (whether it be the magnitude obtained, the numerical value yielded, or both). In fact, in his discussion, Hacking refers to a different meaning of checking with respect to a calculation, when he introduces the "virtue" of "practical certainty" that would be brought about by calculation. Indeed, the certainty in question derives, in his view, from the fact that precisely "each individual step can be checked." The algorithm to compute the area of a circle quoted above from *The Nine Chapters* makes clear that the type of "checking" Hacking has in mind in this case does not apply to the algorithm *qua* algorithm, but to its execution. Indeed, the execution is what draws on the signs that represent the mathematical objects dealt with and proceed by mere "transformation of signs." We can think of a written procedure to carry out a multiplication for instance. A checking procedure in this case verifies that each of the successive manipulations conforms with the rule followed. It bears on the actions and not on the result of the global computation. Moreover, the efficiency of a checking of this kind will depend on the nature of the signs chosen for the execution and also on the type of manipulation put into play. In any event, Leibniz's idea to gain "practical certainty" for procedures of proving to which Hacking refers relied on this precise dimension of a practice of computation. We return to this issue below.

However, as many ancient sources illustrate, checking is not the only way of approaching the correctness of an algorithm. In fact, nowhere in the commentaries on *The Nine Chapters* do we find actors – actors here referring to the characters of the past whom we observe and thus being opposed to the observers whom we are – checking the numerical result of an algorithm. By contrast, they systematically seek to establish that the procedures given in *The Nine Chapters* actually yield the magnitude sought-for. In addition, for cases like the area of the circle, they discuss in which respect the numerical value obtained corresponds to it. The certainty thereby obtained is thus different from the "practical certainty" Hacking envisages. Let us examine the operations that such proofs put into play to this effect.

2.3 Practices of Proving the Correctness of an Algorithm and the Issue of Understanding

The oldest extant texts attesting to proofs of the correctness of algorithms are arguably mathematical tablets written in cuneiform script and dating from the first centuries of the second millennium BCE. These texts illustrate practices that highlight a key operation involved in proving, in cases where the proof concerns the correctness of a statement that unfolds in time, as that of an algorithm does. However, to understand this point, readers must apply a specific type of reading to these tablets, which Jens Høyrup was the first to advocate and which is correlated with the fact that the formulation of these texts is unexpected.

As Høyrup (2002, p. 11–13) emphasizes, the earliest translations of problems and procedures borne by a tablet such as BM 13901 shared a characteristic feature. This feature can be illustrated with the translation of its first problem published by François Thureau-Dangin in 1936, which reads as follows:

> J'ai *additionné* la surface et (le côté de) mon carré : 45'.
> Tu *poseras* 1°, l'*unité*. Tu *fractionneras* en deux 1° : 30'. Tu *multiplieras* (entre eux) [30'] et 30': 15'. Tu *ajouteras* 15' à 45': 1°. 1° *est le carré de* 1°. 30', que tu as multiplié (avec lui-même), de 1° tu *soustrairas*: 30' est le (côté du) carré. (my emphasis)

Thureau-Dangin, like other historians at the time, interpreted the problem as corresponding to the equation $x^2 + x = 45'$, and, in his view, the list of operations placed after its statement was the rule to follow to determine the value of the unknown.

The crucial feature of Thureau-Dangin's interpretation and, thus, of his translation is this: for him, the terms that I have italicized referred to arithmetical operations that had to be applied in succession, first, to the numerical values given by the statement of the problem and, also then, to those gradually determined along the text. In this way, the algorithm eventually yielded the value sought-for.

In a sense, Thureau-Dangin's interpretation reads in the text a formulation close to Hacking's example of a calculation. However, Høyrup raises an objection to this reading: if the terms italicized simply referred to arithmetical operations, why, he asks, in a given text of procedure, do we have several terms referring to the same operation? His answer to this question is this: in fact, each of these terms points out not only an arithmetical but also a geometric operation. Høyrup (2002) thus offers an interpretation of all the terms referring to operations in these texts of problems and procedures that convey both meanings. As a result, his translation of the same text in BM13901 reads as follows:

1. The surface *and* my confrontation: 45′ is it. 1, the *projection*,
2. you posit. The *moiety* of 1 you *break*, 30′ and 30′ you *make hold* each other.
3. 15′ to 45′ you *append*: by 1, 1 is the *equalside*. 30′ which you have made hold
4. in the inside of 1 you *tear out*: 30′ the confrontation. (I have emphasized the same terms as above)

Moreover, Høyrup restores the sequence of geometric operations to which the terms italicized refer as shown on Fig. 1.

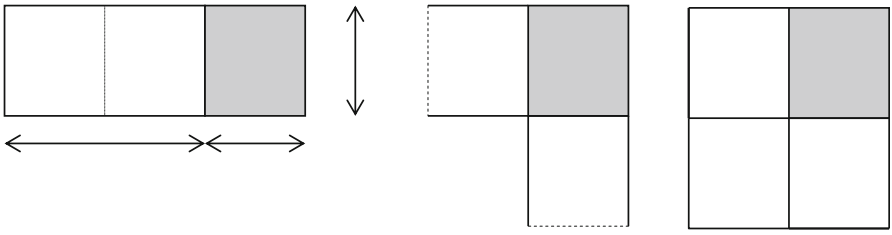


Fig. 1 "A possible geometric procedure for BM 13901 #1. Distorted proportions" (Høyrup 2002: 13)

In other words, in Høyrup's view, in the context in which BM 13901 was composed, a practitioner of mathematics read the technical texts by means of which cuneiform algorithms were written down, using a double interpretation. First, the terms referring to the operands and those prescribing operations designated, respectively, numerical values and arithmetical operations on them. Second, the same terms also indicated, respectively, a geometrical configuration and geometrical transformations that allowed the practitioner to associate a meaning with the operations carried out. This second facet of the interpretation enabled the user of the text to understand what the list of operations was carrying out and why the algorithm was correct.

The practitioner using such texts thus knew at the same time which operation to carry out and the reason why this operation was to be used. In contrast to a text of calculation such as Hacking's, for which the formulation yields no understanding, texts such as those we find on cuneiform mathematical tablets – if we adopt the reading practice restored by Høyrup –conveyed simultaneously a prescription and an understanding.

Two key points must be emphasized.

First, in such cases, the text of an algorithm delivered at the same time a procedure and the reasons of its correctness. However, for the proof of the correctness to be expressed in this way, the formulation of the algorithm had to possess the property that the list of operations could be interpreted step by step or sub-procedure by sub-procedure. I designate texts of this kind as "transparent." They imply a practice of formulating proofs that is quite specific. Below, I will consider fundamental reasons why not all texts of algorithms are transparent in this sense.

The fact is that we do find transparent texts for procedures in ancient cuneiform sources as well as in Chinese sources such as *The Nine Chapters*. In the latter case, the commentator makes explicit the meaning read in the formulation of the algorithm, and hence we can observe an ancient reader engaging with such texts precisely in this way (Chemla 1991). Interestingly, in this context, the commentator refers to the meaning of a step or a sub-procedure using the technical term of "meaning/intention *yi* 意." This actor's terminology, which points out a key dimension of the proof, designates the intention presiding over the use of an operation or a list thereof. By derivation, it also designates the global reasoning that accounts for the correctness of the procedure (Chemla 2012c).

In fact, if we return to the cuneiform tablet mentioned above, we observe that the text refers to the reasons of the correctness of an algorithm using another property of the formulation. Indeed, the terms employed for both the operands and the operations designate at the same time a computation to be carried out and the reason why it should be: This represents a second facet of the practice of proof to which these documents attest. In *The Nine Chapters*, we also find texts that use terms referring to operations in the same double way. These texts for procedures also illustrate a practice of writing down algorithms, for which the readers must apply specific reading to get from the text the specific list of operations to be applied depending on the case dealt with (Chemla 2015).

The second key point is this: in the case of the cuneiform tablet discussed above, Høyrup puts forward a hypothesis regarding the kind of device actors used to shape the meaning of the algorithm. He suggests that they used diagrams in the plane that could be transformed into other diagrams. In a sense, in the *Meno*, Socrates also makes use of a diagram of this kind: he puts together four equal squares around a point, emphasizing that the figure obtained is four times the original one. He then introduces diagonals in each of the four squares, interpreting them as cutting each square into two equal parts. Socrates can then consider the figure shaped by the four diagonals as making a space that is the double of the original figure: the evaluation of the extension proceeds by a counting that regularly requires interpretation based on the diagram.

The commentators of *The Nine Chapters* sometimes use diagrams of this kind, for instance, when they deal with the correctness of algorithms computing areas, or square roots, or else algorithms used in relation to right triangles. They also use blocks to deal with algorithms that concern volumes. Finally, as can be established by an observation of the commentaries, mathematical problems also appear to have had the key function of allowing practitioners to formulate the meaning of steps or sub-procedures of algorithms in the context of proofs of this kind (Chemla 2009). We will see an example thereof in the next section.

The foregoing analysis has borne on texts of algorithms independently of the way in which they were executed. However, seventh century subcommentaries on the Confucian classics contain descriptions of executions of arithmetical operations that also illustrate ways in which texts can weave the operations to be applied and the reasoning accounting for their correctness (Zhu 2016, 2022). To this end, the subcommentators introduced regularly geometrical interpretations of quantities, which allowed them to make the meaning of steps in the execution explicit. Interestingly, during the same decades, the mathematical canon that was taught at the Imperial University and that included *The Nine Chapters* as one of its main textbooks inculcated ways of executing the same operations in a completely different way: the execution bore on a decimal place-value numeration system and was carried out in a purely formal way – much in the same way as we carry out arithmetical operations in a written way. These two sets of texts nicely illustrate the tension that exists between the understanding of a calculation that can be acquired during its execution (the former texts) and, by contrast, the practical certainty that can be achieved through calculating formally on the basis of a symbolic notation (the latter).

2.4 Understanding and Proving as Actors' Categories

So far, we have approached the issue of understanding a computation from an observer's viewpoint. However, we have evidence attesting to the fact that actors also intended to gain understanding through proving. We will now examine a document of this kind to analyze the type of understanding actors sought in a

given context as well as what this tells us about their motivation for proving the correctness of algorithms.

The document in question is part of the commentary ascribed to Liu Hui on the algorithm given to multiply fractions. The passage of the commentary we are interested in starts with the following declaration: "If here one makes use of the formulation of a field with length and width, it is difficult to have (the procedure) *be understood with a greater generality* (*nan yi guang yu* 此田有廣從, 難以廣諭, my emphasis.)"[6]

The "formulation of a field with length and width" refers to the mathematical problem in the framework of which the multiplication of fractions is dealt with in *The Nine Chapters*, that is, a problem asking to compute the area of a rectangular field with fractional width and length (4/7 *bu* and 3/5 *bu*, respectively, to which, below, I refer as *a/b* and *c/d*). The problem is followed by a general and abstract procedure. From the commentator's perspective, this problem does not allow him to convey an "understanding" of the algorithm "with a greater generality." Indeed, this declaration is followed immediately by the statement of a sequence of three equivalent problems in the context of which the meaning of the steps of the procedure are made explicit. Let us outline how these problems are used to prove the correctness of the procedure and how we may interpret the understanding in question.[7]

The procedure under consideration reads as follows:

> Multiplying parts
> Procedure: the denominators being multiplied by one another make the divisor. The numerators being multiplied by one another make the dividend. One divides the dividend by the divisor.
> 乘分
> 術曰: 母相乘爲法, 子相乘爲實, 實如法而一。(Chemla and Guo 2004, 170–171)

In modern terms, the procedure prescribes multiplying the fractions *a/b* and *c/d* as follows:

$$\frac{a}{b}.\frac{c}{d}=\frac{ac}{bd}$$

To establish the correctness of the procedure, the commentary in which we are interested first introduces a problem solved by the last operation of the procedure (dividing *ac* by *bd*). The problem reads as follows – I insert symbolic notations between brackets for the sake of a better understanding:

[6] (Chemla and Guo 2004, 170–171). My emphasis.

[7] The connection between mathematical problems and the issue of "understanding" the algorithm through proving it also occurs in other passages of the commentary, for instance at the beginning of chapter 8, where we read: "It would be difficult to *understand* (the procedure) with abstract expressions (*kongyan*), this is *why* one deliberately linked it to (a problem of) *millet* to eliminate the obstacle. 以空言難曉, 故特繫之禾以决之。" (Chemla and Guo 2004, 947, my emphasis). On this issue, see (Chemla 2009).

> Suppose that one asks: 20 (*bc*) horses are worth 12 (*ac*) *jin* gold. If one sells 20 (*bc*) horses and if 35 (*bd*) people share this [the gain], how much does a person get? Answer: 12/35 (*ac/bd*) *jin.*
> To solve it, one must follow the "procedure for directly sharing" (NoT: division between quantities combining integers and fractions) and take 12 *jin* gold as dividend and 35 people as divisor.
> 設有問者曰, 馬二十匹, 直金十二斤。今賣馬二十匹, 三十五人分之, 人得幾何。荅曰: 三十五分斤之十二。其爲之也, 當如經分術, 以十二斤金爲實, 三十五人爲法。[8]

The commentator goes on by introducing a second problem, equivalent to the first one, but whose solution requires using all the operations of the procedure for multiplying fractions. Let me quote this problem and its solution, before commenting on them:

> Suppose that, *modifying* (the problem), one says: 5 (*b*) horses are worth 3 (*a*) *jin* gold. If one sells 4 (*c*) horses and if 7 (*d*) people share this [the gain], how much does a person get? Answer: 12/35 (*ac/bd*) *jin*.
> To solve it, one has to *homogenize* these quantities of people (*bd*) and gold (*ac*), and they (the quantities transformed) then all conform to (the quantities in) the initial problem and they are entered into the (procedure for) "directly sharing".
> 設更言馬五匹, 直金三斤。今賣四匹, 七人分之, 人得幾何。荅曰: 人得三十五分斤之十二。其爲之也, 當齊其金、人之數, 皆合初問入於經分矣。

To begin with, the commentary focuses on how the second problem can be solved. "Homogenizing," which is put into play here, is a technical term that belongs to a pair of operations: "homogenizing and equalizing," which the commentaries have introduced earlier and which recur in several proofs for crucial reasons, to which we return. What is meant here is this:

b horses are worth *a jin* gold.
c horses are sold by *d* people

hence
bc horses are worth *ac jin* gold.
bc horses are sold by *bd* people.

We thus see that horses are "equalized" to bring the two statements into communication. Accordingly, to warrant the validity of each statement, *a* is also multiplied by *c* while *d* is multiplied by *b*. These correlated transformations explain the relation between the two operations of equalization and homogenization and also the name of homogenization given to the latter. Now, as the commentary emphasizes, this global transformation reduces the problem to the first problem, and hence the sequence of

[8] For the sake of simplicity, when I translate Chinese passages from *The Nine Chapters*, I use so-called "Arabic numerals." It would be more appropriate, but also more cumbersome, to translate using English words expressing quantities. In the context of a discussion of proof, this nuance does not seem to me to matter.

the two homogenizations and of a division completely solves the second problem. However, the commentator continues:

> If this is so, "multiplying the numerators by one another to make the dividend (*ac*)", is *like homogenizing* the corresponding (quantity of) gold; "multiplying the denominators by one another to make the divisor (*bd*)" is *like homogenizing* the corresponding (quantity of) people. *Equalizing* the corresponding *denominators* makes 20 (*bc*). *But the fact that the horses are equalized plays no role. One only wants to find the homogenized (quantities) and this is all.*
> 然則分子相乘爲實者, 猶齊其金也; 母相乘爲法者, 猶齊其人也。同其母爲二十, 馬無事於同, 但欲求齊而已。(my emphasis)

Indeed, in the context of the second problem, we see that the homogenizations correspond exactly to the operations of multiplying *b* by *d* and *a* by *c,* that is, in other words, to the first two operations in the algorithm given to multiply fractions (which the commentary quotes *verbatim*). To put it differently, the complete algorithm for the multiplication of fractions thus solves this second problem, and now, the sequence of problems and related procedures allows us to see why the procedure solves this problem correctly. Finally, to conclude his commentary, the commentator points out that:

b horses are worth *a jin* gold.
c horses are sold by *d* people.
is equivalent to
1 horse is worth *a/b jin* gold.
c/d horse is sold by 1 person.

The third statement of the problem is now identical to the original problem in *The Nine Chapters.* The former problem – the second one considered, the commentator insists – is stated using integers and the latter – the third one, using fractions. Beyond the change of formulation, they amount to the same. However, looking at the operations, solving the problem with fractions from this new perspective has embedded the multiplications yielding *ac* and *bd* into a wider set of computations, which highlight their meaning as "homogenizations." As the commentator immediately stresses, the computation of *bc* is not needed to *solve* the problem. It is only required to highlight a meaning of the algorithm for the multiplication of fractions. In other words, this computation is needed not for the *solution* of the problem, but for the *proof* of the correctness of the algorithm solving it.

We see that the key point in the change of problem is precisely to enable this "equalization" to become visible. As a result, we grasp that bringing to light this additional operation is essential to the "understanding" of the algorithm as the commentator conceives of it. Moreover, we see how this understanding derives from the proof. In fact, several proofs allow the commentators to show that the same formal pattern of equalization and homogenization underlie different procedures and account for the correctness. Through highlighting the pair of operations "homogenizing and equalizing" – or similar formal strategies – in the proofs of the correctness of different algorithms, the commentaries bring to light that, beyond the

diversity of the procedures given to solve all kinds of problems, the same small number of formal strategies are put into play (Chemla 1991): these formal strategies have thus a greater generality than the algorithms themselves.

I interpret the second term that commentators use to refer to the "meaning" of a procedure, namely "*yi'* 義,"[9] as designating precisely the underlying formal strategy that accounts for its correctness and that at the same time reveals its connection with other procedures—*pace* Biot. "Homogenizing and equalizing" is one such formal strategy, and showing that it underlies the algorithm for the multiplication of fractions is what I interpret as yielding "understanding with a greater generality." We have seen the role played by the new problem to make this formal strategy visible. In the commentaries of *The Nine Chapters*, seeking to prove the correctness of the algorithms stated in the classic appears to aim primarily at finding out the fundamental formal strategies underlying them all. From this viewpoint, the connection between actors' motivation for proving and the "understanding" they seek to acquire appears clearly.

2.5 Proof as Calculation and Why Not All the Algorithms Are Transparent

The arguments developed in the previous sections have illustrated in various ways why, when proofs aim at proving the correctness of algorithms, they are entwined with computations. Our last example, in particular, has showed how it was through embedding the operations prescribed by the algorithm into a broader set of computations and through interpreting these computations in the semantic field provided by a problem that the correctness was established (Chemla 2009). However, in this case, the mathematical work put into play operations *qua* operations, and not their execution. This example makes clear why the fact that there might have been connections between proof and computation is to be distinguished from the aim of securing "practical certainty" for proving, in contrast with what has been suggested by Hacking. However, this issue calls for a deeper analysis, which I will present on the basis of a problem and an algorithm found in *The Nine Chapters*.[10]

The eleventh problem in chapter 5 of *The Nine Chapters*, titled "discussing works," and the procedure given in relation to it read as follows:

> Suppose one has a truncated pyramid with a circular base, the circumference of the lower circle of which is 3 *zhang*, the circumference of the upper circle of which is 2 *zhang*, and the height of which is 1 *zhang*. One asks how much the volume is.
>
> Answer: 527 *chi* 7/9 *chi*.

[9] I transliterate *yi'* and not *yi* to distinguish the second "meaning" from the first one discussed above.

[10] The text analyzed below is fully edited and translated in (Chemla and Guo 2004: 424–427). A much more detailed analysis than I can give here can be found in Chemla 2012c).

Procedure: The circumferences of the upper and lower circles being multiplied by one another, then multiplied each by itself, one adds these (the results); one multiplies this by the height and divides by 36.
今有圓亭, 下周三丈, 上周二丈, 高一丈。問積幾何。
荅曰: 五百二十七尺九分尺之七。
術曰: 上下周相乘, 又各自乘, 并之, 以高乘之, 三十六而一。

A commentary is appended to this procedure, and it establishes its correctness: I will highlight some of its characteristic features to illustrate aspects of the proof of the correctness of algorithms that have not yet been discussed so far, returning eventually to the issue of practical certainty.

Note that the commentator assumes that *The Nine Chapters* uses throughout the ratio of 3 to 1 for that of the circumference of a circle to its diameter. The proof of the correctness is thus given in the framework of this assumption.

If we call C_s (resp. C_i) the circumference of the upper (resp. lower) circle and h the height of the truncated pyramid, the sequence of operations given in *The Nine Chapters* to determine its volume can be represented as follows:

$$\begin{matrix} C_i \\ C_s \end{matrix} \xrightarrow[\text{Sum}]{} C_i C_s + C_i^2 + C_s^2 \xrightarrow[\text{by } h]{\text{Multiplication}} (C_i C_s + C_i^2 + C_s^2)h \xrightarrow[\text{by 36}]{\text{Division}} (C_i C_s + C_i^2 + C_s^2)h/36$$

What matters most for us here is this: the commentator's proof proceeds along three interwoven lines of argumentation. First, Liu Hui establishes an algorithm, for which he proves that it yields the desired volume, that is, the volume of the truncated pyramid with a circular base. However, this algorithm differs from the one given in *The Nine Chapters*, whose correctness needs to be established. This is where the second and third lines of argumentation are put into play. The commentator's second line of argumentation is the main focus of this section: Liu Hui transforms this algorithm *as such* into the algorithm the correctness of which is to be proved. For this, Liu Hui applies valid transformations to the algorithm taken as a list of operations, thereby modifying it progressively into other lists of operations, without affecting its result. In the following, we shall make clear what such transformations can be. Thirdly, in doing so, the commentator simultaneously accounts for the form of the algorithm as found in *The Nine Chapters*, by making explicit the motivations he lends to its author for not stating the algorithm as he or she most probably first obtained it, but instead changing it: we refer to this facet of his proof as the third line of argumentation. This whole process provides an analysis of the reasons underlying the correctness of the algorithm.

Let us concentrate on the parts of the commentary under consideration that matter to us in this section.

The first step in Liu Hui's reasoning is to make use of an algorithm for which the correctness was established in relation to the problem placed immediately before the one under examination. This algorithm provides the means to compute the volume of a truncated pyramid with a square base when one knows the sides of the upper square (D_s) and lower square (D_i) as well as the height h. It can be represented as follows:

Multiplications Division
Sum by 3
Multiplication by h

$$\begin{matrix} D_i \\ D_s \end{matrix} \longrightarrow (D_i D_s + D_i^2 + D_s^2)h \longrightarrow (D_i D_s + D_i^2 + D_s^2)h/3$$

On this basis, Liu Hui states a first algorithm (algorithm 1) which determines the volume of the truncated pyramid with a square base circumscribed to the truncated pyramid with a circular base which is considered. To do this, given the fact that the value of π lent to *The Nine Chapters* is 3, it suffices to divide by 3 the upper and lower circumferences of the truncated pyramid with a circular base to get the related diameters, which are the sides of the upper and lower circumscribed squares. From a syntactic viewpoint, the only transformation needed to make use of the algorithm above in this new context is thus to begin with the upper and lower circumferences and to prefix the text of the algorithm with two divisions by 3. These operations change the given circumferences into the corresponding diameters, the lengths of which are respectively equal to the lengths of the sides of the upper and lower circumscribed squares. This is how Liu Hui starts shaping an algorithm that will eventually compute the volume of the truncated pyramid with a circular base: this belongs to what I have called above the "first line of argumentation." This first algorithm (Algorithm 1) can hence be represented as follows:

Divisions by 3 Multiplications, sum, Division by 3
Multiplication by h

$$\begin{matrix} C_i \\ C_s \end{matrix} \longrightarrow \begin{matrix} D_i = C_i/3 \\ D_s = C_s/3 \end{matrix} \longrightarrow [\tfrac{C_i}{3}.\tfrac{C_s}{3} + (\tfrac{C_i}{3})^2 + (\tfrac{C_s}{3})^2]h \longrightarrow [\tfrac{C_i}{3}.\tfrac{C_s}{3} + (\tfrac{C_i}{3})^2 + (\tfrac{C_s}{3})^2]h/3$$

Liu Hui can assert the "meaning *yi*" of the result – as the volume of the truncated pyramid with square basis circumscribed to the pyramid under consideration – using the previous problem and algorithm as well as the numerical values employed by *The Nine Chapters* as corresponding to a value of π. This is precisely the virtue of the "first line of argumentation" in Liu Hui's proof of the correctness of the investigated algorithm. However, the next step in the proof belongs rather to both the second and the third lines. Indeed, Liu Hui will transform the algorithm whose meaning is established, *qua* list of operations, into another algorithm that yields the same magnitude and the same numerical value, but follows a different sequence of operations. This step makes us encounter the aspect of proof that is the main focus of the present section.

Indeed, the division of C_i and C_s by 3, which give respectively D_i and D_s, yields as a rule numerical results under the form of an integer (resp., a_i and a_s) increased by a fraction (resp., $b_i/3$ and $b_s/3$). In algorithm 1, these quantities subsequently enter into three multiplications, in the context of which they are multiplied by each other. These operations imply inserting at this point the algorithm that *The Nine chapters* gave for multiplying integers with fractions, the correctness of which has been

previously established. The first step of the inserted algorithm consists of transforming the said quantities into, respectively, $(3a_i + b_i)/3$ and $(3a_s + b_s)/3$ and then, as we have seen in the previous section, to multiply the numerators by each other, the denominators by each other, and finally to divide the result of the former by the result of the latter.

The key observation is this: dividing the circumferences by 3 gave numerical values corresponding to the sides of the circumscribed squares. Entering these numerical values into the inserted algorithm meant, to begin with, computing numerators through a multiplication by 3. From a numerical viewpoint, the multiplication by 3 cancels the effect of the division by 3 – precisely because the results of division are given as exact. The numerators computed therefore return in fact to the values of the circumference. Multiplying the numerators with each other is strictly identical to multiplying the circumferences with each other. Restoring the division by 3 and the opposed multiplication allows the practitioner to understand the meaning of the computation. Deleting these two operations makes the computation more straightforward. As long as one postfixes a division by 9, one can thus eliminate the first two operations of algorithm 1 and compute exactly the same magnitude (the volume of the circumscribed truncated pyramid with a square base) and its numerical value using algorithm 2, which can be represented as follows:

	Multiplications Sum Multiplication by h		Division by 9		Division by 3	
C_i C_s	———>	$(C_i\,C_s+C_i^2+C_s^2)h$	———>	$(C_i\,C_s+\,C_i^2+C_s^2)h/9$	——>[	$(C_i\,C_s+\,C_i^2+C_s^2)h/9]/3$

Algorithm 1 enabled the practitioner to determine the meaning *yi* 意 of the result: this is what is at issue in restoring in the first place the division by 3 and the opposed multiplication. However, when one computes, these operations are useless, and they can be deleted. Algorithm 2 derives from algorithm 1 by rewriting the list of operations in a formal way (this is typical of the second line of argumentation on which I lay stress here). The fact that this rewriting gives the algorithm the property that the computation becomes more direct belongs to the justification Liu Hui gives to explain why the algorithm was rewritten – I have referred to this part of the reasoning as the third line of argumentation. Note that the formulation of algorithm 2 implies that other transformations of the list of operations *qua* list of operations have been carried out on algorithm 1. One such transformation was swapping the multiplication by the height and the division by 9: this is again typical of the second line of argumentation. Elsewhere, Liu Hui accounts for the correctness of this transformation by reference to the fact that the results of the division are exact. In addition, Liu Hui makes clear that this avoids getting fractions as results and then computing with fractions (third line of argumentation).

To transform an algorithm computing the volume of the circumscribed truncated pyramid with a square base, such as algorithm 2, into one that computes the volume of the truncated pyramid with a circular base, one needs to append to the list of

operations a multiplication by 3 – which is assumed to be the value for π used in *The Nine Chapters* – and a division by 4. Doing this, we obtain algorithm 3, which performs the task to be fulfilled, namely, computing the volume of the truncated pyramid with a circular base (first line of argumentation). Algorithm 3 can be represented as follows:

Multiplications Sum Multiplication by h — Division by 9 — Division by 3 — **Multiplication by 3**

C_i C_s ——> $(C_i\,C_s+C_i^2+C_s^2)h$ —> $[(C_i\,C_s+C_i^2+C_s^2)h/9]/3$ —> $[[(C_i\,C_s+C_i^2+C_s^2)h/9]/3].3$

Division by 4

—> $[[[(C_i\,C_s+C_i^2+C_s^2)h/9]/3].3]/4$

The "meaning *yi*" of the result of Algorithm 3 is clear. However, if one used it to compute, computations would be cumbersome. This is why, once again, Liu Hui relies on the list of operations to transform it in a formal way into an equivalent list, more straightforward with respect to the computations prescribed. In doing so, the goal, from this point onwards, will be to transform this algorithm, *qua algorithm*, into the one for which the correctness is to be established, that is, the one provided by *The Nine Chapters* for the volume of the truncated pyramid with a circular base.

Here, Liu Hui hence resumes reasoning along the second line of argumentation. Considering the list of operations obtained by the last transformation, he remarks that the algorithm yielding the circumscribed truncated pyramid with a square base ended by a division by 3, whereas the transformation first appended to it a multiplication by 3. Once again, these two operations hence cancel each other out. This transforms algorithm 3 into algorithm 4, as follows:

Multiplications Sum Multiplication by h — Division by 9 — ~~**Division by 3**~~ — ~~**Multiplication by 3**~~

C_i C_s ——> $(C_i\,C_s+C_i^2+C_s^2)$h —> $[(C_i\,C_s+C_i^2+C_s^2)$h/9$]$~~/3 —> $[[(C_i\,C_s+C_i^2+C_s^2)h/9]/3].3$~~

Division by 4

—> $[(C_i\,C_s+C_i^2+C_s^2)h/9]/4$

The commentator clearly considers the operations that follow each other as a *list* and carries out a transformation of this list as such. The correctness of this operation, which is applied to a list of operations to produce another list that is equivalent to it, depends on the fact that the results of the division are exact – a point the commentator emphasizes (Chemla 1997/1998, 2012c). The latter transformation modifies the list of operations without altering the meaning or the value of the result. We meet here with the same phenomenon as above. Bringing to light the opposed multiplication and division was crucial to interpreting the meaning (*yi*) of the result. However, when viewing the list of operations as a means for computing, the two

operations appear unnecessary. This is how Liu Hui progressively accounts for the shape of the algorithm found in the Classic.

Exactly the same remarks apply to the subsequent transformation, which continues along the second line of argumentation, while applying a different operation to algorithm 4. Indeed, instead of carrying out the divisions by 9 and then by 4 in succession, the final transformation suggests "dividing in combination" (*lianchu*). This implies transforming the end of algorithm 4 into the multiplication of the two divisors by each other and dividing by the product. With the expression of "dividing in combination," we meet with a technical term that recurs regularly in the commentaries but is not to be found in *The Nine Chapters*. We may account for this by noting that it is a designation of the division typical of the mode of proving the correctness of algorithms on which this section concentrates.

Two successive divisions were accounted for, each being shown to be necessary for their own reasons. As above, Liu Hui had to dissociate them to bring to light the meaning of the result of the algorithm he shaped. However, viewing the list of operations as a means for computing leads to modifying the way of carrying out the computation and transforming the end of the algorithm. Liu Hui thereby accounts for the form of the algorithm given by *The Nine Chapters*, by highlighting that the two operations were grouped into a unique division. The technical term chosen for this division refers to the motivation of the effected transformation. As a consequence, algorithm 4

$$C_i, C_s \text{—} \text{—} > \left(C_i\, C_s + {C_i}^2 + {C_s}^2\right)h \text{—} > \left(C_i\, C_s + {C_i}^2 + {C_s}^2\right)h/9 \text{—} > \left[\left(C_i\, C_s + {C_i}^2 + {C_s}^2\right)h/9\right]/4$$

is transformed into the algorithm

$$C_i, C_s \text{—} \text{—} > \left(C_i\, C_s + {C_i}^2 + {C_s}^2\right)h \text{—} > \left(C_i\, C_s + {C_i}^2 + {C_s}^2\right)h/36$$

which is equivalent to it and identical to the desired algorithm. This was what was to be obtained: the correctness of the procedure provided by *The Nine Chapters* is established. The way in which the proof was conducted shows clearly how the activities of shaping an algorithm and proving the correctness are intertwined.

However, there is more. The algorithm obtained following the first line of argumentation is transparent in the sense put forward in Sect. 2.3. If the final algorithm no longer has this property, the reason is clear: the transparent list of operations was rewritten *qua* list of operations, in order to improve the convenience of the computation.[11] We hence meet here for a fundamental reason why not all algorithms manifest this property of transparency.

I have suggested to designate the type of proof illustrated in this section as "algebraic proof in an algorithmic context." Proofs of this kind are characterized by the articulation of the three lines of argumentation I have distinguished. However,

[11] On the transformations that can be applied to a list of operations, and how the commentators of *The Nine Chapters* account for their correctness, see (Chemla 2012c).

clearly, the second line of argumentation is the part of the proof that is specific to algebraic proofs.

The designation of this proof as algebraic can be understood by comparison with a modern rewriting of it, which would read as follows:

$$V = \frac{\left[\frac{C_i}{3}.\frac{C_s}{3} + \left(\frac{C_i}{3}\right)^2 + \left(\frac{C_s}{3}\right)^2\right].h}{3}.\frac{3}{4}$$

$$= \frac{\frac{\left[C_i.C_s + {C_i}^2 + {C_s}^2\right].h}{9}}{3}.\frac{3}{4}$$

$$= \frac{\left[C_i.C_s + {C_i}^2 + {C_s}^2\right].h}{9}.\frac{1}{4}$$

$$= \frac{\left[C_i.C_s + {C_i}^2 + {C_s}^2\right].h}{36}$$

The modern rewriting of the proof has replaced lists of operations formulated discursively, as we find them in *The Nine Chapters* and its commentaries, with lists of operations written as formulas. Notwithstanding this new way of representing lists of operations, in this proof, like in the proof analyzed above, lists of operations are rewritten into other lists to yield eventually the formula sought-for. Both types of proof further share the same fundamental reason for their correctness: the fact that divisions (and in other cases, root extractions) are given exact results. These are the two key reasons to consider them as similar. For these two types of reason, computation is central to proofs of this kind.

However, the precision that in the case of the commentaries on *The Nine Chapters*, the "algebraic proof" is carried out "in an algorithmic context" points to a key difference between the two. Indeed, the commentators draw on lists of operations formulated under a discursive form and carry out their rewriting in a discursive way. This is precisely the point where we can return to the issue of "practical certainty." Indeed, the introduction of notations of the type illustrated by the modern rewriting of our proof allowed practitioners to transform operations carried out discursively into operations that were performed on a notation comparable to the inscription of numbers using a decimal place-value notation. On these supports, their transformations only required "manipulating signs" and "transform[ing]" them.

The proof by the commentator highlighted a first link between proof and computation, since proofs proceeded through producing an algorithm executing a task and rewriting lists of operations carrying out this task formally. Now, the reformulations of lists of operations into formulas highlight a new connection between proof and computation. Indeed, in this change in the practice of proving, proofs have become manipulations of signs using rules similar to the execution of an operation relying on, e.g., a decimal place-value notation of number. So now the link

between proof and computation is established at the level of computation *qua* execution. The latter practice sheds light on *how* proof gains practical certainty. However, this dimension should not overshadow the multiple other ways in which proofs are tied with computation. Indeed, this change represented only the establishment of a new relationship grafted on an older practice in which proof and computation already had intimate links.

3 Conclusion

To conclude, *pace* Biot, Whewell, and Hacking, the ancient sources on which we have dwelled in this chapter shed light on several facets of the relationship between proof and computation that, so far, have not been discussed widely. It is probably not by chance that the ancient sources that proved useful for our project dealt with proofs of the correctness of algorithms. However, the conclusions they have allowed us to draw appear to be valid beyond this context.

First, proofs have appeared to be related to computations by the mere fact that some of them make use of computations. Second, we have seen that for some actors, proofs of this kind yielded understanding, however, in specific ways. Indeed, in some contexts, we have seen that proofs seek to highlight formal strategies that are more general patterns of computation. In the eyes of these actors, such a practice is what yields understanding. Gaining deeper understanding is a complex endeavor that actors undertake, and we might expect that a historical approach to this endeavor would be useful to interpret it in a nuanced way. However, for us, these remarks also raise the more specific question of how computations contribute to understanding – a question that we leave open for the moment. Finally, we have seen that proofs present ties with computation in at least two ways, depending on the facet of computation considered: computation as transforming magnitudes through operations and computation as producing a numerical value for a magnitude through the execution of these operations. Interestingly, numbers are involved in different ways in each of these facets: giving exact results to division and root extraction – more generally, drawing on the algebraic structure of field to which the numbers dealt with belong – is essential for some key operations on lists of operations carried out in the former sense of computation to be valid, while how one notates numbers and defines rules of rewriting on the notations is crucial to the second. This opens on a research program in the history of mathematics that would broaden these views and also inquire in a historical fashion into how these ties were established.

References

Biot J-B (1841) Compte-rendu de: *Traité des instruments astronomiques des Arabes,* traduit par J. J. Sédillot. Journal des Savants 513–520:602–610, 659–679

Charette F (2012) The logical Greek *versus* the imaginative oriental: on the historiography of 'non-Western' mathematics during the period 1820–1920. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 274–293

Chemla K (1991) Theoretical aspects of the Chinese algorithmic tradition (first to third century). Historia scientiarum 42:75–98+errata in the following issue
Chemla K (1997/1998) Fractions and irrationals between algorithm and proof in ancient China. Studies in History of Medicine and Science. New Series 15(1–2):31–54
Chemla K (2009) On mathematical problems as historically determined artifacts. Reflections inspired by sources from ancient China. Hist Math 36(3):213–246
Chemla K (ed) (2012a) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge
Chemla K (2012b) Prologue: historiography and history of mathematical proof: a research program. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 1–68
Chemla K (2012c) Reading proofs in Chinese commentaries: algebraic proofs in an algorithmic context. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 423–486
Chemla K (2015) Describing texts for algorithms: how they prescribe operations and integrate cases. Reflections based on ancient Chinese mathematical sources. In: Chemla K, Virbel J (eds) Texts, textual acts and the history of science. Springer, Dordrecht, pp 317–384
Chemla K, Guo S 郭書春 (2004) Les neuf chapitres. Le Classique mathématique de la Chine ancienne et ses commentaires. Dunod, Paris
Gillings RJ (1972) Mathematics in the time of the pharaohs. MIT Press, Cambridge, MA. Quoted on the basis of the reprint, 1982, Mineola, N.Y: Dover
Hacking I (1973) Leibniz and Descartes, proof and eternal truths. Dawes Hicks lecture on philosophy. Proceedings of the British Academy LIX:175–188. https://www.thebritishacademy.ac.uk/documents/2191/59p175.pdf. Reprinted in (Hacking 2002: 200–213)
Hacking I (2000) What mathematics has done to some and only some philosophers. In: Smiley T (ed) Mathematics and necessity. Essays in the history of philosophy. For the British Academy by Oxford University Press, pp 83–138. https://www.thebritishacademy.ac.uk/publishing/proceedings-british-academy/103/
Hacking I (2002) Historical ontology. Harvard University Press, Cambridge, MA
Hacking I (2009) Scientific reason. National Taiwan University Press, Taipei
Hacking I (2014) Why is there philosophy of mathematics at all? Cambridge University Press, Cambridge
Høyrup J (2002) Lengths, widths, surfaces. A portrait of Old Babylonian algebra and its kin, Sources and studies in the history of mathematics and physical sciences. Springer, Berlin/London
Peet TE (1923) The Rhind mathematical papyrus. British museum 10057 and 10058. Introduction, transcription, translation and commentary by T. Eric Peet. The University Press of Liverpool limited, London
Whewell W (1837) History of the inductive sciences from the earliest to the present times, vol 1. J. and J.J. Deighton, Cambridge
Zhu Yiwen 朱一文 (2016) (available online 23 April 2015). Different cultures of computation in seventh century China from the viewpoint of square root extraction. Hist Math 43:3–25. https://doi.org/10.1016/j.hm.2015.03.002
Zhu Yiwen 朱一文 (2022) Another culture of computation from 7th century China. In: Chemla K, Keller A, Proust C (eds) Cultures of computation and quantification in the ancient world. Numbers, measurements, and operations in documents from Mesopotamia, China and South Asia. Springer, Dordrecht, pp 555–602

Section IX

Ontology

Bharath Sriraman and Michael N. Fried

Ontology in the History and Philosophy of Mathematical Practice: An Introduction

Michael N. Fried

Contents

1 Introduction 2166
2 Ontology and Mathematical Practice 2168
3 History of Mathematics, Practice, and Ontology 2170
4 Ontology, Practice, and Mathematics Education 2174
5 Conclusion 2175
References 2176

Abstract

This very short introduction will first outline how ontological investigations and questions of practice go together. The second section will bring in the next pole of this entire book, history of mathematics. How do ontology, practice, and history go together? Is this a forced marriage or one born in true love? That is, do these three belong together in some very basic way? One chapter in the section argues that the philosophy of mathematical practice intersects with the history of mathematics essentially. If the first section of this introduction is persuasive that will also show that the history of mathematics intersects with ontological questions as well, specifically, that historical questions are very much about the ontological commitments of past mathematicians. The last section will take up what might seem a very different theme, namely, how these things matter to mathematics education. This section is connected closely to the second section since mathematics education has taken an interest in the role of history of mathematics in mathematics education, and much insight can be gained by seeing the issues in relation to the ontology-practice-history triad. The question of education is also important since teachers must consider what it is exactly that they are teaching, or, to adapt the title of

M. N. Fried (✉)
Program for Science and Technology Education, School of Education, Ben-Gurion University of the Negev, Beer Sheva, Israel
e-mail: mfried@bgu.ac.il

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_123

Reuben Hersh's well-known book, "What is mathematics really about?" And the "aboutness" of mathematics is really the ontological question.

Keywords

Ontology · Ontological commitment · Practice · History of mathematics · Mathematics education

1 Introduction

The chapters that make up this section are, on the face of it, quite diverse in focus. For many readers, it may not be evident that all belong together in a section called "ontology." This diversity owes partly to the broad spread of ontological investigations in general and partly to the wide range of meanings and manifestations of mathematical practice. As for the latter, this entire work amply demonstrates just how much is subsumed under the name "practice" and how these many forms of practice contribute to what we call "mathematics." Such are not only those forms of practice arising directly from a socio-institutional world where one must do things to get promoted at a university, to get published by leading journals, and, generally, to gain recognition. As important and relevant those forms of practice may be, mathematical practices, as has been shown in this volume, also include the devices we employ to persuade one another. Questions of proof are thus also questions of practice, for example, whether proof uses forms of algorithms to convince other practitioners or not (see Azzouni's chapter below). Even what one might deem purely logical reasoning can be viewed in terms of mathematical practice (e.g., the use of counterfactuals in reasoning discussed in Baker's chapter below). So, from the start, one should expect a great variety of topics and focuses in each part of this entire work.

But this still does not explain what these might have to do with ontology – and, here, just as with practice, the problem is only complicated by the range of questions and themes subsumed under the name "ontology." I should say a word about that. The most basic motives of modern ontological investigations still exist as they did for Aristotle, even though Aristotle did not use the Greek-inspired word "ontology" and even though we do not necessarily accept his own ontological conclusions. Those motives are rooted in the questions of what is and of the categories that apply to what is. But the *way* one pursues these questions depends much on what one imagines one is looking for in the first place. This is no less true in the circumscribed field of mathematics nor is the point particularly modern. The treatment of mathematical ontology in the Platonic tradition, for example, where mathematical objects – triangles, spheres, and numbers – were universal things existing apart from anything that happened to be triangular or spherical or a pair, albeit, on the scale of being, intermediate between the objects which are the source of images and the pure ideas, the *eidē*, naturally gave its distinctive color to ontological investigations. And as one

moves through the spectrum, through nominalism which denies universal objects altogether, to a kind of physicalism, like that of Mill where mathematical statements refer to physical things and mathematical concepts are derived from perceptible experience, mathematical ontology becomes a different kind of investigation altogether. And there are others, for example, ontological investigations inspired by structural semiotics (see Roy Wagner's contribution in this volume). But despite all this diversity, we can say this. Adapting the title of Reuben Hersh's well-known book (Hersh 1998), what an ontological investigation has always asked, now and in the past, is what is it that mathematics is *really* about?

In answering that question, what distinguishes many modern investigations of ontology (and what made Hersh's book so interesting to so many) is precisely their emphasis on forms of practice, that is, on *what mathematicians actually do*. In other words, for the philosophy of practice, the question of what mathematics is about is precisely the question what is it that mathematicians are doing when they do those things one designates as mathematical. Contrary to the emphasis in the past, the philosophy of practice thus ostensibly shifts the question away from the question of mathematical objects as the main subject of ontological inquiry. But it is more subtle than that, for our very practice may require us to make ontological assumptions of the old type, make assumptions, among other things, about mathematical objects and their objecthood. On the other hand, it may still be true that the theories driving our practices are accepted because we make assumptions about the nature of mathematical entities. So, we cannot simply remove ontology from practice as a different can of worms. The two are entwined.

This very short introduction, accordingly, must say a bit more about how ontological investigations and questions of practice go together. We will give our direct attention to this in the first section. In the second section, we will bring in the next pole of this entire handbook, history of mathematics. How do ontology, practice, and history go together? Is theirs a forced marriage or one born in true love? That is, do they belong together in some very basic way? One main theme in Marco Panza's chapter is that, indeed, the philosophy of mathematical practice intersects with the history of mathematics *essentially*. And if the first section is persuasive that will also show that the history of mathematics intersects with ontological questions as well, specifically, that historical questions are very much about the ontological commitments of past mathematicians. Finally, in the last section, we will take up what might seem a very different theme, namely, how these things matter to mathematics education. This section is connected closely to the second section since mathematics education has taken an interest in the role of history of mathematics in mathematics education, and much insight can be gained by seeing the issues in relation to the ontology-practice-history triad. More generally, however, the discussion of education brings us back again to our leitmotif inspired by the title of Hersh's book, for how we teach mathematics and what we expect students to know are unavoidably connected to how we answer the questions what mathematics is really about and what it is we hope students will do when they really do mathematics.

2 Ontology and Mathematical Practice

If one accepts a Platonic position regarding mathematical ontology in its usual form (see Panza for a *less* usual version), or just that of a hardcore realist, the general connection between ontology and practice is not all that difficult to state. For since mathematical objects, according to this view, exist in a form that is wholly fixed and unchanging, mathematical practice must only be adapted to the task of drawing out their equally fixed and unchanging properties. Practice, in this setting, can change of course: it can expand, become more refined, or even change direction. However, such changes are governed, and their effectiveness is measured by the presumed unchanging character of mathematical objects. For example, the power of using complex analysis in number theory was established when it yielded a proof of the prime number theorem, that is, when it could be shown to shed light on a fundamental fact about the distribution of fundamental mathematical objects, prime numbers. This, granted, is in many ways an oversimplification: for example, one can wonder about the sense in which Hadamard and Vallée Poussin's proofs were viewed as deep precisely because they used complex numbers and Erdos and Selberg's later elementary proof, precisely because it did not. Still, the point is only that when objects like prime numbers and their properties are thought to be given and fixed, *practice is subordinated to object*.

Interestingly enough, one can argue that the view traditionally *opposing* Platonism, nominalism, also puts objects before practice in its own way. For just as Platonism takes universal abstract objects – the kinds of things we usually think of as mathematical objects – as real existing things, nominalism takes only the individual concrete objects before us as real – the kind of things we usually think of as things. But since nominalism does commit to these individual concrete objects, once again our practice becomes measured by what it does with respect to *these* things. Once again, there is a standard to which practice is subordinated. But, of course, nominalism as a philosophy of mathematics has great difficulties, and not the least because it is hard to say what mathematical practice really is about: what does it mean to say that a practice is mathematical when the best you can do to identify a mathematical thing as mathematical is by putting one's finger on a certain similarity between concrete individual objects (which is not too far from what Mill did)? It is very hard to avoid slipping into a kind of Platonism after all.

It must be pointed out that there are variants of nominalism in the philosophy of mathematics (see Bueno 2020). One comes out of the thought of Jody Azzouni (though it is not the main thrust of his chapter in this volume) and has come to be known as deflationary nominalism (2004). This does take mathematical discourse seriously, that is, it does not question that mathematical practice is about something. In particular, the claim is that one can use quantifiers such as $\forall$ and $\exists$ without insisting that they say "For all such [existing] things" or "There exists a such or such," that is, without committing oneself to a set of existing mathematical objects that one then surveys to see whether something is true or all of them or just some of them. In a way, the claim is saying that one can speak seriously of "all" or "some" without the addition of "of them." It must be said, however, that whether deflationary

nominalism truly does allow us to speak about mathematical truths literally without committing ourselves to the existence of real mathematical objects has been debated (see, for example, Kovacs 2019).

The general tenor of these arguments comes close to another variant of nominalism, one that also rejects the existence of abstract universal mathematical objects but accepts that it makes sense to speak mathematics. This is fictionalism. Fictionalism, faithful to its name, treats mathematics like a work of fiction. This does not mean that mathematics, according to this view, is something frivolous – a work of mathematics is as serious to the mathematician as *King Lear* is to a reader of Shakespeare. Just as "true" and serious things can be said about the fictional Lear, "true" and serious things can be said about sets, functions, and categories. Taking mathematical truths literally means, then, precisely that, *literally* as in literature. But even in literature, one does wonder how it is that fictions tell us something "true." Is it not because while the characters are invented they somehow address aspects of real people? Is there a parallel to mathematical fictions? That is to say, is there in fact a hidden commitment to an ontological claim about real objects or real relations that, in the end, makes the fiction true and useful to us? Distinguishing fictionalism from nominalism, where abstractions are also just convenient words having no ontological status, Bueno (2009) says:

> . . .fictionalism is an *agnostic* view; it doesn't state that mathematical objects don't exist. Rather, the issue of their existence is left open. Perhaps these objects exist, perhaps they don't. But, according to the fictionalist, we need not settle the issue to make sense of mathematics and mathematical practice. (p.63)

Bueno's point that fictionalism does not take an ontological stand one way or another goes farther than simply making ontology inessential for mathematical practice to proceed: by allowing mathematical practice to be free from an ontological domination, we can now shift our attention to mathematical practice as *primary* or at least be on an equal standing with ontology. This may not be the central issue for fictionalism as such, but it opens the way to a different kind of relationship between practice and ontology.

As we already remarked in the opening of this introduction, the shift towards practice as primary or at least not secondary in mathematical philosophy allows for a much richer relationship between what mathematicians do, the practice of mathematics, and the existence, characteristics, and properties of mathematical objects, that is, mathematical ontology. Rather than being subordinate to ontology or, alternatively, unaffected by ontology, practice now may have an active role in shaping an ontology or at least in choosing an ontology. One can think of this in terms of Quine's ontological commitment (e.g., Quine 1953) which makes our ontology in mathematics dependent on the requirements of a mathematical theory. Although, granted, Quine's original formulation of this emphasized more the ontology our use of logical quantifiers committed us to, we can look at the commitment in a more active way, that is, what we assume to be the case to *do* mathematics according to theory we have adopted – and we always do mathematics according

to *some* theoretical position. Those assumptions, however, may again be about what is a mathematical object. The difference is that we come to the question *from the position* of working according to a particular theory or according to a set of procedures that we view as valuable, fruitful, or conducive to knowledge – itself partly determined by our theoretical position. With regards to practice, the two important points are: (1) that adherence to a theory – as opposed to, say, belief in a theory – is part of what we mean by adherence to a set of practices and (2) adherence and commitment are actions rather than consequences of discovery, something we do, not something that has come to us in some way (though, admittedly, if it is something we have been taught, it may not strike us as subject to adherence or commitment).

Taking our commitments and ways of doing mathematics as formative of a mathematical ontology rather than simply the result of it, naturally, raises the problem of the objectivity of mathematics – what has always been thought of as the hallmark of mathematics. Objectivity is unambiguous from a Platonist perspective where mathematical objects exist unchanging in some extra-human world and are universally intelligible. But once they are placed squarely in a human world, must one write "true" with quotation marks as we did in our description of fictionalism above? That may be the case. Yet, "true" still does have some meaning as "true," albeit with a somewhat weaker sense than it does in a Platonist view. Byers' term "weak objectivity," which he discusses in his chapter below, is useful in this regard. He grants that "objectivity" requires independence from anything "merely subjective," personal prejudice or extra-mathematical cultural forces, religion, race, politics, etc. "Of course," he goes on to say, "some people consider mathematics itself to be a culture but within mathematics the criterion of independence from arbitrary opinion might serve as a minimal way to differentiate subjectivity and objectivity – we could call this *weak objectivity*" (Byers, in this volume, section 13). Weak objectivity allows mathematics both to be dependent on human thought and practice, thus weak, and yet not arbitrary, thus objective.

3 History of Mathematics, Practice, and Ontology

Granting that what mathematics is about is dependent on human thought and practice means that mathematics, in this sense, not only has a history, a chronology of discoveries and historical figures, but is itself also a historical subject. Understanding what mathematics is is understanding different modes of thought and different forms of practice. The idea of practice, in fact, embodies thought. This is made clear by Kitcher's definition of practice, which sees practice as a complex of five components: L – the language of the practice; M – the set of metamathematical views; Q – the set of accepted questions; R – accepted reasonings; and S – accepted statements (Kitcher 1984, pp.163–4). More than specific mathematical techniques or modes of argument that we simply use, we begin to see how mathematical practice also entails what we accept as legitimate ways in which we can *speak* about ideas in mathematics or about mathematics itself and, ultimately, how we *think* about mathematics. And since that changes and develops, mathematics is necessarily, to that

extent, a historical phenomenon. With practice understood this way, Kitcher goes on to say that "Mathematical knowledge develops through the rational modification of the mathematical practices" (p.165). But of course such modification happens from the inside, that is, within the context of practice. This is perhaps why Panza in his chapter below says that "More than having a history, mathematical activity *is (or generates)* a history [emphasis added]." Arguing for the essential connection between philosophy and history of mathematics (as did Kitcher), Panza accordingly says trenchantly that "the philosophy of mathematical practice is. . .the philosophy of an historical fact."

This essential connection, notwithstanding, one should not *identify* philosophy of mathematical practice and history of mathematics. This is particularly important with respect to ontology and mathematical practice. In their attempt to understand the mathematical practice of the past, which we now see includes mathematical thinking of the past, historians of mathematics must engage, as it were, in the very ontological commitments made by their historical subjects. Unlike philosophers of mathematics, they cannot stand outside those commitments. In order to understand Euclid, historians need to try and think like Euclid; in order to understand Monge or Möbius, historians must try and think like Monge or Möbius.

This is really the idea of *reenactment* which Collingwood identifies as distinguishing historical thought from other forms of scientific thought (see Collingwood 1992, part V). The nexus that leads Collingwood to this begins with the simple observation that the field of history, the occupation of historians, concerns the history of human affairs. From here, he points out that unlike natural sciences (even natural *history*!), the history of human affairs has an outside and an inside. The outside is observable and physical, the external facts of an event. "By the inside of an event," he says, "I mean that in it which can only be described in terms of thought" (Collingwood 1992, p.213). Collingwood certainly appreciates the outside and recognizes that that aspect of historical events is essential, but he emphasizes,

> [The historian's] work may begin by discovering the outside of an event, but it can never end there; he must always remember that the event was an action, and that his main task is to think himself into this action, to discern the thought of its agent. (p.213)

It is here (though not only here) that Collingwood says famously that "All history is the history of thought" (p.215). And so the job of historians is to rethink the thoughts of historical figures in the historians' own mind. Needless to say, Collingwood's idea of reenactment has been criticized on several fronts as a way of interpreting historical events. In the case of intellectual history, like the history of mathematics, rethinking the thoughts of the historical mathematician is also problematic, partly because our own ways of thinking, our own prejudices, as Gadamer would say, cannot be entirely put aside. That said, in intellectual history, thinking the thoughts of our subject is certainly what one must *try* to do. At very least, one must assume that our subjects' thoughts are different from ours so that we must suspend as much as possible thoughts that clearly belong to our own time; it may not be entirely possible, but it is also not impossible. To the extent that it is not impossible, we are able to learn

from history, and what we learn, according to Collingwood, is thus something about the powers and limitations of our own minds (see Collingwood 1992, p.218).

Whether or not we can truly understand the mathematics of the past, we can seriously *mis*understand the past if we do not enter into the ontological commitments of past mathematicians and suspending our own. It should be emphasized that misunderstanding the past is not the same as making claims that are "wrong," as if one were speaking only of "facts," for example, whether or not there was a red house on the corner 10 years ago or not. The whole question of "rightness" or "wrongness" is one very much connected to one's ontological commitment, for it depends very much on *what* one is truly talking *about*. To make this more concrete, let us consider a historical example in some detail: let us consider the place of ontological commitment and thus also of mathematical practice in reading a historical work, the *Conics* of Apollonius of Perga.

In a classic book on the history of mathematics, one that in many ways I like and find useful, Carl Boyer (1985) writes:

> Apollonius, like his predecessors, derived his curves from a cone in three-dimensional space, but he dispensed with the cone as promptly as possible. (p.164)
>
> . . .
>
> After Apollonius had derived from a stereometric consideration of the cone the basic relationships between what we should now call the plane coordinates of a point on the curve—given by the three equations $y^2=lx-b^2x^2/a^2$, $y^2=lx$, and $y^2=lx+b^2x^2/a^2$—he derived further properties from the plane equations without reference to the cone. (p.165)
>
> . . .
>
> . . .whereas today we invariably refer a conic to a pair of mutually perpendicular lines or axes, Apollonius generally used a pair of conjugate diameters as equivalents of oblique coordinate axes. (p.166).

Are these statements *wrong*? Suppose a student who understands conic sections and other curves to be defined by their equations came to her mathematics teacher and said, "I have found a wonderful geometrical visualization for the conic sections" or even "a geometric derivation of the curves and their properties" and then presented her teacher with something like Apollonius's *Conics*. What would we say? Would we say that, in fact, she got it wrong? We would have to check. We would first establish some basic rules: length is a number, or a symbol for a number, x or y; a rectangle or an area is a product of such numbers or symbols, xy or x^2 or y^2; an area taken a certain number of times, say, m, is mxy; a ratio is a fraction, x/y, and so on. One would have to look at the geometry, see that it is correct, and, with these rules at hand, that the geometrical statements lead to some form of the familiar equations and relations. One would have to admit that, in this sense, it is not "wrong" to say that the geometric picture is truly a visualization or derivation of the "true" relations.

What we want to say is that it is anachronistic, historically wrong. But what is the content of the anachronism? Surely it is not just that our student is, presumably, writing in English or, in any case, not in ancient Greek. Nor, allowing that she is writing in English, that she puts things in her own words, in her own forms of

English expression. But is she putting things in her own terms? For her, it may be that the algebraic expression xy is no different than the "rectangle contained by segments AB and BC." For her, these may well be perfectly acceptable statements, perfectly within the bounds of acceptable metamathematical assumptions, and so on. For her. But, for Apollonius – and, one assumes, for those with whom he shares his thoughts – in his own inner view of conic sections, our student's way of seeing conic sections may be utterly foreign and foreign because of the kinds of ontological commitments they respectively adopt.

The primacy of magnitude or number and the primacy of relation and the unhesitating willingness to see $x, y, xy, x^2 + ay^3$ as genuine objects to be manipulated no differently than 3, 4,and 5 – these are marks of mathematical practice since algebra assumed its place at the heart of the subject. In this regard, what Descartes writes towards the end of his 1628 *Regulae* (Descartes 1970) signals the revolution of analytic geometry as pointedly, if not more, than what he writes in *La Géométrie* written almost a decade later, namely, that the product ab can be written as a line or some extended magnitude and can be multiplied by another magnitude c to produce another magnitude, and, conversely, that a magnitude given as a line can be written as an area (Rule XVIII). Whether x is a line or an area is, for him, *only a matter of interpretation*. The shift away from what x and y are is even clearer earlier in the same work (Rule IV) where Descartes shows that the kind of mathematics he aims for, the *mathesis universalis*, considers only order and measurement (i.e., relation and magnitude) and "that it make no difference whether it be in numbers, figures, stars, sounds or any other object that the question of measurement arises" (Descartes 1970, p.13). For him – and for us – mathematics is not about those objects as such but about pure relations between magnitudes, the latter are, for us, the "true" relations.

The ancient mathematicians certainly did care whether one is speaking of numbers or figures or stars. Mathematics, for them, is about such things. In at least one instance, namely, in Diophantus's *Arithmetica*, there are symbols; however, they serve as placeholders – unknowns, "blanks to be fill" – or word abbreviations of operations or number combinations. This is not to belittle what Diophantus did. I only want to stress that he is still focused on concrete numbers rather than, to use Viète's expression, operations on "species of numbers" where x or y may be treated as "things" in themselves. The ancient mathematicians could not speak about an x or a y without saying what they were speaking about – a line, an area, a number. Even in the most abstract book of Euclid's *Elements*, Book V, about ratio and proportion, one always reasons about lines taken as magnitudes. Symbols, what there are of them, are never taken as objects. For this reason, J. Klein saw the concerns of the ancients as being explicitly ontological while those of the moderns (i.e., since Viète, roughly) are symbolic. Indeed, he goes farther and says: "From now on [i.e. after Viète] the fundamental *ontological* science of the ancients is replaced by a *symbolic* discipline whose ontological presuppositions are left unclarified" (Klein 1968, p.184, emphasis in the original).

The ontological status of the conic sections, for Apollonius, is in their being sections of a cone. Even if, as Boyer correctly points out, the cone largely vanishes

after Book I of Apollonius's *Conics*, that does not mean that the cone was somehow inessential in light of the "equations" that he "derives" from it. First of all, he does not dispense with the cone altogether: it reappears at the end of Book I and at the end of Book VI. Its reappearance at the end of Book I is significant for there he uses it to show how to construct a conic section, for it is not enough to say that the points satisfy a certain relation but that the curve arises from a cone. So the cone is there implicitly even when it does not appear explicitly. Indeed, it is for this reason that the entire work begins with the definition of a conic surface and cone and then launches into a "mini-treatise" on the cone and how it is sectioned (see Fried and Unguru 2001 for more details). Writers such as Boyer and Heath do not forget Apollonius's cone, but they trivialize it, for their own ontological commitment does not require that a curve come from a geometric situation but only that it represents an algebraic relation. Their misunderstanding of Apollonius is not a misunderstanding of the relations that they can derive from his geometry but of his own ontological commitments. And these commitments are part and parcel of his geometric way of doing mathematics. To understand historically is to see clearly that the way mathematicians have done mathematics in the past was not a matter of mere style or peculiarity of language but an expression of what mathematics was about for them: their practice – Apollonius's sectioning cones, for example – and their sense of mathematical ontology are inseparable components of their mathematical thinking.

4 Ontology, Practice, and Mathematics Education

Let us conclude this introduction to the section on ontology in the history and philosophy of mathematical practice by saying a few words about mathematics education. Indeed, should any of this matter to mathematics educators? I think it should. As mentioned at the outset, mathematics educators do have to ask themselves the ontological question: what is it that mathematics is about? After all, this is also a question that students ask. Students rarely understand the complexity of the question, and, sometimes this is so of their teachers as well.

On the other side, situating our own discussions about practice and ontology in an educational setting can put our own questions in a new light. In that connection, one recalls Dewey's remark towards the end of *Democracy and Education* (2004) that, "Education is the laboratory in which philosophical distinctions become concrete and are tested" (p.316). It is thus not by accident that Byers' discussion below about mathematics as process and creative rationality revolves around the learning of mathematics, that is, the coming to know mathematics.

For him, though, coming to know mathematics is not a purely epistemological issue but an ontological one: what mathematics is about, as he sees it, arises out of our doing mathematics and the subsequent reification of our activity. In other words, learning mathematics is creating mathematics; it is not merely having students adapt themselves to an already crystalized body of knowledge. Students make mathematics as mathematicians do, and that insight, Byers stresses, provides a common ground for educators and people who think about mathematics, mathematicians, to shed light on one another's concerns. As he puts it in part 4 of his paper:

> Though mathematicians and math educators do not normally communicate with one another because they live in different disciplinary compartments, educators are well placed to understand the difference between being content or process oriented. Educators and teachers need to focus on process because it is evident that learning is a process. Today, influenced by a view of knowledge as mere data, there is a tendency to think of learning as the accumulation of facts and techniques. But this is not what I mean by learning or doing mathematics and has little to do with becoming fluent with a particular mathematical world.

In a way, what Byers says about the focus on process and the way doing and creating mathematics run hand-in-hand is very much along the lines of what was known and much talked about in the 1990s in educational theory as radical constructivism – and, needless to say, mathematics education research was particularly influenced by it. Radical constructivism, according to Ernst von Glasersfeld (1989), whose name is associated with it, is grounded in Piaget's constructivist thought which, like Byers, sees process and particularly operations as primary. More distantly but surely, it is, he says rooted in Vico's eighteenth century thought which held that what we know is what we have made – and, for Vico, mathematics is no exception to this: "Geometrica demonstramus, quia facimus" ("We demonstrate geometry because we make it"), as Vico said in his address, *On the Study Methods of Our Time* (Vico, (1708)1990, p. 23), significantly a work reflecting Vico's philosophy in thoughts about education. Von Glasersfeld's reference to Vico, as his epigraph taken from Piaget and Garcia's *Psychogenesis and the History of Science* (1983), is telling, for Vico's idea that what we know is what we have made necessarily raises history to a high level as a form of knowledge, and Vico did just that. Byers too recognizes that reframing mathematics in terms of process, *doing* mathematics, makes considerations of history essential for understanding mathematics and the learning of mathematics.

Conversely, in trying to clarify what it meant to have an historical understanding of mathematics above, I too resorted to an educational context where a student presents an Apollonian geometrical development of conic sections to her teacher. The point was to show that the close interrelationship between practice and ontology and the kinds of ontological commitments that arise from it make statements about "rightness" and "wrongness" problematic when they refer to historical mathematics: one must take into consideration both the practices, for example, geometrical or algebraic practices, and the ontological commitments of the historical mathematicians. In a way, the lesson that ontology and practice cannot be separated is just a realization of the standard wisdom in all historical work that content cannot be separated from form.

5 Conclusion

Viewing mathematics from the perspective of practice and ontology – and, therefore, history – can certainly enlighten teachers and others who deal with mathematics as to what mathematics is "really about." This provides also a solid justification for bringing history of mathematics into the mathematics classroom (see Fried 2014

for the general issues involved). More often than not, however (luckily not always!), the justification is vitiated by the tendency to see history is only something to be *used* (Fried 2001), or, as Jankvist put it, treat history as a tool rather than a goal (2009). The problem with seeing history as a tool only is that when one asks what it is a tool for, one comes to a rather static view of mathematics with an unproblematized ontology: one has to teach the subject of quadratic equations in two unknowns – taken as an almost Platonic category – and then looks for any historical context which can be construed to be about this, for example, Apollonius's *Conics*. There are no equations in the *Conics*, but one can say they are there, *mutatis mutandis*. Using history in this way makes form independent of content, and it fails to recognize how practice is interconnected with ontological commitments: it certainly does not take into account ontological commitments of the past. This makes it bad history, but more importantly, by not taking into account ontological commitments of the past, we and our students may not be fully aware of our own ontological commitments and how those are interacting with our own mathematical practice. Taking history quite seriously, as Byers and Panza do in their essays, that is, *not* merely as a tool for something else but as an opportunity to explore the process of mathematics and what says about mathematical ontology, can inform mathematics education in ways that would make it deeper. Thinking about mathematics education and about what messages about mathematics we convey to students is also, as I have said, allows us a way of confronting the philosophy of mathematical practice and mathematical ontology.

References

Azzouni J (2004) Deflating existential consequence: a case for nominalism. Oxford University Press, New York

Boyer CB (1985) A history of mathematics. Princeton University Press

Bueno O (2009) Mathematical fictionalism. In: Bueno O, Linnebo Ø (eds) New waves in philosophy of mathematics. Palgrave Macmillan, Basingstoke, pp 59–79

Bueno O (2020, Fall) Nominalism in the philosophy of mathematics. In Zalta EN (ed) The stanford encyclopedia of philosophy. https://plato.stanford.edu/archives/fall2020/entries/nominalism-mathematics/

Collingwood RG (1992) The idea of history. Oxford University Press

Descartes R (1970) The philosophical words of descartes (trans and eds: Haldane ES, Ross GRT), Vol. 1. Cambridge

Dewey J (2004) Democracy and education. Dover Publications, Inc

Fried MN (2001) Can mathematics education and history of mathematics coexist? Sci Educ 10: 391–408

Fried MN (2014) History of mathematics and mathematics education. In: Matthews M (ed) History, philosophy and science teaching handbook, vol I. Springer, pp 669–705

Fried MN, Unguru S (2001) Apollonius of Perga's Conica: text, context, subtext. Brill Academic Publishers

Hersh R (1998) What is mathematics really? Vintage

Jankvist UT (2009) A categorization of the "whys" and "hows" of using history in mathematics education. Educ Stud Math 71(3):235–261

Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press

Klein J (1968) Greek mathematical thought and the origin of algebra (trans: Brann E). MIT Press
Kovacs DM (2019) Deflationary nominalism and puzzle avoidance. Philos Math 27(1):88–104
Piaget J, Garcia R (1983) Psychogenese et histoire des sciences. Flammarion
Quine WV (1953) From a logical point of view. Harper Books
Vico G (1990) On the study methods of our time (trans: Gianturco E). Cornell University Press
von Glasersfeld E (1989) Cognition, construction of knowledge, and teaching. Synthese 80(1): 121–140

The Algorithmic-Device View of Informal Rigorous Mathematical Proof

Jody Azzouni

Contents

1 Mathematics Is Strange 2180
2 Effective Recognition Procedures 2182
3 Why Derivational Accounts of Mathematical Proof Don't Work 2185
4 Algorithms 2189
5 Recognition Procedures in Games 2195
6 Language-Based Algorithmic Systems and Algorithmic Devices 2199
7 Some Observations About Algorithmic Devices, Conceptual Engineering, Synchronistic Effects, and Imperatives 2202
8 Diagram-Based Formal Systems; Shin's Venn-II 2209
9 Informal Reasoning by Means of Algorithmic Devices 2214
10 The Generality Problem 2219
11 An Algorithmic Device in Natural-Language Mathematics: Euclidean Diagrams 2224
12 Phenomenologically Faithful Embeddings of Algorithmic Devices in Formal Algorithmic Systems 2236
13 Inference Packages (Embodied Algorithms) and Phenomenological Faithfulness 2241
14 Inference Packages and Shin on Perceptual Inference 2246
15 Ideology and Formal Systems 2251
16 Conclusion 2256
References 2257

Abstract

A new approach to informal rigorous mathematical proof is offered. To this end, algorithmic devices are characterized and their central role in mathematical proof delineated. It is then shown how all the puzzling aspects of mathematical proof, including its peculiar capacity to convince its practitioners, are explained by algorithmic devices. Diagrammatic reasoning is also characterized in terms of algorithmic devices, and the algorithmic device view of mathematical proof is compared to alternative construals of informal proof to show its superiority.

J. Azzouni (✉)
Department of Philosophy, Tufts University, Medford, MA, USA
e-mail: Jody.Azzouni@tufts.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_4

Keywords

Informal mathematical proof · Diagrams · Turing computability · Algorithms · Euclidean geometry · Venn diagrams · Formalization · Conceptual engineering · Games

1 Mathematics Is Strange

A mere couple thousand years ago, mathematics seemed innocent enough. (At first glance, anyway.) You sketched out some fairly simple diagrams, and you talked them up in ancient Greek. You said things like, "look over here," "look at that," and then you drew a conclusion or two (Plato 1963). Somehow, nevertheless, it all amounted to something like nothing else we do – like nothing else we do *on Earth.* The conclusions drawn had an exactness impossible even by careful measurements (interior angles of triangles were proved to be *exactly* 180° – nothing slightly more or less, which is what always happened if, instead, the angles of a triangle were measured directly)[1]; furthermore, the results shown this way seemed to be necessarily true (they *couldn't be* otherwise), and so the process ("reasoning") by which these results were achieved seemed independent of our usual tools for learning about things – our senses. Equally odd was that these conclusions were often unexpected and sometimes undesirable. Many of them were results no one saw coming; some of them were even results that no one *liked* after they arrived – for example, the famous (and scandalous) fact that the length of the diagonal of a triangle isn't a ratio of the lengths of the adjacent sides of that triangle … from which it can be shown that $\sqrt{2}$ isn't rational.[2] Call this *the epistemic strangeness of mathematical proof.*

Philosophy was born, perhaps, from the desire to explain the weird and bizarre – less melodramatically, from the need to explain *odd cases.* This used to be called *wonder*, a "what on Earth is going on *here*?" Shortening a long descriptive history of (very ingenious) explanations of the epistemology of mathematical proof, many of them *supernatural* – such as mathematical proof being our recollection of eternal entities that we perceived directly in a previous life, or as involving innate ideas that have been stamped conveniently into our souls by God, or proofs being sequences of thoughts true by virtue of their conceptual structure alone (syllogistic sequences, "analytic truths"), or proofs as involving the addition of content from "intuition" ("synthetic a priori truths"), or proofs being lists of sentences true by the meanings of the words within them, etc. – a post-mid-twentieth-century conceptual revisiting of the artificial ("formal") languages that were invented by Frege (1879) at the end of

[1] Also *very* striking was that these exact results arise from methods applied to diagrams – compass and straightedge – that themselves don't intrinsically involve exact measurements: no ruler is used, for example.

[2] A rumor passed down from ancient Greek times is that the person who discovered this was *murdered.*

the nineteenth century seemed to *finally* offer a conclusive explanation. Derivations in such *artificial* languages are effectively recognizable patterns of reasoning.[3]

That is (and here's one sort of explanation gestured at in the last sentence above), if the many mathematical proofs in natural languages just *are* (disguised) derivations in artificial languages – if they're abbreviations of such derivations, for example – then this can explain all the odd epistemic qualities mathematical proofs in natural languages have. It can explain, for example, the impression of necessity: Our senses play only as much of a role as they seem to with ancient Greek diagrams (only enough to enable us to manipulate the language entities involved, and nothing more), and the results aren't necessary or eternal, but only as necessarily true as the premises the proofs start from. If logical inference is to be characterized as formal derivation in these artificial languages, then the epistemic qualities of mathematical proof become purely matters of logic. In particular, the strange unchanging and perfect eternal entities that Plato used to explain the apparent properties of mathematical results (exactness and necessity) are no longer needed: If an argument that mathematical entities exist is to be given, it can no longer arise via an inference to the best explanation for the epistemic qualities of mathematical proof; instead, it must result from the requirement, when applying mathematics to empirical science, that one necessarily uses phrases like "there are numbers. . .," coupled with the linguistic claim that such a phrase commits us to the existence of numbers.[4]

Call a *derivational account of informal rigorous mathematical proof*, any account that takes the epistemic qualities of ordinary mathematical proof to derive from a relationship (of some sort) between ordinary mathematical proofs and formal derivations in one or another (or a family) of artificial languages.

[3]I'm condensing some intellectual history here. "Effective recognizability" as relevant to an explanation for the epistemic strangeness of ordinary mathematical proof emerged after the discovery of formal languages (it emerged, perhaps, even some *decades* after the Church/Turing/Kleene discovery of the mathematical characterization of effective recognizability). The possible explanation Fregean languages initially seemed to offer for epistemic strangeness was instead a development of a much earlier view (one already in, to some extent, Descartes, and certainly in Leibniz) that mathematics is just "logic" *once all the missing inferential steps are explicitly given ("filled in") in the corresponding derivations.* There still is an issue, of course, about the epistemology of logical inference: how challenging this issue is corresponds to how much apparent mathematical content has been pressed into the "logical" principles. In any case, the explanation doesn't work because it has the same problems faced by *any* explanation of the epistemic properties of ordinary mathematical proofs in terms of their relations to derivations in artificial formalisms (see Sect. 3 below).

[4]That is, I'm describing proponents of what have come to be called *indispensability arguments.* There is a large contemporary literature on these arguments. For the version of them that influenced their recent application to mathematics, see Quine (1953). See Azzouni (2009a) for an indispensability argument with weak premises that establishes only the truth of applied mathematics and not the existence of mathematical *objects*; see Panza and Sereni (2015) for a categorization of the whole family of indispensability arguments according to the strength of the premises used in those arguments, and for a survey of this literature as of that date.

2 Effective Recognition Procedures

Before discussing the drawbacks of derivational accounts of mathematical proof, I should first discuss in more detail why formal derivations of formal languages seemed and continue to seem so illuminating about (or, at least relevant to) the epistemic properties of informal rigorous mathematical proof. The key point is that formal derivations come with *effective syntactic recognition procedures.* That is, anyone (including computers or mechanical calculating devices) can acquire a set of rules by which she or he or it can determine after a finite amount of time, *whether or not* a sequence of strings of vocabulary items is a proof. In describing these recognition procedures as "syntactic," I'm pointing out that a person can execute such a procedure successfully without knowing what the sentences "mean"; indeed, such syntactic recognition procedures don't require knowledge that what's being recognized are "sentences" of a language. So far, I've only mentioned well-known truisms about formal languages. I now compare natural languages with formal languages, and how they differ on the recognition procedures they come with. What I say next isn't quite as truistic; some of it is even controversial.

Natural languages also have recognition procedures. What's required of *every* language to begin with (ones we can *use*, anyway – one way or another) is that there be recognition procedures that enable the sorting of vocabulary tokens (the "alphabet") into type-classes. Vocabulary types are sensorily accessible classes of physical items, for example, the type of items, "e," that include those that occur in the paper you're reading. These may be electronically composed of pixels or instead they may be chemical stains on a background medium. We idealize these type-classes of *physical* vocabulary items by assuming that there is a decision procedure (looking "by eye"[5]) that determines what type-class *any* candidate physical item belongs to. Implicitly excluded (call these *exclusion idealizations*), therefore, are all those many physical tokens that can't be decided by eye because (among other things), they're too blurry or badly shaped. The very coherence of physically executing *any* algorithm (by humans, computers. . .) *requires* exclusion idealizations.[6]

Both natural languages and many formal languages – not all formal ones – have recognition procedures for primitive vocabulary. Consider now the recognition procedures for syntactic classes of items other than primitive vocabulary items – in particular, for sentences. The recognition procedures available for these syntactic classes of items of formal languages go beyond those of natural languages in two ways.

[5]"By eye" is metaphor; other human (or nonhuman) senses can be used instead of vision.

[6]This is true of physically real Turing machines: a written symbol in a cell that a Turing-machine head "scans" can't be so lightly printed that the Turing-machine head is unable to tell whether or not the cell is blank. At least when it comes to humans, these recognition procedures are extremely complex, neurophysiologically speaking. If one physically built such a thing, scanning could be designed to be mechanically quite simple, as it isn't for us. Even so, there would still be cases where the physical device failed to scan a cell successfully; this is a fact about any Turing machine that exists in space and time. Indeed, this is true of any actual calculating device whatsoever: there will always be "mishaps."

First, although we *can* assume that those competent in natural languages have (and use) recognition procedures for the sentences of those languages, we can't state these procedures explicitly.[7] We *can* state in mathematically precise ways the recognition procedures for the sentences of many formal languages – in particular, Frege's original language, and the many formal languages currently studied, including the first-order predicate calculus.[8] We can describe and mathematically study, of course, many formal languages – some with infinitely long sentences, but not those alone – that don't have recognition procedures for their sentences or even, say, for their primitive vocabulary (because the set of types of primitive vocabulary in a formal language can be infinite too).

Many formal languages have a second property that's often treated as a constitutive element of those languages. They can have syntactic-transition rules from sentences (of such-and-such syntactic forms) to sentences (of so-and-so forms), as well as *axioms*.[9] Consider the set of finite sequences of sentences of a formal language, where each sentence is either an *axiom* or follows from earlier sentences in the sequence by the application of a syntactic-transition rule to those earlier sentences. Call this a *derivation.* If there is a recognition procedure for the syntactic-transition rules, and for sentences, then there is a recognition procedure for the set of derivations. Call an *interpretation* of a formal language a function that takes all the sentences of a language to truth values (true or false), where the axioms are mapped to true, and so that the syntactic-transition rules take true sentences to true sentences.[10] A formal language (individuated to include axioms and syntactic-transition rules) is *consistent* if it has an interpretation; the foregoing stipulates soundness as a requirement of an interpretation.

The formal languages I've just finished describing are special insofar as the derivations in them have been characterized purely syntactically. This isn't the only possibility. The "semantics" of such languages – what the sentences mean – has been restricted to truth values that are assigned to sentences; but of course richer semantic characterizations of the sentences (and syntactic characterizations of other items of formal languages – predicates and names for example) can be given. The now-traditional way of doing so for classical languages is to supply a Tarskian model for a formal language, in which a domain of objects is assigned as the ranges of the

[7]It seems, and this is a typical assumption of Chomsky-influenced linguists, that competence in a natural language simply *is* the possession of a recognition procedure for its grammatically correct sentences. What such recognition procedures look like is a contested empirical question; it's clear, in any case, that we neither have conscious access to them – as individual users of natural languages (but only to, at best, some of their results when candidate sentences are presented to us) – nor do we, as a matter of linguistic science, know much about their properties (let alone what they are).

[8]Indeed, the grammars (syntax) of most of these languages are extremely simple – stated by a handful of easily understood rules.

[9]It's possible to drop axioms altogether; replace them with syntactic-transition rules from the empty set to sentences.

[10]I momentarily consider interpretations that are richer than the mere assignment of truth values to sentences – so this isn't a *definition* of "interpretation."

quantifiers, individuals from that domain are assigned to individual constants (names), and classes of *n*-tuples of objects from the domain are assigned to *n*-place predicates (Tarski 1933).[11] This isn't, of course, the only way to assign semantic properties to the sentences of formal languages. Regardless, semantic-transition rules can be formulated for formal languages (given one or another characterization of the semantic properties for the sentences of such a language), and a resulting notion of a derivation can depend on semantic properties of sentences instead of (or in addition to) the syntactic properties of such sentences. The resulting set of derivations will have a recognition procedure (or not) depending in part on whether there is a recognition procedure for the relevant semantic properties assigned to sentences.

It *seems*, as I mentioned, that semantic properties (of *some* sort) are intrinsically involved in informal rigorous mathematical proof. This empirical question is entangled in part with the broader empirical question of whether reasoning in natural languages is (purely) syntactic or not. Syllogisms are long-known examples of *bits* of natural-language reasoning that can be characterized syntactically. Nevertheless, it's not obvious that when nonprofessionals reason syllogistically that they are reasoning syntactically (i.e., independently of the perceived content of the sentences in such syllogisms).[12] It's also not obvious that reasoning doesn't occur purely syntactically in cases where it doesn't appear to: that depends on exactly what the grammar of natural languages turns out to be – what aspects of what (pretheoretically speaking) seem to be meaning properties of items of natural languages turn out to be such and which don't (anaphora, e.g., seems to be a syntactic matter in natural languages; so seem aspects of noun/verb agreement). In this respect, informal rigorous mathematical proof – inference steps in informal proofs – couched as they are in natural language, may also (or not) turn out to be based (at least some of the time) on meanings.

Regardless, we can now characterize the family of derivational accounts of informal rigorous mathematical proof with a little more precision. The idea is this: Informal rigorous mathematical proofs have the epistemic properties they have because there is a family of formal languages, where each informal proof is associated with a derivation or (possibly) a class of derivations (which may be from more than one formal language), and where the epistemic properties of informal rigorous mathematical proofs are explained by this association. For example (and, again, this crude example can't be right, but it illustrates the idea), any informal rigorous mathematical proof is an abbreviation of a particular derivation, and mathematicians (subconsciously) appreciate this fact.

[11]In nonclassical settings – for example, paraconsistent ones – generalizations of the Tarskian approach to model theory have been developed. So too, for intuitionism, multivalued logics, and other "logics."

[12]For reasons from the psychological literature to think otherwise, see Thompson et al. (2011).

3 Why Derivational Accounts of Mathematical Proof Don't Work

Mathematical proofs were expressed for centuries in natural languages – long before Frege's invention of artificial ones. And, despite the fertile infusion of formal tools into mathematics during the twentieth century – both as methods of proof but almost more importantly, as subject matters *of* mathematical study – mathematical practice continued and continues to occur in natural language.[13] We can ask why; and one straightforward answer places a heavy burden on derivational accounts of mathematical proof. This is that it's proofs, after all, in *natural languages* that convince mathematicians of the theorems that are established by these proofs. Furthermore, and this is an important datum too, it's *not* that mathematicians are convinced (finally) when these proofs are explicitly transliterated into their formal cousins. That, pretty much, never happens: "Oh wow, *now* I see why this is true." There would be no issue if the word "abbreviation" that I used in Sect. 1 was literal – that the version of the derivational account that worked was one where informal rigorous mathematical proofs were abbreviations of formal items – but if that were the case, then Frege's invention would not have been the original and amazingly creative invention it was.

It's important to realize that when I describe Frege's invention as that of an *artificial* language, I'm not exaggerating: he didn't (as a superficial glance might imply) lift out a couple of ordinary-language words or phrases, the German words or phrases for "and," "or," "there is," for example, along with a couple of syntactic rules of German and exhibit *them* as his invention. The syntax (and the semantics) of these artificial languages (i.e., what currently passes for "logic" in contemporary studies) is nothing like that of *any* natural language. Mathematical proof seems epistemically strange; but we can't explain that strangeness by treating mathematical practice as *actually occurring* in languages *that didn't exist* until (about) 1879. Explanations of mathematical practice have long trafficked in inferences to the best explanation of it being our sensitivity to supernatural entities of one sort of another – but perhaps this is really taking things too far.

[13]There is no book, no article, of mathematics that's entirely written in one or another formal language – not one in first-order logic, not one in any higher-order logic, not one in any of the numerous modal logics, etc. The proofs that occur in mathematical articles and books are rarely completely or even partially formalized; even in *logic* books and textbooks, completely formalized proofs occur only as *examples.* Frege, on the other hand, really intended his invention to replace then current mathematical *proof practices.* He saw ordinary language as flawed in a way that he took to create serious problems for mathematical proof. As the subsequent failure of mathematicians to adopt his formalism, or anyway, to switch to one or another formal language when doing mathematics shows, this is wrong. Exactly why it's wrong is one topic of this chapter.

Call *Phenomenology* the requirement that any derivational account explain why formal transcriptions of informal rigorous mathematical proofs don't ever have a property that the original informal rigorous mathematical proofs sometimes have: the property of inducing an experience of recognizing that this is a good proof – based (partially or sometimes entirely) on what the sentences of the proof *say.* It's important not to overstate the datum *Phenomenology* labels. The experience of (pretty much) any informal rigorous mathematical proof is that of *heterogeneous reasoning*: some inferential steps involve an "aha"-epiphany (Feferman 2012) that's one probable source of early modern philosophical views about the a priori; some, however, involve complex semantically obscure mechanical manipulations ("calculations") that are recognized to be valid manipulations because they follow rules that have been shown to be truth preserving (or that we've assumed are); some are taken on the authority of the author or as understood-to-be-background assumptions in the field; some involve significant and rich inferential moves (although this is often tacit, unexplored, and sometimes unrecognized); some are, strictly speaking, meta-proof considerations (e.g., "the other cases are the same"; "the indices on these operators are countable"). A new notion is required to enable the analysis of the heterogeneous nature of reasoning in ordinary rigorous mathematical proof – that of an *algorithmic device* (see Sect. 6 and later sections for discussion).[14]

Leaving the option of abbreviation aside, it wouldn't help if we instead hypothesized that natural-language proofs function successfully as epistemic guides for mathematicians by providing semantic or syntactic directions – "indications" of the corresponding formal proof – where such guidance is posited to be the psychological machinery (conscious or unconscious) by which mathematicians become convinced of theorems by the informal proofs they give. For, unfortunately for derivational accounts, nothing like this looks right either.[15] The initially slow evolution of specialized terminology in mathematics (the invention of place-holder Arabic numerals, algebraic notation, scientific notation, etc.), simultaneously accompanied with an evolving notion of "rigor," can give the impression that mathematical language slowly crawled over the eons toward an artificial-language endstate; Frege's "genius" was just to speed up the process by jump-starting the first of subsequently many such artificial languages in which the heretofore buried structure of mathematical proof could be finally fully exposed.

[14]A terminological conflict: Barwise (1993) uses the phrase "heterogeneous reasoning" to label formal-reasoning systems with both language-based inference patterns and diagrammatic ones. This usage has become influential in this literature. My usage (Azzouni 2017a, Sect. 9) describes something different: the character of typical informal rigorous mathematical proofs, as I describe them in the passage this footnote is appended to.

[15]What follows in the rest of this section is distilled from many discussions about mathematical proof over the past 20 years or so. Among the relevant papers and books are: Hersh (1997), Rav (1999), Azzouni (2006), Rav (2007), Azzouni (2009a, b), Lavor (2012), Tanswell (2015), and many others. Avigad (2019) and Hamami (2019), however, are recent attempts to defend derivational accounts. I discuss these a little in Sect. 15.

Several considerations, however, undercut this view of the terminological evolution of mathematics. There is, most importantly, that the phenomenology of mathematical proof – our experience of validity – doesn't accord with this picture. Mathematicians, pretty uniformly, learn to become convinced of the validity of mathematical proofs in ordinary language, by being trained to appreciate a certain degree of rigor in these proofs – *as they stand.* Such training does involve acquiring a facility with specialized terminology – a terminology that changes over time. Nevertheless, it's clear that the recognition of the validity of proofs often (but not always) involves the earlier mentioned "aha"-epiphany of recognizing that a certain result follows from others. This experience, as I've mentioned, *isn't* preserved when such proofs are transliterated into their purely formal versions but it does occur outside of mathematical practice; it's one we all feel with *certain* items of valid reasoning – for example, with (certain) instances of modus ponens, or when we see that from "All dogs are good pets," and "All good pets are creatures with large carbon footprints," it follows that "All dogs are creatures with large carbon footprints."[16]

That is, this "aha"-epiphany – among all of us – seems to *depend on the perceived meanings of the sentences in question and not on the syntactic properties of those sentences*; it doesn't occur when someone, applying rules, recognizes that a string of symbols follows by those rules from an earlier set of strings of symbols. This is why the formalization of an informal rigorous mathematical proof lacks the epistemic qualities that such proofs (sometimes) possess in ordinary language.[17] We no longer rely on our appreciation – whatever it comes to – of the meanings of the words and sentences we're using. We've switched to the mechanical manipulation of strings according to preassigned rules.

Phenomenology isn't the only challenging aspect of informal rigorous mathematical proof that derivational accounts face. Tanswell (2015, 297–298) requires of any derivational account that it explain (or at least be compatible with) five aspects of informal rigorous mathematical proof as it occurs in natural languages – what he calls *Rigor*, *Correctness*, *Agreement*, *Content*, and *Techniques*.

[16]This validity-experience, however, is frailer (at least in some cases) than many philosophers realize. It can be interfered with by our appreciation of the *content* of sentences – for example, by the conclusion striking us as implausible (see, e.g., Thompson et al. (2011), for studies that show this). I hypothesize that we need to consciously *learn* to distinguish intuitions of validity from intuitions of soundness – something that officially happens in classes on "critical thinking." The distinction, that is, isn't naturally made when we evaluate (in real time) arguments in day-to-day living: we're usually just interested in whether we should believe the conclusion on the basis of what we've heard.

[17]This will turn out to be an *overstatement.* More carefully: formalizations *sometimes* lack the epistemic qualities of the informal proofs they correspond to. They don't if they're *phenomenologically faithful.* See Sect. 13.

- *Rigor*: Informal mathematical proofs must meet conditions of rigor that (implicitly and explicitly) hold of informal proofs. This "rigor" seems different from the rigor exhibited by derivations (the latter is only that the derivation be fully written out according to formal rules). Notably, informal conditions on rigor change over time.
- *Correctness*: Mathematicians, perusing informal proofs, recognize them to be correct and incorrect intrinsically – "intrinsically" is understood here only in the sense that mathematicians don't recognize these properties of proofs by comparing them to formal transliterations.
- *Agreement*: Mathematicians, upon perusing candidate informal rigorous proofs, largely agree with one another about whether the proofs are suitably rigorous and on whether these proofs are correct or incorrect. Strikingly, mathematicians from different proof-practice backgrounds (e.g., classical as opposed to intuitionistic ones), and who understand the respective proof practices, agree on this as well. As noted earlier, this doesn't seem to occur by a process of recognizing the derivations (valid or invalid) that these proofs are associated with.
- *Content:* Any derivational explanation must explain how the perceived content of an informal rigorous mathematical proof – what the sentences of that proof are experienced to *say* – determines which formal proof(s) it indicates.
- *Techniques*: Any derivational account must explain the role of informal techniques of proof that aren't obvious transcriptions of formally licensed inference steps in a way that's compatible with the posited epistemic role of formal derivations.

Coupled with *Phenomenology*, these legitimate demands on any derivational explanation are a high bar to meet. I'm going to assume – within the framework of this chapter – that a derivational account that meets all of these demands adequately isn't possible.[18] I'm hardly alone in drawing this conclusion, of course; most of those philosophers who have studied informal rigorous mathematical proof agree. If so, epistemically speaking, informal rigorous mathematical proof is sui generis in one sense: we can't explain its epistemic properties in terms of formal derivations in

[18]See Azzouni (2017a) for discussion of how well various derivational accounts fare in meeting these demands – better than one might expect. I draw the conclusion, nevertheless, that (at the end of the day) they can't be met by such accounts. It's here, however, that issues about the large size of formal transcriptions (as compared to the original informal proofs) arise; this poses a challenge, for example, to the derivational account being able to explain *Agreement.* It's also here that issues arise about the explicit rules governing derivations (including ones of logic) needing to be treated as only tacitly available to ordinary mathematicians – and *essentially* tacit until the emergence of Frege's explicit characterization of logic.

artificial languages that the informal proofs supposedly correspond to.[19] We must look elsewhere for an explanation. I'll start my particular take on this by noting an historical irony. If I'm right about what I claim in the rest of this chapter, then Frege's invention is a garden path (or rabbit hole) – at least as far as the epistemology of mathematical proof is concerned. Formal languages, it turns out, have nothing directly to do with the epistemology of informal rigorous mathematical proof. Formal derivations play no more of a role in explaining the properties of ordinary mathematical proofs – implicitly or explicitly – than strange Platonic objects do.

4 Algorithms

I slightly overstated things in the concluding paragraph of the last section. The study of formal languages *isn't* a garden path or rabbit hole – not with respect to resolving the puzzles of informal rigorous mathematical proof (although it's still true, I'll argue, that formal languages – *directly* – play no role in the epistemology of mathematical proof). What turns out to be extremely important to resolving the puzzle of the epistemic strangeness of ordinary mathematical proof is something that the progenitors of formal logic (Frege, Russell, Hilbert,. . .) didn't anticipate. This is that the study of formal languages would directly lead, in the course of hardly more than a half century, to Gödel's incompleteness theorem (Gödel 1931), which, in turn, would enable (provably equivalent) characterizations of the crucial notion of an algorithm (Church 1936; Kleene 1936; Turing 1936). My claim: it's algorithms – and not formal languages (of any sort) – that are the key to understanding the

[19] I use a qualifier here ("in one sense") and earlier when characterizing Tanswell's correctness condition ("only in the sense that . . .") not to foreclose on the view, held by many of these philosophers (but not me) that sociological vectors explain *Agreement* among mathematicians – more generally, that what we may describe as conformity among mathematicians about mathematical proof is to be explained along the same lines that conformity in other areas (e.g., dietary practices) is to be explained. Although I think such sociological vectors do play a role in what mathematicians write down (for example) in giving an informal proof because what they assume as known by their colleagues shifts over time, or because of inertial effects in styles of proof, I don't think these factors play a very large role otherwise. Indeed, the centuries-long perception that mathematical reasoning is special in a way atypical of reasoning, generally, is a perception of its evident stability vis-à-vis social factors *in contrast to* reasoning elsewhere. (Manders (2008, 67) writes that "Euclid . . . Apollonius and Archimedes, are virtually without error: their every result has a counterpart in modern mathematics, even if subsumed in patterns of claims and proofs recognized much later." Absence of "error" is code for: successfully resists sociological intrusions.) In the course of this chapter I'll invoke intuitively effective recognition procedures that are available for (much of) informal rigorous mathematical proof to explain this. See what follows in the rest of this chapter.

epistemic strangeness of informal rigorous mathematical proof. Seeing exactly what the role of algorithms is in that practice will solve this ancient philosophical problem.[20]

I'll start with a brief description of Turing's well-known characterization of algorithms via what have come to be called (deterministic) *Turing machines*.[21] A Turing machine is an idealized rule-governed writing device – a Turing-machine head – located on an infinitely long tape that's divided into linearly ordered adjacent cells. Each cell is either blank ("□") or has a squiggle ("~") written on it. The Turing-machine head occupies one cell at a time; and it's governed by a finite number of rules which dictate all its possible actions and nonactions, depending on (a) what

[20]A *lecture*: Following the rhetorical lead of Lewis (1991, 57–59), many philosophers, for example, Burgess and Rosen (1997, 34) glibly laugh (or otherwise engage in something that's an approximation of laughter) at the idea that philosophers of mathematics can provide insights that might have an impact on the practice of mathematics or even yield something of interest to mathematicians. This purported laughter is silly. To make the first obvious point: philosophical concerns needn't be mathematical concerns. Sometimes, as with the philosophical puzzle of the epistemic strangeness of mathematical proof, this is because the philosophical wonder is *occasioned* by a specific practice, in this case, mathematical proof. What makes mathematical proof an occasion of philosophical concern is a *contrast fact:* mathematical proofs seem different from arguments in other areas of discourse as I indicated in the opening paragraphs of this chapter. Because the professional mathematician needn't be concerned with what arguments look like in other areas of discourse, and because she needn't be concerned with what it is that makes mathematical proofs different, these concerns, strictly speaking, aren't *mathematical* ones. So too, issues about mathematical ontology are philosophical concerns, not because the mathematicians' use of "there is" in their own practice should be challenged in some way, but because (again) it's a usage that offers puzzles across *all* discourses. (It can turn out for purely linguistic reasons that the mathematicians' use of "there is" is metaphor or a terminological extension of the ordinary phrase; discovering this and publicizing it isn't, by virtue of that, to necessarily recommend a revision in mathematical practice.) Recognizing philosophical puzzles and what's making them *philosophical* puzzles isn't easy – not even (apparently) for people who are officially philosophers.

But a rather more straightforward observation should be made, especially after noting one of Lewis' remarks, mentioned in the last paragraph (59) – italics his: "I'm moved to laughter at the thought of how *presumptuous* it would be to reject mathematics for philosophical reasons. . . . If [mathematicians] challenge [a philosopher's] credentials, will you boast of philosophy's other great discoveries: [and now Lewis gives a list of purportedly bizarre claims made by one philosopher or another]?" Response to Lewis: Classical physical theories of solar combustion were refutable in the late nineteenth century on the basis of the fossil record and other geological evidence: the Earth was clearly more long-lived – by a large magnitude of order – than the Sun could possibly be, given the then contemporary physics. Imagine a Lewis-style physicist snobbily turning his nose up at an argument (by a geologist, say) against classical physics *solely because* she's some upstart from outside a "genuine science" who dares challenge accepted physical theory. Alas (for such hierarchical views of knowledge gathering), that's how confirmation *works:* evidence for something can come from *anywhere* – even (dare I say it) from philosophy.

[21]Turing (1936, 118) uses the word "automatic." I should note that my presentation of Turing machines will be slightly eccentric in several ways – one of which is that I'll use meaningless symbols instead of the usual numerals "1" and/or "0." I do this to accentuate certain philosophical points about the interpretation of Turing formalisms that I subsequently make. Standard presentations of Turing machines, incidentally, are everywhere (especially in numerous textbooks) – if the reader wants or needs a (standard) refresher course in Turing-machine formalisms.

state S_i the machine head is in, and (b) what's written on the cell the machine head is located at. The machine head is capable of triple-actions. First, it erases/writes-in (or doesn't) the cell it's at, then it moves to the left/right (or doesn't), and then it shifts to a new state (or doesn't); what it does at any cell is dictated by the "state" it's currently in. Call the finite list of rules that describe what a Turing-machine head does, given the states $S_1, \ldots, S_n$ it's in, the *machine table* of that Turing-machine head. Let R mean, "move right," L mean "move left," and B mean "stay in place." Here is a sample machine table for a Turing-machine head with three states:

$$(S_1, \square) \hookrightarrow (\sim,L,S_1); (S_1, \sim) \hookrightarrow (\square,R,S_2);$$

$$(S_2, \square) \hookrightarrow (\sim,L,S_1); (S_2, \sim) \hookrightarrow (\sim,B,S_3)$$

This fully describes what this machine head will do, given its location on any cell, and it's being in any state. (I'll now individuate machine heads by the machine tables they obey.) The machine table reveals that if the machine head is in state S_1, and it's on a cell which is blank, it will write a "~" on that cell, move the left adjacent cell, and remain in S_1.[22] If this machine head is in state S_2 on a cell on which a "~" is written, it will leave the "~" in place, not move, and change from the state, S_2 to the state, S_3. There are no instructions for this machine head when it's in S_3; thus, the machine head will, as it's described, "halt."

Call any tape with a distribution of "~"s in its cells a *configuration.* Given a machine head located on a particular cell of that tape (what can be called a *Turing-head configuration*), call an *input* the configuration before the machine head initiates a finite sequence of actions; call the *output* the configuration that results after the machine head has completed a finite sequence of actions. Consider a set of Turing-head configurations, where the configurations are all different. Call this an *algorithmic operation* – that is, an algorithmic operation is a set of differing configurations, each of which has a Turing-machine head (with the same machine table) located on one cell of that configuration. Call the resulting modification by the actions of the machine head after a finite set of steps an *algorithmic calculation.* If the machine head halts, that's an *implementation* of the algorithmic operation.

I've deliberately described the operations of Turing machines as manipulating *meaningless* marks on tapes; consequently, the algorithmic operations (as just defined) are meaningless as well. They're open of course to being interpreted – and it's how widely they can be interpreted that makes the Turing-machine formalism so fundamental to our notion of calculational/effective processes (and, in particular, to the effective recognition procedures of formal languages that I described in earlier sections of this chapter). Let's make the (various) interpretations of Turing machines explicit. Consider a set of Turing-head configurations (consider, that is, an algorithmic operation). An *interpretation* of this algorithmic operation is a

[22]Sometimes, as I illustrated earlier, the machine head is described as "scanning" the cell it's located on. But this talk implicitly builds more machinery into the machine head than is needed.

mapping of its configurations to configurations of objects *along with* a mapping of its machine-head actions to corresponding actions on the corresponding configurations of objects.

Example: Identify "~" as the numeral "1" (interpret "~" as standing for the number 1). Next, single out a set of configurations that can, using this identification, represent the counting numbers (e.g., "~~~" is 3; "~~~~~~" is 6); indeed, single out a set of configurations that represent pairs of such counting numbers (e.g., "~~~□~~~~~~"). Consider the set of Turing-head configurations, where the Turing-machine head is located on the blank cell between strings of "~"s that represent these pairs of counting numbers. We can design a machine table (that halts with respect to these pairs) that can be proven to generate a single representation of a counting number that's the sum of any pair of numbers (represented by an initial pair of represented numbers in this set of Turing-head configurations); we can design a more complex machine table (that also halts with respect to these pairs) that can be proven to generate a single representation of a counting number that's the product of a pair of counting numbers represented by a Turing-head configuration in this set. These are typical exercises in any textbook on (or with chapters on) Turing machines.

Second example: Now consider a mapping of the same set of Turing-head configurations, not to numbers and operations on numbers (as just illustrated), but instead to bead configurations on a physical abacus,[23] and to (certain) physical operations of moving those beads. This is an interpretation of that set of Turing-head configurations too. For that matter, so is any mapping of that set of Turing-head configurations to any numerical representation of numbers (and manipulations of those representations). Interpretations of Turing-head configurations, that is, can be to notational objects or to worldly objects; correspondingly, the actions of a Turing-machine head can be interpreted as manipulations of Platonic *abstracta*, idealized operations on notional objects, or as (also idealized) physical actions on real objects.

The Turing-machine formalism (along with all interpretations of it) is unexpectedly robust along *several* dimensions. I'll discuss this in the rest of this section.

To begin with, describe as *Turing-machine formalisms,* a certain class of variations of the above description of Turing machines. For example, consider Turing-machine heads that scan more than one cell at a time – finitely many, let's say, or ones that can recognize more than one symbol in a cell, finitely many, let's say (or even countably many).[24] And now call *mechanical procedures* those sets of operations (on any collection of objects) that (as above) a Turing-machine formalism can be interpreted as. It has been proven, strikingly, that increasing the powers of a Turing machine in the above ways doesn't lead to an increase in the set of mechanical procedures that Turing machines (under an interpretation) can execute: these remain the same. It's furthermore quite striking what kinds of powers machine heads

[23]I'm going to be fairly spotty in describing abacuses; it's easy to find out about these things in articles in print and on the web.

[24]Consider, in particular, Gandy machines (see Gandy 1980).

must be granted in order to enable the class of mechanical procedures to become strictly larger than what can be managed by Turing machines. If, for example (Abramson 1971), "extended Turing machines" are stipulated that are able to write and read uncountably many symbols, the resulting class of mechanical procedures is strictly larger than those executable by conventional Turing machines. The class of mechanical procedures a "mechanical" device is capable of is strictly larger as well, if, unlike Turing-machine heads, it can complete infinitely many actions in finite amounts of time.[25] It appears that the notion of Turing-machine computability isn't only an upper limit on *human* computation (as Turing originally meant it to be) but also an upper limit on designed-machine (e.g., computer) computation – at least if the physical processes of our universe are what they appear to be.[26]

To continue the discussion of the robustness of the Turing-machine formalism, let's now (temporarily) restrict ourselves to interpretations of Turing-machine algorithms as computing operations on counting numbers. These numerical algorithms can be characterized in other mathematical ways – in terms of being definable from certain classes of functions for example.[27] These various characterizations have been proven to be equivalent to the set of Turing-machine algorithms (as defined above). With a little bit of argument (turning partly on the fact that any interpretation of a set of Turing-head configurations can be routed through an intermediate interpretation in terms of numbers), the same set of mechanical procedures is picked out by all of these characterizations.

Second, consider Church's thesis. (Or, alternatively described, the Church-Turing or the Turing-Church thesis.) A preliminary "intuitive" notion: call a function of the natural numbers an *intuitively effective function* if it's one that can be carried out by human(s) using pencil and paper. Church's thesis is: A function of the natural numbers can be calculated by an intuitively effective method if and only if it is computable by a Turing machine (subject to the above interpretations). "Intuitively effective," so described, *really is* an "intuitive" notion. The use of "intuitive" means various things to various philosophers and logicians[28] – but for current purposes it can be taken to indicate a hand-waving (unanalyzed and nonmathematical) characterization of an open-ended class of what strikes us as humanly possible calculations.

[25]See Copeland (2019) for further discussion and citations of the growing literature on this.

[26]All bets about this are off, of course, if the study of gravitational singularities and the like, where time – in particular – seems to have unexpected properties, allows the design (at least in principle) of devices that can complete infinitely many calculations in what amounts to finite amounts of time.

[27]In particular, the early publicized characterizations in terms of lambda-definable and recursive functions have been joined by further (provably equivalent) characterizations in terms of register machines, canonical and normal systems, combinatory definability, Markov algorithms, and others.

[28]Turing (1936, 135), notably, uses the word "intuition" in relation to his statement of the thesis, writing, that "no attempt has yet been made to show that the 'computable' numbers include all numbers which would naturally be regarded as computable. All arguments which can be given are bound to be, fundamentally, appeals to intuition, and for this reason rather unsatisfactory mathematically." Warning: "intuition," as used here – following Turing's usage – can be misleading. See the parenthetical remarks in footnote 114.

In mathematical practice (specifically, in the study of recursive functions and effective computability), Church's thesis is often used as a proof-shortcut, to avoid the tedium of exhibiting a Turing-computable function for an intuitively effectively calculable numerical function, for example, in Rogers (1967) where a numerical operation is described (or recognized) as humanly calculable and then via Church's thesis that operation is deduced to be Turing computable. The intuitive notion of an effective procedure is idealized in various ways. In particular, we imagine the humans doing these calculations as having sufficient time, tenacity, tools (pencils and paper), and abilities to continue the task (without errors) as long as needed for the particular algorithmic calculations involved. Call this a *persistence idealization*.[29]

A way of attempting to *prove* – in the (informal rigorous mathematical sense) Church's thesis is to argue that the hand-waving ("intuitive") description of an intuitively effective procedure can be characterized by (is coextensive with) a mathematical characterization (of some sort), and then use that characterization to prove Church's thesis. There are several versions of this strategy. Kripke's (2013) version relies on "Hilbert's thesis":

> Every (human) computation can be formalized as a valid deduction couched in the language of first-order predicate calculus with identity.

"Human computation," however, as used here, is every bit an intuitive notion as "intuitively effective method"; thus, crucially, Hilbert's thesis relies (as Church's thesis does) on an identification of an "intuitive" notion with a mathematically precise notion – namely (using my terminology) an identification of "humanly computational" with "derivable in the language of the first-order predicate calculus with identity." The reason that these empirical theses are needed is because a branch of (pure) mathematics, algorithm theory, is being applied empirically to human (and machine) computation. It's *always* an empirical question (a "thesis," if you will)

[29]Here's a generalization of "intuitively effective numerical function" – beyond cases of numerical functions – to any method whatsoever (I'm borrowing this formulation from Copeland (2019)):
M is an intuitively effective method if:

1. *M* is set out in terms of a finite number of exact instructions (each instruction is expressed in a finite number of symbols);
2. *M* will, if carried out without error, produce the desired result in a finite number of steps;
3. *M* can (in practice or in principle) be carried out by a human being unaided by any machinery except pencil and paper;
4. *M* demands no insight, intuition, or ingenuity, on the part of the human being carrying out the method.

I treat the above description as allowing that the instructions in question can be written in a natural language. I also take it to be topic-neutral: *M* can be a method that's applied to anything whatsoever. (I'm not claiming this is intended – or not intended – by Copeland.) I'll take up the topic neutrality of the intuitive notion explicitly in Sect. 5.

whether or not a branch of pure mathematics can be applied successfully to the entirety of an empirical domain that we try to apply it to.[30]

In any case, there is a related family of equally intuitive theses sometimes described as Church's thesis, and sometimes distinguished from it (see Copeland (2019) on this). Consider, in particular, the "maximality thesis": All functions that can be generated by machines (working in accordance with a finite program of instructions) are computable by intuitively effective methods (Gandy 1980). An argument deriving this thesis about "machines" is, in my view, possible along the lines I indicated in the material footnote 26 is appended to; I won't pursue this any further now. In any case, evidence for Church's thesis (and its relatives) continues to accrue by the explicit exhibition of particular Turing algorithms (directly or through equivalences) as able to execute specific intuitively effective procedures and those that machines (of various sorts) carry out. This, as I indicated, is a different dimension along which the applications of Turing-machine formalisms are robust.

There is yet a third dimension, along which Turing-machine formalisms, treated as characterizations of mechanical calculability, have proven strikingly robust. It's one already indicated by a central aspect of the proof of Gödel's incompleteness theorem – that it turns on a reinterpretation of computable functions of number theory as functions of the syntax of a language. This *hints* of the vast potentiality for the interpretations of algorithmic operations – a vast potentiality that's being executed (as I write) by developments in computer technology. For intuitively effective procedures are undertaken by humans on all sorts of objects (real and unreal) in all sorts of ways that are only slowly being interpreted in what amounts to Turing-machine formalisms – for example, a robotic hand picking up a paper cup without crushing it, or all the numerous (and increasingly realistic) video and virtual-reality games and films.

5 Recognition Procedures in Games

I turn now to developing a point I made in footnote 29 – that intuitively effective methods can be stated in ordinary language. Indeed, this is why Hilbert's and Church's theses are "intuitive" theses: ordinary reasoning and ordinary effective methods are invariably couched in natural language – and transliterating such to a formal setting is precisely to apply a branch of pure mathematics to these natural-language activities (as I indicated at the end of the last section). Coupling this point with the observations in the last section about interpreting algorithmic operations allows us to see how ubiquitous intuitively effective methods are in our lives. Let's start with a discussion of games. Many games (although not all games) essentially involve elements and configurations of elements with recognition procedures that players (and observers) of those games rely on. Chess, and other board games, are

[30] I return to a discussion of Church's thesis, generalizations of Hilbert's thesis and formal systems, in Sect. 15.

obvious cases (and so are card games). Recognition procedures are required to distinguish chess pieces, the squares on the board (and the spatial properties of those squares, e.g., adjacency relations), and to recognize admissible moves. Other games, despite involving "luck," are similar. In many board games, the pieces are moved along designed paths after throwing dice (to determine the acceptable number of steps a piece can take). Without appropriately designed intuitively effective recognition procedures, these games would be impossible to play – we'd be fighting over whether an *appropriate* move had been made; we'd even be fighting over what move, exactly, we were fighting over.[31]

Central to the Turing-machine formalism seems to be – to use a commonly used distinction – that the algorithms it describes are digital and not analog. Mechanically recognizable operations are required to be step-by-step ones, where each step is clearly distinguishable from every other step. Turing (1936, 136) makes this point explicitly:

> We may suppose that there is a bound B to the number of symbols or squares which the computer [human calculator] can observe at one moment. If he wishes to observe more, he must use successive observations. We will also suppose that the number of states of mind which need be taken into account is finite. The reasons for this are of the same character as those which restrict the number of symbols. If we admitted an infinity of states of mind, some of them will be "arbitrarily close" and will be confused.[32]

And yet, many games, unlike the board games I've just described, seem to be intrinsically analog: they take place physically, and continuously, in space and time. Such games, baseball for example, appear nevertheless to have intuitively effective recognition procedures. One is inclined (at first glance) to say, of any set of admissible moves in this game – hitting a pitcher's throw appropriately with a bat – that it's an *infinite* set of possible moves. After all, the variations of possible movements in space and time are infinite. How then can a game like baseball – and any game that similarly takes place continuously in space and time – be amenable to an application of the Turing formalism?

There are at least two mistakes involved in this line of thought. The first is that a certain kind of idealization – central to all algorithms that humans actually carry out but eliminated (usually) from presentations of such algorithms – is being

[31]It's been claimed against me (more than once) that games are at best analogous to Turing-definable algorithms because games have rules that are underspecified or even inconsistent. Tanswell (forthcoming) thinks these defects "are ironed out by social features of game-playing." It's, however, a ubiquitous feature of algorithms that there is no general test for "bugs" (there is no general test for inconsistency): they must be *spot-corrected* when discovered. In no significant sense does this introduce "social determinants" into practices, like games, that are governed by algorithms.

[32]Notice that the generalizations of Turing's apparatus (e.g., in Abramson 1971) mentioned in Sect. 3 are ones that shift to an essentially transhuman framework in exactly the ways that Turing rules out. For example, that the purported human can distinguish uncountably many symbols or that the purported human can complete an infinite number of discrete actions in a finite amount of time.

misdescribed. Consider the symbol "~" that I characterized all Turing-machine heads as using. Were an actual machine involved, the tokens of this symbol would be a somewhat ill-defined set of physical items – ill-defined because no matter how we designed Turing-machine heads, they will fail to sort *every* otherwise scannable physical token into one or the other class. In the same way, a chess piece can be placed on a chess board so that it's unclear (to the humans playing the game) which of two neighboring squares it's on. For that matter, the borders between squares on a chess board can be fuzzy, or badly designed chess pieces are possible where humans easily confuse the pawns with kings. We idealize the items and configurations of the items in a game (using what I called in Sect. 2 "exclusion idealization") by simply treating such confusing items as not occurring in the game. A last example is the intuitively effective recognition procedures associated with the derivations of formal languages: these too, when the derivations in such formal languages are actual ones that people (or computers) are constructing, involve exclusion idealizations so that by assumption there are no tokens so badly written that the recognition procedure for those derivations can't be executed.

The second mistake is to take the apparent analogness of our experience of space and time (and, in particular, our actions and perceptions of such actions) at face value. Given a finite amount of space and time, and a possible action across that space (and in that time), such as waving a hand – despite the impression otherwise, we *cannot* distinguish infinitely many variations of movements. What we are capable of seeing, and distinguishing, are only finitely many such movements. This is true of anything we can do in space and time. The analog appearance of our own behavior (or its appearance in the behavior of others), consequently, is an illusion.[33]

This shows that that the appearance of analog practices in such games doesn't belie an application of a generalized Church-Turing thesis to them: The intuitively effective recognition procedures that such games exhibit can be taken to be ones amenable to the Turing notion of algorithmic calculations.

The other crucial point that games illustrate in abundance are intuitively effective recognition procedures that are written in *natural* languages. We can use Hilbert's thesis, if we wish, to draw the conclusion that these can be formalized – indeed, to draw the conclusion that these games, in their entirety, can be formalized. And by this, I mean, transliterations to language-based formal systems.[34] Several connected points must be made about this, however. Our grasp of the intuitively effective recognition procedures used in such games – intuitively effective recognition procedures that children grasp – are ones *grasped directly,* and not via an implicit

[33]See Azzouni (2017a), especially Sect. 3, on this. To show the amenability of these games to Turing formalisms, in my view, it must be shown that the intuitively effective recognition procedures at work in these cases involve an unseen grid-structure as the field against which the admissible moves in these games are being evaluated. The details are in the article just cited.

[34]"Formal" systems, as I illustrate in Sect. 8, can be diagrammatic as well as language based.

(psychological) application of Hilbert's thesis. The neurophysiological/psychological facts that underlie these capacities are ones to be explained (that *are* being explained) in terms of visual-system capacities, saliencies, and the like. Our abilities to physically navigate our world are what underwrite the intuitively effective recognition procedures used in games. This is one reason, among others, why these recognition procedures are only partially described by the "rules" that officially govern such games. Usually taken for granted, and left unmentioned, are the physical details of the elements of such games (e.g., baseballs) that are often only partially specified in terms that explain the successful use of recognition procedures during these games; usually, tokens of them are simply exhibited. The same is true of the primitive alphabet for formal systems. These are, in practice, just listed – that is, the intuitively effective recognition procedures are *assumed* – using tokens, for example: the constants, a_1, a_2, a_3..., the individual variables, x_1, x_2,..., etc. I'll label these, in Sect. 13, *embodied algorithms*, and say a bit more about them then.

When, however, such games are transliterated entirely to language-based systems, the idealized physical properties of the objects and configurations of them that are *used* by the intuitively effective recognition procedures associated with those games must be treated as part of the subject matter to be described; they must be explicitly characterized and treated axiomatically. This is because language-based formal systems use only intuitively effective recognition procedures for distinguishing items of an alphabet: we're required only to see (by "eye") linear concatenation relations between tokens, to be able to distinguish the distinct primitive vocabulary tokens, one from another, and to be able to recognize which they belong to.[35] So, since the other intuitively effective recognition procedures, the ones that enable us to recognize the relations among the chess-board squares, are set aside, these relations must be explicitly described by terms that are axiomatized and treated as governing the "pieces on squares." (I'll say more about this in the next section.)

I turn now to making points about ordinary rigorous mathematical proof that are analogous to those that I've just made about games. Just like a certain subfamily of games, ordinary rigorous mathematical proof involves numerous intuitively effective recognition procedures that mathematicians grasp directly and not via formal transcriptions of those procedures into the medium of formal languages. This is why the derivations occurring in the family of artificial languages that are studied by logicians are irrelevant to the epistemic properties of ordinary mathematical proof. In going forward, I assume a generalized version of the Church-Turing thesis: I assume that all intuitively effective procedures *are* Turing computable.

[35]The other sorts of linear syntactic-relation patterns we – or some of us, anyway – "just see" are helpful but not required by most linear alphabet formal systems we study.

6 Language-Based Algorithmic Systems and Algorithmic Devices

Straightforward examples of language-based algorithmic systems are *axiom systems.* These have been studied extensively. A primitive vocabulary is given and syntactic rules for constructing well-formed expressions, including sentences, are defined. On this basis, in turn, a subset of the sentences are singled out as axioms, the finite sequences of sentences that are *acceptable derivations* are defined, and finally, the theorems (the sentences that appear on any line of an acceptable derivation) are defined as well. All of these have intuitively effective recognition procedures in the ways described earlier – the definitions I've just alluded to are designed to enable these intuitively effective recognition procedures to piggyback on earlier ones (e.g., the intuitively effective recognition procedure for theorems is designed on the basis of the one for acceptable derivations). Intuitively effective recognition procedures in formal systems are nested on one another in the way that the intuitively effective recognition procedures for primitive vocabulary, sentences, acceptable derivations, and theorems are nested here.

Example (Church's P_1)[36]: The primitive vocabulary are parentheses, (,), a two-place sentential connective, $\supset$, a one-place sentential connective, $\neg$, and an infinite list of sentential variables, p, q, r, s, p_1, q_1, r_1, s_1, p_2, q_2, ….[37] All and only the sentences are defined like so: the sentential variables are sentences; if P and Q are sentences, then $\neg$P and (P $\supset$ Q) are sentences. The axioms are: $(p \supset (q \supset p))$, $((s \supset (p \supset q)) \supset ((s \supset p) \supset (s \supset q)))$, and $(\neg\neg p \supset p)$. The syntactic-transition rules are two (I use the symbol "$\vdash$" to indicate syntactic transitions): For all sentences, P and Q, P, (P $\supset$ Q) $\vdash$ Q; if b is a sentential variable that appears in P and if P* is the sentence that results from the substitution everywhere in P by a sentence Q, then P $\vdash$ P*. A derivation is a finite sequence of sentences where each sentence is either an axiom or follows from earlier sentences in the sequence by a syntactic-transition rule.[38]

There are two ways to think about syntactic-transition rules. One is to think of them as structuring every *acceptable derivation:* each sentence in an acceptable derivation is either an axiom or follows from earlier sentences in that derivation by a syntactic-transition rule. The other is to think of them as methods of generating acceptable derivations from other acceptable derivations. Given any acceptable derivation, we can generate a new (augmented) acceptable derivation by introducing a new last sentence that is either an axiom or a sentence that follows from sentences in the original derivation by a syntactic-transition rule.

[36]From Church (1956, 78), I've modified P_1 by introducing an abbreviation that Church (1956, 78) gives later.

[37]Notice that there is an implicit intuitively effective recognition procedure in place for these, turning on an antecedent intuitively effective recognition procedure that's available for numerals.

[38]This apparently austere system is equivalent (via definitions) to the full two-valued classical sentential calculus with all definable (two-valued) connectives.

The point I want to stress about the intuitively effective recognition procedure for derivations in this example (and this is typical of many axiom systems) is that this recognition procedure directly operates in terms of the intuitively effective recognition procedures for whole sentences and not via any intuitively effective recognition procedures on the elements that such sentences are composed of. The intuitively effective recognition procedure for whole sentences of course depends on intuitively effective recognition procedures for the primitive alphabet – as I noted earlier. I'm here stressing a different point, however: Syntactic-transition rules are possible that rely not on whole sentences but instead directly rely on certain expressions *in* those whole sentences. Consider the alternative algorithmic formal system in the next example.

Second example: Same as the previous example, except that there are no axioms, and in addition to the two syntactic-transition rules given in that example, we have three additional syntactic-transition rules. For any sentences P and Q,

$$Q \vdash (P \supset Q);$$
$$(S \supset (P \supset Q)) \vdash ((S \supset P) \supset (S \supset Q))$$
$$\neg\neg P \vdash P.$$

The first two syntactic-transition rules are focused on the connective, $\supset$; the third focuses on the connective $\neg$; it's the occurrence of these connectives in certain syntactically indicated locations in sentences that license these syntactic transitions. In this sense these syntactic-transition rules are centered on a particular syntactic element and not on the sentence the element occurs in. For any expression e, call such syntactic-transition rules *e-centered syntactic-transition rules.* This characterization can be generalized to families of expressions E. Consider, for example, those Peano syntactic-transition rules that are E-centered on the constant 0 and the successor function′.[39]

Now consider an expression e and a language (*any* language, natural or otherwise) in which e appears. I call any set of e-centered syntactic-transition rules an *algorithmic device.*

The individuation conditions on algorithmic devices are importantly weak. Apart from the fact that the above definition allows algorithmic devices to be nested within one another, it's not required of an expression e governed by a set of e-centered syntactic-transition rules that constitutes an algorithmic device A that it be *fully* governed by those syntactic-transition rules. In the second example, just above, both $\neg$ and $\supset$ are governed by the original two syntactic-transition rules as well as by the new ones I then introduced. So too, an expression e, in a formal system, can be governed both by e-centered syntactic-transition rules and other syntactic-transitions rules (e.g., axioms) which aren't e-centered.

[39]For that matter, consider Church's P_1. This is an E-centered algorithmic device, where $E = \{\supset, \neg\}$.

The conditions on something being an algorithmic device are weak in a second sense. Although I've introduced the notion with respect to language-based formal systems, such devices can (and do) occur elsewhere. They show up, specifically, in natural languages and, indeed, they show up everywhere in mathematical practice, as it occurs in natural languages (as I'll indicate in the sections to follow). But, for the moment, consider the old philosophical example, "bachelor," where the meaning of this word is stipulated to mean: unmarried male. We have, according to this aspect of the meaning of the word, the following syntactic-transition rule P |- P*, where P* results from P by zero or more replacements of "bachelor" with "unmarried male."[40] More generally, the meanings of words can be taken (in some cases) to license syntactic-transition rules in arguments. This is compatible with those meanings being governed by other conditions – sufficient-condition application rules of some sort (for example) – which don't involve syntactic-transition rules.

So notice: consider any expression that seems governed by a small set of meaning-transitions (or meaning postulates) that enable the recognition of a class of truths or inferences containing that expression. Such a set of meaning-transitions is a kind of algorithmic device – not by any means the only kind of algorithmic device, of course. And notice, as well, that any expression that's so governed is usually governed by all sorts of other conditions too – both generalizations that involve many other words, and more restricted conditions that govern that expression along with others. The last point to stress is that it's come to be recognized that natural-language expressions – that we otherwise understand the meanings of – don't usually have definitions (necessary and sufficient conditions in terms of other expressions). They have, at best, some necessary conditions and some sufficient conditions. Some of these are encapsulated in the conditions of an algorithmic device; but others may be an implicit (or explicit) understanding of patterns of things we apply such expressions to (e.g., all blue things).

One thing – when studying logic – that makes "Gentzen" or "natural deduction" rules more user-friendly than axiom systems is that in this respect they're modeled on our impression of the meanings of natural-language expressions. We experience the natural-language meanings of expressions as localized to (and due to) those very expressions. This is why, say, a characterization of the "meaning" of the truth-function "&" like so:

$$p\&q \vdash p,$$
$$p, q \vdash p\&q,$$

seems to capture via the *meaning* of "&" the logical properties of "&" – as it operates in derivations of sentences from other sentences.

[40]The syntactic rule must be finessed to avoid counterexamples involving expressions like "Bachelor Buttons," which is the name of a flower. (Sadly, "Unmarried Male Buttons" isn't the name of that flower – or, I think, any flower.) I skip this.

For current purposes, the importance of algorithmic devices is that such items – focused on new expressions or, as it were, modifications of devices already governing old expressions – can be added to any language (formal or otherwise, one that already has one or another proof or derivation system, or one without) so long as the expressions the added devices are focused on (still) syntactically fit in with that language. For example, natural languages have noun phrases, verb phrases, and the like. If the new terms an algorithmic device is focused on are actual noun phrases or verb phrases, all that's required of them is that they operate syntactically (outside the confines of the application of the algorithmic device itself) like other noun or verb phrases already in the language – in particular, as these occur in natural-language sentences.[41]

7 Some Observations About Algorithmic Devices, Conceptual Engineering, Synchronistic Effects, and Imperatives

Imagine a natural language in which we have in place intuitively effective procedures for recognizing some instances of "Q follows from P" facts. Or, imagine that the methods we have aren't intuitively effective. The introduction of algorithmic devices can augment, make easier, or even introduce (to begin with, that is) an intuitively effective recognition procedure for a class of "Q follows from P" facts. Some algorithmic devices (or, sometimes, some transitions of configurations licensed by the device – see Sect. 9) should be seen, essentially, as "black boxes": sentences P, Q, and R (say) are transformed (by an algorithmic procedure) into notational or physical items that the device can operate on; after it does so, the resulting item(s) are transformed (by another algorithmic procedure) back into a sentence, S say. We then draw the conclusion, on the basis of this, that: "S follows from P, Q and R." In other cases, the connections between natural-language words – nouns and verb phrases, etc. – and the inner workings of the algorithmic device are more intimate (as it were). We can describe the transitions of the device in terms of transitions between sentences. The algorithmic devices I'll go on to discuss later will exemplify these differences.

A familiar first illustration: consider the numeral expressions of a natural language ("one," "two," . . ., "eleven," "twelve," . . .). These appear in natural-language sentences as both noun phrases ("one is less than two") and adjectivally ("There are six apples in Granok's basket"). Consider the introduction of a notational calculational device that provides shortcuts – indeed a *method* – for adding (Arabic notation

[41]This introduction of terminology into natural language – grafted from an algorithmic device – needn't be restricted to nouns and verb phrases. Whole sentences possessing what otherwise amounts to an alien grammar are introduced as *idioms* – whole syntactically-unparseable blocks with meanings. This is how – for example – those interpreted sentences of a formal language, untranslatable into natural language, are nevertheless expressed there. (Consider, e.g., those first-order sentences with long sequences of quantifiers that have no syntactic cousins in natural language.) I can't get any further into this now.

and the tableau layout for addition that utilizes, e.g., the base 10 properties of that notation). The inference of "There are seventeen apples in both Mordesh's basket and Granok's basket" from "There are seven apples in Mordesh's basket" and "There are ten apples in Granok's basket" is licensed by this calculational device.

In cases where there is already an antecedent proof or derivation practice, augmenting that practice with an algorithmic device in this way can induce far-reaching global effects on what can be proved or derived. A melodramatic example is when the supplementation of an already in-place derivation system with a new algorithmic device renders the augmented class of derivations trivial: *everything* can then be shown. But, there are many less melodramatic, indeed *routine,* cases.

Describe the introduction of an algorithmic device A into a language – formal or otherwise – with an already-present derivation/proof practice as *syntactically conservative* if, given an augmented derivation/proof D* of a sentence S, there is an antecedent derivation/proof D of S.[42] Often – very often – the addition of an algorithmic device into an already-present proof practice *isn't* syntactically conservative.[43] Failures of syntactic conservativeness – which, with regard to proof, are *successes* not failures – are nevertheless surprising because of our (always in place) local expectations about the noneffects of words (and the rules governing those words) on sentences in which those words don't appear. We don't expect that the introduction of a set of concepts that aren't "about" a subject area can induce new implications that are *purely* about that subject area.

An illustratively surprising result of this sort (Harris 1982) in a purely logical setting is that if we have an antecedent formal language governed by intuitionistic connectives, and we add classical negation (without placing syntactic restrictions on where that negation can appear), so that the system now has both syntactically unrestricted intuitionistic and classical negations, all the *other* connectives become classical. In particular, results like $(((p \supset q) \supset p) \supset p)$ – Peirce's law – and $((p \supset q) \vee (q \supset p))$, as well as (in the predicate calculus – here, let "$(p \Leftrightarrow q)$" stand for "$(p \supset q)$ & $(q \supset p)$"), $((\mathrm{F} \vee \forall x\mathrm{G}x) \Leftrightarrow \forall x(\mathrm{F} \vee \mathrm{G}x))$ and $((\mathrm{F} \supset \exists x\mathrm{G}x) \Leftrightarrow \exists x(\mathrm{F} \supset \mathrm{G}x))$ can now be shown. None of these are intuitionistically valid. This happens *precisely* because these results can be shown via the newly introduced classical negation and they suffice to render "$\supset$," and the other heretofore intuitionistic connectives, classical. This phenomenon is *typical* in

[42]Notice that built into this definition is that the sentence S doesn't contain any *new* terminology. Not required (and in fact rarely available) is that, given a syntactically conservative augmented derivation/proof of a result, a derivation/proof of that result using the original methods can be *easily* shown. In practice, mathematicians are drawn to augmenting their methods when that enables them to easily show new results – regardless of whether those results can be shown with great difficulty or at all by methods they already have. (Indeed, the question of whether or not innovations are conservative or not in the above sense was rarely entertained until modern times. I say "rarely" because, actually, I know of no cases where it *was* entertained – but I hesitate to be definitive about this. In the case of complex numbers or infinitesimals, the worries were ontological, not proof-theoretical – for example, are these weird numbers real or *imaginary?*)

[43]Or, more commonly, no one knows.

mathematics – it's one reason why pure mathematics is so interesting: new concepts introduced into a subject area yield surprising and unexpected results about the antecedent concepts that don't (or don't seem to) follow from reasoning about the antecedent concepts alone.[44] The phenomenon, I should stress, shows up in natural languages too – even in those cases where our concepts are governed (rather weakly) by meaning conditions alone.[45] Call these *synchronistic effects*. Lesson: synchronistic effects are everywhere in mathematics, and, as well, everywhere in our inference practices *outside* mathematics.[46]

Algorithmic devices, as I've described them, are very flexible in two senses. Consider a device that's centered on a set of concepts (terms), c_1, c_2, etc. – for example, a set of natural-language numerals (finite or not). The scope and effects of any such algorithmic device centered on $\{c_1, c_2, \ldots\}$ can be changed by (i) supplementing or changing its rules directly, and apart from this (ii) by introducing other algorithmic devices that are also centered on $\{c_1, c_2, \ldots\}$, on different terms, or both. The usual (informal) impression that there are specific concepts (e.g., the counting numbers) associated with these terms, and that they have certain fixed meanings is usually unaffected by these supplementations of and to algorithmic devices (again, consider the case of counting numbers) which aren't seen as changing their meanings.[47] Strikingly, as I mentioned earlier, the distinction between such

[44]See Azzouni (2000), where I discuss this phenomenon in relation to "relevance intuitions" and recommend it as a crucial element of the intrinsic interestingness of (pure) mathematics.

[45]Example? Consider a natural language in which the set of natural-language numerals is finite.

[46]The label "synchronistic" in "synchronistic effects" is meant to stress not so much simultaneous impact as (apparently) unrelated impact. Field-style nominalist programs, incidentally, are motivated by the (false) belief that the impact of applied mathematics to empirical topic areas can be shorn of synchronistic effects – and that, indeed, this is a *virtue* of mathematical applications. It's neither a virtue nor is it even possible. Because of synchronistic effects, the introduction of new mathematics to an antecedent subject area (mathematical or otherwise) always yields *surprises* – ones in the empirical case that are open to empirical refutation. That is, we can learn that a branch of mathematics should *not* be so-applied empirically. See Detlefsen (2008) for a discussion of the historically long and influential requirement – proponents include notables such as Aristotle and Bolzano – that epistemically good mathematical proofs should be pure (not involve alien or incongruous notions). See the discussion of the work of the mathematician Brunfiel, in Hafner and Mancosu (2008), who insists on pure proofs, and the discussion of Hilbert's analysis of Desargue's theorem in Hallett (2008). Notice what the purity requirement entails, with respect to incomplete axiomatizations. We must replace the proofs using the "alien" vocabulary with ones totally in the language of the unaugmented system. Why assume that the new axioms – in the familiar vocabulary – will be easy to find or that the resulting proofs will be particularly attractive – in *any* sense, other than in the sense of the sheer absence of alien vocabulary?

[47]One source (perhaps *the* source) of the relative insensitivity of meaning intuitions to changes in what I shortly call inferential scope is that such intuitions – especially in the mathematical case – are ontologically focused: the number words, for example, are experienced as referring to particular objects with such and such properties. Therefore, the embedding of such number-words (and rules governing them) into a context with other words (governed by different rules) is seen as simply the putting of numbers – counting numbers, say – into a context with other objects (e.g., rational numbers) which leaves the properties of the counting numbers unchanged. This (falsely – by means of a kind of use/mention error coupled with some sort of intrinsic-property confusion) gives the impression that doing so will be (inferentially) conservative.

supplementations enabling easier and/or quicker proofs (e.g., calculations) and those supplementations introducing new results that can't be shown otherwise, wasn't made – until modern times – except insofar as it was recognized that a change to an algorithmic device or the introduction of a new one into a practice could render the whole thing inconsistent.

Call the *inferential scope* of an expression, or of the concept associated with that expression, the set of truths *intuitive-effectively* determinable by the collection of inferential tools available to practitioners at a time (by "inferential tools," I mean, for example, licenses to infer based on meanings, such as "B is a brown house" implies "the outside of B is brown," logical principles, if any, algorithmic devices, if any, etc.).[48]

The heart and soul (as it were) of *conceptual engineering* in mathematics occurs via the development and/or supplementation of informal rigorous algorithmic devices. One important lesson: Most of the literature on conceptual engineering focuses on meaning and meaning change, and on issues of definition.[49] But "meaning" so-called and definitions – even in mathematics – are (largely) red herrings. I've already explained in the last few previous paragraphs why meaning is a red herring. Here's why definitions are too. Definitions are *local* in the following sense: they provide necessary and sufficient conditions for an expression *in terms of other expressions*.[50] In this way the inferential scope of a defined expression is linked to the inferential scopes of the other expressions occurring in the definition. But these, as I've illustrated, are always directly open to shifts in referential scope – even by the introduction of new expressions governed by new algorithmic devices, and so *that's* where the engineering action is.

Let me summarize this way: The contours of our concepts are always inferentially specified by our actual (intuitively effective) inferential tools; that is, the scopes of these concepts change as we change the tools we use to determine the truths that the expressions corresponding to these concepts occur in. *This* is conceptual engineering. Much of it is invisible to us because we understand our concepts and corresponding expressions *not* to change in their meanings even when the actual contours of their inferential scopes shift drastically.

I turn now to an important symptom of the presence of algorithmic devices in informal rigorous mathematical proof. Tanswell (forthcoming) has noticed that

[48]I do *not* mean, notably, *intuitive-ineffective* semantic determinations of classes of truths, such as "intended models" or intuitively ineffective "tools" such as validity defined in second-order terms using "standard semantics" – these truths are *not* intuitively effectively determinable.

[49]See, e.g., Cappellan (2018), Tanswell (2018). My work on this topic doesn't so focus. See, e.g., Azzouni (2006).

[50]Thus, definitions can range from the apparently trivial ("a triangle is anything that's a triangle") to the relatively "deep" – where the impression of depth arises from how significant or illuminating the expressions being used in the definition appear to be: in every case, trivial or not, a definition gives "necessary and sufficient conditions."

imperatives occur a lot in mathematics.[51] Indeed, the form of many informal rigorous mathematical proofs is that of *recipe-instructions,* of the form: Do *U*; do *V*; notice *W*; and then notice *X*.[52] Often, imperatives occur seamlessly with indicatives in a proof, as in the example of a proof from Rudin (1970, 357) that Tanswell (forthcoming) quotes (I've changed some of the lettering in what follows):

> Theorem: If A is a commutative complex algebra with unit, every proper ideal of A is contained in a maximal ideal. If, in addition, A is a Banach algebra, every maximal ideal of A is closed.
>
> Proof: The first part is an almost immediate consequence of the Hausdorff maximality principle (and holds in any commutative ring with unit). Let I be a proper ideal of A. Partially order the collection $\mathfrak{R}$ of all proper ideals of A which contain I (by set inclusion), and let M be the union of the ideals in some maximal linearly ordered subcollection $\wp$ of $\mathfrak{R}$. Then M is an ideal (being the union of a linearly ordered collection of ideals), $I \supset M$, and $M \neq A$, since no member of $\mathfrak{R}$ contains the unit of A. The maximality of $\wp$ implies that M is a maximal ideal of A.

Tanswell (forthcoming) suggests that the practice of using imperatives is *so* ubiquitous – especially in the context of diagrammatic proofs – that it requires a view of mathematics that centralizes imperatives. (He calls it the *recipe model of proofs*.) The first point to make is that – syntactically speaking, at least – the presence of imperatives is a relatively shallow matter. They can always be replaced with "if . . . then" indicative-locutions. This can be seen already in the above quotation, in which the imperatives may be naturally rewritten like so:

> If the collection $\mathfrak{R}$ of all proper ideals of A which contain I (by set inclusion), is partially ordered, and if M is the union of the ideals in some maximal linearly ordered subcollection $\wp$ of $\mathfrak{R}$, then M is an ideal (being the union of a linearly ordered collection of ideals), $I \subset M$, and $M \neq A$, since no member of $\mathfrak{R}$ contains the unit of A.

I described imperative-to-indicative translations as relatively shallow (grammatical) matters. Someone may resist this characterization for the following reason. Consider a (cooking) recipe Tanswell (forthcoming) gives:

1. Preheat the oven to 170°/gas.
2. Peel the bananas and lay them in a snug-fitting heatproof dish. Finely grate the orange zest and put aside, then halve and squeeze over the [bananas the] juice of 1½ oranges.
3. Drizzle with the honey and sprinkle over 2 teaspoons of the cinnamon, then roast in the oven for around 20 minutes, or until golden and soft, then set aside to cool.
4. Place the bananas and their syrup in a blender, then blitz with the milk and yoghurt.

[51] Indeed, several philosophers of mathematics have noticed and stressed the use of action terminology in mathematics. Larvor (2012) and De Toffoli and Giardino (2015) are three. They think its presence is significant for philosophically understanding informal rigorous mathematical proof. I disagree – what follows indicates why.

[52] Permissions seem to come up as well. See the discussion of the postulates of Euclidean geometry forthcoming in Sect. 11.

5. Add in a handful of the zest, the juice of the remaining orange half, a pinch of sea salt and the rest of the cinnamon, then churn in an ice-cream machine. If you don't have one, freeze the mixture in a suitable container for about 3 hours; every hour, take it out, whip with a whisk, then return to the freezer.
6. Serve in bowls with a scattering of orange zest and a fine grating of dark chocolate.

This can be rewritten, letting "(*n*)" stand for what follows each numeral *n* in the above list of commands:

If you (1), you (2), you (3), you (4), you (5), then you can (6).

Or we can introduce a stylized list (corresponding to how recipes are given):
If you:

(1),
(2),
(3),
(4),
(5),
then you can (6).

Rewriting Tanswell's recipe in indicatives, that is, reads as naturally as a list of commands does, but comes across a bit long-winded because (grammatically speaking), we can give ("bark out") commands, one after the other, in single sentences, just as in recipes; but we can't do the same with the list of "if" conditions followed by a "then" conclusion. The use of commands in recipes, however, is a mere stylistic fact about how recipes are given in cookbooks; it may even be due to the fact that a complex set of conditions are more easily surveyable if they're broken up into single sentences, and so one naturally puts them into command-form even though in no real sense is a recipe a list of *commands*. One can certainly imagine a different recipe-tradition, where recipes are instead given as predictions of what will happen under certain circumstances. In such cases, "if" conditions would be written, stylistically, as single sentences. Indeed, it can be argued, as I just have, that – semantically – that's what recipes actually are: they're predictions. (Consider the old organized-crime joke: *B*: Do this or a horse's head in your bed tomorrow morning. *C*: You're telling me what to do. *B*: No, I'm not; I'm predicting the future here.)

This point aside, algorithms (as I indicated in Sect. 4) are very naturally couched in terms of commands; machine tables look like lists of commands. It should be no surprise, therefore, that if an algorithmic device is (implicitly or explicitly) present, that presence will be usually – if not invariably – syntactically marked by imperatives.[53]

[53]That said, as Transwell's quotation from Rudin (1970) shows, and as many of the examples from De Toffoli and Giardino (2015) also illustrate, sometimes instructions on how to manipulate notation are ontologically disguised as descriptions of actions on *abstracta*. Again, a motivation for this may be that it's easier to understand single sentences (in command form) than it is to understand long "if, if, . . . if, then" sentences. In addition, it should be noted, this is an illustration of the ubiquitous use/mention conflations that occur everywhere in mathematics – right from Euclid's work on, as I illustrate in Sect. 11.

We do many things with the sentences of the natural languages we speak; in particular (and perhaps fundamentally), we express our thoughts and describe states of affairs; neither of these uses of language directly involve inference or inferential devices. The focus in this chapter is what we do with language entities by the machinery of intuitively effective inferential tools: these license the expression (speaking, writing) of certain language entities given the expression of other language entities – often sentences from sentences, but this isn't just a matter of sentences. In any case, we carry out such inferences in ways that (sometimes) phenomenologically appear rule-governed. Central examples – the ones under study here – are some of the cases where the process of inferring one sentence from another is seen as the antecedents semantically implying the consequents. These goings from one item of language to others are syntactic *functions* (their *domains* are pieces of language which they map to other pieces of language – what's described as their *ranges*). There are, of course, many ways we go from one sentence (or piece of language, words, for example) to others – other than by inference. In any case, in the foregoing, I've been focusing on syntactic functions – and in particular, on syntactic functions from sentences to sentences that have intuitively effective recognition procedures. Call these *syntactic algorithmic functions*; call the particular syntactic algorithmic functions that (at least in part) govern inference *inferential algorithmic functions*.

It should be obvious by this point in the chapter that an inferential algorithmic function can occur in a language or context that hasn't (or doesn't appear to have) other inferential algorithmic functions. Indeed, it's possible for there to be an E-centered algorithmic device in a language that has no other inferential algorithmic functions. Return, for example, to the case of categorical syllogisms, and consider the expressions, "all," "no," "not," etc. that occur in such syllogisms; let these expressions be our family E, and let the syntactic method(s) of recognizing valid syllogisms be the algorithmic device in question. Recall that I mentioned (in Sect. 2, footnote 12) that – empirically speaking – we don't know whether natural-language reasoners actually use the intuitively effective syntactic recognition procedures (that are in principle) available to determine when they're faced with a valid syllogism. Imagine, however, natural-language reasoners who *do* use an intuitively effective recognition procedure for categorical syllogisms, but who are otherwise unable to reason effectively. Imagine, that is, that these language users have no intuitively effective recognition procedures for inference patterns *other than* the one they use for categorical syllogisms. Such language-users – even if they have languages that otherwise are as rich as natural languages – can't *reason* except via categorical syllogisms. I hypothesize that such reasoners would strike us as strangely unable to grasp certain things that strike the rest of us as "obvious."[54]

[54]This is because, as we now understand, syllogistic reasoning is quite restricted in its inferential scope. However, an important cost in understanding "logic" more broadly is the loss of something that was (tacitly) assumed about reasoning until modern times – that there are *decision* procedures for recognizing validities and contradictions, and indeed (this was believed by Hume), except for issues of length, they're trivial ones.

I define an *algorithmic inferential system* as one in which *all* the inferential functions have intuitively effective syntactic recognition procedures. A first-order language which is Gödel-complete is one where the *semantic* implication function for the sentences of that language is coextensive with an intuitively effective syntactic recognition inference procedure. A higher-order language, with a model theory (an interpretation), according to which its syntactic deduction rules are Gödel-complete (Henkin models, as they're called), also has a semantic implication function that's coextensive with an intuitively effective syntactic recognition procedure. A higher-order language, however, which (relative to what's described as "standard model theory") is Gödel-incomplete, is one where the implication relation lacks *any* intuitively effective recognition procedure, let alone a syntactic one.[55]

8 Diagram-Based Formal Systems; Shin's Venn-II

As the discussion of games in Sect. 5 illustrated, just as we manipulate pieces of language (using intuitively effective recognition procedures), we can manipulate pieces of *anything* using intuitively effective recognition procedures. And similarly, there are functions, and indeed algorithmic functions, with domains (and ranges) governing some of the ways that we manipulate *things* – be those things baseballs, chess pieces, robotic hand-movements, items of notation, etc. Call a *context* the set of objects (and configurations of objects) that a set of algorithmic functions is defined on. A context can be the context of an algorithmic system or of an algorithmic device.

Diagram-tokens are a kind of object and so diagrammatic algorithmic systems and diagrammatic algorithmic devices are subcollections of object-based algorithmic systems and object-based algorithmic devices.[56] Examples of diagrammatic devices are Euler and Venn diagrams (as introduced originally by Euler and Venn), Euclidean diagrams, as practiced by ancient Greek mathematicians, and so on; examples of object-based algorithmic devices are physical tools like abacuses, various mechanical calculating devices, etc.

The "formal" in "formal language-based algorithmic system" conveys that the system in question is mathematically *fully-explicit* – in particular, the primitive vocabulary is given, and the rest of its syntax is explicitly defined on that basis (regardless of

[55]See Van Benthem and Doets (1983) for details. Notice that the absence of intuitively effective recognition procedures for sentence-functions turns on those functions being characterized in some other manner – as I suggested, for example, semantically. This will be significant for the later discussion of mathematics because, unlike games, the items in mathematics – ordinary-language sentences, sentences containing specialized terminology, and algorithmic devices (diagrammatic and otherwise) – are usually (explicitly and implicitly) interpreted. See what follows on this, but especially Sect. 10.

[56]Actually, all notation-tokens are, in this sense, objects (or, in the case of utterances, events). Nevertheless, I'll continue to distinguish between algorithmic devices using what I call objects – and here I'll be alluding to devices like abacuses (but also computers) – and those based in notation that we manipulate, such as diagrams.

whether certain kinds of syntactic items, like sentences, are, as a result, intuitive-effectively recognizable). Call a *language-based formal system* one that defines all syntactic objects in terms of the primitive vocabulary and (one-dimensional) concatenation alone. Nothing, in principle, prevents *equally formal* systems being based on syntactic constructions *other than* one-dimensional concatenation, for example, adjacency constructions in two or three dimensions (what we might call "higher-dimensional concatenation").[57] Call such systems *formal diagrammatic systems.* As with language-based formal systems, some of these formal diagrammatic systems are algorithmic and some aren't. [58]

My use of the distinction between one and higher-dimensional syntactic (concatenation) relations to distinguish non-diagrammatic from diagrammatic formal systems may seem obviously wrong – especially to those who are familiar with Shin's (1994, chap. 6) discussion of the distinction between diagrammatic and non-diagrammatic representation, as well as Giaquinto's (2007) discussion of this. Giaquinto focuses on whether conventions or resemblances are being utilized to determine interpretations – and to what degree; Shin focuses on whether new syntactic devices are introduced by a system of representation – ones that require conventions to use.

Apart from this, it may seem clear (to the reader) that a representation of individuals sitting between one another in space by locating their names one-dimensionally on a line, so as to imitate their relative spatial locations, is diagrammatic (Shin (1994, 156) – despite being one-dimensional – and it may seem, by contrast that the integral notation of the calculus, for example:

[57]I claim, but won't argue here – see Azzouni (2017a) – that if the resulting syntactic classes of items (sentences or admissible diagrams, and syntactically defined parts of these) are to be intuitive-effectively recognizable then there must be an implicit grid – in one, two or three dimensions – that the primitive items appear in; the syntactically defined items must be various (intuitive-effectively recognizable) classes of distributions of the primitive vocabulary across the grid. One approach to this places a lower-limit on the size of the areas of the cells of the grid; in one-dimensional concatenation systems (the ones we are most familiar with – alphabet systems), this is done, typically, by choosing a specific *font* for (all of) the primitive vocabulary items that appear in the (implicitly present and implicitly sized) cells. Underlying grid-frameworks aren't always explicit – they're not in Miller's (2007) Euclidean diagrammatic system FG, but they're made explicit in Mumma's Euclidean system Eu by means of arrays, which are the official backgrounds against which diagrams appear (see, e.g., Mumma 2010). I discuss these two formal diagrammatic systems in more detail in Sect. 11.

[58]I claim (see my comments on Feferman (2012) in Sect. 5 of Azzouni 2017a) that there are *two* diagrammatic systems in play in the proof of the Schröder-Bernstein theorem. One isn't algorithmic because one diagram that's alluded to in the proof is *actually* infinite. This diagram isn't *used* by mathematicians but only referred to. There is another diagram – a series of diagrams, actually – that are unbounded in size (although none of them are actually infinite). A finite number of these *are* used in the proof, and the pertinent properties of these diagrams are intuitive-effectively recognizable. There are many such examples of what should be called two-tiered diagrammatic proofs (the proof used to show what's called "Koch's snowflake" is another), ones which induce a structured semantics where the diagrams actually used by mathematicians belong to an algorithmic system but such diagrams are interpreted as referring to actually infinite diagrams in a system that isn't algorithmic; in turn, these actually infinite diagrams are interpreted as standing for functions (and sets) of various sorts. I can't get any further into this now.

$$\left[\int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\, dr\, d\theta\right]^{1/2}$$

isn't diagrammatic, despite being two-dimensional (neither, one might think, is Frege's two-dimensional notation diagrammatic).[59] I'm tabling further discussion of this until Sects. 13 and 14 because it involves subtle complications involving the interpretation of diagrams; both Shin and Giaquinto's approaches to the diagrammatic/language-based distinction, that is, involve interpretations of diagrams and sentences. Restricting myself solely to dimensional considerations as they play out syntactically yields a rough-and-ready distinction between notations it's important to explore. (Whether it tracks the actual intuitive distinction between diagrams and items that aren't diagrams is relatively unimportant.) I'll consider – at the end of Sect. 14 – whether a principled distinction is available that's more in accord with intuitions about when a representational system is diagrammatic or not diagrammatic. I'll there suggest that the question isn't semantic as it, perhaps, appears to be, but is best taken to be a psychological one that requires empirical study to resolve.

In the rest of this section, meanwhile, I'll illustrate a fully-explicit formal diagrammatic algorithmic system: Shin's (1994) algorithmic system Venn-II. "Fully-explicit" means here that the diagrammatic system is as syntactically explicit (as mathematically explicit) as any typical interpreted language-based formal language. In the section that follows, I'll consider what stands in (in informal settings) for the semantic apparatus that's used to characterize formal systems – for example, model theory. After that, I'll consider some specific cases of algorithmic devices that don't occur in formal algorithmic systems – although they can be embedded in them (this is described in Sect. 12); included among these cases are those where algorithmic devices occur in natural languages; this is what happens in ordinary mathematical practice. One important case is the (ancient) use of Euclidean diagrams – where the details of the practice are somewhat controversial and complicated.

Venn-II, as Shin's choice of nomenclature partially indicates, is a formalization of Euler/Venn/Pierce diagrams.[60] The primitive vocabulary (Shin (1994, 115) calls them "primitive objects") are closed curves, rectangles, complete shadings of

[59]Interestingly, Giaquinto (2008a, 39) suggests matrices provide "a significant example of visual operations on symbol displays. . .." These strike me, on the contrary, as involving "visual operations" to exactly the same extent that one-dimensional catenation systems do – that is, only adjacency is involved and nothing more (e.g., part/whole spatial perceptions *aren't* involved, nor do other perceptual powers for recognizing two-dimensional visual patterns – that don't directly involve adjacency – seem involved: recall how one multiplies and add matrices). In any case, the difference between one and two dimensions looks significant to other philosophers even if (according to me) they misdescribe its import.

[60]Shin (1994) presents two formal diagram systems. Venn-I corresponds to categorical-syllogism reasoning; Venn-II is an expansion of Venn-I to a formal diagrammatic system that she proves is equivalent in expressive strength to a version of first-order monadic logic (one in which the empty set can be a model). There are other formalizations of these diagrams and various formalizations of extensions of them as well. One interesting extension are "spider diagrams" – see Howse et al. (2005) and, more recently, Stapleton et al. (2013), where there are references to earlier work.

(certain) spatial regions, and the diagrammatic elements, ⊗ and —, that appear within those regions. Well-formed diagrams are defined in terms of these elements (116), and persistence and exclusion idealizations are introduced so that we're presumed able to intuitive-effectively recognize (distinguish between) distinct diagrams containing any finite number of closed curves $C_1, \ldots, C_n$ contained properly within any number of rectangles (connected by lines), where, for any rectangle, containing a subset of these curves, $C_{i+1}, \ldots, C_{i+m}$, their intersections in the diagrams, and whether or not the symbols, ⊗ and —, appear in the regions defined by these curves, are also intuitive-effectively recognizable.[61]

The next step is to provide intuitive-effectively recognizable *diagram-transformations* (121–124): these are acceptable (intuitive-effectively recognizable) transformations of diagrams into other diagrams. The intuitively effective recognition procedure for diagram-transformations is based on the well-formed-diagram conditions, which in turn are based on the recognition procedures for the primitive vocabulary. The "obtainability" relation of a diagram from a set of diagrams ("⊢") is defined via diagram-transformations (125). This obtaining relation is an intuitively effective syntactic one; it corresponds to the derivation-relationship of sentences from other sentences definable in formal language-based algorithmic systems.

As I mentioned before, unlike games (or most games, anyway), diagrammatic algorithmic systems and diagrammatic algorithmic devices usually have interpretations. This is one aspect of such diagrams that enable them to have such a significant role in mathematical practice. In the case of Shin's Venn-II, one particular interpretation is pretty natural.[62] It's a model theory where sets are assigned to rectangles, and subsets of those sets are assigned to the closed curves in the rectangles.[63] Using

[61]Shin (1994, 60, footnote 7) here relies on Polythress and Sun (1972) for the result that, given any arbitrary number of closed curves, a configuration exists (an "atomic diagram") in which for every subset of these closed curves, there is a visible region generated by the intersection of all and only the curves in that subset. Assuming a lower bound on the (tacit) cells that such curves are drawn across, this means that such configurations (and the curves themselves) are arbitrarily large. The evident implausibility of these diagrams being intuitive-effectively recognizable (we surely can't "recognize" such diagrams of arbitrary complexity – not even those containing only a small number of such closed curves) involves exactly the same sort of persistence idealizations that are invoked to allow that people, with pencils and paper, can continue a computation in time as long as needed. We are, in this case, idealizing the intuitively effective recognition procedures by allowing arbitrarily large diagrams that a person "scans" progressively (while retaining in her memory what she has seen). Exclusion idealizations are also at work when, for example, Shin (1994, 82) rejects diagrams containing partial shadings as "ill-formed."

[62]Interpretations for algorithmic systems, however, are never uniquely determined – no matter how "natural" any specific interpretation may be. This is equally true of diagrammatic systems; and so this can be confusing (especially where certain interpretations seem natural because of "resemblances," e.g., the supposed resemblances between drawn triangles and real triangles). I'll discuss this later, in Sect. 13, when I turn to debates about the "rigor" of diagrams in mathematical practice.

[63]I'm omitting niceties in describing how Shin uses this model theory to interpret the diagrams and their elements of Venn-I and Venn-II; among such niceties are ones that involve identifications of the sets that a model assigns to rectangles and curves in different atomic diagrams that are connected to one another by lines; this is needed, for example, to enable to Venn-II to express (some) disjunctive expressions.

this model theory, a notion of a diagram "being true in a model," can be defined, and via that, a notion of implication ($\models$) between diagrams can be defined as well (one collection of diagrams implies another diagram if every model that every diagram in the collection is true in, is one that the second diagram is true in as well); in terms of these notions of implication and obtainability, soundness and completeness are shown (i.e., for all diagrams D_1 and D_2, $D_1 \models D_2$ iff $D_1 \vdash D_2$).[64]

As the above model theory indicates, the diagrams of Venn-II can be associated one-to-one with the sentences in a formal language, specifically a first-order monadic language L.[65] In particular, the formal language-based system L and the formal diagrammatic system Venn-II have exactly the same metalogical properties with respect to the model theory they share.[66] Furthermore, as I indicated, this isomorphism between the diagrams of Venn-I and Venn-II and the sentences of L is intuitively effective. That is, we have an intuitively effective (decision) procedure for reading off from any diagram of Venn-I or Venn-II a sentence of one or another (interpreted) formal language L which "says" (in terms of the interpretations of both the language and the diagrammatic system) what the diagram "says." We can similarly depict what any sentence of the formal language L says with a diagram that captures its "content" exactly.

The model theory for L – traditionally understood – provides interpretations for all the sentences of L. That is, the model theory for L indicates what the sentences of L say. Further, given a specific model M of L, we can speak of the sentences of L being true or false relative to M. The same is true for the diagrams of Venn-II. Each diagram, using the same model theory, says something, and what it says (relative to a specified model M) is either true or false. The model theory for Venn-II and for

[64]See Shin (1994, chap. 5). The fully-general result, in terms of infinite collections of diagrams, that enables a compactness result for the implication relation of models, is given by Hammer and Danner (1996).

[65]Again, avoiding certain niceties, that the procedure of associating sentences with diagrams is intuitively effective turns on how the interpretation for a diagram (in terms of sets) is given – this yields a normal-form theorem: each atomic diagram corresponds to a conjunction of existential and negations of existential sentences describing the (finite) intersections of the sets corresponding to all the closed curves in a diagram, and a diagram (composed of atomic diagrams connected by lines) corresponds to a disjunction of these conjunctions. The reverse procedure turns on giving an intuitively effective method of transforming any sentence of the language to its normal-form equivalent. I stress that these procedures are decidable.

[66]On this, see Shin (1994, specifically Chap. 5), Hammer (1994), and Hammer and Danner (1996). One metalogical property the two systems (naturally) share: the theorems of both systems are decidable and not merely effectively enumerable.

Venn-I (for that matter) thus provides interpretations for the elements of the diagrams that also, as I indicated, yields a notion of truth-in-a-model for the diagrams.[67]

Relations between sentences and diagrams when formal systems like the above are in play are not only decidable, but involve metalogical properties such as completeness. Relations between sentences and diagrams are useful in mathematical contexts – and generally – even when such metalogical properties are missing, as they usually are when algorithmic devices are employed in natural-language contexts. Here, usually, model theory is absent as well – and so model theory *isn't* the tool by which mathematicians interpret diagrams. Nevertheless, we can (and do) legitimately speak of interpreted diagrams or, more generally, of interpreted configurations of objects (a certain sequence of executed movements on an abacus, say) as being true or false – or as being implied by later interpreted configurations. The interpretations in question, however, are ones given by specified sentences of natural language – or sentences in a natural language augmented with (technical and specialized) vocabulary items that nevertheless respect the natural-language syntax this vocabulary occurs in. I turn to developing this important point in the next section.

If we combine Shin's diagrammatic formal system together with a language-based formal system, the result is an example of a *formal* system of "heterogeneous reasoning" in Barwise's (1993) sense (Shin 1994, 188–189; Hammer 1994): a *formal* algorithmic system that has both diagrammatic and language-based elements. Although this kind of combination of systems – diagrammatic tokens along with language-based tokens – is atypical of most formal systems currently studied, the heterogeneity it exhibits is endemic in informal rigorous mathematics. I turn to attempting a description of how similarly "heterogeneous reasoning" works in informal settings.[68]

9 Informal Reasoning by Means of Algorithmic Devices

Consider the use of an algorithmic device **A** in a natural language context. As I indicated above, informal versions of the semantic/syntactic tools in play in formal systems like Venn-I and Venn-II arise here too, although notably they don't have all

[67]For example, Shin (1994, 118) gives a definition of a set assignment (a mapping of sets to the rectangles and closed curves of a diagram) *satisfying* a well-formed diagram. Well-formed diagrams, however, correspond to *sentences*; there is nothing in her definition of satisfaction involving variables as they occur in the formal language-based systems (e.g., those studied in Tarski (1933)). Atomic diagrams are given interpretations by mapping the curves to subsets of the domain of the model, and mapping the elements to the sets in question being empty or nonempty (see Shin 1994, Sect. 3.31). These correspond to sentential relations between sets, for example, *The intersection of A and B is empty*; *the intersection of A and B is nonempty*. The additional compositional facts – about diagrams – that Shin's definition of the satisfaction relation turns on involves the linking of atomic diagrams by lines (which corresponds to the disjunction of what the atomic diagrams represent).

[68]See Sects. 8 and 9 of Azzouni (2017a). I there described algorithmic reasoning as occurring on the "surface" of informal rigorous mathematical practice. This is a metaphor that I'm explicating here in terms of algorithmic devices.

the properties that their formal cousins have. I here sketch out what things look like in the informal setting.

Start with the observation that a natural-language algorithmic device **A**, as in formal settings, allows the intuitively effectively recognition of well-formed diagrams D_i.[69] As before, there is an intuitively effective, recognizable obtainability relation, $\vdash$, that **A** licenses between its diagrams. In addition, there are intuitively effective recognition methods of reading ordinary sentences *off of* the diagrams of **A**, $D_i \rightrightarrows S_j$, and reading sentences *into* diagrams of **A**, $S_j \rightarrow D_i$. Specifically, given a diagram D_i and a sentence S_j (from a certain set of designated sentences), we can often change that diagram into another D_j, where for all sentences S_n, we can recognize that if $D_i \rightrightarrows S_n$, then $D_j \rightrightarrows S_n$, and furthermore, we can recognize that $D_j \rightrightarrows S_j$. I'll call $\rightrightarrows$ and $\rightarrow$, respectively, "reading-out" and "reading-into" procedures.

These procedures may help themselves to language-elements that are contained in diagrams (e.g., letters that appear in the diagrams and that label diagrammatic elements of it) but they needn't. It also needn't be the case that, generally, one can mechanically generate all the sentences that can be read off of a diagram – although that's true of a lot of algorithmic devices (e.g., Shin's Venn-I and Venn-II, and, I think, the informal Euclidean-diagram practice). What's assumed here is just that it's mechanically recognizable for any sentence S_i and diagram D_j, whether or not $D_i \rightrightarrows S_j$ or $S_j \rightarrow D_i$.[70]

Here is the first important contrast with Shin's formal diagrammatic system – and with formal diagrammatic systems generally: there is no assumption of uniqueness in the reading-off and reading-into relations – nor need there be "normal-form" sentences that can be described as what captures (uniquely) what the diagrams of an algorithmic system **A** "say." That is, given a diagram D_i, there can be many different sentences, S_j, S_k, ..., with appreciatively different (although compatible) contents, such that $D_i \rightrightarrows S_j$, $D_i \rightrightarrows S_k$, ... Similarly, given a sentence S_i, there can be many diagrams D_j, D_k, ... (also with appreciatively different contents), where $S_i \rightarrow D_j$, $S_i \rightarrow D_k$, ... All that's assumed is that the reading-into and reading-off relations have intuitively effective procedures. In many cases, these reading-into and reading-off relations are additional decision procedures available to diagram users: given S_i and D_j, it can be determined whether, or not, $S_i \rightarrow D_j$ or $D_j \rightrightarrows S_i$.

[69]For ease of exposition, I'm describing all algorithmic devices as if they're diagrammatic, rather than more general object-based devices. What I say here can be applied easily to the broader case, although I won't (systematically) do that explicitly.

[70]Notice that we can often intuitive-effectively recognize, of two diagrams, D_i and D_j that if $D_i \rightrightarrows S_n$, then $D_j \rightrightarrows S_n$, without necessarily being able to recognize all the sentences S_j of which it's true that $D_i \rightrightarrows S_n$ and $D_j \rightrightarrows S_n$. For example, D_j might result from D_i by the addition of a single diagrammatic element to D_i (e.g., a line or cross) so that it's easy to tell how D_i and D_j differ in their sentential contents.

Define the *content* of a diagram to be all the sentences S_n such that $D_i \rightrightarrows S_n$. (I'll notationally indicate the content of a diagram D_i as Content (D_i).) As noted already, that a given sentence is in the content of a given diagram is intuitively effectively recognizable.[71]

The content of a diagram, importantly, can sometimes look *empty.* In certain standard language-based formal derivation systems, for example, this happens because syntactic-transition rules allow transitions of formal sentences to well-formed expressions with free variables – that don't strictly speaking, say anything – and back again. In the case of mechanical devices that enable calculations, this can happen because certain machine-state transitions can be to and from states that don't nicely correspond to sentences (natural-language or otherwise). Diagrammatic algorithmic devices can operate similarly. Call such syntactic-transitions between diagrams *semantically null* transitions. The algorithmic systems we'll be concerned with will be ones in which any such diagram has, nevertheless, an interpretation – a content; this will be imposed on it by syntactic-transition rules that generate such diagrams from earlier ones. (See the description of soundness results, immediately below.)

The content of a diagram is treated as the *interpretation* of a diagram. This is usually taken to be due to antecedent interpretations of the elements of the diagram, for example, that diagrams of triangles represent triangles, closed curves represent sets, and the interpretation of their syntactic locations in the diagram – for example, that a point-element within a triangle-element represents the presence of a point topologically within a triangle. That is, the interpretation of the diagram is taken to result compositionally from the syntactic placing of these elements in the diagram. There is much more to say about this which I can't do here.

In any case, we usually have the following important *soundness* results that hold of algorithmic devices **A** that are under study in this chapter:

For all diagrams D_i and D_j in **A**, D_i, $\vdash D_j$ only if

(a) For all $S_n \in$ Content (D_i), $S_n \in$ Content (D_j),
(b) For all $S_n \in$ Content (D_j), Content (D_i) $\models S_n$,
(c) for any semantically null transition from D_i to D_j (D_i, $\vdash D_j$), Content (D_i) = Content (D_j).

[71] Contents of diagrams here are sets of sentences. This isn't the only option. We may want to treat the contents instead as "what the sentences say": in other words, the sets of configurations of elements that the sentences are about. Doing it this way or my way doesn't matter, except insofar as the approach indicated in this footnote generalizes straightforwardly to model-theory – whereas my approach involves extra steps. The reader who impatiently wants to get past *sentences* to items that the sentences are about can substitute configurations of elements for sentences in my above definition of content. This is fine except that it introduces extra expository steps in my forthcoming discussion of how practitioners shift from diagrams to natural-language results that such diagrams are taken to show, for example, in the case of Euclidean geometry, Proposition 1 in Book I, that an equilateral triangle exists that has a given straight line as one of its sides.

Some comments about the algorithmic devices these soundness results hold of.

First, as I said, this is a *soundness result* for the proofs generated by the algorithmic device **A**, coupled with the reading-into and reading-off procedures. Unsurprisingly, for algorithmic devices in informal rigorous mathematical contexts generally, we don't have completeness results. These come up – when they're available – when such devices and the informal mathematics they occur in are embedded in one or another formal algorithmic system that's formally interpreted as well (e.g., in relation to a model theory).[72] I'll say more about this in Sect. 12.

Second, I don't think (and so I don't assume) that such soundness results were *shown* in informal settings before modern times. Rather, they were empirically established (or, perhaps more accurately, assumed). In addition (but related to this), practitioners tacitly employed reading-into and reading-off procedures without recognizing much about the properties of these procedures. They recognized that there were potential issues with these procedures only insofar as they grappled with unwarranted-generalization problems. These amount to (ancient and modern) questions about what the elements in a diagram can be taken for stand for (which, exactly, mathematical triangles diagrammatic triangles in a Euclidean diagram can be taken to stand for; what mathematical functions diagrammatic curves can be taken to stand for; etc.); related to this, what the reasoning with specific diagrams can be taken to have shown, generally. (I discuss this in further detail in Sect. 10.) Apart from this, with physically instantiated algorithmic devices, empirical assumptions are needed to show soundness results. This is because persistence and exclusion idealizations must be assumed as successfully operative when these devices are used.

Third (as I've already implicated in the second remark just made), the reading-in and reading-off procedures are crucial to the use of algorithmic devices in informal rigorous mathematical proof. This is because the theorems shown by any execution of an algorithmic device **A**, is read off of the "final state" of that device. In general, any theorem yielded from a diagram by a reading-off procedure is of the form, $S_1, \ldots, S_m \models S_n$, where $S_1, \ldots, S_m$ have been read-into the sequence of diagrams at various earlier points; and S_n is read-out from the final diagram.

Fourth (as I indicated in the discussion in Sect. 7, and especially in footnote 41), the natural-language that the reading-in and -out procedures rely on is invariably terminology-enriched.

This is already obvious in the later Euclidean tradition where "triangle," "line," and so on refer to abstracta, and not to irregularly shaped physical items, that the original Greek words surely referred to. The sentences read off of a diagram can also be, syntactically speaking, "whole cloth" items: alien jargon that functions idiomatically vis-à-vis natural language (i.e., items that are unparceable as far as natural language is concerned).

[72]Of course, completeness results – relative to certain given model theories – may not be possible, as Miller (2012, 34–35) indicates is the case with Euclidean diagrammatic formal systems and certain (intended) models.

Because of the crucial role of reading-into and reading-out procedures in these proofs, there is an important generalization of the notion of "heterogeneous reasoning" beyond cases of formal systems to cases where algorithmic devices occur in informal rigorous mathematics proof (to cases of mathematics, that is, that occur in natural languages). I'll call this *informal heterogeneous reasoning*.[73] The idea is straightforward. Suppose we have an algorithmic system **A** that licenses transitions among diagrams D_i; suppose also that we have ways of recognizing implications ($\models$) between sentences of natural language (be those sentences natural-language ones or sentences composed – in part – of supplementary and specialized mathematical vocabulary). A proof (relying in part on an algorithmic device) can look something like this:

$$S_i \models S_j, S_j \rightarrow D_k, D_k \vdash D_l, D_l \rightrightarrows S_m, S_m \models S_n,$$

where we take, as a result, $S_i \models S_n$.[74] This is justified by a soundness result, of the above form, for **A**.

Example: Suppose the relevant sentences of a natural language are ones like "ten plus seventeen is twenty-seven." Let D_i be the configurations of an object-based or language-based algorithm **A**, such as an abacus, or an (Arabic) numeral system for addition or multiplication. The elements of D_i are interpreted in natural language, say by the assignment of the elements of a diagram, "1," "2," etc. to the natural-language numerals, "one," "two," etc. or instead, by, say, configurations of the elements of a physical algorithmic device (bead patterns in an abacus, for example) being assigned to natural-language numerals. And we (consequently) interpret the diagrams, and (possibly) certain subconfigurations of those diagrams, as having truth values as a result.

Consider an *abacus.* Here are some typical "word problems" we can solve using it:

The auditorium has 75 normal chairs and 20 high chairs. How many chairs does the auditorium have in total?

A bricklayer stacks bricks in two rows with ten bricks in each row. On top of each row, there is a stack of six bricks. How many bricks are there in total?

Serena buys two packs of tennis balls for $12 in total. All together there are six tennis balls. How much does a pack of tennis balls cost? How much does one tennis ball cost?

In each case, we extract an implicit arithmetical question that the word problem poses, and then, using a reading-in procedure, we generate a (finite) sequence of configurations of abacus beads; in the first word problem, for example, the arithmetical statement to be shown is: "Seventy-five chairs plus twenty chairs is X chairs," where X answers the question, and thus the word problem. Call the

[73] I'm here generalizing Barwise's (1993) use of the phrase to algorithmic devices in natural language. See footnote 14 above.

[74] In general, of course, many sentences occur, both as antecedents, and in the stages where sentences are read into diagrams.

antecedent description the calculation, and call the substitution of a natural-language numeral for "X" in "X chairs," "X bricks," etc. *the result*.[75] The calculation is coded into the algorithmic device as a sequence of configurations of abacus beads, and the result is read off from the final configuration. Introducing the usual exclusion and persistence idealizations, we imagine that abacus devices can be arbitrarily large in size (and yet humans can still execute operations on them). There is an intuitively effective recognizable class of arithmetical statements (involving the ordinary-language addition, subtraction, multiplication, and division), where the calculation can be intuitive-effectively coded as a sequence of configurations of an abacus, and where the last configuration corresponds to the result. This, therefore, is an example of heterogeneous informal reasoning, where there are both language-based inferences (of, for example, word-problem statements into arithmetical statements) and abacus inferences. The abacus results, in this way, are returned to an ordinary-language arithmetic practice (such as it is).

We can imagine, instead, that the algorithmic devices in question are notational (short division, multiplication, etc.). Notice the result: Via the reading-in and reading-off procedures, algorithmic devices of all sorts can easily augment a sentence-based informal mathematical practice. In particular, the invention of a new notation (that provides an algorithmic device – or several) can be easily added to an already existing mathematical practice in natural language with preexisting algorithmic devices. It's then empirically established that the new algorithmic devices both enable the proofs of apparent new results or they introduce ease in yielding theorems that are already accessible otherwise. That the algorithmic device doesn't render the practice trivial by yielding inconsistent results (either internally or in relation to results established by other means) is also – and must be – empirically established.[76]

I need a label for the characterization I've just given: that informal rigorous mathematical practice involves the use of algorithmic devices (either language-based or not) along with reading-in and reading-out procedures. Call this the "algorithmic-device interpretation of informal rigorous mathematical proof," or (for short) the *device view.*

10 The Generality Problem

The diagrams we make, look at, and study are tokens of syntactic objects (of types) – that is, they're tokens of types of *physical* items (sketches on paper, in sand, fleeting utterance-events, etc.). Call *the generality problem* the question what interpretation

[75]Often, natural language is supplemented, for example, in this case with new numerals, so that the reading-out procedures are still ones into natural language – although now supplemented.

[76]I mean "empirically established," not in the sense that a mathematical proof can't be given for this; I mean it in the sense that we don't have a decision procedure for discovering things like this: not having unearthed an inconsistency *so far* doesn't mean an inconsistency isn't there to be unearthed.

is to be given to the elements of the diagram, and thus, compositionally, to the diagram itself. In the case of formal algorithmic systems, this is the question of what model or set of models the diagrammatic or language-based formal system is to be interpreted in terms of. *The straightforward solution* to the generality problem in the formal context takes Gödel *completeness* results to be the guide to answering this: Choose a model theory to interpret a formal system **A** where the validities and the theorems match.

This, when philosophers study formal languages, usually *isn't* taken to be an acceptable solution to the generality problem. On the one hand, there is *the uniqueness problem.* There are many "unintended" models for any formal theory – models of "nonstandard arithmetic" for first-order Peano arithmetic or the many unintended models for first-order ZFC, for example. It's felt by many practitioners that, despite our having no syntactic grip on any specific model in these cases, there nevertheless *is* an intended model – that is, there is a unique model that we (somehow) *intend* the axioms we write down to capture.[77] But, not only do the syntactic consequences of first-order Peano arithmetic fall short of extending over all the validities (theorems) of this intended model, any augmentation of Peano arithmetic that has an effective syntactically characterizable consequence relation falls short the same way (Gödel's incompleteness theorem).[78] Relatedly, it's felt by many proponents of certain logical systems – for example, higher-order logics – that a model theory vis-à-vis which their syntactically specified theorems are complete doesn't capture the intended meanings of the terms in those systems. According to the "intended" meanings of the quantifiers of such systems, for example, any effectively specified set of syntactically characterized theorems are incomplete.[79]

[77]This model, in the case of Peano arithmetic, is called "the standard model" – for example, by Kaye (1991).

[78]For the record, I'm skeptical there is genuine content to "intended model" in any case in which characterizing it outruns our syntactic resources. I can't get further into this now; see, instead, Azzouni (2018) for discussion of the case of arithmetic – which is a case where intended-model intuitions are especially strong. Notice what the problem is. We feel we understand – grasp the meaning of – the intended extension of a notion when it's given by negating a term whose extension we *do* grasp, for example, the consistent formulas of first-order logic (which are all those which *aren't* first-order validities or contradictions). On the other hand, there is certainly a sense in which we don't understand – don't grasp the meaning of – the extension of a notion if it isn't recursively enumerable. I omit further discussion of this very delicate matter except to observe that it's connected to the discussion – in Sect. 7 – of how our intuitions of meaning are insensitive to what I there called inferential scope.

[79]Nomenclature: I think, for the reasons just given, that the generality problem should be called, instead, the intended-model problem. The latter label more sharply focuses what's at issue, both in formal and informal settings. But the "generality problem" is the label Mumma and other philosophers in this literature use – so I'm sticking with it instead.

I intend to focus on that version of the generality problem faced by a specific informal diagrammatic system: the generality problem of ancient Euclidean diagrammatic practice. This version of the generality problem gets a lot of attention, both now and over the centuries. But it's important to indicate first that it's something of an illusion that a rather different form of interpretation is at work in formal settings – where models are supposed to be used to interpret formal languages – than in the case of informal mathematical practices with algorithmic devices, where, according to my characterization of them, reading-in and reading-off procedures into natural language are used instead.

The reason is one that was made clear long ago by Davidson's (1986) characterization of how Tarski's theory of truth is supposed to be used in interpreting languages: The semantics for any language – logical or otherwise (i.e., one in which there is a nonlogical vocabulary that's allowed to vary across models, or one in which there isn't any such nonlogical vocabulary) – occurs in an *interpreted* metalanguage.[80] That means that the meanings of the elements of sentences and diagrams, and the meanings of the syntactically composed sentences and diagrams occur via transliterations of them into the antecedently meaningful language of the metalanguage. And *that* means that an interpretation of a formal algorithmic system, specifically, a formal diagrammatic system, is – strictly speaking – a formalization of what I described in the last section as the reading-in and reading-off procedures that are applied to algorithmic devices.

More precisely, the formalization in question is accompanied by an interpretation of the algorithmic system that's a transliteration of a bit of informal mathematics – diagrammatic or otherwise – and including some algorithmic devices, reading-in and reading-off procedures, as well as (usually) some accompanying informal natural-language reasoning. The algorithmic system and its interpretation is specified in ordinary language (as always happens in informal rigorous mathematics) that has

[80]That the interpreting language be, strictly speaking, a distinct language – a "metalanguage" – isn't required. That's just the traditional way of doing it. For example, it's possible to have one language in which the interpretation of a part of that language take place in another part, and where the two parts are syntactically isolated from one another. I should add that in classical settings the background interpreted metalanguage is usually (a version of) set theory – that is, a mathematical subject matter that's ordinary informal rigorous mathematics (in a natural language). See what follows.

been supplemented with specialized vocabulary – in standard cases, with set-theoretic vocabulary.[81] If the metalanguage – in which this mathematical specification of the syntax and semantics of the algorithmic system takes place – is, in turn, formalized then the same process (of interpretation) can be repeated.

In a purely informal setting, the generality problem coalesces onto the reading-in and reading-off procedures; and this version of the problem has troubled mathematicians (with respect to diagrams) right from ancient times until the present. The Euclidean version of the problem is this: how is a diagram to be understood – to what do the elements of the diagram refer? Consider a drawing of a triangle, for example. One (natural) interpretation of it is that it represents a class of triangles that (roughly) match the *drawing's* angular properties: the acute triangles are referred to by drawings of acute triangles; oblique triangles are referred to by drawings of oblique triangles; and certain properties of other drawn figures – *and not others* – play similar scope-indicating roles. Call these *resemblance interpretations* of the diagrams in a diagrammatic practice.[82] Euclid's proofs don't seem to respect any obvious resemblance interpretation systematically: sometimes certain properties of a figure seem involved; sometimes other ones do. The generality problem (in the Euclidean setting) is made complicated by the fact that what determines the scope of

[81]Model-theoretic approaches to interpretation (and Davidsonian approaches to interpretation, therefore) cannot, contrary to Davidsonian claims (Davidson 1986), be taken as providing the theoretical framework for explanations of how we understand languages and of how the meanings and references of these languages are facilitated – that is, what the machinery is by which the terms of a language acquire the referential scopes that they have (e.g., that "everything" refers to everything; that "gold" has the extension it has). Call the latter language-world questions; call the former language-mind questions. The reason Davidsonian approaches don't answer language-world and language-mind questions is that, as I've just mentioned, the mechanism by which Tarskian-inspired model theories interpret languages is by providing compositional-element-to-element *correlations* between target-language elements and metalanguage elements that provide the interpretations for targeted elements by simply giving them the referential and meaning properties of the correlated metalanguage elements. Doing this can't answer language- world and language-mind questions unless we *already* have answers to those questions with respect to the metalanguage elements. That is, it *defers* answers to these questions for a targeted language to whatever answers we have for the metalanguage – none as far as the Davidsonian approach is concerned since that approach bizarrely takes the targeted language questions as *answered* by these correlations. Notice the objection *isn't* the truistic complaint that, if a theory of how language elements refer to the world is given, such a theory needs to be given *in* a language which embodies (or has) language-world connections and is one that we understand. The point, rather, is that no illumination of language-world and language-mind questions is given by "explaining" that how "pomme" manages to refer to all and only the pommes is that it happens in just the same way that "apple" manages to refer to all and only the apples. An analogous point holds of the question of how Capucine understands "pomme"; it doesn't help to say that she understands it in just the way that I understand "apple." See Azzouni (2017b) for an attempt to spell out what's required for a genuine attempt to answer language-world and language-mind questions. I'll also note that to complain this way about truth-conditional semantic theories isn't to complain about their ability to capture the compositionality of languages: they *successfully* capture compositionality phenomena.

[82]Philosophically refined versions of resemblance interpretations are still taken seriously – by for example, Shin and Giaquinto. I discuss them in Sect. 12.

a diagram are the reading-in and reading-out procedures; but they're *implicit*.[83] The result is centuries-long controversies over whether Euclid provides sufficient diagrams to cover all the relevant cases that seem to fall within the scope of his established results – whether, in fact additional diagrams are needed for particular diagrammatic proofs to successfully prove results that he seems to understand as having a certain generality (e.g., as about all triangles, as opposed to some obeying such and such specific conditions – for example, acuteness)[84].

The generality problem for informal diagrammatic practices seems to have a different flavor than the one faced by language-based systems. This is because there are additional properties that we're visually sensitive to above and beyond those visually available in one-dimensional concatenation systems. In addition to adjacency, for example, there are topological properties (containedness) that it seems that reading-in and reading-off procedures in some diagrammatic traditions can (and do) exploit. In those diagrammatic practices, additionally, mathematicians "just see" visually available relations that are often interpretationally correlated with properties of the mathematical items the diagrams are about.[85] This "just seeing" is (largely) why the reading-in and -off procedures exploiting what's "just seen" are tacit. Nevertheless, the generality problem is the same one that's faced by language-based systems; this is because nothing requires the visual properties of a diagrammatic system to be exploited by the reading-in and reading-off procedures: it's a convention to so-interpret syntactic properties – diagrammatic or otherwise – *regardless* of whether the procedures, that exhibit these conventions and that provide the interpretations, exploit resemblances or not.

When an informal diagrammatic practice is formalized, one explicitly "solves" the specific version of the generality problem that diagrammatic practice faces – for example, the Euclidean diagrammatic practice – via *two* moves, versions of both of which are used when formalizing *any* informal inferential practice. The first move is to mathematize the (physical) diagrams themselves by identifying them with one or another kind of mathematical object. Shin (1994), for example – recall Sect. 8 – mathematizes Venn-diagram drawings by characterizing them as closed curves

[83]Mumma (2010, 256–257) states the generality problem for the informal rigorous Euclidean practice this way: "The generality problem arises with Euclid's proofs because the diagram used for a proof is always a *particular* diagram. Euclid clearly did not intend his propositions to concern just the figure on display beside the proposition. They are applied in subsequent proofs to other figures, which are not exact duplicates of the original. And so, for Euclid, consultation of the original diagram, with all its particular features, is somehow supposed to license a generalization. But Euclid leaves the process by which this is done obscure."

[84]See Heath 1956, 245–246.

[85]These have come to be called "free rides" – a phrase introduced by Barwise and Shimojima (1995). I'll eventually characterize these in terms of what I'll call "embodied algorithms." See Sect. 13 below. Also see Azzouni (2013a), where I discuss the free-ride phenomenon, although without using that nomenclature.

within rectangles (that can be connected by lines) with certain mathematical properties.[86] This corresponds to our making metamathematically explicit the syntax of a formal language – as when the sentences of a language we speak are identified with certain abstract objects. As I described it (in Sect. 8), well-formed syntactic items are characterized, the syntactic-transition rules are given, and so on. The second move is to provide a model theory (a semantics) for the resulting language-based or diagrammatic formal system. Although the diagrams are recognizably identified with well-formed syntactic items, the reading-in and reading-out procedures are (relatively invisibly) encapsulated into the interpretations given by the correlations of the mathematical objects standing in for the actual diagrammatic devices (drawn circles, etc.) to model-theoretic items.

I now turn to a closer discussion of the formalization of the Euclidean diagrammatic practice.

11 An Algorithmic Device in Natural-Language Mathematics: Euclidean Diagrams

Everyone admits that Euclidean geometry, the *Elements*, is the most influential reasoning practice in human history.[87] Consider Book I of Euclid's the *Elements*. A post-Fregean temptation – one that's still alive among contemporaries – is to think of Euclid's approach as *the* earliest axiom system: a pioneer version of the formal language-based systems that emerged with Frege, but occurring in natural language, and in which diagrams play a purely heuristic role.[88] Views about exactly *how* the

[86]This is often implicit, but important: The abstracta that in the formalization replace the physical items we actually use in the informal practice – for example, the actual drawings – are chosen to preserve (at least in principle) the intuitive-effectiveness of the recognition procedures of the informal rigorous mathematical practice. Preservation of certain properties of the phenomenon being mathematicized is a typical precondition on the application of branches of mathematics; in general relativity, for example, we choose a mathematical object (a manifold) that replicates – at least up to current observational thresholds – the properties of spacetime.

[87]Mumma (2008b, 256, italics his) puts it nicely: "[The *Elements*] represented the limit of mathematical explicitness. It served as *the* paradigm for careful and exact reasoning." This is seen by Manders (1995), Mumma, and others studying the practice, to pose an historical puzzle: why the relatively widespread repudiation as being *obviously* inadequate of the Euclidean proof-practice by later nineteenth-century mathematicians? I provide my answer to this question in what follows.

[88]Note the suggestion: diagrams play a purely heuristic role *in the informal rigorous Euclidean practice.* That the practice can be formalized one way or another in a pure language-based system (e.g., in the formal system of Avigad et al. (2009)) is irrelevant because it follows from Hilbert's thesis that this can always be done.

reasoning with this system works – in particular, what role the diagrams have – has shifted over time. Plato's early characterization treats the diagrammatic elements as crucial to reminding us what the proof elements (words and diagrams) refer to, and therefore why the results diagrammatically shown are true.[89] Later views of reasoning treat a perfect item of reasoning – a proof – as a completely explicit series of inferences without gaps (no "missing steps"); this view is present in Descartes[90] but remained programmatic until Frege. Frege, however, sustained the picture by assuming, naturally enough that the (artificial) language that proofs occur in is one with a fixed semantic interpretation, and where each step follows immediately (by rules) from earlier ones. Once the uniqueness of the possible interpretations of artificial languages (or of the introduced terminology in natural languages) is stripped from the picture, what remains (indicated by the phrase "by rules") are intuitively effective recognition procedures for proofs. The internal syntactic structure of the language and the interpretations of the logical operators across models that it embodies are all that remain to fix the interpretations of its sentences and that of the other language elements.[91]

Until it was recognized that diagrammatic proofs – specifically, the transitions between diagrams in a diagrammatic proof – utilize intuitively effective recognition procedures just as language-based proofs do (and thus are equally open to formalization), the temptation remained to treat our recognition of the inferential properties of proofs as matters of "intuition" to be eliminated. So, I'm hypothesizing that a certain picture of proof (gaplessness) and the impression that gap-free proofs are

[89]My interpretation of Plato's view of reasoning is that it's akin to perception: "seeing" in the sense of understanding and "seeing" in the sense of visually perceiving aren't ambiguously different meanings of "see" as they appear to be. (This is how I'm stating Plato's position in English; I'm certainly not making any sort of claim about the corresponding Greek words.) This shortly mutated into a view where reasoning was seen as a matter of deducing particulars from universals, although, I think, this deducing continued to be regarded as a kind of "seeing" – the *universals*, in particular.

[90]Descartes (1931, 19) writes of deductions that they must be "scrutinized by a movement of thought which is continuous and nowhere interrupted. . .." I'm tempted to credit Descartes with originating this view of proof, but this is a matter for historians to determine, not me: I don't know whether there were earlier expressions of the claim or (as important) whether Descartes knew of and was influenced by them. The idea, of course, may originate in the impression that inferring is like seeing.

[91]Thus: the emergence of unintended interpretations.

only possible in language-based media were decisive in turning mathematicians against the use of diagrams.[92]

There are several contemporary formalizations of Euclidean geometry. In this section, I'll focus on two: Miller's FG and Mumma's Eu.[93] I'll first give a short overview of how I regard the original Euclidean practice – this involves a bit of speculation, of course.[94] Then I'll turn to sketching how Miller and Mumma formalize the Euclidean practice; this naturally raises issues about which properties

[92]Mumma (2010, 257–259) suggests that what instead turned mathematicians against diagrams, and, specifically, Euclidean diagrams, was the generality problem.

A different challenge to my historical hypothesis is posed by something Frege (1879, 5) writes of his own version of mathematical proof: "To prevent anything intuitive [Anschauliches] from penetrating here unnoticed, I had to bend every effort to keep the chain free of gaps." This makes it sound like the gapless ideal for proof was introduced by Frege to avoid "intuition," and not that the programmatic gapless model of proof was appealing (and motivating him) in its own right, and thus was what was undermining the acceptability of "intuition." My interpretation is saved, however, if Frege's use of "anschauliches," is read – as I think it should be – not as only a label for pictures but as standing for *anything* unarticulated that enables a mathematician to "see" his way from p to q without using the intermediate steps that are presumed present in a "fully-explicit" canonical proof. "Intuition," when used to describe the mathematical insights that diagrams enable involves *several* elements that weren't clearly distinguished. One is the historically unrecognized role of intuitively effective recognition procedures directly involved in informal diagrammatic proof; but a second element (that I discuss in further detail later) is the implicit interpretations of diagrams via the reading-in and reading-off procedures.

Still another apparent challenge to my historical interpretation of Frege being primarily focused on gapless proof is that he also seems to explicitly commit himself to purity views of mathematical proof, as Detlefsen (2008, 187) illustrates with a quotation. Frege's purity condition, however, is specifically focused on the exclusion of "geometrical elements" from arithmetic. And, further, it also seems that Frege is thinking that the "geometrical elements" are ones that aren't *logical* – he writes, in the quotation that Detlefsen gives, "the task of deriving what was arithmetical by purely arithmetical means, i.e. purely logically, could not be put off." This quotation, therefore, is compatible with the gapless logical ideal of proof still being Frege's central motivation, since he doesn't commit himself to a more general form of purity of *mathematical* subject-matter. See footnote 46 for examples of philosophers and mathematicians who extoll purity goals in mathematical proofs, where those goals are driven by subject-area considerations and not logic.

[93]See Miller (2007, 2012), Mumma (2008a, b, 2010, 2019). There is ongoing controversy about Eu, and so this system is, as a result, still somewhat in flux. Miller (2012) shows that the earlier version of Eu (Mumma (2010) – and elsewhere) is inconsistent, unsound, and has several other infelicities as well. Mumma (2019) addresses the issues Miller raises by substantially modifying the system. (Although Mumma has not done so, the two systems should really be distinguished by different names.)

[94]This discussion further supplements the picture I gave of the Euclidean practice in Azzouni (2004). I wrote that paper in 1998, for the *New Trends in the History and Philosophy of Mathematics Conference* held at the University of Roskilde, Denmark, August 6–8, 1998; the conference papers, however, weren't published until 2004 – thus my neutrality vis-à-vis mathematical ontology in the paper (which is the same version I gave at the conference) instead of my by-2004 adopted nominalism. I should add that both in this chapter, and back in 1998, I'm not and wasn't attempting a suitably careful historical analysis of the Euclidean practice. See, instead, for this, Manders (1995) and Manders (2008). I'm avoiding, for example, the complex question of diagrammatic proofs-by-contradiction that he (1995) discusses with such care.

of informal rigorous proofs can be successfully transferred to their formalized cousins, and which can't – in particular if and when such formalizations replicate the phenomenological properties of the original proof practice.

There is an implicit (and sometimes explicit) assumption by those who formalize informal rigorous diagrammatic practices (e.g., Miller, Mumma, Shin, and others) that the informal reasoning practice is *captured* by these formalizations – that the formalizations reveal the reasoning implicit in the informal practice. Mumma (2008b, 256, footnote 1) writes, for example: "One should keep in mind that my criticisms [of Miller (2008)] are closely connected to my views on how my work and Miller's compare with respect to capturing Euclid's reasoning." At least insofar as the implicit reading-in and reading-off procedures are concerned, this is wrong: these informal procedures are invariably indeterminate in various respects that *cannot* be captured by any formalization – simply because formalizations decide interpretational issues that are indeterminate as far as the informal rigorous practice is concerned. Apart from this, both Millers' and Mumma's formalizations fail to replicate the informal practice as I'll indicate in what follows: specifically, these formalizations don't capture the case-structure of the original Euclidean geometry because they substitute different diagrammatic cases for Euclid's; more dramatically, they substitute a *mathematical basis* that diagrams are constructed from that's phenomenologically alien to our experience of the physical diagrams themselves. I'm not making a point about the interpretations of the Euclidean diagrams and those that Mumma and Miller give; I'm making a point about their metamathematics: how they characterize the *syntax* of the diagrams.[95] I now illustrate this.

Consider a preliminary observation about Mumma's approach. Mumma (2010) adopts Manders' (1995) distinction between exact and co-exact properties of diagrams. This (roughly) is a distinction between the metrical properties – the "exact" properties – of the figures in a diagram (the lengths of lines, the areas of triangles, the relative sizes of corresponding figures, etc.) and their topological properties, the "co-exact" properties (what regions are defined by figures, what elements are within and without figures, etc.).[96] Mumma claims – following Manders (1995) – that the original practice involved two proof systems, a diagrammatic one and a language-based one; both philosophers claim that only the co-exact properties of diagrams are used in the (strictly) diagrammatic Euclidean proof-system; the exact properties are either stipulated of the figures or proven in language-based proofs. (Mumma (2008a,

[95]Shin's approach, described in Sect. 8, is much closer to the original informal Venn-diagram practice, and retains much of the phenomenological qualities of the original informal diagrams – despite the fact that she introduces closed curves for circles (not *all* of her diagrams retain the relevant phenomenological qualities, but *some* do; I discuss this further in Sect. 12).

[96]See Manders (2008, 69) for a nice description of his exact/co-exact distinction. He adopts the term "appearance" for the set of co-exact conditions on a diagram. I'll suggest later in this section, at least as far as the experience of these diagrams is concerned, that this is misleading terminology. Even if the proof-machinery operates solely via the co-exact conditions of a Euclidean diagram, our *appreciation* of what it shows – including how much its results hold of diagrams different from the one being seen, turns on our perception of how it appears: its *geometry*.

italics his) writes: "The key observation is that Euclid's diagrams contribute to proofs *only* through their co-exact properties.") I've not determined whether this claim is true of the original proofs in Euclid – it certainly looks like it might be – but, in any case, the informal rigorous Euclidean practice isn't obviously divisible into two proof procedures in this way. Manders' observation about co-exactness, if right, is a substantial *discovery* about Euclidean proof-practice.[97]

Mumma (2010) also takes Eu to "solve" the generality problem (the question of what the diagrammatic elements in Euclidean proofs refer to), and criticizes Miller's earlier approach on the grounds that *its* solution to the generality problem increases the number of diagrams that any diagrammatic proof requires; he denies that the original practice could have operated in this way since Euclid never considers so many cases – and more importantly, couldn't do so (Mumma (2010), 263–264, footnote 4) because of their "staggering" number and because most of them "are not physically realizable."

The first point to make is to object to a mutually held assumption of Mumma and Miller. Formalizations of any informal rigorous mathematical practice always involve well-defined interpretations (models are given), and so they always "solve" the generality problem one way or another. But any such "solution" invariably deviates from the original practice because there is no solution (in the informal rigorous practice itself) to the generality problem *that original practice* faces: the scope of a diagram – the particular theorem it establishes – and, relatedly, the references of its elements are given by the reading-in and reading-out procedures, and (since the interpretations these give occur only tacitly in the practice, and even tacitly in the natural-language sentences that supply the interpretations) the informal practice is necessarily indeterminate in referential scope. I'll discuss this further in Sect. 13. I'll add now, however, that a similar generality problem occurs with respect to language-based informal rigorous mathematical proofs – when specialized terminology is involved. This is because specialized mathematical terminology can (and usually does) referentially *mutate* over time: consider the notorious case of the notion of a "function," during the course of several centuries.

The original Euclidean practice, of course, doesn't explicitly exhibit the neat distinctions *I* drew in the last section *either* – between diagrammatic transitions, on the one hand, and the reading-in and reading-off recognition procedures, on the other. This is hardly surprising since, in my description above, the reading-in and reading-off

[97]Although the insight is based on an observation commonly made through the centuries: Euclidean proofs avoid exact measurements (marked rulers) of the figures that are diagrammatically manipulated. One supporting consideration in favor of Mander's two-proof-procedure suggestion – diagrammatic and language-based – is that slash-marks (which are diagrammatic) seem to be a nineteenth-century innovation (Manders 1995, 97, footnote 14). On the other hand, another interpretational possibility is the conventionalist construal of the diagrams (see the definition of the "conventional option" shortly below) coupled with a meta-diagrammatic language in which *all* the proofs take place. According to this interpretation, there are no "diagrammatic proofs": rather, in the meta-diagrammatic language, one has the resources to "point to" aspects of diagrams – both exact and co-exact features. (Diagrams are, as it were, "quoted" in the metalanguage.) I entertain this interpretational hypothesis in Azzouni (2004); I claim that it wasn't available to ancient Greeks as a possible interpretation of their Euclidean practice – even if it was *true* of that practice – because of their implicit philosophical views about (correspondence) truth.)

procedures (which supply interpretations for the diagrams and their elements) are sharply distinguished from the syntactic obtainability relations among diagrams; by contrast, talk of diagrammatic elements and talk of the interpretations of those elements are, generally, not distinguished in Euclid's *Elements* – less charitably put, they're systematically *confused*. Consider the following definitions:

1. A **point** is that which has no part.
2. A **line** is breadthless length.
3. The extremities of a line are points.
4. A **straight line** is a line which lies evenly with the points on itself.
5. A **surface** is that which has length and breadth only.[98]

These (with the exception of (3)) purportedly describe the *mathematical* objects of study. Items which have no parts, or are breadthless, or have length and breadth only, can't be the actual diagrammatic items we draw and so they're *abstracta*. Call this *the Platonic option.* This option, that is, treats diagrammatic elements as referring to abstracta, and it treats the definitions as attempting to determinately *fix* the referents of these pictorial elements, drawn points, lines, etc. that appear in diagrams. For those who take diagrams to refer to Platonic entities, these definitions must fail as a result (worse, they're pointless): one cannot in this way provide interpretations; one can, at best, restrict the interpretations of primitive terms via axioms, which interlock those interpretations to one another in various ways.

This problem is quite general. The primitive terms – and now I use the phrase "primitive term" broadly to cover both the words in language-based sentences and the pictorial elements in diagrams – cannot be defined in this way, as we now know. Where referents are physically available, they can be indicated by language-exit rules (gestures and the like); when they aren't, these referents can only be fixed by axioms and inference rules (diagrammatic or otherwise) that constrain – but generally can't determine uniquely – possible interpretations. And this exhausts what can be done *within* a language. Within the confines of another language (a "metalanguage") a model can be given – this includes explicit interpretations of the primitive terms, which are given with the model. The *generality problem*, in the context of an informal rigorous mathematical practice, is the problem of how the primitive terms of that practice are given interpretations.

A second option for characterizing the role of these definitions takes them to be meant to describe diagrammatic items in a way that indicates *some* of the conventions governing the drawing of diagrams; and how, as a result, we're to understand such items. In this case the subsequent proofs involve claims – in natural languages – about the conventionalized diagrams themselves. I'll call this *the conventional option*.[99]

[98]Heath (1956, 153); his numbering, his boldface.

[99]Leibniz writes: "[Geometrical] figures must also be regarded as characters, for the circle described on paper is not a true circle and need not be; it is enough that we take it for a circle" (quoted by Manders (1995, 80). See Azzouni (2004) for further discussion of these two options and the evidence for why the Platonic option was eventually universally adopted despite the *Elements* retaining aspects of a conventional diagrammatic-practice, for example, in the statements of some postulates, as I indicate shortly.

On this view, diagrammatic elements are (exactly) like chess pieces; it's *understood*, that when a chess piece is moved, it's subsequently located in one and only one square even if it's been placed sloppily on the board (or even if the board is designed so badly that it can be hard to tell where a piece is). This involves understanding *physical* chess pieces conventionally – by imposing laws of movement on them via stipulations: these are laws of movement they'll otherwise violate. It *isn't* – notice – a matter of treating chess pieces as standing in for (or referring to) chess-piece abstracta that actually have the properties that chess pieces are conventionally treated as having (that's a version of the Platonic option – again).

In any case, the *Elements* both describes diagrammatic syntactic-transitions of diagrams in terms of what can be done to a diagram with respect to the above visual points, lines, etc. but runs this together with assumptions or claims about the interpretations of those visual items when they appear in diagrams, as in the postulates:

1. To draw a straight line from any point to any point
2. To produce a finite straight line continuously in a straight line
3. To describe a circle with any center and distance.[100]

Although naked infinitives are hard to interpret (and although I'm relying on Heath's *translation* when I say this and interpret the words as I do in the next few lines – assuming, that is, that what I say is true of the ancient Greek), the postulates (1) and (2) seem to describe permissions for possible actions that can be taken on diagrams; they don't seem to be characterizations of what the diagrammatic items, points and lines, refer to (i.e., to the mathematical objects: points and lines); they aren't, to put it another way, existence claims about what the elements of the diagrams depict. This "action-interpretation" of these postulates follows from the use of the phrases "to draw" and "to produce." On the other hand, (3), because of the word, "describe," seems, in addition to apparently describing an admissible action that may be taken to construct a diagram, to be offering an interpretation of what that diagram (of a drawn circle with a particular center and distance) is to stand for. According to the Platonic option, these definitions together with the postulates exhibit various use/mention confusions; according to the conventional option, there aren't such confusions (except insofar as interpretational issues are intruding in what otherwise would be a pure description of the syntactic conventions of a diagrammatic practice, and the nature of the syntactic idealizations it involves).

Recalibrating the informal Euclidean practice so that it accords with the device view (as it was described in Sect. 9) entails very little by way of changes to the look of the actual practice. One needs to, first, distinguish the diagrammatic/notational elements from the interpretations of those elements, second, to describe (in an

[100]Heath (1956, 154); his numbering, his spelling.

interpretation-free way) the syntactic-transitions embodied in that practice, and third, to indicate the interpretations of the diagrammatic/notational elements. That's not changing much.

I claim this is generally the case with the device view – in contrast to current formalizations of informal mathematical proofs, especially those of the Euclidean practice. I feel safe, thus, in regarding the device view as providing a theoretical *schematic* that enables *descriptions* of ordinary informal mathematical practices – diagrammatic and otherwise – despite its drawing a sharp distinction between actions on elements (of diagrams or languages) and how those elements are interpreted, a distinction which isn't exhibited directly in those practices themselves (as the discussion of the Euclidean case indicates). Reformulating an informal mathematical practice according to the device view should leave all the proof-theoretic flaws (if any) in the practice intact. For example, if (as many have illustrated) Euclidean diagrams can be ambiguous or misleading, the presence of these ambiguous or misleading diagrams should survive the practice being recast according to the device view. Second, it should also leave intact interpretational flaws – for example, failures to (fully) solve the generality problem. In this sense, if only because formalizations are consistent and determinate, they impose normative constraints on the informal practice that the practice itself needn't exhibit.[101]

The upshot: Attempts to formalize informal rigorous mathematical proof practices pose challenges that go quite beyond what's required to frame that practice correctly and according to the device view. In particular, with respect to diagrammatic reasoning, since (almost invariably) there is relevant reasoning occurring in the natural language apart from the syntactic transitions from diagram to diagram, that reasoning must be formalized as well. In addition, the reading-in and reading-out procedures must be treated both as supplying interpretations – so they are implicit guides (for example) of the model theory that is to be used to interpret the formal system – and as providing crucial inferential structure between the mathematical content expressed by ordinary-language sentences and that expressed by the diagrams licensed by the algorithmic devices.

[101] I'm not saying, of course, that an informal mathematical practice need be flawed in various ways; but whether it is or not is usually clear to practitioners at the time that the practice is living, and not because they're aware of one or another formalization that they're comparing the informal practice with. This is true, for example, of the ancient debates over how many diagrams should appear in a diagrammatic proof, and whether some diagrams are "ambiguous." This is similarly true of the debates over the acceptable manipulations of the notation for infinite series in the eighteenth and nineteenth centuries. It's important to realize that these debates are (often) resolved in the informal practice itself by notational reforms of various sorts along with changes in the rules for manipulating the syntax of an algorithmic device: these changes can be characterized in ways that don't require discussions of formalizations. (There is an exception, of course, once formalizations themselves become part of the mathematical toolkit among ordinary mathematicians; then a formalization may itself replace the previous informal rigorous mathematical practice – at least in some respects.)

Miller and Mumma formalize Euclid's *diagrammatic* practice in (quite distinct) formal diagrammatic systems. That means, to begin with (as I indicated earlier), that in both cases the physical diagrams themselves are identified with certain kinds of mathematical objects – certain kinds of two-dimensional abstracta – that can be (intuitive-effectively) recognized as having certain properties and not others. Right at this point both formalizations deviate in a significant way from the Euclidean (book I) practice; this – all by itself – faults claims that either resulting formal system "captures" the reasoning embodied in the original practice.

First, recall that the figures in the original Euclidean practice are constructed *only* by using straightedge and compass.[102] Actual physical drawings – I stress again – can be identified (preliminarily to representing an informal practice formally) with mathematical objects of any sort; the mathematical objects chosen will vary, of course, in how adequately they capture the original practice. But to identify physical drawings with any particular set of abstracta involves decisions of certain sorts that bear directly both on the question of how the figures in the drawings are being interpreted, and specifically, on how those figures are being interpreted to pick out certain items in the formal domain *and not others.* The geometrical properties of the actually drawn figures don't all by themselves dictate an answer to what abstracta they should be identified with. For example, we may choose to identify the physically drawn lines with various nonlinear curves of all sorts – not just straight lines – on the grounds that such physically drawn lines (especially when drawn badly, as they can be) *really do* approximate nonlinear curves of various sorts more closely than straight lines.[103] The same point can be made about the curves generated using a compass – they needn't be literally circular but only approximately circular, and we may take them to better approximate something noncircular. Miller's FA and Mumma Eu are formal systems based on an interpretation of Euclid's figures that identifies them syntactically with what at first glance appear to be quite broad classes of mathematical objects.

Miller, for example, defines the notion of a "nicely well-formed diagram" in terms of four primitive symbols (compare this to Shin (1994) – the strategy is the same one): frames, dots, unbroken lines, dash-lines, and slash-marks.[104] Any diagram

[102]In addition, points can be placed into the diagram – pretty much anywhere you want to drop them (subject, of course, to resolution considerations that prevent points from being too close to one another). This doesn't falsify the claim I just made about *figures.* In the Euclidean context, you *can't* drop a bunch of points in succession and use *them* to design a nonlinear/noncircular curve (that goes through all of them). The automatic existence of various (infinitely many) one-dimensional curves going through a set of points goes quite beyond the Euclidean system. Any one-dimensional curve that connects *any* two points *must result* from the operation of straightedge or compass.

[103]Miller (2007, 3) stresses right at the beginning of his study, that the "different kinds of line segments" that appear in Euclidean diagrams "do not have to really be straight."

[104]"Dash lines," are broken lines, for example, -----, as opposed to unbroken lines, for example, ——. I'm taking a few expository shortcut liberties in my description of Miller's approach, for example, "marked diagrams" are defined, not by the introduction of slash-marks as part of the primitives for diagrams, but as an extension of diagram equivalence to "marked diagrams" (see Miller 2007, 31).

occurs within a frame and it's composed of a finite number of dots, unbroken lines, and dash-lines – the unbroken lines represent Euclidean lines (although they needn't be straight) and the dash-lines represent arcs of circles (although they needn't be circular arcs). The definition of nicely well-formed diagrams includes a list of conditions that force unbroken lines and dotted lines to behave roughly (that is, topologically) like Euclidean lines and circles. (For example, one condition is that two "circles" intersect no more than twice.) To replicate Euclidean reasoning, these diagrams need these and additional characterizations. The result is that, again roughly, Miller's "nicely well-formed" diagrams can be associated with the original Euclidean diagrams. Miller then designs the actual proof-system in terms of topologically equivalent diagrams.[105]

Mumma's Eu, on the other hand, proceeds in a strikingly different way.[106] Instead of correlating Euclid's physical diagrams with a domain of rectangles containing arbitrarily nonlinear curves (except for obeying certain conditions, as mentioned above) and dashed curves (also, generally, noncircular, except for obeying certain conditions), etc., Mumma's domains are finite rectangles defined by square arrays of evenly spaced hollow ("○") and solid ("●") dots. The dots can be connected by lines, and some of those lines represent Euclidean lines; the finite lines, however, that are part of various convex polygons (together) represent circles. Like Miller, Mumma starts with an initial characterization of diagrams (call them "pre-diagrams") but then, like Miller does with *his* pre-diagrams, Mumma places an equivalence relation on them. *A* completion (nonunique) of a diagram D is one, D*, that's like D except that it has additional evenly spaced hollow and solid dots – enough so that every intersection point between the drawn lines and rays occurs as a dot. A definition of equivalent diagrams in terms of completions of them is then given by Mumma.

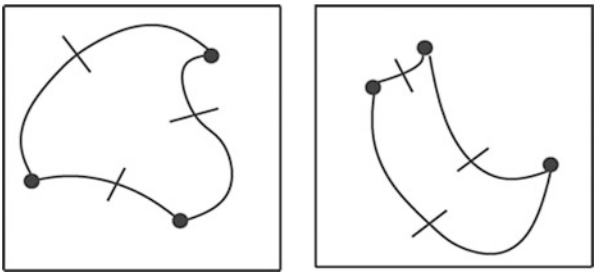
The result still deviates greatly from the informal rigorous diagrammatic practice – the experience of working in Mumma's system, specifically, deviates phenomenologically from Euclid's original diagrammatic practice. Why? For two reasons. The first is relatively insignificant: the diagrams we work with in Mumma's system involve visually recognized lines and polygons – not circles. Circles are conventionally represented by polygons. But the important difference should be visible, even given my rather sketchy description of the approach: the visibility of the array of hollow (and solid) dots that diagrams occur on. This transforms proofs in the Euclidean context – where diagrams are built upon one another by dropping points arbitrarily into what appears to be a manifold, and by drawing lines between points,

[105]This equivalence relation is technically involved – using a mapping between nicely well-formed diagrams and planar graphs. In terms of this mapping, two diagrams are defined as equivalent if they have the same corresponding graph structure. The different cases that must be considered in one of Miller's proofs then correspond to how many diagrams with differing graph structures become involved. It's worth pointing out, when comparing the reasoning in Miller's system with that of the original Euclid that the computer (which carries out the reasoning in Miller's system) works not with diagrams but instead with planar graphs. I make something of this shortly.

[106]I'm here following Mumma's (2019), where he lays out the modified system nicely.

and otherwise, by drawing lines arbitrarily and by drawing circles arbitrarily as well – into something experientially very different. The array-structure of the Euclidean context – in ordinary Euclidean proofs – as I said, isn't directly visible; it's required only by virtue of the incapacity of the practitioners to resolve arbitrarily small differences between drawn figures.

This suggests that Miller's identification of Euclidean figures with curves results in something phenomenologically much closer to the informal Euclidean approach than Mumma's is: it's one, recall, where figures are (apart from certain conditions) are arbitrarily nonlinear and noncircular – despite being "lines" and "circles." The manifold-flavor, that is, of the original diagrammatic proofs seems preserved. Nevertheless, Miller's approach isn't phenomenologically closer to Euclid's original practice than Mumma's reincarnation of it is. One reason (there are several) is because of something about the original Euclidean practice that's right on the surface. As Mumma (2008b, 259) points out in his review of Miller (2007), rather strange – from the Euclidean perspective – diagrams can appear in Millers' system. Mumma's example (his Fig. 3) – of "two ways to represent an equilateral triangle in **FG**" – follows:



Mumma (2008b, 262), further criticizing Millers' approach, writes (italics and boldface his):

> . . . when a case-heavy **FG** formalization is laid beside Euclid's original version, the original does not appear deficient. Rather, the multitude of cases generated by the rules of FG appear excessive. The geometric differences recorded by a case-branching often do not seem material to the issue the proof decides.

This criticism, however, largely misses the point. It primarily focuses on the sheer number of cases; but a formalization captures the phenomenology of an informal rigorous mathematical practice not by (even somewhat closely) matching the *number* of cases that appear informally; what's important is what *kind* of cases are involved: do the cases that are required by a formalization of Euclid's approach track the spirit of why the particular diagrams in the original practice come up? More important is this consideration: does the phenomenology of the original diagrammatic practice match (at least to some extent) the phenomenology of the formalized practice?

Mumma's criticism focuses on the fact that Millers' approach multiplies the number of cases that a Euclidean proof requires in a way that goes beyond the cases that the Euclidean approach seems to consider – even when that approach is supplemented with additional diagrams, as occurred later in the geometrical tradition. But the way that Millers' approach deviates from the Euclidean approach, although indicated by Mumma's Fig. 3, isn't stressed by him. This is that it treats the geometric information a diagram supplies as irrelevant to the proofs. It may be true that the geometric information a diagram supplies is irrelevant to the proof procedures that Euclidean geometry licenses but it's surely not irrelevant to how we come to *understand* these proofs because of how the diagrams *look.* Miller's approach is designed to facilitate a certain computer implementation, CDEG.[107] But the computer doesn't manipulate Miller's *diagrams*; it manipulates the *graphs* that the diagrams determine nonuniquely. Thus, the proof-practice Miller's formalization licenses amounts to this. A Euclidean diagram D is represented by a Miller diagram, M – more accurately, D is represented by a finite set of Miller diagrams. M, in turn, isn't itself manipulated according to a proof procedure; rather, M is used to generate a graph G, and it's G that's manipulated proof-theoretically. In turn, then, G is either transformed into programming code (which can be printed out), or into a canonical Miller diagram M*. This, it should be clear, is very far from the original Euclidean practice. In particular, Miller's approach cannot be used to explain what it is about Euclidean diagrams that makes the results they license so "intuitively" *comprehensible*.

In fact, neither Miller's nor Mumma's approach has even a hope of managing the replication of the understanding/insight that the Euclidean practice provides to its practitioners; this is because of the initial way both approaches mathematicize Euclidean diagrams. Return to the original practice for a moment. Three and only three sorts of primitive diagrammatic elements show up: points, lines, and arcs of circles. *That's it.* Diagrammatically speaking, there are no other entities that appear *in* diagrams. Of course it's true that what's drawn (physically) can unpleasantly deviate from what we (conventionally) take points, lines, and circles to *resemble.* But the expected (and allowable) response, if the intention is to formalize the practice in a way that captures the original reasoning, isn't to formalize the original proof-processes where badly drawn figures that deviate from what looks like circles and lines *become crucial to the cases that a proof must accommodate.* Rather, the response is: "that's drawn badly; I can't see what's going on; can you make a better one?"[108]

This means that, in a formalization of the *syntax* of a formal Euclidean diagrammatic system, the appropriate domain (the appropriate *alphabet*, as it were) should

[107] I'm here relying on Miller (2007, Sect. 3.5).

[108] See Manders (2008, 71) on how the original Euclidean practice tried to avoid "sensitive situations." This dovetails with what I've earlier called exclusion idealizations (defined in Sect. 2).

be one in which *only* circles, lines, rays, and points appear, and where these items are restricted in their sizes (and how they interact with one another) so that they are (by eye) surveyable.[109] The formalization approach I'm sketching treats deviations from circularity and linearity not as diagrammatic phenomena to be captured or replicated in the syntax of the resulting formalization but instead, as it was in the original practice, as items to be *excluded*.[110] The cost of not doing this – as I've indicated – is to "replicate" the informal practice by a formalization that doesn't respect the phenomenological qualities of the original practice – this matters because those phenomenological qualities are what underlie our experiences of understanding proofs, specifically the "aha"-epiphany. I now turn to an explicit discussion of this and its relation to the generality problem.

12 Phenomenologically Faithful Embeddings of Algorithmic Devices in Formal Algorithmic Systems

The phenomenological flavor of a proof is a somewhat elusive matter (and so it's dangerous to venture saying anything about it). Nevertheless, it's, importantly, what I think philosophers are tracking when they describe certain proofs as explanatory and others as unexplanatory; the same phenomena – the phenomenology of a proof (how the proof is experienced) – is also being tracked by mathematicians when they describe certain proofs (or, as commonly, certain subject areas – with their typical proofs) as beautiful or ugly. I'm not going to dwell on issues about explanation in mathematical proof – this is too tangled and complex to take up even in a chapter this big. But I do want to describe the phenomenological aspects of informal proofs enough to give a rough characterization of how – and to what extent – formalizations capture these qualities of informal proof practices.

[109]This characterization of the syntactic domain is, of course, subject to the persistence and exclusion idealizations given in Sects. 2 and 4, and later. Despite appearances (i.e., the phrase "by eye"), I'm describing a purely mathematical exclusion condition on circles, lines, rays, and points: that is, we translate an empirical fact about average-resolution capacities into numerical lower-limit conditions on admissible circles, lines, and rays.

[110]I don't know of, at the moment, any formalization of the original Euclidean practice that approaches it the way I'm suggesting here. I don't see any reason, however (offhand), to rule it out. I'm gambling in saying this, however, because an approach to formalization can always be ruled out by unexpected technical obstacles. Illustration: The original Venn diagrammatic practice used overlapping circles. In order to generalize the approach to finite collections of sets (recall Sect. 8, specifically footnote 61), Shin – relying on a result due to Polythress and Sun (1972) – replaces the circles of the original practice with closed curves. I should add that I don't claim that the resulting formal system I'm sketching an approach to won't be heterogeneous, rather than purely diagrammatic. It might be, and it might respect some descendant or other of Manders' exact/co-exact distinction. But the informal practice underdetermines exactly what belongs to the language-based part of the formal system and what belongs to the diagrammatic part – for example, exact-numerical identifications between figures and their parts might be indicated diagrammatically by slash-mark notation or by language-based characterizations accompanying the diagrams.

Recall the initial discussion of the "aha"-epiphany described by Feferman (2012) in Sect. 3. Recall, for that matter, the constraint on derivational accounts, *Phenomenology*, that such accounts must explain why formal transcriptions of informal rigorous mathematical proofs don't replicate that experience. The first point is that, although formalizations are often described as lacking the epistemic "aha, now I see why this is true" epiphany that's sometimes experienced with an informal proof, this simply isn't true of formalizations, *in general*: they needn't be so lacking. Other philosophers who contrast informal rigorous mathematical proof with formal derivation disagree with me.

Giaquinto (2008a, 27) writes, for example, "...there is a reason to avoid going formal: in a formalized version of a proof, the original intuitive line of thought is liable to be obscured by a multitude of minute steps." It should be clear from the foregoing that, even though it's a requirement on a formal system that its syntax be completely given, it's not a requirement that the syntactic transitions licensed by the formal system involve minutia.[111]

Rav (2007, 315, and elsewhere) suggests that formalizations *must* fall short in relation to their informal cousins because built into the appreciation of informal rigorous proofs and not their formalized cousins are the meanings of the terms used in the sentences of that sort of proof. He writes: "I hold that mathematicians' manner of reasoning and inferences are based on *meanings* and an informal notion of truth that a formal deduction calculus cannot capture" (italics his).[112] Rav (1999, 12) some years earlier puts his claim about the experiential distance between informal proofs and formalizations of them far more melodramatically (italics and spelling his):

> Once we have crossed the *Hilbert Bridge* into the land of meaningless symbols, we find ourselves on the shuffleboard of symbol manipulations, and as these symbols do not encode meanings, we cannot return *via* the Hilbert Bridge and restore meanings on the basis of a sequence of symbols representing formal derivations. After all, it is the very purpose of formalisation to squeeze out the sap of meanings in order not to blur focusing only on the

[111]This is an impression many have from a familiarity with formal language-based systems that involve, say, a lot of parentheses, or where the transitivity and symmetry of "&," and other connectives aren't built into the syntactic recognition procedures. All that's required of a formal system, however, is that its syntactic transitions be effectively recognizable. Thus, a formal system can be substantially modified by "shortcuts" and abbreviations," that make it far more user-friendly – as long as these are open to intuitively effective recognition procedures. I should add that in past work I've also contrasted formal and informal proof in just this way (mea culpa).

[112]A position in the same neighborhood is expressed by Larvor (2012, 723) when he writes: "The benefit of viewing inference as action is that we can see how the *subject-matter* of informal arguments shapes and contributes to inferences" (italics mine), and by De Toffoli and Giardino (2015, 332–333) who explicitly follow Larvor in this, writing: "In Rolfsen's proof, we saw that among the permissible actions on the pictures are continuous transformations. These are part of the background material in the sense that any topologist knows immediately that these transformations *can be interpreted* in terms of homeomorphisms. The validity is thus based on the 'practice' ..." (again, italics mine). Notice that De Toffoli and Giardino describe "continuous transformations" as permissible actions on *pictures*. This is a confusion.

> logico-structural properties of proofs. Meanings are now shifted to the metalanguage, as is well known.

This misleads how much informal phenomenology – specifically the experience of *meanings* – can occur in a formal setting. First point: meanings aren't, in any case, shifted to the metalanguage, as Rav writes – rather the meanings of the object language expressions are *given in terms of* the metalanguage (recall the discussion of this in Sect. 10). Second: In any case, our ability to psychologically invest formal terms with meanings should already be obvious from the centuries-long practice mathematicians have had with the routine introduction of new terminology into mathematical settings. Indeed, the point should be clear from the ubiquitous practice already in place with ordinary words – we (collectively) both modify the meanings of words and phrases already in public possession and routinely introduce new terms and phrases by indicating their meanings. What's striking, in addition, is that we can have the experience that a meaningful expression has been introduced (and that we grasp it) without being given a definition but just an illustration or two – typically all that's presented in dictionaries. Many of us continue to learn what we experience as new meaningful expressions all our lives.[113] Nothing prevents us doing the same thing to new terminology that's been introduced in a formal setting. (Indeed, we *regularly* do this; Rav's remark, quoted above about Hilbert's Bridge is especially surprising in light of the fact that, even when focusing on syntax, it's rather hard not to interpret certain logical expressions – notably "&" and "$\exists x$," among others – as having *fixed* meanings derived from natural language.)

At this juncture, it's natural to attempt a preliminary characterization of the phenomenological closeness of a formal system to an informal practice that it's taken to formalize. Consider an algorithmic device **A** (diagrammatic or language-based) that's employed in an informal rigorous mathematical practice. Describe that device as *embedded* in a formalization **F** if

(i) There is one-to-one mapping *M* of the elements of **A** to the syntactic elements of **F**. (*M* doesn't have to be "onto" – **F** can, and usually does, have additional elements.)

[113] I discuss our experience of meaning, in natural languages in Azzouni (2013b). Especially pertinent is a common experience that we have – with respect to certain expressions – of their being meaningful even though we don't know what their meanings are (Azzouni (2013b), 93). It's true that there are cases where we cannot experientially dislodge a meaning that we experience a word to have and replace it with another (e.g., "cold" for "hot") as I there discuss (Azzouni (2013b, Sect. 1.4). Nevertheless, we're quite good at introducing new words and, in specific contexts, like mathematical ones, of even setting aside the meanings that words apparently have – for example, stripping away the impression that "and" has a temporal dimension (e.g., "John brushed his teeth and got out of bed"), as we routinely do in mathematics.

(ii) Given any two elements D_1 and D_2, of **A**, given an intuitively effective-recognizable obtainability relation $\vdash_{\mathbf{A}}$, for that practice, and given that syntactic transitions $\vdash_{\mathbf{F}}$ for **F** are defined, then if $D_1 \vdash_{\mathbf{A}} D_2$, we have $M(D_1) \vdash_{\mathbf{F}} M(D_2)$.

So, given a formalization **F** and an algorithmic device **A**, call the mapping M that obeys (i) and (ii) an *embedding* of **A** into **F**. We say, further, that M is *phenomenologically faithful* to **A** if the experience of the inference, accompanying a syntactic transition $D_1 \vdash_{\mathbf{A}} D_2$, is relatively similar to the experience of the inference accompanying the syntactic transition $M(D_1) \vdash_{\mathbf{F}} M(D_2)$.

Why have I introduced the qualification "relatively similar" into the definition of "phenomenologically faithful"? Because (I take it) if we change the notation an inference is couched in – even minimally – we, as a result, change the experience at least somewhat: A = B, therefore, B = A is an experience of an inference that feels (slightly) different from $x = y$, therefore, $y = x$. There is, however, a bit more to say about the experience of inference – and the psychological mechanisms – that will give more content to the above characterization of "phenomenologically faithful." I'll turn to that in the next section.

I do need, at the outset, to stress two important aspects of phenomenological-faithful embeddings of algorithmic devices into formalizations. One is that the process of so embedding these devices invariably involves the inclusion of additional proof procedures (in the formalization) that don't possess the phenomenology of the original practice. This is the case no matter how carefully we choose the diagrammatic abstracta to match the original physical diagrams; and *this* is because the phenomenology itself is often fragmentary in its reach (our "aha"-epiphany, e.g., occurs only with fragmentary aspects of any informal proof-practice – let alone the extension of such a practice that results from formalization). A second point is that because we sometimes have free-ride experiences with respect to an informal rigorous proof, the actual algorithmic device involved in such a proof isn't itself – even informally – characterized. The algorithmic devices involved in such visual proofs remain tacit even in the informal mathematical practice. I'll round out this section by giving examples of both of these cases.

Start by recalling the original Venn diagrams, and comparing them to Shin's (1994) formalizations. Our "aha"-epiphany with the original diagrams – immediately seeing ("free riding") certain implications only occurs when a small number of circles are involved (about three). Furthermore, the experience is sensitive to circularity: change the figures involved into different (weirdly shaped) closed curves and the experience vanishes (one has to work out the implication visually – take a few moments to "see it"). That is, the immediate-recognition property doesn't work with other sorts of figures, as the reader can easily determine. As I'll discuss in the next section, this is because the particular "inference packages" we employ to recognize

the implications of Venn diagrams are (perceptually) content-sensitive: they work only with certain sorts of figures, and visual patterns, and not others. Thus, any formalization – especially an ambitious formalization (e.g., one, like Shin's, which is designed to capture, say, full monadic reasoning) – and even one that's otherwise diagrammatic, will replicate the phenomenological experience we have with the original diagrams only with a few of the formalized diagrams.[114]

I'll now illustrate the second kind of case with the diagrammatic proof of the "mutilated chessboard." Here's a description of it due to Black (1946, 157) that I've borrowed from Tanswell (2015, 305):

> An ordinary chess board has had two squares – one at each end of a diagonal – removed. There is on hand a supply of 31 dominos, each of which is large enough to cover exactly two adjacent squares of the board. Is it possible to lay the dominos on the mutilated chess board in such a manner as to cover it completely?

And here's the solution, quoted from Gardner (1988) – but I'm again taking the quote from Tanswell (2015, 305):

> It is impossible . . . and the proof is easy. The two diagonally opposite corners are the same color. Therefore their removal leaves a board with two more squares of one color than of the other. Each domino covers two squares of opposite color, since only opposite colors are adjacent. After you have covered 60 squares with 30 dominos, you are left with two uncovered squares of the same color. These two cannot be adjacent, therefore they cannot be covered by the last domino.

This proof and many other diagrammatic proofs[115] do *not* occur against an informal proof-theoretic background the way, for example, that the ancient Euclidean diagrammatic practice (and the Aristotelian syllogistic tradition) do. But it's not difficult to see what such an informal proof procedure would look like – it's not difficult to *create* one. An informal rigorous mathematical two-color-domino practice (on an ordinary chess board) would involve the placing of domino pieces on the board. The results – some immediately seen and others shown – would be various generalizations about what sorts of patterns of domino placements on the chess board are possible, and which aren't. Another family of results – some immediately seen and some not – is about which initial placements (of a few dominos) can be "completed," into a full covering of the board with dominos (with certain color patterns) and which can't. Lastly, one can consider results of domino-tiling various mutilated chess boards, as in Black's posed

[114]The problem is already visible in informal practices. That certain proof procedures can be characterized quite broadly using persistence idealizations runs well beyond when such proofs are convincing and understandable. (So my use of "intuitive," in Sect. 4 and elsewhere throughout this chapter, in, for example, "intuitively effective function" has to be understood as a somewhat specialized usage of the word – introduced because of Turing's own use of the word (see footnote 28).) Part of the phenomena here is described by Manders (1995, 107, and elsewhere) as "appearance control."

[115]for example, the proof in footnote 118

problem. In turn, these just-described informal practices can be embedded in one or another formal diagrammatic system that's phenomenologically faithful to these posited informal-proof practices. These can – Hilbert's thesis tells us this – be otherwise formalized in various language-based formalizations.[116]

13 Inference Packages (Embodied Algorithms) and Phenomenological Faithfulness

It's undeniable that the use of certain diagrams enormously facilitates our abilities to reason to and from various mathematical results these diagrams are taken to depict. Related to this is that diagrams also seem to enable us to understand the reasoning involved, and often to provide an understanding of the content of the results and assumptions (that the results follow from) that language-based formal transcriptions of that reasoning fail to provide. One approach to explaining this focuses on the apparent similarity between the elements of certain diagrams and what they're diagrams of – what I earlier called a "resemblance interpretation." If *A* pictures *B*, then we can use *A* to understand aspects of *B*. A photograph, for example, or a drawing, replicates certain aspects of the thing photographed or drawn, and by directly considering the photograph and the drawing, we can understand (or realize) something about the things themselves.

Despite the temptation to centralize visual similarity – resemblance – in an explanation of the just-described value of diagrammatic proof, pressing a resemblance interpretation looks bizarre almost on the surface; and it has led to serious mistakes among professional mathematicians when evaluating the value of diagrammatic proofs; these mistakes continue to affect contemporary philosophical discussions of diagrams. To start with, consider the case of Euclidean figures: points and lines. Given what these are supposed to be, it isn't even *sensible* to describe them as "resembling" physically drawn points and lines. Apart from the fact that no physically drawn figure lacks the dimensions that the mathematical objects, points and lines, officially lack, there is the related problem that the physical items have –

[116]Tanswell (2015) and Tanswell (forthcoming) press against derivation accounts of informal rigorous mathematical proof (Sect. 1) his "underdetermination" problem – that such accounts have the problem that attempts to capture the epistemic properties of informal proof via formalizations face the fact that there are equally good (equally close) formalizations that nevertheless involve drastically different mathematical ideology. Although derivational accounts (Sects. 1 and 3) have been rejected here, Tanswell (forthcoming) frames the underdetermination problem in a way that threatens a different claim that I've made in this chapter – that in general, phenomenologically faithful formalizations for informal rigorous mathematical practices exist. He writes of the mutilated chess board proof, in particular, that it "is not obviously algorithmically checkable, so if it is claimed to be then we would like to see the algorithm that does so given explicitly . . .." His suggestion is that if it isn't, then any embedding of the proof in an algorithmic system will face underdetermination. The remark that a system of recognizing patterns of two-color dominos on what amounts to a finite array isn't "obviously algorithmic" is puzzling: it looks exactly like other cases of intuitively effective procedures that – via Church's thesis – are Turing computable.

geometrically speaking – *the wrong* "*shapes*." Notice the point. This isn't a criticism of one or another theory of "abstraction" as a theory of understanding – that we *understand* what a line is, say, by abstracting a certain approximate aspect of drawn lines: pure "straight-line directionality," or whatever. It isn't even to criticize a view of understanding that claims we understand mathematical straight lines via "approximations" of them by visually thin lines. Some theory or other like this about how we understand points and lines might be true. The point, rather, is the resulting objects – *abstracta* – even if we can *understand* what they are supposed to be, aren't *seeable.* Despite whatever value a theory of abstraction has as a theory of understanding, it has none as an explanation of the value of diagrams of lines and points taken as "visually resembling" actual points and lines.

Similar complaints can be made about the diagrams that appear in other diagrammatic practices. Shin's diagrams don't visually resemble (in any sense) classes or their relationships to one another: two circles that overlap don't *visually resemble* the relationships two sets have when they have elements in common – this isn't even true if the overlapping region is shaded. Finally, recall (footnote 112, Sect. 12) that De Toffoli and Giardino (2015, 332–333) treat pictures and what's pictured as so similar that they feel they can describe diagrammatic actions on pictures as continuous transformations.

Mistaken resemblance interpretations are very old ones, and their implicit operation is one of the central factors behind the late nineteenth-century repudiation of visual methods of proof. Unaware of the diagrammatic constraints on Euclidean diagrams that were necessary to making Euclidean diagrammatic proof coherent, pictures of functions – printed drawings of curves in books and articles (let's be accurate about what the diagrams in question *were*) were treated as depicting functions, *of any sort*, by sheer visual resemblance. More accurately, drawings of functions – ink on paper – were treated as yielding "intuitions." Intuitions, so called, can only be understood in this case as *visual imaginings* of the curves in question – the scope of those visual imaginings supposedly telling us what sorts of curves are and aren't possible. The implicit idea was this: diagrams elicit our visual faculties. But when we so-visualize functions, we can get it very wrong. This, however, is the wrong picture of diagrams of functions (or the diagrams of any mathematical object, really). We cannot "visualize" functions any more than we can "visualize" lines – and for the same reason. Functions – if characterized geometrically – are one-dimensional[117]: that means we *can't* and *don't* visualize *them*.

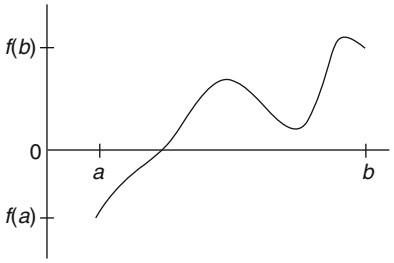
Giaquinto (2008a, 35) channels this nineteenth-century mistake about the role of visualization in mathematics. He writes:

> When it comes to analysis mathematicians of the nineteenth and twentieth centuries were right to hold that visual thinking rarely delivers knowledge Visualizing cannot reveal what happens in the tail of an infinite process. So visual thinking is unreliable for situations in which limits are involved, and so it is not a means of discovery in those situations, let alone a means of proof.

[117]Even "space-filling" curves are, nevertheless, one-dimensional!

The letter of this remark is correct: We cannot design a diagrammatic proof procedure that will handle the tails of infinite processes.[118]

But this remark occurs nested (as was also true of nineteenth- and twentieth-century objections to diagrammatic proofs by mathematicians) among misapprehensions about (informal) diagrammatic proof. Consider this well-known informal diagrammatic proof of the intermediate zero value theorem:



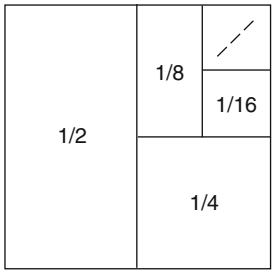
Giaquinto (2008a, 35) criticizes Brown (1999) for suggesting that "reflection on the diagram [above] suffices to prove the Intermediate Zero Theorem." Giaquinto (2008a, 36) notes there are many functions visually indistinguishable from a continuous smooth function defined on the reals, as above, of which this theorem is false.

But speaking this way of "visual thinking" (it should be clear) overlooks that an informal rigorous diagrammatic practice that treats diagrams of functions, as above, as referring to (and only referring to) smooth continuous curves, *does show* the intermediate zero theorem for that class of curves.[119] Diagrammatic practices must be curtailed sharply in what mathematical items are to be treated as the referents of their elements, and thus, what generalizations they can be taken to have shown. This

[118]Actually, no. Giaquinto's remark, about the inability of diagrammatic reasoning to capture "the tails of infinite processes," is *false*, as the following proof that ½ + ¼ + . . . = 1, indicates:

Proof:

Proof:



What's true is that no diagrammatic proof-procedure can be designed that adequately handles all the tails of infinite processes – more accurately, that handles all possible functions – as "function" is understood in the contemporary mathematical setting.

[119]The informal proof, of course, can be successfully extended to smooth curves with a finite number of discontinuities, and other classes of functions. I omit further discussion of this.

can't be done on the basis of vague impressions of "resemblances" between diagrams and mathematical abstracta.

The same point can be made against Giaquinto (2008a, 31) when he writes:

> Missing an untypical case is a common hazard in attempts at visual proving. A well-known example is the proof of Euler's formula V – E + F = 2 for polyhedra by 'removing triangles' of a triangulated planar projection of a polyhedron. One is easily convinced by the thinking, but only because the polyhedra we normally think of are convex, while the exceptions are not convex. But it is also easy to miss to miss a case which is not untypical or extreme when thinking visually. An example is Cauchy's attempted proof (Cauchy 1813) of the claim that if a convex polygon is transformed into another polygon keeping all but one of the sides constant, then if some or all of the internal angles at the vertices increase, the remaining side increases. . .. The frequency of such mistakes indicates that visual arguments often lack the transparency required for proof; even when a visual argument is in fact sound, its soundness may not be clear, in which case the argument is better thought of as a way of discovering rather than proving the truth of the conclusion.

This, as earlier, is to use the generality problem – which, recall, is a problem *all* proof-procedures face, diagrammatic *and* language based – against a diagrammatic proof procedure which, otherwise, can be syntactically characterized coherently. There is nothing wrong with taking the above diagrammatic proof-procedure to show Euler's result for *convex* polygons. What follows from the frequency of such mistakes isn't that "visual arguments often lack the transparency required for proof" – in particular, that it isn't obvious that they're sound. Soundness turns crucially on interpretation – and what this shows is that interpretations must be chosen with care; this is especially the case when an informal practice is being embedded in a formal practice. One thing these failures *do* show is that diagrammatic proofs often have restricted domains of soundness that are narrower than the range of abstracta that mathematicians are concerned to show results about.[120] But there is nothing obvious, about diagrammatic proofs, that shows this will always be a problem, when comparing such proofs to language-based proof procedures.

My point is that Giaquinto's analysis presupposes (something nineteenth-century mathematicians presupposed as well) resemblance interpretations at work: visual proofs represent functions by drawings that resemble these functions. But this is wrong, and as a result the wrong interpretations are imposed on diagrammatic practices – sometimes (as with the case of functions), the drawings of curves on a page are being confused with the functions they're being taken to depict, with the result that visualization-proof methods are saddled with the expectation that they're responsible for depicting functions for which no syntactic diagrammatic system has ever been characterized, even informally. At root, as I've mentioned, are confusions between the diagrams themselves (in this case, drawings on paper) and functions that can't be seen. The use of the word "intuition," as used in "it's intuitive that a function can't be everywhere continuous but not differentiable anywhere" enables a

[120]This is especially the case with the notion of a function, which from the eighteenth century to the nineteenth century, drastically changed in what it was (implicitly and explicitly) taken to refer to.

confusion between pictorial methods of proof and the items being studied. Lesson: Vague intimations of visual resemblance don't show that a cogent diagrammatic practice has been established; worse, such vague intimations don't sensibly provide interpretations when a diagrammatic practice is cogent.

I've complained about one error that the visual-resemblance views make: confusing a method of *understanding* abstracta with a method of *visualizing* abstracta. This is a failure to realize that a visual process can be used to *understand* something without it supplying a *visualization* of that something.[121] But there is a second way that focusing on visual resemblances – as the machinery that governs the implicit interpretations for diagrammatic elements – goes wrong. This is that there is an implicit *suggestion falsi* that a diagram works its magic by capturing structural resemblances between pictorial elements and the abstracta they depict.[122]

This second important mistake goes far beyond diagrammatic proofs; it infects views of (scientific) modeling generally. Although there is a narrow class of models that can be described as ones that capture aspects of what they model purely by virtue of the models embodying structural resemblances to the items modeled, this is generally not the case.

Consider a computer simulation of a certain dynamical process – the flow of water through a pipe, for example. It would be a mistake to think that we can rely on the simulation to predict certain aspects of that process – when we can – because what happens "in" the computer "resembles" the flow of water through a pipe. That's, frankly, pretty absurd – although this absurd suggestion is made natural by the misleading label: simulation. What happens with a successful computer simulation is quite different: the computer simulation has been designed with certain aspects of water flow in mind – more specifically, it has been designed by taking into consideration certain events that occur in (or during) the flow of water through a pipe, and by ignoring many other ones. Water flow, in particular, is an extremely complex (and badly understood) process where a great deal is going on. It would be a naïve analysis to divide – as it were – water flow into levels: the quantum level, a higher level (where the water flow is approximated by continuum-fluid idealizations), etc. What's actually going on is that certain effects in and to the water are being taken account of and other ones are being ignored, where these effects cross

[121] See Azzouni (2020) for further discussion of this distinction between imagining and conceiving.

[122] See Giaquinto (2008b). It's perhaps no surprise that Giaquinto takes the cognition of abstracta-structure seriously as an explanation of the value of diagrammatic proof: it's a natural move for someone who, to begin with, thinks that resemblance to abstracta is how diagrammatic elements refer to those abstracta. See Resnik (1997), for an earlier book-length argument for this kind of epistemic position vis-à-vis abstracta. Notice that a nominalist must reject this picture of the understanding of proof in any case: I'm not presupposing my nominalism, however, in raising any of the objections of this paper.

"levels" in intricate and subtle ways.[123] The hope, sometimes dashed and sometimes fulfilled, is that the aspects of a phenomenon (water flow) that are being taken account of – by approximations, usually – are ones that will enable the model to capture certain causal effects of that phenomenon (e.g., in the case of water flow the average pressure on the inner pipe surface) while ignoring other ones (e.g., in the case of water flow, changes in temperature of the inner pipe surface). As I said, it would be naïve to impose a structure-picture on such models and what they model – although this is, imaginatively, extremely natural to do. Unfortunately for what's psychologically natural here, the actual interaction (causal and otherwise) between what's being ignored by a model and what's being taken account of is too complicated for that. The process can be *imagined* – misleadingly, to repeat – as one in which the water, in the case of water flow is, at a certain "level" being structurally modeled by the computer simulation. But this, it must be stressed again, isn't accurate to what's being retained and what's being ignored by the model – for example, what's being ignored aren't simply the quantum-mechanical, discrete, elements of the water-flow phenomenon.

Instead, and this alternative characterization of the computer simulation directly applies to diagrammatic proof as well, there is a body of *statements* – "predictions," in the case of water flowing through a pipe. And the model is a piece of computer machinery (that yields statements of the same form as the predictions) that we hope will replicate those predictions that are *true*.[124]

Summary: presumed visual resemblances between diagrammatic elements and the abstracta they supposedly resemble fool mathematicians and philosophers into thinking, in the case of drawing curves, that these resemblances are the proof-theoretic and interpretational machinery of diagrammatic proof – that is, that visual resemblances supply both intuitively effective recognition procedures and visually determined (by resemblance) interpretations for diagrammatic elements. These are mistakes.

14 Inference Packages and Shin on Perceptual Inference

Sun-Joo Shin (1994, Chap. 6) offers a more sophisticated theory of diagrammatic reasoning than resemblance interpretations do. Central to her view are the roles of *perceptual inference* and *convention* which she uses to explain in what ways, and to what extent, any specific formal system is diagrammatic and in what ways and to what extent it isn't. The difference between the use of perceptual inference and the use of convention (when looking at something), on her view, is that conventions

[123]Consider the fact that gravity is relatively weak if you're small enough. Your biological mechanisms for handling your recognition of "up" and "down" can't turn on gravity as it can for bigger things. Similarly, in engineering design, certain effects, wind, torque, etc. can be ignored when the items (bridges, shoulders) are certain sizes, and not otherwise. How, and to what degree, causal effects can be ignored or minimized by a model is badly captured by thinking in terms of "levels."

[124]Exactly how inferential practices capture the supposed properties of mathematical entities will be described in terms of our capacity to "simulate algorithms" – where I'm using "simulate" in the same misleading way it's used above. See Sect. 14.

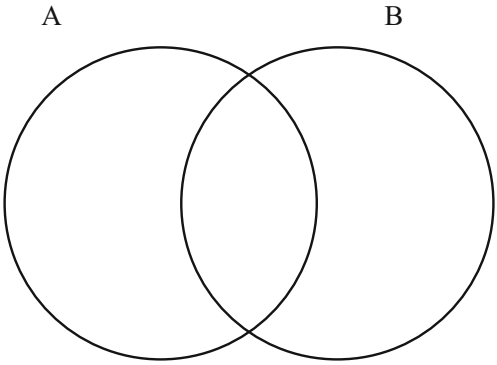
need to be learnt; on the other hand, when we use perceptual inference, we exploit aspects of a diagram that we don't have to learn how to use. Importantly, she thinks, the respective amounts (as it were) of conventionality and perceptual inference involved in our ability to navigate linguistic texts, diagrams, and pictures comes in degrees. A photograph has no conventional elements, nearly enough: we use and only use perceptual inference to recognize what visual elements in a picture refer to (and, therefore, how a photograph depicts things to be). A pure language-based piece of text, using symbols, is entirely conventional. Diagrams are in-between – usually involving both conventional elements and perceptual inferences.

The examples Shin explicitly gives, and others that we can easily apply her distinction to, make that distinction tolerably clear. A depiction of the seating arrangements of Tom, Susan, and Mary like so:

Tom Susan Mary,

involves convention insofar as "Tom" refers to Tom, "Susan" to Susan, and "Mary" to Mary; it involves perceptual inference insofar as the side-by-side seating arrangement of Tom, Susan, and Mary is recognized to be as depicted above.

Similarly, in Venn-I, it's a convention that the closed curves and rectangles represent sets, and – specifically – that their interiors represent the interiors of those sets.[125] It's equally conventional, in a language-based setting, to introduce capital letters, A, B, . . . for names of sets, lower-case letters, a, b, . . ., for names of elements in sets, and to introduce a symbol, $\in$, for membership. Shin takes the distinctive roles for perceptual inference and for convention in Venn-I, as opposed to a language-based class system to contrastively emerge when we consider relations between sets. In the language-based case, we need to introduce a *new* syntactic device to conventionally represent intersection: $\cap$, so that "A $\cap$ B" thus can represent the intersection of A and B. In the case of Venn diagrams, we don't need to do this. Consider, for example, the following figure:



Shin claims that, given the previous conventions, the interpretation of the common spatial area of the circle associated with A and the one associated with B *must be* the

[125] Shin (1994, 169) writes: "In Venn-I we adopt a convention that assigns a set to a basic region."

intersection of the sets A and B. This isn't, therefore, the introduction of a new syntactic convention; instead, we use perceptual inference, along with the earlier-mentioned conventions, to see what spatial area in the diagram represents the intersection of the sets.

Shin (1994, 161–162) similarly argues that if we already have in place a convention that (two-dimensional) boxes represent sets and dots represent individuals, then to represent the set-theoretic membership relation, one need only place a dot in a box. She adds, "Putting down an object *in* another object does not involve any new syntactic device. A spatial arrangement between a dot and a box represents a nonspatial relation, 'being a member of'" (italics hers).[126]

There's an important subtlety that Shin's remarks – as quoted – overlook. If we interpret dots and rectangles (and closed curves) – the lines of the figures, that is – as Shin suggests, then we need an additional convention that dictates that the *interiors* of squares and circles represent where the members of the corresponding sets are located. Even if we represent *area-filling* rectangles (as opposed to the borders alone) as sets, it's still open to us to conventionally dictate an item as a member of a set only if it appears *outside* the interior of the figure rather than inside.[127] Putting a figural element inside another figural element *is* a piece of (two-dimensional) syntax that needs to be interpreted, and the interpretation of which goes beyond the sheer convention that a two-dimensional figure (area-filling or otherwise) represents a set. We can see the point I'm making if we consider the typical language-based notation, Pa, in which a piece of syntax (concatenation) is interpreted as predication. The situation is exactly the same one.

Nevertheless, I think Shin's approach to diagrammatic reasoning is largely correct, insofar as she recognizes the role of what she calls perceptual inference. I now introduce some alternative terminology and make a few other changes to her approach. Perceptual inference involves a number of *cognitively embodied algorithms,* what I call *inference packages*.[128] Inference packages are modularized and content-specific.[129]

---

[126]She continues:

> An isomorphism between this spatial arrangement and the relation "being a member of" is not as perceptually obvious as an isomorphism between a spatial arrangement in a diagram [as in the Tom, Sally, Mary example above] and the relation "being to the right of." However, this isomorphism is more perceptually obvious than any linguistic symbol that a linguistic representation adopts, since no extra convention involving syntactic devices is required.

In adapting elements of her position, I'll remove talk of isomorphisms.

[127]For manageability, the shadings and dots, etc. to be placed outside squares could be conventionally quite close to the outsides of the figures.

[128]I borrow, in part, this discussion of "inference packages," and other aspects of visualization, from Azzouni (2005).

[129]Exactly how inference packages are modularized is a subtle empirical matter that I have to avoid saying much more about. There is more than one visual mechanism at work in visual perception – the evidence for this is that we naturally recognize some visual patterns and not others, and that these recognitional abilities differ greatly between individuals. Exactly what neurophysiological resources are at work, and how they operate in tandem is undergoing (currently) intense study. I *don't* assume one-to-one relationships between inference packages and visual-perception capacities.

A nice illustration of one, however, is our capacity to imagine triangles on a plane or a sphere. We recognize that if we move the triangles rigidly about on a plane that doesn't change their areas or change the magnitudes of their interior angles. If we imagine an equiangular triangle growing in area in a way that preserves its equiangularity, the sum of its angles won't change. This isn't true, we immediately recognize, if we instead imagine a triangle on a sphere. In this case, moving the triangle about rigidly won't change the magnitude of its interior angles, or the areas of those triangles. But if we grow an equiangular triangle in a way that (again) respects its equiangularity, the magnitudes of those angles change: It gets "more curved" and its angles gain magnitude. One can also imagine blowing up the sphere – like a balloon – uniformly, although keeping the triangle the same in its area; in this case its angles flatten and lose magnitude. Triangles moving around on an ellipsoid act differently still: if we imagine moving a triangle from the flattest part of the ellipsoid to one of its ends, we can see how the magnitudes of the interior angles of the triangle will increase and decrease.

Involved here are various "perceptual inferences": we can see some facts about the angles, lines and areas of the triangles immediately (they're "free rides"); others we see quickly, or we can work them out with a step or two. I call these inference *packages* because, from an axiomatic point of view, a number of different relations between the properties of these figures (that can be distinguished axiomatically) are bundled together. I stress again that these packages are content-specific: the inferences govern exactly what we imaginatively experience them as governing: certain sorts of visual phenomena. I hypothesize that there is more than one inference package that we can apply to two-dimensional visualizations; this is a matter to be determined empirically. Included among them, however, are the part/whole inferences about spatial regions that we see at a glance: for example, when a spatial area is part of another area, and when they're disjoint. It's important to realize that we have inference packages with respect to one-dimensional figures as well as with respect to three-dimensional figures, and that these aren't the same. (Some of them, at best, overlap.) We effortlessly recognize a number of relations alphabetic items have when in concatenated strings. Certain patterns "leap out" at us; others don't. The same is true of the inference packages with respect to three-dimensional objects located in space. These differ from two-dimensional ones simply because we can imagine two-dimensional objects as *on* surfaces, and as inheriting their properties from those surfaces. Three-dimensional objects are "in" space, and not experienced as "on" anything. This affects what kinds of inference packages are available with such space-filling figures.

Diagrammatic conventions can be chosen in ways that run in parallel with our inference packages or in ways that work against them – even though the diagrammatic conventions in both cases can be equally coherent. Imagine, for example, a systematic diagrammatic Venn system, involving rectangles, closed curves, diagrammatic elements, ⊗, —, and shadings, such as Shin's, but where the

representation conventions for regions are inverted: for example, to indicate that a region R has no members, we must shade all the regions *outside* R.[130] This, coupled with the appropriate conventions for the other elements, yields a perfectly consistent system; it's – logically speaking – impeccable. It will, however, be a very *user-unfriendly* system because it works against our inference packages which operate naturally via spatial wholes and *their* spatial parts. This isn't to say we couldn't get used to such an alternative diagrammatic practice – and that some of us would even find ourselves as able to work with it with some comfort; it's to say there would be a learning curve.[131]

In describing this imagined alternative to Shin's Venn-I as user-unfriendly, I'm making a phenomenological point: to use such a diagrammatic system for proofs feels differently from so-using Venn-I (at least with respect to those diagrams in Venn-I that phenomenologically resemble the diagrams in the original Venn practice). When we engage in a proof-procedure that, relatively directly, uses our cognitively embodied inference packages *we feel we understand what's going on*: we *get* it. That's true with all cognitively embodied inference packages: the inferentially packaged reasoning we engage in is one we always feel we understand: we feel we recognize why it must be right.

A simple phenomenological rule for proof procedures – formal and otherwise: The more closely we can use our cognitive-embodied inference patterns to simulate the processes involved in carrying out a proof, the more we experience that proof as one we understand or "get." To illustrate, consider sets. Sets, recall, are supposed to be abstract objects that have members. We can imagine a set as a kind of box in which its members are contained – but that's a spatial metaphor that only goes so far. Sets aren't in space and, in particular, they aren't located where their members are. The set containing the six umbrellas, the one that I have in my possession now, and the five that I lost last week, isn't located at (or where) the umbrellas it contains are. There is a sense in which "sets," therefore, make no sense. What sorts of things can these be? But the reasoning about sets can be imitated – up to a point – by reasoning instead about objects in boxes. More pertinently to Venn diagrams, reasoning with sets can be imitated – up to a point – by reasoning about regions and their interiors. And so, reasoning about sets can be executed by a notational device that exploits our ease with reasoning about regions and their interiors.

To stress again: inference packages (cognitively embodied algorithms) always come with content; they are always interpreted – and indeed, the interpretations are

[130]Again, with a convention, as described in footnote 127, to make markings outside squares manageable.

[131]Consider Polish notation. Ordinary alphabetic formal systems, using parentheses and the like, operate in parallel with our one-dimensional containment inference packages. The result is that we can read that notation easily, and learn it quickly. Polish notation flouts these inference packages – and so is "intuitively unappealing."

always *sensory*.[132] But when mathematicians use these inference packages they invariably *reinterpret* them. I've just illustrated one way that this is done: by using them to manipulate diagrammatic notation that our cognitively embodied algorithms are at ease with, although those diagrammatic elements (officially) refer to something else.

Let me return briefly to the question, raised in Sect. 8, about the distinction between when notation is diagrammatic and when it's not. The foregoing implies, I think, the following result. There are distinctive visual capacities that enable distinctive inference packages with respect to one-, two-, and three-dimensional visual experience. When a two-dimensional notation doesn't take advantage of the distinctive inference packages available with two-dimensional notation (but not one-dimensional notation), but only relies on, say, adjacency, that notation isn't diagrammatic. Thus, I suggest, matrices and integral notation aren't diagrammatic notation. When notations rely on distinctively two-dimensional inference packages, as Euclidean and Venn notation do, they are diagrammatic. This distinction requires an empirical underpinning, which I can hope will be supplied in the future; it doesn't turn – notice – on semantic considerations, that is, it doesn't turn on interpretation. (I think this is an advantage of it.)

15 Ideology and Formal Systems

Before (finally) concluding this chapter, I'd like to revisit the interrelationships between Church's thesis, Hilbert's thesis, formal systems, and informal rigorous mathematics, while at the same time discussing some recent work by Avigad and Hamami that defends a particular version of the derivational account.

Twentieth- and twenty-first-century mathematics is extremely rich – and in large part a lot of that richness is a result of mathematical developments subsequent (and in contrast) to the logic that emerged with Frege and was joined with set theory by Russell and Whitehead, and others. In particular, major parts of contemporary mathematics are the alternative axiom systems that have emerged *in contrast to* Zermelo-Frankel set theory with the axiom of choice (ZFC) that's couched in the language of classical first-order predicate logic. Intuitionism is well-known: it has generated a great deal of *informal rigorous* mathematics – in (intuitionistic) analysis, in particular. But, of course, there are also numerous "alternative logics" and some of these alternatives – paraconsistent approaches, among others – are actively pursued not just as logical disciplines but in terms of the mathematics they also enable. There are also explicit alternatives to ZFC, including ZF + the axiom of indeterminacy. Again, these have led to additional pursuits in bread-and-butter mathematics, for example, alternatives to standard measure theory.

[132] Although: they're not always *visual*.

There is no choice, I would have thought, but to be a *pluralist* as far as contemporary pure mathematics is concerned. It's simply a fact of the contemporary mathematics that these numerous alternatives – alternative to classical logical frameworks, alternatives to standard Zermelo-Frankel set theory with the axiom of choice – exist. Call the formalized pursuit of these alternatives *axiomatic pluralism*.[133]

Much – if not most – of this mathematics *isn't* restricted to the deduction of theorems within formal systems, as I've already indicated; rather, it's bread-and-butter informal rigorous mathematics. That's the context, actually, in which such axiomatic systems are studied. More obviously, it's in the context of informal rigorous mathematics that the knock-off studies I described above, in one or another nonstandard framework, a particular logic and/or set-theoretic framework, or something else (e.g., category theory) occur. Understanding informal rigorous mathematical proof as a matter of mechanical recognizability – either according to the algorithmic-device view or some other – enables an accommodation of this pluralism, which I'll call *informal mathematical pluralism.* This also accommodates the datum that mathematicians can recognize good proofs across mathematical traditions.

Informal mathematical pluralism is why, however, that the derivational accounts I've been concerned with in this chapter, and in earlier work, have been *ideology-neutral* accounts. The formalizations in question that are supposed to explain the epistemic properties of informal rigorous mathematics can't be restricted to frameworks governed by standard classical logic and ZFC. Indeed, foundational projects for mathematics – ones dedicated to transcribing all mathematics into one or another *specified* formal system (e.g., ZFC, Russell-Whitehead type theory), haven't been in focus at all.[134] A formalization of arithmetic with an ideology of *numbers* (an intended model of numbers) suffices to provide the formalization needed for an informal rigorous *arithmetical* practice – on the derivational views I've been concerned with. Church's thesis (going beyond functions of natural numbers via reinterpretations of numbers) accommodates axiomatic pluralism. The version of Hilbert's thesis I borrowed from Kripke (2013) in Sect. 4, however, requires changes to enable its application beyond the language of first-order predicate calculus with identity. Notice: Although any Turing machine formalism can be *represented* in the language of first-order predicate calculus with identity – that's a mathematical result about Turing machines – this result doesn't straightaway yield a first-order-predicate-calculus-with-identity version of Hilbert's thesis.

These preliminaries aside, recent work by Hamami (2019) defends what he describes as "orthodoxy," or "the standard view," and he traces it (historically) back to Mac Lane and Bourbaki. This is that "a mathematical proof P is *rigorous* if and only if *P* can be *routinely translated* into a formal proof," where such a proof

[133]I committed myself to axiomatic pluralism (without using that label) in Azzouni (1994, Part II).

[134]One may well doubt, in light of the richness of contemporary mathematics, whether any straightforward ideologically nonneutral foundational project is possible. Such can't be ruled out, of course, because there can always be rather unexpected approaches – based on relative consistency proofs – to embedding branches of mathematics in one another, and therefore grounding such in specified axiomatic systems.

"is (or could be) written out in the first order predicate language L($\in$) as a sequence of inferences from the axioms ZFC, each inference made according to one of the stated rules."[135] As I've just indicated, I think this version of a derivational account cannot accommodate contemporary mathematics because it isn't ideologically neutral. Let us, accordingly, broaden it. Further, let us distinguish (as Hamami (2019) and Avigad (2019) both do) an epistemic/descriptive version of the derivational account from a normative version of that account. The epistemic/descriptive account is required to use the existence of formalizations to explain the epistemic properties of informal rigorous mathematical proof. The normative version instead defines a mathematical proof as "rigorous" if it can be routinely translated to a formal proof. No explanation of the epistemic qualities of informal mathematical proof is required.[136]

Consider the descriptive version of the derivational approach, and recall Tanswell's condition *Content* (Sect. 3). The task *Content* poses to descriptive derivational approaches, pretty much, is to require them to show how the materials, *already present in informal proofs*, indicate their formal transcriptions. I attempted, in Azzouni (2004) and Azzouni (2006), to defend this approach by giving an open-ended set of informal-mathematical tools – inferential abbreviations and shortcuts, meta-proof considerations, etc. – many of which mathematicians have learned, that I hoped would provide sufficient content to informal proofs to show how they indicate the relevant derivations.[137] It strikes me that, insofar as Hamami (2019) considers the descriptive strategy, this is what he offers as available for this – although he stresses more strongly than I did the idea that mathematicians come to the task with certain background competences.

Here is Avigad's (2019) excellent (and approving) description of Hamami's account of informal proof:

> According to Hamami, we accumulate not just theorems but inferential patterns and procedures throughout our mathematical education, and we put them to good use towards understanding an informal proof. When a competent practitioner reads such a proof, the inferential cues and context trigger algorithmic processes that expand the inferential steps to smaller ones. These, in turn, are recognized as applications of prior theorems or lemmas, or at least patterns in memory that have been previous justified.

[135]The first quote is from Hamami (2019); the second is his approving quotation of Mac Lane (1986, 377). "Rigorous," here is specialized terminology; it doesn't mean what it means in the phrase, commonly used (and used by me throughout this chapter), "informal rigorous mathematical proof."

[136]Hamami (2019) writes: "[A] *descriptive account* of mathematical rigor provides a characterization of the mechanisms by which mathematical proofs are judged to be rigorous in mathematical practice; *a normative account* of mathematical rigor stipulates one or more conditions that a mathematical proof ought to satisfy in order to qualify as rigorous" (italics his).

[137]Strikingly, the meta-proof considerations include those where shifts occur during the course of a proof to explicitly notational considerations – for example, that the indexes on such operators have such-and-such properties. These notational considerations are ones, of course, about notation occurring in the informal rigorous proof.

The algorithmic-device view agrees that "inferential cues and context trigger algorithmic processes," but these needn't (and often don't) lead to expansions of inferential steps; they often lead to just seeing the result. More important, they usually *don't* point *past* informal proving to a formal transcription.[138] That's because algorithmic competence is at back of all of this, but algorithmic competence, as I've been stressing throughout this chapter, *doesn't* epistemically *require* formalization. Church's thesis, coupled with the mathematical result that any Turing-machine computation, can be represented formally in a first-order predicate language *does not* translate to an epistemic requirement of a formalization of this sort explaining our grasp of any intuitively effective bit of reasoning.

Avigad (2019), on the other hand, although officially accepting Hamami's approach, attempts to give the *normative* version of the derivational account some teeth. He tries to show, that is, that a number of the virtues that good informal mathematical proofs (and proof practices) have – for example, being based on sound analogies, reasonable generalizations, and a wide collection of examples – supply *good guarantees* that an informal proof will have a formal correspondent. Avigad means to show that the existence of corresponding formal derivations is thus an appropriate standard for informal proof practices: the gap between formal derivations and informal proofs isn't a reason to reject derivations-as-standards precisely because good informal proof-practices *raise the likelihood* that an informal proof has a formal correspondent.

I worry (as I have for a long time), however, that "good guarantees" – raising likelihoods in these ways – aren't good enough to explain the epistemic strangeness of informal rigorous mathematical proof. That is, informal mathematical proof practices are *too good* to be explained this way. Recall the quotation from Manders (1995) in footnote 19; for that matter, here's a quotation from MacKenzie (2001, 322–323)[139]:

> . . .many rigorous arguments in ordinary mathematics have been replaced successfully by formal proofs, using automated theorem provers and proof checkers, especially the AUTOMATH and MIZAR systems . . .. What is most remarkable about these many replacements of rigorous arguments of mathematics with formal, mechanized proofs . . . [is that] it is a conservative process. Applied to programs, hardware designs, and system designs, efforts at formal, mechanical proof frequently find faults and deficiencies that have not been detected by other means . . .. Applied to rigorous arguments within mathematics, however, efforts at mechanized proof nearly always suggest at most the need to remedy matters that a mathematician would regard as basically trivial, such as typographic errors or failures to state the full range of conditions necessary for a theory to hold.

[138] They certainly *don't* point past informal rigorous mathematical proofs with one kind of content – arithmetic, say – by virtue of being "decomposable" (Hamami's phrase) into proofs with a strikingly different ideology, that is, that of ZFC.

[139] I'm under the impression that, vis-à-vis various formalization projects, and their impact on informal rigorous mathematics, nothing has changed in the intervening years.

But Avigad's general-virtue approach to mathematical proof *should yield* the same result for programs, hardware designs and system designs that Avigad hypothesizes it yields with respect to informal rigorous mathematical proof. It clearly doesn't.[140]

I've suggested in earlier work (Azzouni 2009b) – in a way related to Avigad's (2019) approach to a normative role for formal derivations – that transcribability to a formal derivation has, in the contemporary setting, become a norm for informal rigorous proof. I want to end this section by revisiting considerations that cut against that idea.[141] The problem is that there are informal rigorous mathematical proofs that are *counterexamples* to the normativity thesis. Consider, for example, the proof in footnote 118. (Visual proofs like that one are the best counterexamples, but there are many similar ordinary-language-based ones.) Phenomenologically – notice – this proof is *utterly convincing* as it stands. Furthermore, there is no sense in which it looks like it needs to be completed or filled in. It's true of many informal rigorous mathematical proofs that they look like they're missing steps; but this is *not* one of those.[142] That is, neither epistemically nor normatively does this proof and many other informal rigorous proofs (those in the Euclidean tradition, e.g., Venn diagrammatic proofs – but *many* other informal rigorous mathematical proofs as well) need supplementation of any sort. Further, although Hilbert's thesis indicates that these proofs can all be formalized (and indeed, as I've illustrated by citation from the literature, many of them *have been*), they themselves don't indicate the existence of formalizations that, in turn, justify why they're true: their content, that is, does nothing of this sort.[143]

The epistemic process, rather, is the exact *reverse* of what normative and descriptive derivational accounts hypothesize. The intuitively effective procedures such proofs exhibit right on their surfaces, when preserved formally, simultaneously preserve the epistemic qualities (the phenomenology) of those informal proofs. The formalization inherits, that is, what it is about the informal proof that convinces us – what *justifies* our being convinced of the result of the proof. It's not, that is, that the formalization reveals what's convincing about that proof or that the formalization justifies that proof. This is enough to show that – at least with respect to *many* informal rigorous mathematical proofs – derivation accounts are intrinsically misleading.

Nevertheless, compatibly with how I've just described the phenomenology of some proofs that formal derivations correspond to informal proofs might *still* be a (general and newly acquired) norm. Whether such a correspondence *is* a norm is,

[140]Compare this response to Avigad with my earlier (Azzouni 2009b, 16, footnote 12) response to Rav (2007, 317).

[141]I first raised this objection to my earlier norm view in Azzouni (2013c) – but this discussion goes beyond that one.

[142]That to formalize this proof would involve something like translating visually seen spatial areas into complex axioms about areas isn't to "fill in" the original proof. It's not even to "make explicit" what's tacit in the proof, as it stands. That's an interpretation that treats, by fiat, a language-based transcription of something we see visually, as therefore (on those grounds alone), "more explicit."

[143]The fact that other proofs of the same form may be misleading is irrelevant.

ultimately, a sociological matter. Consider, however, the possibility of a flourishing diagrammatic tradition that – contrary to Hilbert's thesis – can't be transcribed to a language-based formalization. More dramatically, consider a flourishing informal diagrammatic tradition (involving proofs, say, as phenomenologically compelling as the one in footnote 118) that can't be formalized *at all* – diagrammatically or otherwise. Rather than there being a norm that would fault this proof-tradition, it seems, rather, that generalized versions of Hilbert's thesis, and, in fact, Church's thesis, would be *overthrown.* Indeed, that these are regarded as *theses* is, in fact, the sociological evidence that formalizability of any sort *isn't* a norm of informal rigorous mathematical proof.

16 Conclusion

I've covered far too much territory for an *article.* I'll nevertheless try to summarize (some of) what I've done here in a couple of concluding paragraphs. The problem this chapter opened with is the epistemic strangeness of mathematical proof. This is a (psychological/sociological) claim about informal rigorous mathematical proof – that there is a widely shared experience of "it must be this way" that proofs in mathematics sometimes have (and that long proofs exhibit, usually, only in part) as well as "aha – I see why this must be" epiphanies. That is, mathematical proof, in full, is a complex heterogeneous experience, and that experience (involving short-cuts, invocations of authority, "aha"-epiphanies, and so on) is hardly a seamless impression of a priori inference. The algorithmic-device view explains this phenomenology: algorithmic devices – diagrammatic and language-based – exploit our ability to execute some intuitively effective recognition procedures – ones that are cognitively embodied in us and that we can apply to mathematical problems.

I'll put the point more picturesquely: A mathematical proof is like a stew with chunks of reasoning governed by algorithmic devices plus lots of other – and some looser – stuff. Usually, but not always, algorithmic devices are derived from recognition procedures we normally use (to navigate our world visually, for example) and then we feel the "aha"-epiphany when we apply them. But we don't otherwise – for example, if we use an abacus for certain calculations, or engage in a difficult computational maneuver with integrals.

We cognitively embody many kinds of algorithms. When applying these to the topics they naturally apply to, we effortlessly experience inference processes, often as immediate inferential insights about things. Successful diagrammatic practices notationally hijack these processes and thus enable us to apply these inference packages to subject areas that such inference packages wouldn't otherwise apply to.

The algorithmic-device view of mathematical proof, unlike derivational approaches, explains the epistemic properties of informal rigorous mathematical proof in terms of the mathematical machinery that's *visible* in informal rigorous mathematical proof *itself.* It doesn't explain these properties via a transmutation of informal rigorous mathematical proof to formal algorithmic systems – diagrammatic or otherwise. In turn, other properties of informal rigorous mathematical proof

(e.g., Tanswell's "agreement," "correctness") are explained by the fact that the phenomenology of informal rigorous mathematical proof is based on inference packages and various conventions that are held in common and applied in common by appropriately trained mathematicians.

Nevertheless, transmutations of informal rigorous mathematics into one or another formal system are always possible, via generalizations of Hilbert's thesis. In particular, any natural-language reasoning with diagrammatic algorithmic devices can be formalized in a heterogeneous formal system. Any such formal heterogeneous system, in turn, can always be recast as a language-based formal system because all diagrammatic operations on higher-dimensional grid systems can be recast in one-dimensional concatenation systems – and indeed, as Turing machine calculations. Phenomenologically faithful embeddings of informal rigorous mathematical proof into formal systems retain the apparent content of the mathematical statements – at least insofar as there remains an intended model for the formal systems in question. Respecting this apparent content also explains why the formal systems in question are always based on something in the neighborhood of classical bivalent logic. That classical mathematics is couched in a logic in the neighborhood of the classical first-order predicate logic despite its taking place in natural language which is *not* governed by such a logic is the result of a complex historical process.[144] That means the process is contingent: we could (in the future) find ourselves – collectively – deserting classical logic, as we, in effect, have already with respect to pure mathematics. So too, we might desert ZFC as we, in effect, also have with respect to pure mathematics. This latter possibility turns, largely, on the empirical-science needs that mathematics meets. But discussion of that involves a different set of considerations I can't pursue now; my only point – again – is to stress that our philosophical understanding of ordinary rigorous mathematical proof, certainly as it is and has been pursued in this and the last century, requires axiomatic and informal mathematical pluralism.

References

Abramson FG (1971) Effective computation over the real numbers. Twelfth annual symposium on switching and automata theory. Institute of Electrical and Electronics Engineers, Northridge

Avigad J (2019) Reliability of mathematical inference. Synthese. https://doi.org/10.1007/s11229-019-02524-y

Avigad J, Dean E, Mumma J (2009) A formal system for Euclid's elements. Rev Symb Log 2(4):700–768

Azzouni J (1994) Metaphysical myths, mathematical practice: the ontology and epistemology of the exact sciences. Cambridge University Press, Cambridge

Azzouni J (2000) Applying mathematics: an attempt to design a philosophical problem. Monist 83(2):210–227

[144] I tell this story in Azzouni (2005) and Azzouni (2009b).

Azzouni J (2004) Proof and ontology in Euclidean mathematics. In: Kjeldsen TH, Pedersen SA, Sonne-Hansen LM (eds) New trends in the history and philosophy of mathematics. University Press of Southern Denmark, Odense, pp 117–133
Azzouni J (2005) Is there still a sense in which mathematics can have foundations? In: Sica G (ed) Essays on the foundations of mathematics and logic. Polimetrica International Scientific Publisher, Monza, pp 9–47
Azzouni J (2006) Tracking reason: proof, consequence, and truth. Oxford University Press, Oxford
Azzouni J (2009a) Evading truth commitments: the problem reanalyzed. Log Anal 52(206):139–176
Azzouni J (2009b) Why do informal proofs conform to formal norms? Found Sci 14:9–26
Azzouni J (2013a) That we *see* that some diagrammatic proofs are perfectly rigorous. Philos Math (III) 21:323–338
Azzouni J (2013b) Semantic perception: how the illusion of a public language arises and persists. Oxford University Press, Oxford
Azzouni J (2013c) The relationship of derivations in artificial languages to ordinary rigorous mathematical proof. Philos Math (III) 21:247–254
Azzouni J (2017a) Does reasoning evolve? (does the reasoning in mathematics evolve?). In: Bharath Sriraman (ed) Humanizing mathematics and its philosophy: essays celebrating the 90th birthday of Reuben Hersh. Springer, Cham, pp 253–289
Azzouni J (2017b) The rule-following paradox and its implications for metaphysics. Springer, Cham
Azzouni J (2018) Deflationist truth. In: Glanzberg M (ed) The Oxford handbook of truth. Oxford University Press, Oxford, pp 477–502
Azzouni J (2020) Conceiving and imagining: examples and lessons. In: Moser K, Sukla AC (eds) Imagination and art: explorations in contemporary theory. Brill, Leiden, pp 281–303
Barwise J (1993) Heterogeneous reasoning. In: Mineau G, Moulin B, Sowa J (eds) Conceptual graphs for knowledge representation. ICCS 1993. Lecture notes in computer science (Lecture notes in artificial intelligence), vol 699. Springer, Berlin, pp 64–74
Barwise J, Shimojima A (1995) Surrogate reasoning. Cogn Stud 4(2):7–27
Black M (1946) Critical thinking. Prentice Hall, Upper Saddle River
Brown JR (1999) Philosophy of mathematics: an introduction to the world of proofs and pictures. Routledge, London
Burgess JP, Rosen G (1997) A subject with no object. Oxford University Press, Oxford
Cappellan H (2018) Fixing language: an essay on conceptual engineering. Oxford University Press, Oxford
Cauchy A (1813) Sur les polygones et les polyèdres. Gauthier-Villars, Paris
Church A (1936) An unsolvable problem of elementary number theory. Am J Math 58:345–363
Church A (1954) Introduction to mathematical logic, vol I. Princeton University Press, Princeton
Copeland BJ (2019) The Church-Turing thesis. The Stanford encyclopedia of philosophy (ed: Zalta EN), Spring 2019 edn. https://plato.stanford.edu/archives/spr2019/entries/church-turing/
Davidson D (1986) Inquiries into truth & interpretation. Oxford University Press, Oxford
De Toffoli S, Giardino V (2015) An inquiry into the practice of proving in low-dimensional topology. In: Lolli G, Panza M, Venturi G (eds) From logic to practice: Italian studies in the philosophy of mathematics. BSPS, vol 308. Springer International Publishing, Cham
Descartes R (1931) Rules for the direction of the mind. In: The philosophical works of Descartes (trans: Haldane ES, Ross GRT). Cambridge University Press, London, pp 1–77
Detlefsen M (2008) Purity as an ideal of proof. In: Mancosu P (ed) The philosophy of mathematics. Oxford University Press, Oxford, pp 179–197
Feferman S (2012) And so on . . .: reasoning with infinite diagrams. Synthese 186:371–386
Frege G (1879) Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. In: van Heijenoort J (ed) From Frege to Gödel: a source book in mathematical logic, 1879–1931. Harvard University Press, Harvard, pp 5–82 (1967)
Gandy R (1980) Church's thesis and principles for mechanisms. In: Barwise J, Keisler HJ, Kunen K (eds) The Kleene symposium. North-Holland, Amsterdam, pp 123–148

Gardner M (1988) Hexaflexagons and other mathematical diversions. University of Chicago Press, Chicago
Giaquinto M (2007) Visual thinking in mathematics: an epistemological study. Oxford University Press, Oxford
Giaquinto M (2008a) Visualizing in mathematics. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 22–42
Giaquinto M (2008b) Cognition of structure. In: Mancusu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 43–64
Gödel K (1931) On formally undecidable propositions of Principia Mathematica and related systems I. In: van Heijenoort J (ed) From Frege to Gödel: a source book in mathematical logic, 1879–1931. Harvard University Press, Harvard, pp 596–616 (1967)
Hafner J, Mancosu P (2008) Beyond unification. In: The philosophy of mathematical practice (ed) Paolo Mancosu. Oxford University Press, Oxford, pp 151–176
Hallett M (2008) Refletions on the purity of method in Hilbert's Gundlagen der geometrie. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 198–255
Hamami Y (2019) Mathematical rigor and proof. Rev Symb Log. https://doi.org/10.1017/S1755020319000443
Hammer E (1994) Reasoning with sentences and diagrams. Notre Dame J Form Log 35(1):73–87
Hammer E, Danner N (1996) Towards a model theory of diagrams. J Philos Log 25(5):463–482
Harris JH (1982) What's so logical about the 'logical axioms'? Stud Logica 41:159–171
Heath TL (1956) The thirteen books of Euclid's elements, vol I, 2nd edn. Dover, New York
Hersh R (1997) What is mathematics, really? Oxford University Press, Oxford
Howse J, Stapleton G, Taylor J (2005) Spider diagrams. LMS J Comput Math 8:145–194
Kaye R (1991) Models of Peano arithmetic. Oxford University Press, Oxford
Kleene SC (1936) Lambda-definability and recursiveness. Duke Math J 2:340–353
Kripke S (2013) The Church-Turing thesis as a special corollary of Gödel's completeness theorem. In: Copeland BJ, Posy CJ, Shagrir O (eds) Computability: Turing, Gödel, Church and beyond. MIT Press, Cambridge, MA, pp 77–104
Larvor B (2012) How to think about informal proofs. Synthese 187(2):715–730
Lewis D (1991) Parts of classes. Basil Blackwell, Oxford
Mac Lane S (1986) Mathematics: form and function. Springer, New York
MacKenzie D (2001) Mechanizing proof: computing, risk, and trust. The MIT Press, Cambridge, MA
Manders K (1995) The Euclidean diagram. In: Mancuso P (ed) Philosophy of mathematical practice. Oxford University Press, Oxford, pp 80–133
Manders K (2008) Diagram-based geometry practice. In: Mancuso P (ed) Philosophy of mathematical practice. Oxford University Press, Oxford, pp 65–79
Miller N (2007) Euclid and his twentieth century rivals: diagrams in the logic of Euclidean geometry. CSLI Publications, Stanford
Miller N (2012) On the inconsistency of Mumma's Eu. Notre Dame J Form Log 53(1):27–52
Mumma J (2008a) Ensuring generality in Euclid's diagrammatic arguments. In: Stapleton G, Howse J, Lee J (eds) Diagrammatic representation and inference. Springer, Cham, pp 222–235
Mumma J (2008b) Review of Nathaniel Miller, Euclid and his twentieth century rivals: diagrams in the logic of Euclidean geometry. Philos Math (III) 16:256–281
Mumma J (2010) Proofs, pictures, and Euclid. Synthese 175:255–287
Mumma J (2019) The Eu approach to formalizing Euclid: a response to "On the inconsistency of Mumma's Eu". Notre Dame J Form Log 60(3):457–480
Panza M, Sereni A (2015) On the indispensable premises of the indispensability argument. In: Boston studies in the philosophy and history of science: from logic to practice, vol 308. Springer, Cham, pp 241–276
Plato (1963) Meno (trans: Guthrie WKC). In: Hamilton E, Cairns H (eds) Plato: the collected dialogues. Princeton University Press, Princeton, pp 353–384

Polythress V, Sun H (1972) A method to construct convex connected Venn diagrams for any finite number of sets. Pentagon 31:80–83
Quine WV (1953) On what there is. In: From a logical point of view (1980). Harvard University Press, Cambridge, MA, pp 1–19
Rav Y (1999) Why do we prove theorems? Philos Math 7(3):5–41
Rav Y (2007) A critique of a formalist-mechanist version of the justification of arguments in mathematicians' proof practices. Philos Math 15(3):291–320
Resnik MD (1997) Mathematics as a science of patterns. Oxford University Press, Oxford
Rogers H Jr (1967) Theory of recursive functions and effective computability. McGraw-Hill, New York
Rudin W (1970) Real and complex analysis. McGraw-Hill, New York
Shin S-J (1994) The logical status of diagrams. Cambridge University Press, Cambridge
Stapleton G, Howse J, Thompson S, Taylor J, Chapman P (2013) On the completeness of spider diagrams augmented with constants. In: Moktefi A, Shin S-J (eds) Visual reasoning with diagrams. Springer Basel, Heidelberg, pp 101–133
Tanswell FS (2015) A problem with the dependence of informal proofs on formal proofs. Philos Math 23(3):295–310
Tanswell FS (2018) Conceptual engineering for mathematical concepts. Inquiry 61(8):881–913
Tanswell FS (forthcoming) Go forth and multiply! On action, instructions and imperatives in mathematical proof (ed: Joshua Brown, Otávio Bueno), volume on Azzouni. Springer, Cham
Tarski A (1933) The concept of truth in formalized languages. In: Wooder JH, Corcoran J (eds) Logic, semantics, metamathematics (1983). Hackett Publishing Company, Indianapolis, pp 152–278
Thompson VA, Prowse JA, Turner GP (2011) Intuition, reason, and metacognition. Cogn Psychol 63:107–140
Turing A (1936) On computable numbers, with an application to the Entscheidungsproblem. In: Davis M (ed) The undecidable (1965). Raven Press, New York, pp 115–151
Van Benthem J, Doets K (1983) Higher-order logic. In: Gabbay D, Guenthner F (eds) Handbook of philosophical logic, vol 1. D. Reidel Publishing Company, Dordrecht, pp 275–329

Counterpossibles in Mathematical Practice: The Case of Spoof Perfect Numbers

Alan Baker

Contents

1 Introduction 2262
2 *Reductio* Arguments Are Not Countermathematicals 2264
3 Classifying Countermathematicals 2267
3.1 Metamathematical 2267
3.2 Axiomatic 2267
3.3 Stipulational 2268
3.4 Counterarithmetical 2269
4 Spoof Perfect Numbers 2271
5 Vacuist Responses: Analysis 2273
5.1 Mathematical Practice Is Irrelevant 2273
5.2 "Spoof Perfect Number Theory" Is Not Part of Serious Mathematics 2274
5.3 Mathematically Acceptable Spoof Perfect Numbers Are Interesting 2275
5.4 Vacuously True Counterarithmeticals Do Not Imply That All Numbers Are Spoof Perfect 2277
5.5 Spoof Perfect Numbers Need Not Be Defined Using Counterarithmeticals 2278
6 Paraphrasing Away Counterarithmeticals 2279
6.1 Sigma Functions 2279
6.2 Discussion 2281
7 Nonvacuism and the Evaluation Problem 2284
8 Conclusions 2286
References 2286

Abstract

Philosophical theories of counterfactuals have had relatively little to say about counterfactual reasoning in mathematics. Partly this is because most mathematical counterfactuals seem also to be counterpossibles, in that their antecedents deny some necessary truth. In this chapter, I delineate several different categories of mathematical counterfactual (or "countermathematical") and then examine in

A. Baker (✉)
Department of Philosophy, Swarthmore College, Swarthmore, PA, USA
e-mail: abaker1@swarthmore.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_24

detail a case study from mathematical practice that features counterfactual reasoning about "spoof perfect" numbers. I argue that reasoning about spoof perfect numbers presents both a challenge to philosophical analyses of counterfactuals and a resource for thinking more productively about the role of counterfactual reasoning in mathematics.

Keywords

Counterfactuals · Counterpossibles · Mathematical practice · Perfect numbers

1 Introduction

It has been suggested that a good "philosophical heuristic," when trying to attack a proffered analysis, is to look at extreme cases (Hàjek 2016). When the analysis in question concerns counterfactual conditionals, one class of extreme cases are counterfactuals with impossible antecedents – so-called "counterpossibles" (It has been pointed out to me (by Ralph Wedgewood) that the term "counterpossible" here runs counter to the analogy with "counterfactual." If the antecedent of a counterfactual denies an actual fact, should not the antecedent of a counterpossible deny a possible fact? In other words, a more appropriate term for these impossible counterfactuals is "counternecessary." Unfortunately, current usage is too entrenched at this stage to be open to being changed). As is well-known, the standard possible-worlds analysis of counterfactuals delivers the verdict that all counterpossibles are true. Vacuously true, one might say, since it matters not what the content of the consequent is. Nor is this the only route to vacuity. Syntactic approaches, based on the derivability of the consequent from the antecedent, will also class all counterpossibles as uninterestingly true, if the background logic is classical.

So what? A full analysis of counterfactuals may need to say something about counterpossibles. But does anything substantive hang on what is said? In particular, can it be shown to be a liability for an analysis to classify all counterpossibles as vacuously true? In his 1974 book, *Counterfactuals*, David Lewis devotes several pages to the issue of counterpossibles. Here is an excerpt:

> [O]ne sometimes asserts counterfactuals by way of *reductio* in philosophy, mathematics, and even logic. . . . Their antecedents deny what are thought to be philosophical, mathematical, or even logical truths, and must therefore be thought not only false but impossible. These asserted counterphilosophicals, countermathematicals, and counterlogicals look like examples of vacuously true counterfactuals. (Lewis 1973, p. 24)

In the above passage, Lewis simultaneously coins (at least I am not aware of any earlier usage) the term "countermathematical" for the class of counterpossibles that have mathematical subject matter and makes explicit the implication of his analysis that all counterpossibles are vacuously true. On Lewis's definition, a countermathematical is a conditional whose antecedent denies a mathematical truth. In order to come up with a definition which hinges less on the realism/

antirealism debate in the philosophy of mathematics, let us drop reference to truth and adopt the following definition:

(1) A *countermathematical* is a subjunctive conditional whose antecedent denies an accepted mathematical claim. (I plan to ignore, for present purposes, the question of what degree of agreement across the mathematical community is required in order for a mathematical claim to count as "accepted.")

As is well known, Lewis is a champion of a possible-worlds analysis of counterfactuals. According to his analysis, a counterfactual $A \mathrel{\square\!\!\rightarrow} B$ is true if and only if the closest A-worlds to the actual world are all B-worlds. But if A is impossible, then there are no A-worlds, so the standard analysis makes all counterfactuals with impossible antecedents vacuously true. Note that there are also other routes to this conclusion that have nothing directly to do with possible worlds. For example, consider any analysis of counterfactuals rooted in classical logic which makes the consequent of a counterfactual true if it can be deduced from the antecedent together with appropriately amended auxiliary hypotheses. If the antecedent is contradictory, then every consequent is deducible, regardless of what other hypotheses are added. In what follows, I shall use the term *vacuist* for any analysis according to which all counterpossibles are vacuously true, and *nonvacuist* for any analysis which makes room for some counterpossibles to be true and others false. (Although I shall not pursue the issue in this chapter, I think that most of what I say below against the vacuist position would apply equally well against the converse view that all counterpossibles are vacuously false (or that all counterpossibles lack a truth-value). The point of the examples from mathematical practice is that any uniform classification of counterpossibles is insufficient.)

Pressure on the vacuist account comes from three main sources. However, or so I shall argue, a compelling case against vacuism has yet to be made. First, some philosophers have argued that classifying all counterpossibles as trivially false does violence to our intuitions concerning specific examples (and especially pairs of examples with appropriately contrasting consequents). The problem with this line of objection is that our intuitions about extreme cases tend to be weak to the point of uselessness, and easily manipulated by peripheral contextual factors.

A second source of pressure on vacuism is the increasing use of nonvacuist accounts of counterpossibles to do substantive philosophical work, especially in the analysis of mathematical explanation. Recent work by Sam Baron, Mark Colyvan, David Ripley, Alexander Reutlinger, and others has sought to extend a difference-making account of scientific explanation to cover mathematical explanation. This involves making sense of mathematical facts "making a difference," and a natural way to do this is to consider what would be the case if a certain mathematical fact were different. For this analysis to work, the resulting counterpossibles must be nontrivial, hence it requires some kind of nonvacuist account. This approach has been explored both for mathematical explanation in science (Baron et al. 2017; Baron 2020; Reutlinger et al. 2020) and for intramathematical explanation of one mathematical fact by another (Baron et al. 2020). Although important progress has

been made in fleshing out this analysis, there remains much work to be done in clarifying just how the counterpossibles featuring in mathematical explanations ought to be evaluated. Some general guidelines have been proposed, but without further grounding in mathematical practice there remains the issue of overreliance on vague and only partially articulated intuitions.

This brings us to the third – and in my opinion most crucial – source of pressure on the vacuist view, which is that counterpossibles play a substantive role in actual mathematical practice. If this claim is correct, then it not only casts doubt on the cogency of vacuism by exposing its tensions with practice, but it also bolsters the first two sources of pressure on vacuism. On the one hand, patterns of acceptance and rejection of counterpossibles in mathematical practice can provide a firmer scaffolding for our intuitions about individual cases. On the other hand, this same kind of evidence from practice can flesh out a positive analysis of mathematical explanation within mathematics, and potentially also in science. The problem, however, is that the evidence provided to date of the role played by counterpossibles in mathematical practice is equivocal at best. Although some examples have been given, in many cases – as we shall see – they turn out either not to involve genuine counterfactuals, or not to be significantly linked in to broader aspects of mathematical practice.

My focus in this chapter is on counterpossibles with mathematical content, and in particular arithmetically oriented examples. I shall argue that cases of this sort can be found in actual mathematical practice and that their role in this practice constitutes a *prima facie* challenge to mainstream contemporary accounts of counterfactuals. My principal targets are accounts that are vacuist, in the sense defined above, because such accounts make it particularly difficult to accommodate any interesting role for counterpossibles in practice. But I think that the examples discussed also pose a challenge for existing nonvacuist accounts. Accounts of this latter sort can allow for true counterpossibles and false counterpossibles, but this is not much use in itself if it is not accompanied by workable guidelines for evaluating the truth or falsity of particular examples. Broadly speaking, then, the challenge for the vacuist concerns analysis, and the challenge for the nonvacuist concerns evaluation. ("We can ask two types of questions about counterfactual conditionals: What is the meaning of the statement, and how do we determine whether it is true or false?" (Lewis-Beck et al. 2004).)

I am not the first to look to mathematics for examples of counterpossibles, nor the first to stress "practice-based" examples as a way to put pressure on certain analyses of counterfactuals. However, as I shall argue in the next section, the purported examples of countermathematicals that have been discussed in the philosophical literature to date fail to stand up to scrutiny.

2 *Reductio* Arguments Are Not Countermathematicals

I have argued that one way to advance the debate over the proper analysis of counterfactuals is to find examples of counterpossibles from nonphilosophical practice. But does not the beginning of the Lewis quote from the previous section

indicate one area, straightforwardly distinct from philosophical practice, in which such examples can be found in abundance, namely in proofs by *reductio ad absurdum*? In the first part of this section, I shall argue, against what is the prevailing view in the literature on counterfactuals, that *reductio* arguments are not countermathematicals and thus, more generally, that they are not counterpossibles either. Then, at the end of this section, I shall show that even if I am wrong and *reductio* arguments are counterpossibles, they still cannot be used to put any real pressure on the vacuist position.

The core of my argument against *reductio* arguments being counterpossibles is that they are not even counterfactuals. Counterfactuals are subjunctive conditionals, but *reductio* conditionals do not function as subjunctives. For example, consider the following condensed argument:

(2) Assume (for purposes of *reductio)* that there is a largest prime *p.* Then (*p!* + 1) is prime.

This is true because we can derive the primality of ($p!$ + 1) from known facts about divisibility – in particular, that ($p!$ + 1) leaves remainder 1 when divided by any number less than or equal to *p.* Hence ($p!$ + 1) has no prime factors less than or equal to *p.* This is in sharp contrast to the corresponding claim expressed as a subjunctive conditional. If there *were* to be a largest prime, then (intuitively) all sorts of other established mathematical facts might be different too. But, if so, then we cannot just "automatically" appeal to any such other established mathematical fact in trying to derive a contradiction from the assumption that there is such a *p.* One might argue, for example, that if there were finitely many primes then there would also be only finitely many numbers. In particular, the following might be an acceptable subjunctive conditional,

(3) If there were a largest prime p, then there would be a largest number q such that $q < (p! + 1)$.

But if this is right, then it is false that if there were a largest prime p then ($p!$ + 1) would be prime. It is false because in these circumstances there would be no number ($p!$ + 1). (It is striking that Lewis uses the subjunctive formulation when discussing this exact example in his 1973, p. 25.)

It is enough for present purposes to show that *reductio* arguments are not "disguised" subjunctive conditionals. For, if they are not subjunctive conditionals, they are not counterfactuals, and if they are not counterfactuals then they are not counterpossibles. Are *reductio* arguments "disguised" conditionals of some other sort? I used to think that *reductio* conditionals are more like indicative conditionals. But this now does not seem right to me either. If there *is* a largest prime, then I have been radically misled by the various proofs I have seen which purport to establish the contrary result. So who knows what I should accept as a (mathematical) consequence of this assumption! (This connects with the behavior of indicative conditionals in other contexts. I might accept the subjunctive conditional, "If I hadn't vacationed in

France last summer then I would have vacationed in Spain," but this does not commit me to accepting the corresponding indicative conditional, "If I didn't vacation in France last summer then I vacationed in Spain." If I did not vacation in France, then something very strange is going on because I have lots of vivid memories of doing just that.) (More generally, the role of this conditional in the context of a reductio proof is very different from the normal role of a counterfactual. Roughly speaking, we are interested in the "reductio conditional" only as a means toward demonstrating the falsity of the antecedent. By contrast, our interest typically in a conditional, whether indicative or subjunctive, is in exploring the ramifications of the truth of its antecedent. So, even when there are conditionals in the context of reductio proofs, they do not have the normal role of counterfactuals.)

My thesis that *reductio* arguments do not involve counterpossibles is, I acknowledge, controversial. Fortunately for the broader debate over counterpossibles, there are independent grounds for thinking that the case of *reductio* arguments is not especially useful in settling the issue of how properly to analyze counterfactuals. Consider how an argument against the vacuist analysis of counterfactuals might proceed. The basic form is as follows:

(1) The standard possible-worlds analysis makes all counterfactuals with impossible antecedents vacuously true, since if A is impossible then there are no A-worlds.
(2) *Reductio* arguments in mathematics start from impossible premises. [e.g., the assumption that there is a largest prime number, p].
(3) But not all conclusions advanced in the context of a *reductio* argument are mathematically acceptable.
(4) Hence, there are mathematically unacceptable claims of the form "If P then Q" where P is impossible.
(5) Hence, not all counterfactuals with impossible antecedents are vacuously true, so the standard analysis is untenable.

The typical response to the above argument, for those who favor a possible-worlds-based analysis of counterfactuals, is to make a distinction between what can *truly* be asserted and what can *appropriately* be asserted. In a 1997 paper, Daniel Nolan argues that for a proof to be acceptable it is not enough that it be formally valid; it must also be *obviously* formally valid. Thus a defender of the standard analysis of counterfactuals can argue that the problem with "If $\sqrt{2} = a/b$ then 42 is prime," is not that it is false but that the steps from premise to conclusion have not been made clear. Nolan concludes that "*reductio* proofs do not show the need for nontrivial reasoning about impossibilities" (Nolan 1997, p. 538). (The above line of defense draws a distinction between mathematical acceptability and truth, but not in a way that presupposes any particular background philosophy of mathematics. It should be noted that antirealists about mathematics have independent reasons for rejecting the original argument because on their view mathematical acceptability does not even aim for literal truth.) In this way, the vacuist can hold onto the claim that all countermathematicals are vacuously true, but add that a countermathematical is (mathematically) acceptable if and only if a derivation is provided of its

consequent from its antecedent using other demonstrable mathematical facts. This is in contrast to the nonvacuist, who holds that some countermathematicals are not just mathematically unacceptable but also false.

Can this debate be moved forward, and – if so – how? I think that it can, and one of my main aims in this chapter is to indicate why and how. The key to making progress is to realize that there are other examples of countermathematicals which play a role in mathematical practice but are not part of *reductio* arguments.

3 Classifying Countermathematicals

Recall our earlier definition, that a *countermathematical* (henceforth, CM) is a subjunctive conditional whose antecedent denies an accepted mathematical claim. Before proceeding to look at some case studies, it will be useful to have in place a preliminary classification of different types of CM. I do not claim that the taxonomy that follows is either optimal or exhaustive, but hopefully it will allow for a more focused discussion of different examples.

3.1 Metamathematical

Some mathematical claims are not claims of a particular mathematical theory but claims *about* the properties of a theory. The denial of some established metamathematical claim is a *metamathematical* CM (or, perhaps, a *countermetamathematical*). For example,

(4) If ZF set theory were a conservative extension of Peano Arithmetic, then Goodstein's Theorem would be provable in PA.

It is unclear to what extent metamathematical CMs play a role in mathematical practice. Sometimes, metamathematical CMs are used in *reductio*-style reasoning, for example, showing that if a certain formal theory were complete and consistent then it would be strong enough to embed arithmetic, and hence Gödel's Incompleteness Theorem would be violated. I would be inclined to resist classifying such cases as CMs for the reasons canvassed in Sect. 2. Another reason for distinguishing this category from "genuine" CMs comes from stressing the division between mathematics and metamathematics. Perhaps reasoning of this sort is not "internal" enough to mathematical practice, and too often bound up with broader issues in logic or philosophy.

3.2 Axiomatic

An *axiomatic* CM is a conditional whose antecedent denies an established axiom of an established mathematical theory. For example,

(5) If the Axiom of Choice were false, then there would be sets S and T of incomparable size (i.e., neither can be mapped in a one-to-one fashion onto a subset of the other).

Evaluating the truth or falsity of axiomatic CMs seems relatively straightforward, at least in principle. It is a matter of determining whether the consequent is a theorem of the amended theory produced by replacing the axiom in question with its negation. Jonathan Bennett has argued – convincingly, I think – that evaluation is possible in this sort of case, despite the lack of analytical support, precisely because the axioms of a given theory provide a framework for carving up logical space in a specified manner;

> Consider . . . how counter*logicals* are handled. . . . [A] speaker can say things of the form 'If conjunction were not commutative, then C would be the case,' and be right for some values of C and wrong for others, just so long as he is talking about the power structure of some *system* of logic – some set of rules and independent axioms – and saying that C is a theorem in the system that results from the original one by deleting its commutativity axiom. Such conditionals are saved from triviality by being made relative to some *formulation* of logical truth. (Bennett 2003, pp. 228–229)

Independence does seem to be a crucial feature of what makes axiom-based CMs relatively unproblematic to evaluate. It means they can be dropped (or even negated) without having to alter any other axioms of the theory (Although I agree with Bennett's general point here, the specific example in the given quote does not work quite as he claims. If all that is done is that the denied axiom is "deleted," then the new system will be weaker than the old system. To get any new results, the negation of the deleted axiom must be added in its place. Otherwise, the most we could legitimately say would be along the lines of, "liven if conjunction were not commutative, C would still be the case."). Another way of putting this point is that there is another (logically) consistent theory in the vicinity which can be used as a framework for evaluation.

Even if it is conceded that evaluation of axiomatic CMs is straightforward, there remains the issue of how such CMs are to be analyzed. It is true that this may raise problems for platonists, for example, if the Axiom of Choice is necessarily true, then what can be meant by drawing conclusions from the supposition that it is false? (Nolan 1997) Nonetheless, I doubt that axiomatic CMs will provide good case studies for the debate over counterpossibles because this problem of analysis does not extend across other positions in the philosophy of mathematics. Indeed, even some other versions of mathematical realism, for example, Quinean indispensabilism, will have no problem with conditional claims that deny the truth of established axioms.

3.3 Stipulational

A CM is *stipulational* if its antecedent asserts that a "borderline" case (or cases) of some mathematical concept be reclassified in a different way. One of the clearest

examples of this sort concerns the nonprimality of 1. Here is a sample quote taken from a mathematics textbook (there are also plenty of more technical examples, "Take the Euler Fi-function for example. it can be shown the $Fi[p] = p - 1$ when p is a prime. If 1 was prime that wouldn't be true for $Fi[1] = 1$"):

(6) If 1 were prime, then the Fundamental Theorem of Arithmetic, which says that every whole number greater than one can be written uniquely as the product of prime numbers, would be false.

The most natural way of reading this example, I suggest, is as contradicting not actual mathematical *facts* but actual mathematical *terminology.* Since the mathematical facts can therefore be left untouched, evaluation is again relatively straightforward. Each old theorem mentioning primality (or a property defined using primality) must be retranslated using the new terminology and then assessed to see if it remains true. Some do (for example, the theorem that there are an infinite number of primes) and others do not (for example, the theorem about the uniqueness of prime factorization).

Crucial to this reading is that the reclassified cases be borderline, otherwise the new concept becomes too far removed from the original one. Compare (6) above with.

(7) If 4 were prime, then some composite numbers would have no prime factors.

As things stand, 4 is a paradigmatic example of a nonprime number. If the meaning of "prime" were altered to the extent that 4 is counted as prime, then it becomes unclear what would happen to related terms such as "composite" and "factor," and this interferes with the evaluation of (7).

In light of the above remarks, there is a sense in which stipulational CMs are not really counter*possibles* at all. A platonist, for example, can entertain the antecedent of (6) without supposing that anything is different in the realm of mathematical abstracta. This suggests that even the vacuist can allow for such conditionals to be nontrivial by focusing on change of terminology rather than change in intrinsic facts about mathematics. (Perhaps there is a way to make the distinction between stipulational CMs and other sorts of CM in terms of Kripke's rigid designator/nonrigid designator distinction.) If so, then stipulational CMs cannot provide a definitive challenge to competing analyses of counterpossibles.

3.4 Counterarithmetical

Our discussion of the above two categories suggests that the most effective examples of CMs are likely to be neither axiomatic nor stipulational. A further desideratum is that the CM feature in an area of mathematics in which there is a single, well-established theory, and about which there is little debate among mathematicians

concerning the correctness of its basic results. Penelope Maddy has usefully distinguished areas of mathematics in which there is a presumption of a single correct theory in terms of a background maxim she calls UNIFY (Maddy 1997). In an area not governed by this maxim, such as group theory, it seems that subjunctive conditionals do not generally correspond to counterfactuals. For example, consider the conditional:

(8) If the operator of the group Z_6 were not commutative, then the group would form the symmetry group of an equilateral triangle.

This is best read, I suggest, is simply shifting the focus from an abelian group of order 6 to the (only) nonabelian group of order 6, namely S_3, and making a true assertion about a feature of this latter group.

For our purposes, a good place to look is arithmetic since it seems to have the above feature; moreover, there is little doubt over its core axioms and theorems. (According to Maddy, set theory is one paradigm area of mathematics where UNIFY is operative. However, although there may be a presumption of a single correct theory, there is ongoing debate about which set-theoretical axioms ought to be included (especially in the area of large cardinal axioms).) I coin the term "counterarithmetical" to cover the special case of a CM which pertains to arithmetic. More precisely:

(9) A *counterarithmetical* is a subjunctive conditional whose antecedent denies an accepted arithmetical claim. (Rather than trying to build into the above definition some distinction between CAs and the previous two categories, I shall just accept that some CAs are axiomatic (e.g., "If there were a number without a unique successor then there would be a largest prime number."), and some CAs are stipulational (e.g., (8) above).)

Linking examples of counterarithmeticals to mathematical practice is crucial here, and it is in this emphasis that I part company with what prior discussion there has been of the topic. Various examples of counterpossibles with arithmetical subject matter have been given in the philosophical literature, but these "philosophical" examples tend to bear little relation to mathematical practice and hence must be evaluated almost entirely on the basis of intuitions that are typically fairly weak, even conflicting.

For example, Almerindo Ojeda asserts in passing, while arguing against the standard possible worlds analysis of counterpossibles, that

(10) If 3 were even, 4 would be odd

is intuitively true, but that

(11) If 3 were even, 4 would be prime

is intuitively false (Ojeda 2005, p. 31). Beyond the bare appeal to intuition, Ojeda's argument is brief. Roughly, his idea is that the successor of an even number need not be prime, so if 3 were even this would not "force" 4 to be prime. The problem with using intuitions as the sole data for philosophical analysis is that it provides no guidance when faced with conflicting intuitions. For example, I happen not to share Ojeda's intuition that the second of the above counterarithmeticals is false. In support, I offer the following argument: I accept that if 3 were even then 4 would be odd. But if 4 were odd, then it would not be divisible by 2 (from the definition of oddness). Also, 4 would not be divisible by 3, because 3 would be even and only even numbers are divisible by other even numbers. So 4 would not be divisible by any smaller number except for 1. Hence, 4 would be prime.

I am not claiming that my argument is decisive. In fact my whole point is that one cannot expect either decisive arguments or robust intuitions concerning such "free-floating" CMs. For this reason, it is unlikely that concocting philosophically motivated examples of CMs will provide a firm enough basis for undermining the standard view. This applies not just to CMs but to any counterfactual conditional with an "extreme" antecedent. (Another class of extreme counterfactuals are those whose antecedents deny a law of nature. For an interesting analysis of such "counterlegals," (see Kvart 1986, p. 208). Also, for an attempt to link counterlegals to certain global kinds of mathematical counterfactual (such as "If there were no mathematical objects, then. . ..") via Quinean indispensability considerations, see Baker (2003).) One problem is that our intuitions concerning counterfactuals tend to weaken as the counterfactual conditions increase in breadth (the scope and magnitude of discrepancies from the actual world) and depth (the modal strength of the actual claims being denied). A second problem is that extreme antecedents do not tend to feature in everyday counterfactual claims. Outside of the philosophy classroom, people rarely go around making assertions about what would be the case if there were round squares, or if some objects were both blue and red all over. Hence, there is little in the way of extraphilosophical "commonsense" usage to measure our analyses against.

It is here, I think, that consideration of counterarithmeticals in particular has the potential to cast light on the broader debate over counterpossibles. For – as we shall see in the next section – there are examples of mathematicians asserting counterarithmeticals in the course of normal mathematical practice, in what appears to be a philosophically unselfconscious manner. Looking at the pattern of acceptance and rejection of counterarithmeticals by practicing mathematicians may provide a potentially important body of data from a source external to, and no less reliable than, philosophy itself.

4 Spoof Perfect Numbers

A natural number is *perfect* if and only if it is equal to the sum of its proper divisors, including 1. Perfect numbers have been an object of interest for mathematicians since ancient times, and the Greeks in particular attached mystical significance to

them. Partly, this is because of the rarity of perfect numbers. Just 5 perfect numbers were known to the ancient Greeks, and even with the aid of today's powerful computers fewer than 50 more have been added to that list. There are also several quite basic questions about perfect numbers which remain unsettled. These include whether there are infinitely many perfect numbers, and whether there are any odd perfect numbers.

It is this latter question which provides the starting point for our first example. In 1638, Descartes corresponded with Mersenne concerning the issue of the existence of odd perfect numbers. Descartes argues that there are likely to be odd perfect numbers, writing at one point that

> I think I am able to prove that . . . there are no odd perfect numbers, unless they are composed of a single prime number multiplied by a square whose root is composed of several other prime numbers. But I can see nothing which would prevent one from finding numbers of this sort. For example, if 22,021 were prime, in multiplying it by 918,009, which is a square whose root is composed of the prime numbers 3, 7, 11, 13, one would have 198,585,576,189, which would be a perfect number. (Quoted in Crubellier and Sip 1997)

As Descartes was well aware, 22,021 is a composite number (its prime factorization is $19^2.61$). Focus for the moment on the core claim which Descartes makes here, namely

(12) If 22,021 were prime, then 198,585,576,189 would be perfect.

This is a subjunctive conditional whose antecedent denies an accepted arithmetical claim, in other words, it is a **counterarithmetical.** In recent years, number theorists have become interested in numbers of this sort, which they have termed *spoof perfect numbers.* Specifically, a spoof perfect number is defined to be a number that is perfect if you assume, contrary to mathematical fact, that one or more of its composite factors are prime.

To make the mathematical side of the discussion more tractable, I shall henceforth focus on a much smaller example of a spoof perfect number:

(13) If 4 were prime, then 60 would be perfect.

At first blush, one might wonder what difference it makes whether a certain factor is prime or not. After all, the perfectness of a number depends on the sum of *all* its factors, regardless of primality. The idea of the spoof antecedent condition is that the spoof prime number, in the above case 4, is now assumed not to break down into any smaller factors. Under this hypothesis, there are fewer factors of 60, because only factors expressible as multiples of 3, 4, and 5 are permitted. Thus, for example, 10 would no longer count as a factor of 60. The "eligible" factors of 60, under the hypothesis that 4 is prime, are 1, 3, 4, 5, 12, 15, and 20. These sum to 60, hence 60 is *spoof perfect.*

My core argument against the vacuist account of counterpossibles can now be formulated:

1. Spoof perfect numbers are a mathematically interesting subset of the natural numbers.
2. Each spoof perfect number is defined using a counterarithmetical.
3. If all counterpossibles are true, then counterarithmeticals do not define a mathematically interesting subset of the natural numbers.
4. Hence, not all counterpossibles are true.

The remainder of the chapter will be devoted to potential ways of responding to this argument on behalf of someone who holds a vacuist view of counterpossibles.

5 Vacuist Responses: Analysis

5.1 Mathematical Practice Is Irrelevant

Perhaps, the most extreme response to the above challenge is to dismiss it entirely on the grounds that appeal to mathematical practice ought to carry no weight in the debate over the proper analysis of counterpossibles. I shall be brief here, and not just because this response effectively puts an end to the present discussion. It is a widely held belief among both contemporary metaphysicians and contemporary philosophers of mathematics that mathematical practice is entirely relevant to the evaluation of competing analyses of concepts which appear within the borders of mathematics, which is not to say that practice overrides all other factors. But if a philosophical analysis implies that significant portions of the mathematical community are systematically wrong about some area of mathematical discourse, then this is – and should be – taken to be a serious drawback for the analysis in question.

The earlier discussion of *reductio* reasoning in mathematics as a potential thorn in the side of the vacuist account of counterpossibles was predicated on the assumption that if it were to be shown that mathematicians use counterpossibles in a way that conflicts with vacuist account, then it would be a problem for the latter analysis. And David Lewis, an archetypal vacuist about counterpossibles, has made numerous remarks about the importance of respecting mathematical practice as a condition of adequacy on philosophical theorizing. One of his better-known proclamations is the following,

> Mathematics is an established, going concern. Philosophy is as shaky as can be. To reject mathematics for philosophical reasons would be absurd. . . . I'm moved to laughter at the thought of how presumptuous it would be to reject mathematics for philosophical reasons. (Lewis (1991, pp. 58–59). The intervening passage between the two quoted segments reads, "If we philosophers are sorely puzzled by the classes that constitute mathematical reality, that's our problem. We shouldn't expect mathematics to go away to make our life easier. Even if we reject mathematics gently – explaining how it can be a most useful fiction, 'good without being true' – we still reject it, and that's still absurd. Even if we hold onto some mutilated fragments of mathematics that can be reconstructed without classes, if we reject the bulk of mathematics, that's still absurd.")

5.2 "Spoof Perfect Number Theory" Is Not Part of Serious Mathematics

A second line of counterargument is to not only acknowledge that systematic patterns of reasoning exhibited in mathematical practice do carry weight, at least in principle, but also to argue that the investigation of spoof perfect numbers is not a legitimate part of "mainstream" mathematical practice. If so, then the core argument may perhaps still be dismissed.

It is true that discussions of spoof perfect numbers are uncommon, even within number theory. On the other hand, results and conjectures concerning spoof perfect numbers can be found in "respectable" mathematical publications – for example, in R.K. Guy's influential survey of open questions in number theory (Guy 2004), and refereed journals such as *Integers and Mathematical Computation* (Nielsen (2003, 2015) and Dittmer (2014). For additional recent work focused specifically on spoof perfect numbers, see also Arnaldo and Dris (2017), and BYU Computational Number Theory Group (2020)).

What about links between spoof perfect numbers and other issues of demonstrable significance in number theory (or indeed elsewhere in mathematics)? Links of this sort are one aspect of what G.H. Hardy has termed the "seriousness" of a piece of mathematics. (A chess problem is genuine mathematics, but it is in some way "trivial" mathematics. However ingenious and intricate, however original and surprising the moves, there is something essential lacking. Chess problems are unimportant. The best mathematics is serious as well as beautiful – "important" if you like, but the word is very ambiguous, and "serious" expresses what I mean much better (Hardy 1940, p. 16).) Unsurprisingly, the area to which research into spoof perfect numbers is most commonly connected is research into perfect numbers *simpliciter*, and in particular odd perfect numbers. As already mentioned, whether there exist any odd perfect numbers is an important and open question in number theory. The consensus seems to be that it is unlikely that there are any odd perfect numbers, partly because various increasingly strong constraints have been proved concerning the size and properties of any such number. For example, it is known that if any odd perfect number must be larger than 10^{1500}, it must have at least 101 (not necessarily distinct) prime divisors, and it must have the form $p^{4a+1}\, m^2$, where $p = 1 \pmod 4$ is prime. (For these and further conditions on odd perfect numbers, see Holdener and Rachfal (2019, p. 541). The nineteenth-century mathematician Sylvester is quoted as saying that "the existence of [an odd perfect number] – its escape, so to say, from the complex web of conditions which hem it in on all sides – would be little short of a miracle.") In Sect. 6 below, I shall discuss an example where results concerning spoof odd perfect numbers are used to draw conclusions about "genuine" odd perfect numbers.

A third indicator of the mathematical interest of spoof perfect numbers is the extent to which the basic notion of spoofness has been generalized or modified to yield other potentially fruitful mathematical concepts. One generalization of Descartes' notion is to that of a *multispoof perfect number.* As the name suggests, these are numbers which are perfect if **two or more** of their composite factors are assumed to be prime. For example,

(14) If 4 and 6 were prime, then 840 would be perfect. (This example is due to Wang (2002). The full "prime" factorization is 840 = 4.5.6.7. It is straightforward (though tedious) to verify that the various combinations of these factors sum to 840, as required.)

Moving even further in this direction, a number is *completely spoof perfect* if it can be expressed as a product of composite numbers, and it would be perfect if **all** of these composite factors were prime. For example,

> 390405312000= 4.8.9.10.15.22.46.94.95 is an even spoof perfect number if you suppose incorrectly that all these factors are primes. (ibid)

Reformulating this in conditional form yields the following counterarithmetical:

(15) If 4 and 8 and 9 and 10 and 15 and 22 and 46 and 94 and 95 were prime, then 390,405,312,000 would be perfect.

Another extension that is worth mentioning, partly because it is more radical in nature, is to *negative spoof perfect numbers.* This involves allowing negative numbers as potential prime factors, for example, when we write $6 = (-2.).(-3)$, thus opening up a whole new class of cases. An example of a negative spoof perfect number is given by Voight. $N = 22{,}017{,}975{,}903 = 3^4 7^2 11^2 19^2(-127)$ "is perfect if we are willing to ignore the fact that −127 is negative" (Voight 2003).

Finally, there are various theorems and conjectures concerning spoof perfect numbers, results that we will discuss in more detail in the next subsection. These provide more evidence, at least *prima facie*, that the topic – though small in scope – is of genuine mathematical interest.

Lurking in the background here are broader questions concerning the relation between mathematical practice and philosophy. What force does appeal to evidence from mathematical practice have in attacking or defending philosophical claims about mathematics? And can distinctions be made between different sorts of mathematical practice? Ought we to privilege "serious" academic mathematics over recreational mathematics? Published mathematics over unpublished? Statements made in the course of a proof over "stage-setting" introductory remarks? At this juncture, I will limit myself to a fairly modest claim: that there is enough going on in the mathematical investigation of spoof perfect numbers for it to deserve to be taken seriously as data for the analysis and evaluation of counterpossibles.

5.3 Mathematically Acceptable Spoof Perfect Numbers Are Interesting

What about responding to the core argument by borrowing one of the vacuist strategies from the discussion of *reductio* reasoning, and making a pragmatic distinction between truth and mathematical acceptability? The idea is to take aim

at Premise 1 and argue that, strictly speaking, the set of spoof numbers is not interesting, because (so the vacuist maintains) all counterarithmeticals are true. However, what mathematicians are picking out in their examples of spoof perfect numbers are part of a subset of *mathematically acceptable* spoof perfect numbers. And these may well be quite interesting. What makes a particular spoof perfect number mathematically acceptable? Just as in the *reductio* case, acceptability hinges on the availability of a step-by-step derivation of the consequent (that the number is perfect) from the antecedent (that a particular composite factor is prime).

There are at least a couple of reasons for thinking that this pragmatic strategy will not be as effective here as it was for *reductio* conditionals. First, it turns out that there is always a simple derivation of the consequent from the antecedent available, because – unlike in typical *reductio* proofs – we already have the negation of the antecedent in our stock of basic theorems and results. (Of course, if the claim being proved by reductio has already been proved by some other method, then we may indulge in the pretense that is not available as a theorem that can be introduced during the course of the new proof. Nonetheless, the canonical case of reductio reasoning proceeds from genuine ignorance over the truth-value of the initial assumption.) For example, we can quickly prove that 59 is spoof perfect as follows.

(1)	4 is prime	Assumption
(2)	So 4 is prime or 59 is perfect	1, Addition
(3)	4 is not prime	Theorem
(4)	Hence, 59 is perfect	2, 3, Disjunctive Syllogism

Clearly, this same pattern of reasoning can be applied to prove the literal truth of any claim about spoof perfectness. So the vacuist is in a bind because it is not just the truth of all counterarithmeticals that is uninteresting but also the trivial easiness of proving individual cases.

Second, and more important, present in the mathematical literature are theorems and conjectures concerning the existence of spoof perfect numbers of various sorts. Many of these results concern completely spoof perfect numbers (recall from the previous section that a number is completely spoof perfect if it would be perfect under the assumption that all of its composite factors are prime). For example,

Theorem 1: For a given number of composite factors, the number of even completely spoof perfect numbers is finite.

Theorem 2: Every completely spoof perfect number has at least six composite factors.

Theorem 3: 907,200 is the smallest even completely spoof perfect number.

The details and proofs of these theorems can be set aside for the moment. What is crucial is that they each make claims that are simply false if – as the vacuist

maintains – *all* numbers are spoof perfect (and multispoof perfect, and completely spoof perfect).

Not only does the vacuist position render the above "theorems" literally false, but it also settles in a trivial manner certain conjectures which are taken by the number theoretic community to be open (and interesting). For example,

Conjecture: There is no odd completely spoof perfect numbers.

To summarize, the inclusion in number theoretic practice of theorems and conjectures of the above sort brings the pragmatic strategy of the vacuist into direct conflict with mathematical practice. If all counterarithmeticals are indeed literally true, then the bulk of these theorems are literally false. And no appeal to a mathematical acceptability/unacceptability distinction will get around this fact.

One glimmer of hope does remain for the proponent of a pragmatic vacuist strategy. She could attempt to formalize the notion of "interestingly true counterarithmetical" and then interpret the above theorems as making claims that are implicitly restricted to the class of "interesting" spoof perfect numbers (I owe this suggestion to Bill Childs). However, the burden of proof is then on the vacuist to come up with an adequate characterization of "interesting." I shall postpone discussion of this option until we have some more technical apparatus in place.

5.4 Vacuously True Counterarithmeticals Do Not Imply That All Numbers Are Spoof Perfect

This response on behalf of the vacuist takes aim at Premise 3 of the core argument. The idea is to question the inference from vacuously true counterarithmeticals to all numbers being (trivially) spoof perfect. As we shall see, there is something to this idea, depending on how exactly one decides to define spoof perfectness. But even on a definition which allows for some numbers not to be spoof perfect, even for the vacuist, the resulting set of numbers still seems to have little independent mathematical interest. Hence, the overall core argument is unaffected.

The point of this response is to focus on the precise definitional link between spoof perfect numbers and counterarithmeticals. Recall the sorts of cases that have been presented, for example,

(16) If 4 were prime, then 60 would be perfect.

The weakest way of defining spoof perfectness is to say that a number is spoof perfect if and only if it features in the consequent of a counterarithmetical of the above sort. More precisely,

Definition 1: A number, *n*, is **spoof perfect** if and only if there is some *m* such that (If *m* were prime, then *n* would be perfect) is true.

On this definition, a vacuist account of counterpossibles implies that all numbers are spoof perfect. However, we could also put further constraints on the conditional claim. One obvious constraint is to add that the conditional in question be a counterfactual, in other words that *m* not be prime. In itself, this does not reduce the extent of spoof perfect numbers. But it does if we combine it with a further plausible constraint that *m* be a proper divisor of *n*. Adding these twin constraints yields the following alternative definition:

Definition 2: A number, *n*, is **spoof perfect** if and only if there is some composite *m* such that *m* is a proper divisor of *n* and (If *m* were prime, then *n* would be perfect) is true.

It is worth noting that for the nonvacuist, the only difference between these two definitions is whether actually perfect numbers are counted as spoof perfect since, in Definition 1, presumably only *ms* that are proper factors of *n* will feature in true counterfactuals about the perfectness of *n*. But for the vacuist, the difference between the two definitions is more significant. Whereas Definition 1 implies that all numbers are spoof perfect, Definition 2 restricts the spoof perfect numbers to those which have at least one composite proper divisor (since this is just the condition placed on *m*).

So, to summarize, there is nothing to stop the vacuist from picking Definition 2 and thus making room for not all numbers being spoof perfect. However, it is easy to see that the resulting set consists of all and only those numbers that have at least one composite proper divisor, in other words numbers with at least three prime factors. Even on this refinement of the vacuist position, therefore, spoof perfect numbers must be equated with a much more basically defined set, leaving no reason for mathematicians to characterize them in such a convoluted, counterarithmetical way. I conclude that the refined vacuist position still does not fit with mathematical practice and hence that this response on behalf of the vacuist is insufficient.

5.5 Spoof Perfect Numbers Need Not Be Defined Using Counterarithmeticals

Thus far, we have looked at challenges to Premise 1 and to Premise 3 of the core argument, as well as attempts to dismiss the force of the argument outright. I now want to turn to possible attacks on what is in many ways the crucial claim of the argument, encapsulated in Premise 2:

2. Each spoof perfect number is defined using a counterarithmetical.

In the original formulation of the core argument, this premise was left intentionally vague. In particular, the "is" in Premise 2 could be read (in increasing order of strength) as

. . . **may be** defined using a counterarithmetical.
. . . **is always** defined using a counterarithmetical.
. . . **must be** defined using a counterarithmetical.

The idea behind challenging this claim is to argue that the counterfactual locutions that tend to be used when introducing the notion of spoof perfect number – "if 22,021 were prime," "if 4 were prime," etc. – are rhetorical "fluff." Not only *can* they be dispensed with when giving a more precise and explicit definition of spoof perfectness, but they *ought* to be also dispensed with. If it turns out that precise definitions can be given which make no use of counterpossibles, then the vacuist would seem to be off the hook. Vacuous counterpossibles can then happily coexist with a contentful role for spoof perfect numbers in mathematics.

6 Paraphrasing Away Counterarithmeticals

6.1 Sigma Functions

What has been presented at the end of Sect. 5 is a program for a defense against the core argument, not a defense in itself. Its success will depend crucially on whether counterpossible-free definitions of spoof perfect number, multispoof perfect number, and so on can be formulated. The approach I shall consider here is based on the apparatus of sigma-functions. I have chosen this approach for two reasons. First, the use of sigma-functions is already well-entrenched in the theory of perfect numbers. Second (and relatedly), it is to sigma-functions that mathematicians most often turn when explicating the more intricate aspects of spoof perfect numbers (see Nielsen (2003, p. 9) and Guy (2004, p. 72)).

Definition: The *sigma-sum* of a natural number, n, written $\sigma(n)$, is the sum of all of the proper and improper divisors of n.

It will be useful for subsequent discussion to note the following theorem concerning sigma-sums:

Theorem: If the prime decomposition of a number, n, is $p_1^a \cdot p_2^b \cdot \ldots$, then $\sigma(n) = (1 + p_1 + p_1^2 + \ldots + p_1^a)(1 + p_2 + p_2^2 + \ldots + p_2^b)(\ldots)$.

The proof is too involved to give here, but the result is intuitive if one thinks of the divisors of a number being made up of all possible combinations of the elements of its prime decomposition. It can be seen that multiplying out the σ product yields each divisor as one element of the final sum.

Lemma: n is a perfect number if and only if $\sigma(n) = 2n$

This follows immediately from the above Theorem, since by definition a number, n, is perfect if it is the sum of all its divisors other than itself. Hence, the sum of all its divisors and itself is equal to $2n$.

The basic strategy for defining spoof perfect numbers is to construct a new "sigma*-sum" corresponding to each counterarithmetical antecedent, such as "If 4 were prime." The sigma* function corresponding to this assumption treats 4 as if it were a prime number for the purposes of calculating sums of divisors. This involves first expressing the number whose spoof perfectness we are investigating as the product of powers of 4 together with a product of prime factors. (In order to guarantee uniqueness, the highest power of 4 possible must be included in the factorization. For example, $16 = 22.4$ would not be a legitimate "4-factorization.") For example, $20 = 4.5$; $24 = 4.2.3$; $75 = 3.5^2$; $80 = 4^2.5$, and so on. If the components of such a factorization are $c_1^a.\ c_2^b.\ \ldots$, then $\sigma^*(n) = (1 + c_1 + c_1^2 + \ldots + c_1^a)(1 + c_2 + c_2^2 + \ldots + c_2^b)(\ldots)$. As can be seen, this directly mirrors the Theorem for sigma-sums that applies to (purely) prime factorizations. If $\sigma^*(n) = 2n$, then n is spoof perfect.

For illustrative purposes, let us run through the above procedure with 4 as our "counterarithmetical prime" and 60 as our number being investigated. $60 = 4.3.5$, so $\sigma^*(60) = (1 + 4)(1 + 3)(1 + 5) = 5.4.6 = 120$. So $\sigma^*(n) = 2n$, and 60 is spoof perfect. With this apparatus in place, we can agree to accept as correct those counterarithmeticals which correctly describe a σ*-function of this sort. So "If 4 were prime then 60 would be perfect" is correct, but "If 4 prime then 59 would be perfect" is not correct (The latter counterarithmetical is not correct because the 4-factorization of 59 is 59, so $\sigma^*(59) = 1 + 59 = 60$ ((2.59)). The counterarithmeticals are not to be read literally (since, given the vacuist account, they are all trivially true) but rather as gesturing at the corresponding claims about σ*-functions.

All that remains is to define the absolute notion of a number being spoof perfect, when this is not tied to any particular counterarithmetical claim. If we introduce subscript notation to indicate the composite number whose primality is being assumed, so $\sigma^*_4(n)$ is the σ*-function for n under the supposition that 4 is prime, then we can define spoof perfectness as follows:

Definition: A number, n, is **spoof perfect** if and only if $\exists m$ such that m is a composite proper divisor of n and $\sigma^*_m(n) = 2n$.

What about the other varieties of spoofness that were mentioned earlier? I will not pause at this point to formulate a sigma-function definition of *multispoof perfect number*, though I think it would be quite straightforward to do so. As for *completely spoof perfect number*, this requires extending the σ^* notation to allow for multiple subscripts. Recall that a number is completely spoof perfect if and only if it can be expressed as the product of composite factors and would be perfect under the assumption that all of these composite factors are prime. So, for composite numbers $a, b, c, \ldots$, let the $(a, b, c, \ldots)$-factorization of n be a factorization into the product of powers of $a, b, c, \ldots$ (This involves abandoning uniqueness of the result of the σ*

operation because in many cases there will be more than one way to express a number as the product of given composite factors. For example, consider factoring 1024 into a product of the composite numbers 4 and 16. Not only $1024 = 4.162$, but also $1024 = 43.16$. It can be verified that these two factorizations yield distinct σ*-sums). Then define $\sigma^*_{a,b,c,\ldots}(n)$ to be the σ*-function for n under the supposition that a, b, c, ... are prime.

Definition: A number, n, is **completely spoof perfect** if and only if $\exists a, b, c, \ldots$ such that a, b, c, ... are composite proper divisors of n and $\sigma^*_{a,b,c,\ldots}(n) = 2n$.

Finally, to illustrate the expressive power of the sigma-notation approach, we shall show how one of the theorems concerning spoof perfect numbers that was mentioned in subsection (III) above can be "translated" into the language of σ-functions.

Theorem 2: Every completely spoof perfect number has at least six composite factors.

Theorem 2 (σ* version): $\forall n \; \forall a, b, c, \ldots$ (If a, b, c, ... are composite proper divisors of n, and if $\sigma^*_{a,b,c,\ldots}(n) = 2n$, then the number of composite factors of $n \geq 6$).

6.2 Discussion

It is clear, I think, that the sigma-function approach sketched above provides the vacuist with a powerful tool for resisting the force of the core argument. It uses apparatus that is already quite familiar from mainstream number theory, and it allows definitions and theorems concerning spoof perfect numbers to be formulated without involving any countermathematical elements. No proof has been given that this method of "sigma-paraphrase" can *always* be applied to claims concerning spoof perfect numbers, but there is no reason to think that it cannot, and plenty of reason to view this is a quite general method for eliminating reference to counterarithmeticals from mathematical presentations of spoof perfect numbers.

The vacuist sees the core argument against their position as resting on a claim of indispensability: Spoof perfect numbers *must* be defined using counterarithmeticals, hence if counterarithmeticals are vacuous then spoof perfect numbers are uninteresting. Let us grant for the moment that the vacuist has succeeded in undermining this indispensability claim. Does this eliminate this challenge to the vacuist position?

Talk of indispensability and of paraphrasing things away brings to mind the Quine-Putnam argument for the existence of mathematical objects, and nominalist challenges to the argument such as Hartry Field's *Science Without Numbers* (Field 1980). The presumption of this debate is that the goal for the nominalist is to come up with an adequate, mathematical-object-free formulation of our best scientific

theories. However, there is another strand of antinominalist argument, advocated by neo-Fregeans such as Crispin Wright, which questions whether even a successful recipe for paraphrasing away mathematical objects would bolster the nominalist position. Consider one area where it does seem to be straightforwardly possible to eliminate explicit quantification over mathematical objects, namely in using (finite) numbers to count objects. As is well known, the reference to number in an expression such as.

(17) The number of cows in the field is 3

and can be eliminated, using just first-order logic with identity, by re-expressing (17) as

(18) $\exists x\ \exists\ y\ \exists\ z[Cx\ \&\ Cy\ \&\ Cz\ \&\ x \neq y\ \&\ x \neq z\ \&\ y \neq z\ \&\ \forall\ w\ (Cw \supset (w = x \vee w = y \vee w = z))]$

The nominalist takes this to show that utterance of (17) does not ontologically commit us to numbers. The neo-Fregean reply is that if (18) really has the same content as (17), does this not – conversely – show that (18) *has* ontological commitment to numbers?

Returning to the case of spoof perfect numbers, the analogous question concerns the precise relationship between a counterarithmetical assertion about spoof perfect numbers, such as

(19) If 4 were prime, then 60 would be perfect,

and its "sigma-paraphrase"

(20) $\sigma^*_4(60) = 120\ (= 2.60)$.

The issue of what to say about the relation between (19) and (20) presents something of a dilemma for the vacuist. If, for example, she postulates some strong sort of equivalence between the two, then this would seem to conflict with her broader commitment to the vacuous truth of all counterpossibles. For this would in turn imply that, strictly speaking, all claims of the form of (19) are true, whereas clearly many assertions concerning the value of σ^* functions are false. In other words, the vacuist cannot rationally maintain both that spoof counterarithmeticals are equivalent to claims about sigma-functions *and* that all spoof counterarithmeticals are true.

A second, more appealing option, is to drop any claim of equivalence between (19) and (20) and say instead that (20) gives the literal meaning of what is expressed, in rhetorical and mathematically imprecise language, by (19). In other words, the model for the relation is more like that which putatively holds between an expression such as "Bob had butterflies in his stomach" and its more literal explication, "Bob felt nervous." Does this literal / metaphorical distinction give the right kind of relation between (20) and (19) for the vacuist's purpose?

The answer to this question depends partly on how the vacuist intends her claims to link to mathematical practice. On the one hand, she might be claiming that what mathematicians *really mean* when they assert a spoof counterarithmetical such as (19) is a claim about sigma-functions along the lines of (20). Following John Burgess's terminology (Burgess 1983) (in the context of the platonism-nominalism debate), call this *hermeneutic vacuism.* On the other hand, the vacuist might be claiming that what mathematicians *ought to do* when making assertions about counterpossibles is to replace claims such as (19) with their more mathematically precise analogs such as (20). Call this *revolutionary vacuism.*

Each of these alternatives faces some difficulties. The hermeneutic vacuist must face the fact that many appearances of spoof perfect numbers in the mathematical literature are not accompanied by any mention of sigma functions. Even more seriously, the original example from Descartes **predates** the development of the sigma-function formalism (by Euler) by over one hundred years! What has the hermeneutic vacuist to say about such cases, where mathematicians make spoof counterarithmetical claims without having any awareness of the sigma-function alternative? This would seem to indicate that mathematicians' understanding of spoofness is not *dependent* on translation into sigma notation. (Or should we say that Descartes was writing purely metaphorically in his letter to Mersenne . . .?)

The revolutionary vacuist faces a different sort of difficulty. For, unlike the hermeneutic vacuist, he is advocating that mathematicians change certain aspects of their practice. In particular, they should stop using counterarithmetical locutions in introducing and specifying spoof perfect numbers. But another way of viewing this call for revision is as a *conflict* between revolutionary vacuism and mathematical practice. And – although this is not a fatal flaw – this is enough to cast serious doubts on the cogency of this brand of vacuism.

There is much more to be said here on both sides, but hopefully enough has been done to show the connections between this issue and some broader debates in metaphysics concerning indispensability and ontological commitment. I shall close this section with a separate objection to the sigma-paraphrase strategy, namely that it fails to account for a certain kind of use to which mathematicians put spoof perfect numbers. In a 2003 paper, Pace Nielsen proves an upper bound on the size of an odd perfect number, expressed in terms of its number of distinct prime factors. At the end of the paper, Nielsen discusses Descartes' spoof odd perfect number (and shows its spoof perfectness using sigma-function notation). He then makes the following remarks:

> To demonstrate . . . the weakness of the bound above, we have the following: . . . If we restrict ourselves to numbers that can be expressed as a power of 22021 and other factors relatively prime to 19 and 61, and if we replace σ by σ^*, then the proof of the above theorem [on upper bounds] still goes through. So it should be true that for $N = 3^2.7^2.11^2.13^2.22021$ we have $N < 2^{1024}$. In fact $2^{1024}/N \approx 4.407 \times 10^{317}$. This demonstrates one of the difficulties in proving the non-existence of odd perfect numbers. (Nielsen 2003, p. 9)

What Nielsen does here is first show that the upper bound he has proved gives a very weak result for Descartes' spoof odd perfect number, since the bound is larger

than the number itself by a factor of 10^{317}. He then uses this fact to motivate the claim that the bound is likely also to be very weak for *nonspoof* odd perfect numbers. To be sure, this latter move is not a formal proof of the weakness of the bound. But the use of investigations into the properties of spoof odd perfect numbers to cast light on the properties of genuine odd perfect numbers (if any there be) is not uncommon in this area of mathematics. The question for the vacuist, then, is whether this sort of "heuristic" use of spoof perfect numbers can be justified once everything has been translated into sigma-functions.

There are at least a couple of reasons for thinking this might be a problem for the vacuist. First, the sigma-function approach is rather formalistic in nature. It is true that the relevant $\boldsymbol{\sigma^*_m}$-function can be precisely defined, for each composite number *m*, but why should the values of these $\boldsymbol{\sigma^*_m}$-functions have any bearing on the regular σ-function expressing the sum of divisors? Second, the sigma-function approach abandons any role for the property of being a genuine perfect number in defining the notion of spoof perfect number. This is in contrast with the counterarithmetical approach, which defines spoof perfectness by pointing to a counterfactual situation in which the given number is **literally** perfect. This encourages the view that spoof perfect numbers are closely connected to genuine perfect numbers – they almost were perfect! – and hence provides motivation for heuristic arguments such as Nielsen's which apply conclusions from one to the other.

7 Nonvacuism and the Evaluation Problem

Thus far, I have deliberately focused attention on the vacuist account of counterpossibles because the challenge posed by spoof counterarithmeticals in mathematical practice is most immediate for views which classify all counterpossibles as vacuously true. However, I think that the counterarithmetical examples presented earlier also offer resources for nonvacuist analyses of counterpossibles. In this section, I shall briefly outline why.

To make the discussion more concrete, I shall use a specific nonvacuist analysis to illustrate the general problem. Consider an analysis along the lines proposed by Daniel Nolan and others, which adds impossible worlds to the Lewisian framework (Nolan 1997). The idea is to keep the form of Lewis's definition of counterfactuals, but to save counterpossibles from vacuity by allowing their antecedents to pick out the closest impossible worlds. A counterpossible is false if its consequent is false at one of these closest impossible worlds. In allowing for false counterpossibles, this analysis makes room for nontrivial counterarithmeticals and thus for a mathematically interesting class of spoof perfect numbers. So far so good, but the nonvacuist is not yet off the hook. Ideally, we would like not just a framework for analyzing counterpossibles but also guidance about how to evaluate the truth or falsity of specific examples.

It is important not to set the bar unreasonably high here when we ask the nonvacuist for guidelines for evaluating counterarithmeticals. If the framework is one of impossible worlds, then counterpossibles will be evaluated at those

impossible worlds that are most similar to the actual world. Thus the request is mainly for guidance about how to assess comparative similarity. And the issue of how to cash out similarity is a thorny problem for possible-worlds approaches in general. Nonetheless, there are reasons for thinking that counterarithmeticals present two distinctive extra difficulties, stemming from their contradictory antecedents and from their mathematical subject matter. For the moment, I will focus just on the second of these features.

One approach to giving guidance is to provide rules of thumb for assessing relative similarity between two worlds. David Lewis provides some such rules in a 1979 paper (Lewis 1979, p. 472):

1. It is of first importance to avoid big, widespread, diverse violations of law.
2. It is of second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.
3. It is of third importance to avoid even small, localized, simple violations of law.
4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.

There are several reasons why applying these rules to counterfactuals with mathematical subject matter is problematic. For example, is there an intuitively clear divide in mathematics between "laws" and "particular facts"? Distinguishing laws in terms of their modal status clearly will not work, since "All even numbers are divisible by 2" and "7 is a prime number" are both necessary, if either is – nor will making a distinction in terms of generality, or being universally quantified in form. (Compare, "All numbers between 6 and 8 are prime," which is universally quantified but seems nonlawlike, and "0 is a natural number" which seems both particular *and* lawlike.) A second problem is how to make sense of – or provide analogs for – notions of spatiotemporal matching for mathematical facts. (See Baker (2003, pp. 256–260) for more discussion of how to apply Lewis-style rules of thumb to counterfactuals with mathematical subject matter.)

Perhaps, it will be possible to come up with rules of thumb for assessing similarity between worlds where different mathematical facts obtain. But at present, this is an open question, and – as such – a challenge to the nonvacuist position. Note also that to be useful in the context of evaluating spoof counterarithmeticals, such guidelines will need to answer quite specific questions, concerning, for example, whether a world in which 6 is prime and 6 is a factor of 12 is more similar to the actual world than a world in which 6 is prime and 6 is not a factor of 12. (The difficulty is that if 6 is a factor of 12, then (presumably) $12 = 2.6$. But also $12 = 2.2.3$, so this implies that uniqueness of prime factorization does not hold.)

It may be that some alternative, nonsimilarity-based analysis will be more effective in analyzing counterarithmeticals. Baron et al. (2020), for example, consider an approach inspired by counterfactual analyses of causation in which a distinction is made between "upstream" and "downstream" facts. They briefly consider a number of ways of trying to flesh out these notions in the mathematical case, where no causal notions are in play, but end up acknowledging that they are

leaving "the distinction between upstream and downstream facts at an intuitive level" (Baron et al. 2020, p. 8).

8 Conclusions

I hope to have shown in this chapter how the role of counterarithmetical reasoning in mathematical investigations of spoof perfect numbers presents both a *prima facie* challenge and a potentially valuable resource to philosophical theories of counterfactuals. To be adequate, any such theory must have something to say about counterpossibles – counterfactuals with impossible antecedents. If the analysis implies that all counterpossibles are vacuously true, then the challenge is to explain the systematic patterns of acceptance and rejection of different arithmetical counterpossibles by practicing mathematicians. If the analysis allows for non-vacuously true counterpossibles, then the task is to provide guidelines for evaluating arithmetical counterpossibles that are precise enough to underpin their role in actual mathematical practice.

References

Arnaldo J, Dris B (2017) The abundancy index of divisors of spoof odd perfect numbers. Indian J Number Theory 2017:13–26

Baker A (2003) Does the existence of mathematical objects make a difference? Australasian J Philos 81(2):246–264

Baron S (2020) Counterfactual scheming. Mind 129:535–562

Baron S, Colyvan M, Ripley D (2017) How mathematics can make a difference. Philosophers' Imprint 17:1–19

Baron S, Colyvan M, Ripley D (2020) A counterfactual approach to explanation in mathematics. Philos Math 28(1):1–34

Bennett J (2003) A philosophical guide to conditionals. Oxford University Press, New York

Burgess J (1983) Why i am not a nominalist. Notre Dame J Formal Logic 24:93–105

BYU Computational Number Theory Group (2020) Odd, spoof perfect factorizations. arXiv:2006.10697v1

Crubellier M, Sip J (1997) Looking for perfect numbers. History of Mathematics: History of Problems, Paris, pp 389–410

Dittmer S (2014) Spoof odd perfect numbers. Math Comput 83:2575–2582

Field H (1980) Science without numbers. Princeton University Press, Princeton

Guy, R. (2004) Unsolved problems in number theory. Springer, New York

Hàjek A (2016) Philosophical heuristics and philosophical methodology. In: Cappelen et al (eds) The Oxford handbook of philosophical methodology. Oxford University Press, Oxford

Hardy G (1940) A mathematician's apology. Cambridge University Press, Cambridge, MA

Holdener J, Rachfal E (2019) Perfect and deficient perfect numbers. Am Math Mon 126(6):541–546

Kvart I (1986) A theory of counterfactuals. Hackett, Indianapolis

Lewis D (1973) Counterfactuals. Blackwell, Oxford

Lewis D (1979) Counterfactual dependence and time's arrow. Noûs 13:455–476

Lewis D (1991) Parts of classes. Blackwell, Oxford

Lewis-Beck M et al (eds) (2004) The sage encyclopedia of social science research methods. Sage, Thousand Oaks

Maddy P (1997) Naturalism in mathematics. Oxford University Press, Oxford
Nielsen P (2003) An upper bound for odd perfect numbers. Dermatol Int 3:A14
Nielsen P (2015) Odd perfect numbers, diophantine equations, and upper bounds. Math Comput 84: 2549–2567
Nolan D (1997) Impossible worlds: A modest approach. Notre Dame J. Formal Logic 38:535–572
Ojeda A (2005) Are intensions necessary? Sense as the construction of reference, Semantics Archive
Reutlinger A, Colyvan M, Krzyzanowska K (2020) The prospects for a monist theory of non-causal explanation in science and mathematics. Erkenntnis. https://doi.org/10.1007/s10670-020-00273-w
Voight J (2003) On the nonexistence of odd perfect numbers. In: Katok et al (eds) MASS Selecta: Teaching and learning advanced undergraduate mathematics. American Mathematical Society 293–300
Wang P (2002). http://www.primepuzzles.net

What Mathematicians Do: Mathematics as Process and Creative Rationality

William Byers

Contents

1 Introduction: What Mathematicians Do 2290
2 Looking at Mathematics from the Perspective of Process 2292
3 Mathematics Is What *People* Do 2294
4 Understanding and Learning 2295
5 The Process of Reason 2296
6 Formalism as Process 2297
7 The Discovery of Reason by Doing Euclidean Geometry 2298
8 What Is Rationality? 2299
9 Intuition or Fast Thinking 2300
10 Reason or Revelation? 2300
11 Creativity and Insight 2301
12 Creativity as Reification 2301
13 Strong and Weak Objectivity 2303
14 Contradiction and Ambiguity 2304
15 Logic, Algorithms, and the Human Mind 2305
16 Conclusion 2305
References 2306

Abstract

Mathematics can be viewed through complementary lenses as structure or as process. The latter point of view which is taken in this chapter (and this volume) emphasizes "doing mathematics" and so is tied to learning and creativity.

Mathematics is not just any process but a rational one and this chapter investigates what is meant by reason in mathematics. A look at the way mathematics is actually done leads to an expansion of the usual understanding of reason by integrating of intuition and creativity into the rational process. This leads us to take another look at what we mean by objectivity in mathematics. We identify

W. Byers (✉)
Concordia University, Montreal, QC, Canada
e-mail: wpbyers@gmail.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_3

weak and strong varieties of objectivity (analogous to weak and strong AI) and claim that mathematics is objective in the weak sense but not the strong. As a consequence of this and the central place we give to process, mathematics is inescapably tied to human culture and thus never absolutely objective, true, or eternal

Keywords

Mathematics · Mathematical process · Rational process · Creativity in mathematics · Objectivity in mathematics · Weak objectivity · Strong objectivity · Ambiguity · Reification

1 Introduction: What Mathematicians Do

Mathematics is what mathematicians do! In this statement there is "what mathematicians do" and then there is the mathematics. They are two sides of the same coin, each defining the other, each side indispensable to the other. Without mathematicians there is no mathematics and vice versa. If you try to make logical sense of the statement, it looks perfectly circular, and since we have been trained to frown upon circularity our instinct is to reject it. We normally demand linear arguments from hypotheses to conclusion and this may be the way it needs to be in ordinary mathematical discourse but, in trying to discern what mathematics is all about, perhaps circularity is not the problem that it appears to be. In fact, circularity may actually point us in the right direction. It may help us to see that mathematicians and mathematics, mathematical process and content, are fundamentally intertwined, and thus that any definitive separation is ultimately artificial. They are two perspectives on one and the same phenomenon. It is a matter of convenience and perspective as to which one we choose to consider at any given time.

Remember that the great mathematician Bill Thurston said, "What we (mathematicians) are doing is finding ways for people to think about mathematics" (Thurston 1994), which is circular in much the same way as my statement above yet says something important about the interrelationship of mathematics and mathematical activity. After all mathematics is a creative process in which new results lead to new questions which lead to further progress in a kind of endless spiral.

I spent the good part of a book (Byers 2007) discussing the kind of situation in mathematics in which there are two equally valid perspectives that appear to clash with one another. In the present situation the two perspectives are process and structure. I referred to this kind of situation as ambiguous, an ordinary word that I appropriated to describe a fundamental aspect of mathematics that is normally overlooked. Normally ambiguity is viewed negatively and suppressed but I choose to emphasize its constructive aspects. Coincidentally a terrific book appeared in the same year which was also about such positive ambiguities (Grosholz 2007).

In this chapter I intend to investigate the process/content ambiguity of mathematics and the ambiguity of mathematics and the mathematician. There is a considerable amount of confusion concerning this situation that I hope to shed some light on. Mathematics is not just about stuff; it is also about how the stuff comes about, how it is understood, and how it is learned. Process and content appear to inhabit logically distinct domains but the point of view taken in this chapter is that they cover the same ground and thus should be thought of as complementary – two valid ways of describing the same thing. Each perspective brings its own insights.

However, these two points of view do not currently have the same status. For the most part, mathematics is considered to be identical to its content. Thus it is unconventional, to say the least, to consider the perspective that would take "process," the doing of mathematics, as primary or that we should consider process as being prior to content. Of course "prior" here does not necessarily mean chronologically so much as it refers to ontological status.

Let us consider for a moment what it would mean to prioritize the process dimension over the dimension of structure. It would root mathematics in the activity, the doing, or the learning. How would this work? For one thing it might force us to expand our definition of mathematics to include what today is called informal mathematics. In addition to the usual formal world that consists of objects and structures, there would also be an informal world that forms the ground out of which the formal is created. The formal, then, would be the crystallization of the informal. Objects would arise from process through what is called reification, which is, for example, the process whereby infinite summation becomes a real number or subtraction becomes a negative integer.

Of course, this is not the way we learned to think about mathematics. It is the opposite of Platonism where you have ideal objects at the most fundamental level. In fact it is not the way we think about science nor, for that matter, the way we think about the human condition. We tend to think of objects as primary and some of these objects as elementary, the "atoms" of our theory. We are most comfortable with these sort of "atomic" ways of organizing the world.

The objects whether they are actual atoms or else molecules, cells, people, or numbers are the primary focus. We start with elementary objects and then consider their relationship to each other and the ways in which they combine and interact to produce new, more complex, objects. That is the way we structure the world. Even our language (but not necessarily every language) is structured in this way. Things (nouns) appear to come first and thus have priority over actions (verbs). We imagine that the world has this structure and that language reflects this reality but it is also possible that the reverse is true, that language comes first and the structure of language determines how we think and has given us a theoretical approach in which the world and various scientific disciplines are structured in this, object-oriented and atomistic, way. For the most part this object-oriented way of looking at things has been handed down to us from the ancient Greeks.

2 Looking at Mathematics from the Perspective of Process

The alternative, a process approach, has something in common with the way one looks at things in quantum mechanics in which it is insufficient to think of an electron, say, only as a particle, that is, an isolated object with clear boundaries and a definite location in space-time. It takes the collapse of the wave function to "create" a particle, that is, bring it into being. The wave is the particle before the collapse of the wave function.

There is also a "collapse" in the process of reification. Process, which is open and fluid, collapses, or crystallizes, into the reified object which appears to have definite boundaries or a formal definition and therefore seems to be solid and permanent. The process world is chaotic and that is how things are "at the beginning," in the informal dimension of the mathematical adventure. Process, which has no permanent form, certainly exists, but not in quite the same way as form does. The manner in which it exists resembles in some ways the process of informal thinking or free association where ideas keep flitting by and we have to be very alert to grab the significant ones and write them down so as to give them a kind of permanence. We sometimes do this by giving names to the ideas that we deem to be significant, and this highlights the connection between language and the process of creating structure. When most people use the word "mathematics" they are normally talking about the world of mathematical structure. When I use the word I will mean both worlds: the fluid one of process and the solid quasi-static world of structure.

This fundamental assumption about how mathematics and, more generally, scientific disciplines should be organized is ultimately in question here. Perhaps objects are not the fundamental and permanent building blocks of reality. Objects change – even what we mean by specific mathematical concepts like continuity change; they come into being under certain circumstances and change or even disappear when they are no longer relevant. If change is fundamental and universal, then one key question would be the manner in which objects, including mathematical objects, come into being both for the individuals doing mathematics, for the larger culture of mathematics, and for society in general. Thus the manner in which a process becomes an object, reification, is central.

However, there is also the inverse process, namely, the manner in which structure disintegrates, is modified, or gives rise to process. That process can come from structure is simple. When you have numbers you can then do arithmetic. But how do structures change? How does a simple structure give rise to a more complex one? One way this happens is in response to problematic elements that inevitably crop up in any theory and only appear after the structure is investigated by the mathematician. Thus you cannot divide in the system of the integers. Or there are the famous unsolvable problems in Euclidean geometry such as the trisection of the angle. These problems reveal certain inadequacies of the system in question. Their resolution often requires the creation of a new system that extends the old. But before the new structure can appear it is necessary that the way of thinking that comes with the old system, and in which the theorems are considered to be absolute truths, breaks down. Only this provides a space in which a new way of thinking can emerge. In this new

theory the original truths are seen as relative (to the set of axioms used, for example) and not absolute. A good example of this is the movement from Euclidean to non-Euclidean geometry, or simply making ratios into rational numbers.

Thus there is reification, which is the way structure is constructed from process and what you might call "deconstruction," the melting down of old structures to make room for new. One goes from process to content, the other moves in the opposite direction, from structure to process and then to new structure. Both directions involve mathematical creativity, although deconstruction and reframing are not always recognized as fundamental aspects of mathematics. It is important to bear in mind that both construction and deconstruction are human activities. Reification involves the birth of concepts, and deconstruction is essential to reframing an old concept in a new way.

Let us think for a moment about what it would mean to take "mathematics is process" as a place to begin the discussion of the nature of mathematics. Of course, this would not mean that mathematics is only process. It means that you must start with process (and with an understanding of reason which I go into below) when you describe mathematics. You do not start with mathematical objects such as number systems, for example. You do not even start with axioms, not even with a system of logic such as the propositional calculus. These traditional approaches view process as, for example, the unraveling of the properties of mathematical objects that are implicit in an axiomatic system. Learning and creativity are bracketed off from "real mathematics" and only considered in separate disciplines such as education and psychology. Not even the history of mathematics gets treated as a "mathematical" discipline and so even advanced students tend to be quite ignorant of the history of their own discipline. We tend to ignore process in our teaching so students get little training in problem solving or even training in the process whereby one becomes fluent with a difficult mathematical concept. As a direct consequence of the mathematical community's approach to the nature of mathematics, process is neglected and even reason is not seriously investigated other than to identify it with logic. There is a sense in which all traditional approaches are fundamentally static.

Yet at its very origins mathematics is an activity: counting, measuring, comparing, estimating, and so on. For example, we might be interested in investigating "quantity" or "amount." We begin by trying to answer the question, "Which is more?" or equivalently, "Which is larger?" To consider such questions it is necessary to structure quantity by inventing a language. We require abstract carriers of quantity and that is how we get to the concept of "number." "Number" is a reification of the various activities that were mentioned above.

Even though it may not be possible to go back in the history of mathematics to isolate the first emergence of number, we can look at the development of the number concept in infants which will give us some hints. The situation is surprisingly complex. It seems that infants (and some animals) have a latent "numerosity," which could be defined as a rudimentary ability to recognize number, a built-in capacity to develop the number concept. This does not mean that this latent capacity need necessarily be activated in the same way that other genetic traits can remain latent until some environmental stimulus turns them on so to speak. So the capacity

for number leads to various activities that require an enriched and refined sense of number which leads to a more sophisticated number sense in a spiral of development that ends up with the very complex number systems with which that we are familiar.

It is interestingly that educators have long observed that students learn by doing. Then they learn to give names to bits of mathematical activity, creating in this way a structure which becomes the scene of further, higher level, activity. Thus a fraction starts as a problem in long division and ends up being a number just as the negative numbers arise from the process of subtraction. Thus processes get reified out as conceptual objects. Then in the formal recapitulation we begin with the objects and reintroduce processes of various sorts. We are reluctant to begin with process because objects seem so much more precise and easier to get your hands on. Of course a given process could conceivably be reified in more than one way and lead to different formal systems. Possibly we move in certain directions because of certain tendencies that are built-in.

There are other ideas that flow directly from ordinary activities that find their way into mathematics. Ideas such as "chance, "continuity," "nothing," and "infinity"? All of these arise out of activities, out of solving problems. These and a host of other questions begin at a level that is concrete and experiential; they arise in the ordinary activities of human beings.

Of course, thinking about mathematics in this way may make some people a little uncomfortable, maybe even bring on a feeling of disorientation. We have become habituated to the metaphor of mathematics as an enormous and intricate edifice of eternal truth. In this view even though mathematics may have been created by human beings, nevertheless it transcends the human condition and inhabits a mystical, Platonic realm that lies beyond time and space. We love the idea that even if individual human life is finite, the mathematics that we create will live forever because its truths are eternal and immutable. It is a comforting place to inhabit whereas mathematics as process brings us back to the contingency and uncertainty of life, something that may appear to recede as we contemplate the grandeur of the great mathematical theories.

3 Mathematics Is What *People* Do

To go back to the statement "mathematics is what mathematicians do," note that it has an emphasis on process but also that it emphasizes that mathematics is an activity that is carried on by human beings. The late neurologist Oliver Sacks said, "Science is a human enterprise through and through, an organic, evolving human growth,. . ." (Sacks (2019), quoted in the review of his posthumous book). The same could and should be said about mathematics.

It is particularly appropriate to look at mathematics from this point of view in this chapter since this volume celebrates precisely the change from structure to activity. This change can be traced back to two remarkable books: *Proofs and Refutations* (Lakatos 1976) by Imre Lakatos and *The Mathematical Experience* (Davis and Hersh 1981), by Philip Davis and Reuben Hersh. Both of these books turned our

attention away from the structure and content of mathematics and toward what is actually going on when mathematicians do mathematics. This shift involves moving from a view of mathematics as consisting of static, permanent structures in the form of definitions, theorems, and proofs but also as algorithms, and various methods of computation to a view in which mathematics is always changing and coming into being, to mathematics as a dynamic, creative, and thus human process. This necessitated a critique of the prevailing framework for mathematical activity, namely formalism and Platonism, and this volume is a search for a replacement.

4 Understanding and Learning

Thurston and many others have identified "understanding" as the basic mathematical activity. Mathematicians try to understand the mathematical phenomena that they encounter in their professional lives. But what do we mean by understanding? I recall a comment made by a student concerning some very dry derivation of the consequences of the axioms of a topological space, "I follow the argument but I don't understand it." Following involves a step-by-step process of verifying that each step follows logically from the previous one. Understanding is more global or holistic. It usually happens suddenly and may happen more than once for the same material, that is, there are depths of understanding. It often appears suddenly and one says something like, "Aha! I see what is going on here." Understanding a proof, for example, often means that the idea behind the proof is revealed. Even a fairly complex proof may be the elaboration of an idea which can be stated quite succinctly. But understanding a subtle mathematical idea like real numbers, continuity, or randomness is a significant accomplishment that, in a certain sense, is never complete. There is a continuum of understanding which moves from the first glimpses of understanding by the young student to the sophisticated understanding of the research mathematician.

It is worth emphasizing that both research and learning involve understanding. Doing mathematics involves learning (about) mathematics. We would have a much more profound appreciation for education if we realized that learning at the most profound level involves the same mechanisms as doing research. (Of course not the same level of complexity.) Learning is not the passive accumulation of facts. Learning involves mastering ideas, that is, conceptual understanding. Learning and understanding are both creative activities.

Though mathematicians and math educators do not normally communicate with one another because they live in different disciplinary compartments, educators are well placed to understand the difference between being content or process oriented. Educators and teachers need to focus on process because it is evident that learning is a process. Today, influenced by a view of knowledge as mere data, there is a tendency to think of learning as the accumulation of facts and techniques. But this is not what I mean by learning or doing mathematics and has little to do with becoming fluent with a particular mathematical world.

I can remember a first-year graduate student in one of my classes. I asked him what mathematics he had studied and he answered that he had read such and such a book on analysis and could quote the theorems and proofs from memory. Very impressive! Unfortunately he was unable to answer or even to approach the questions in the first assignment because it turned out that he had never actually done a problem before. This student knew no mathematics at all, and this illustrates what I mean by saying that mathematics is not primarily about content. Both learning and research are processes whereby human beings "find ways . . .to think about mathematics," where "thinking about" could be taken to mean "understanding."

The shift from viewing mathematics as structure to viewing it as process is a radical change in perspective. We were all brought up to consider the content of mathematics as what is real and defines the subject. It I as though every area in mathematics could potentially be written down in some sort of definitive textbook which, of course, would never be complete. In this view a subject would consist of a compendium of facts and research involves adding to this collection. In this view learning would involve getting familiar with those facts in the manner of the graduate student I caricatured. Of course this is not what learning means at all. It really means becoming fluent with the material, understanding the main ideas and concepts, and being able to apply them in an intelligent way. Just as fluency with language implies the ability to produce sentences that may have never existed before so fluency in mathematics means being able to work with situations that one has never encountered previously.

The shift from facts to process is crucial for the philosophy of mathematics because facts seem to be solid, objective, and potentially eternal. Process, on the other hand, carries the connotation of impermanence, of coming into being and then disappearing. Process also seems to have elements of subjectivity. There is always someone who is creating or learning. Of course we could imagine that the agent behind the process is an algorithm or computer program, which gets us into questioning the possible role of artificial intelligence in the process of mathematics. I shall have more to say about this later in the chapter.

5 The Process of Reason

Of course, mathematics is not just any process. It is a rational process and, in particular, it is an intelligent process. People *do* mathematics; they *learn* mathematics. So in order to find a way to talk about mathematics as process we must define what we mean by a rational process. What is reason? I will give a kind of definition a little later on but, whatever reason turns out to be, it must involve intelligence, that is, mind. And, as far as we know, this is why mathematics is a human activity – mathematics is about *people* learning to think about mathematics. As we go along we shall see that this implies what machines do and can do is not mathematics. Artificial intelligence is not rational intelligence; it merely mimics certain aspects of rationality.

It follows that what I have called elsewhere "the dream of reason" (c.f. How Mathematicians Think p. 12 or 84) does not hold, partially, as we shall see, because we misunderstand the word "reason." We identify reason with logic and divorce it from intelligence. Then we imagine that "reasoning" involves following the rules of logic, in particular, that it is rule-based in the same way that a computer algorithm is rule-based.

The traditional dream of reason is, of course, that we can produce mathematical and scientific theories that are certain, objective, and identical to what they model. In particular we imagine that a theory is static and eternally true. We feel that truths can never and will never change. In the dream of reason the theory stands alone liberated from reason, mind, and humanity. Understood in this way the dream of reason is a bad dream because the humanity of the dreamer has been lost or ignored. If this dream is not valid, then there is no "final theory" of physics, economics, or any other science; there is not even a theory of the stock market that works over an extended period of time. In our subject it is clear that mathematical activity will continue without end.

This chapter will propose a new way of thinking about rationality as process. On this basis we might be able to rehabilitate the dream of reason and reclaim it for humanity instead of a way of replacing human beings by machines and human reason by sophisticated, but intrinsically limited, algorithmic processes.

6 Formalism as Process

As a precursor to a discussion of the rational process let us consider the nature of formalism. Formalism identifies mathematics with a process, namely, the axiomatic method that operates within the context of a set of axioms, definitions, and operational rules of logic. It is still the dominant way of thinking about how doing mathematics works and so let us go back and revisit the nature of an axiomatic system and the process of working within it.

Here are a few questions regarding axiomatic systems that reveal the limitations of the way we normally think of it.

1. What does Thurston's comment mean when applied to an axiomatic system, that is, what does it mean to understand such a system as a whole or even to understand some particular theorem that follows from the axioms? In particular where is understanding to be found – is it part of the system or extraneous to the system? In other words, what is understanding and where is it located?
2. Given a choice of some particular set of axioms why are some results interesting while others are not? Why do we often make statements like "that result is deep" or "this result is trivial." Can the basis for the judgments "deep," "trivial," and "interesting" be found within the system or are they outside the system? Or are they merely subjective with no objective meaning? Are these judgments even part of mathematics at all? If they are not, then how do you decide whether or not to publish some article in a prestigious research journal?

3. Suppose you are trying to prove a result which has been found to be plausible empirically. Where does the idea for the proof come from? Does the mathematical idea exist within the system or is it extraneous to the system? In general, what is the relationship between the concepts and ideas of the subject and the axiomatic structure?

These questions and others demonstrate that when we discuss the axiomatic system we focus on certain aspects to the exclusion of others. There is a great deal more going on than mere deduction from axioms. Another way of putting this difference consists in saying that the rational process involves much more than applying logical inference to a set of a priori truths.

What I have to say will be clarified if we take a new look at what is happening in the axiomatic system. In order to make things concrete I shall take a quick look at Euclidean geometry as a prototype of an axiomatic system and give a more complete account of what is going on. What are the answers to the three questions I listed above in the case of Euclidean geometry? What, in short, does it mean to look at Euclidean geometry as a rational process?

7 The Discovery of Reason by Doing Euclidean Geometry

Euclidean geometry is one of the foundational elements of Western civilization and the root of the technological revolution that is currently sweeping over us. I am thinking about Euclidean geometry these days because I am teaching the subject to my 13-year-old grandson in order to give him a little intellectual enrichment. It is not taught in schools any more but friends and colleagues who are scientists and mathematicians all concurred that the subject had been important to them in their intellectual development. Why was that? In other words why study Euclidean geometry?

Most people regard Euclidean geometry as the prototype of a deductive system – definitions, axioms, and theorems deduced through pristine logical thought. They may believe that the subject is algorithmic so that it could be done by a computer. In other words they imagine that Euclidean geometry could be done without human intervention in exactly the same way that some people believe that AI systems can operate independently of human beings. So is Euclidean geometry a matter of pure deductive logic?

The significance of Euclidean geometry to the mathematician goes beyond its theorems. It includes the means through which these results are obtained. Euclidean geometry was such a significant part of the education of scientists, physicians, and mathematicians of my generation because it taught us how to think. In particular it showed us that thinking involves a deep connection between logic, intuition, and insight.

Even though Euclidean geometry is held up as the prototype of a deductive system, which may appear to be abstract, in reality it is concrete and geometric. It

is concerned with the structure of two- and three-dimensional (flat) space. In fact for the Greeks, Euclidean geometry *defines* space. It is not only about proving theorems but also about developing an intuition for the geometry of space and fitting these intuitions into a system of verifiable truths. Doing Euclidean Geometry involves logical thinking but it also involves the nonlogical procedures of intuition and insight that I will describe below.

Imagine that you are studying Euclidean geometry from scratch. Imagine that it is an area of research. "Doing" Euclidean geometry would work like this: First of all you have to think up some geometrical statement (or potential theorem) that is interesting and accessible on the basis of current knowledge (the axioms you have agreed on and the theorems you have established so far). (This is akin to deciding which hypothesis to test in a scientific experiment.) Then one has to guess whether the statement is true or not. If you guess "no," you try to find a counterexample. If you guess "yes," you try to construct a proof. The proof itself does not come into existence through a series of purely logical statements. A proof (or counterexample) is built around some idea (which may turn out to be a geometrical construction). In other words you have to know *why* it is true before you try to prove it. Actually the idea for the proof is usually what you mean by "understanding" the result.

Where does the idea or the understanding come from? Obviously it comes from the mathematician. It is only at this stage, after the idea has revealed itself to you, that you attempt to write down a proof and this is the only step in the procedure that is strictly logical.

Thus doing geometry involves **intuition** (developing the hypothesis and guessing true or false), **insight or creativity** (coming up with the hypothesis and the idea for the proof), and **logical argumentation** (for the purposes of verification and communication).

All three are essential parts of a deductive system, and this is how it feels to work on such system *from the inside*. This is how it feels to *do* mathematics. The further you go the less value you ascribe to the third, logical, part. Thurston said, "When the idea (behind a result) is clear, the formal setup is usually unnecessary and redundant." Creative scientists are basically interested in new insights, original ideas. Creativity is the very essence of science and mathematics.

8 What Is Rationality?

Let us *define* the rational process by abstracting my description of Euclidean geometry. Thus, *rationality involves intuition, insight, and logic.* Logic is just one of the steps not the whole ball of wax. This gives us a working definition of the rational process. Intelligence, on the other hand, is a concept that is very subtle and hard to get your hands on. Maybe we could think of it as the capacity for problem solving and thus for reason. I propose that we substitute "rational process" for "intelligence" whenever we can. (For example, maybe we should talk about "artificial rationality" rather than "artificial intelligence.")

9 Intuition or Fast Thinking

Intuition is thus an essential aspect of the process of reason. The Nobel Prize-winning psychologist and economic theorist Daniel Kahneman is famous for demonstrating that the economic actor is not only logical. In fact according to Kahneman human beings are capable of two kinds of thinking that he calls fast and slow. Slow thinking is what many people think of as logical thinking. They forget that fast thinking, which is involved in intuition is also essential to rational thought. Fast thinking is what the "leader of the free world" calls his "gut" and he is a disastrous example of the damage that can occur when fast thinking is unchecked by slow. You can read the history of rational thought as the attempt to control fast thinking with slow thinking but, in my opinion, that would be a mistake for reason has room for both of these modes. AI takes slow thinking, puts it into an algorithm, and uses a machine to speed it up. It remains a kind of victory of slow thinking over fast thinking but comes with a cost. Rational thinking involves a synthesis of intuition and systematic, rule-based procedures.

10 Reason or Revelation?

I have heard it said that the Ancient Greeks contributed their ideals of reason and harmony to modern civilization. The ancient Hebrews, on the other hand, contributed the ideal of revelation, which you can think of as the voice of God speaking to Moses from the burning bush but, more prosaically, as the discontinuous emergence of a new and powerful idea. In other words, the aha! experience. Anyone who has done any research in mathematics or even mastered some nontrivial mathematical concept like real numbers or limits will have had this moment when things began to come together and make sense.

The confusion I mentioned above concerning the true nature of the rational process comes from reducing reason to logical inference, on the one hand, and ignoring the role of revelation on the other. Logic is rule based, orderly, and predictable whereas creativity is not predictable and does not seem to follow any system of a priori rules. By reducing reason to logic we can easily conclude that the role of the human mind in mathematics could be replaced by some kind of artificial intelligence. The problem with this is that what I (a bit dramatically) called revelation is an essential part of the process of doing mathematics and even of learning mathematics (There is much more about this in my book, *Deep Thinking: What Mathematics Can Teach Us About the Mind*, World Scientific, Singapore, (2015).). The way toward a more complete view of mathematics involves accepting that reason is a lot more than logic and includes both intuition and insight. Only then will it be possible to respond to Poincaré's question – the perpetual mystery of power of mathematics to penetrate so deeply into the reality. The power comes from mathematics' ability to harness all aspects of the rational process.

11 Creativity and Insight

Creativity is not to be confused with the production of what is new. It involves insight, by which I mean the discovery of a new way to "see" some situation. We understand some event or situation by placing in a context, that is, framing it. Then the most basic creative act involves *reframing*, that is, coming up with a new way to understand a given situation or event. So "getting the idea" often means finding a new point of view from which to look at some situation. In science a frame is called a paradigm and reframing is referred to as a paradigm change. I am thinking, for example, of the way Einstein reframed our understanding of gravity.

But to remove the idea of creative reframing from the rarified atmosphere of genius let me supply a more down-to-earth example that applies to everyone.

According to developmental psychologists like Susan Carey (2009) children are born with two primitive but vastly different conceptual systems for number. An early conceptual task for children consists of combining these two systems into their first learned number system – the system of the counting numbers: 1, 2, 3,. . .. That development is a child's first creative learning experience in mathematics. However, because we are here concerned with the rational, I want to focus on the next step children take a few years later – the reframing of "number" from the frame of the counting numbers to the new frame of fractions or rational numbers. Initially fractions are introduced in terms of relative areas. We all remember that two-thirds was given meaning by dividing a pie into three parts and choosing two of them. Notice that this kind of example does not yet make two-thirds into a number but merely a ratio, that is, a relationship between the numbers two and three.

12 Creativity as Reification

The crucial but nonobvious question concerns why 2/3 is a number at all? If it is a number, then by what process does it become one? The child must come to see 2/3 as a single object, that is, she must reify the ratio of two to three into a new kind of number. When this happens with respect to two-thirds the child can do the same thing for other fractions. She has then undergone a total conceptual reorganization, a reframing of her understanding of number. Reification of pairs of whole numbers, which brings the fractions into existence, is nothing less than a paradigm shift and yet almost every child goes through this shift sooner or later. Of course we do not call it reframing, we call it learning. We have all been there but we have forgotten and so for us it is "obvious" that fractions are numbers. We forget what a significant development this was. But there is no reason a priori that fractions should be numbers that extend the system of counting numbers. Moreover there is no reason why the set of ratios should make up a new number system for which the old operations of arithmetic still make sense. You can see what a big deal it is when you think about the extravagant claims that the Greeks made for the rational number

system (which they seem to have identified with reason itself) and how profoundly shocked they were when confronted by the existence of irrationals. The creation of the rational number system in human culture or in the mind of a child is a prototype of an act of creativity.

One crucial point about this example. The two number systems, counting numbers and factions, are incompatible (or incommensurate, as the philosopher of science Thomas Kuhn would have it) with one another in the following sense. If you ask a child who lives in the world of counting numbers how many numbers there are between 2 and 3, she will say none. But a child who made the creative leap to the world of rational numbers will say that there are lots of them. And of course they would both be correct within their respective frames or conceptual systems. In other words correct and incorrect are *relative* terms, relative to context, of course.

Within a conceptual system like the integers or the rational numbers one insists on consistency and thus the absence of things like contradiction and ambiguity.

A well-designed theory has no place for such logically problematic things. However, it is precisely these problematic elements that drive paradigm change, the movement from one conceptual system to another. In Greek geometry there were "unsolvable problems" like trisecting an angle and squaring a circle. To solve them, you need to change the context of the discussion, that is, reframe the problem. An even more subtle problem concerned the status of the parallel postulate in Euclidean geometry. Questioning its status as an axiom and therefore "obviously true" leads ultimately to the development of non-Euclidean geometry.

One does not leave one's present paradigm willingly or easily; you need to be driven out by a problem that you cannot solve in the old system. For the move from the counting numbers to the rationals the problem might be that the division of whole numbers is not closed within the system of counting numbers. In other words though you can divide 7 by 3 in the counting numbers (and get an answer of 2 with remainder 1), the answer is not a number in the system you started with.

This leads me to say that *creativity transcends logic.* It involves the sudden leap to a higher point of view. A problem that was intractable in the original frame becomes transparent when looked at in the new.

One further comment is brought out by this example. In mathematics there are many kinds of numbers: counting numbers, rational numbers, real numbers, complex numbers, and many others. This whole hierarchy is built on the primitive idea of "number." Yet "number" in the abstract is never defined in mathematics. We all have a feeling for number because even a 6-month-old child has two separate conceptual systems for number. Nevertheless this "feeling for number," what I called numerosity earlier, is informal and thus never defined explicitly. The specific kinds of numbers I mentioned are, in comparison, well-defined. In this way the conceptual world of mathematics (and physics) emerges out of an informal world. Number is not special in that regards. What I said also applies to the concepts of time, space, energy, randomness, and many other concepts. In fact all of the building blocks of science have informal roots. Mathematics lives in both formal and informal worlds. Intuition functions very well in the informal world and creative reframing often has

to go back there but logical processes live exclusively in the formal world where things have explicit meanings.

13 Strong and Weak Objectivity

The power of mathematics and science lie in their claim to "objective" truth. In this section I would like to take a few paragraphs to repeat a point that I have made elsewhere (Byers 2017) concerning the nature of objectivity. Most people believe in a kind of objectivity which I shall call "strong objectivity" and define below, which is analogous to what is meant by "strong AI."

We sometimes say that something is "merely" subjective when we mean that it comes from personal prejudice or idiosyncratic opinion. Thus we would object to a mathematical theorem or scientific experiment being influenced by religion, race, or gender. Objective means that such matters, as well as many other cultural factors, do not influence the result. Of course some people consider mathematics itself to be a culture but within mathematics the criterion of independence from arbitrary opinion might serve as a minimal way to differentiate subjectivity and objectivity – we could call this *weak objectivity.*

However, there is another possible meaning of objectivity. Something is objective in this sense if it does not depend on mind or intelligence. You could say that it is objective in this sense if it would continue to be true even if there were no human beings around. This must have been the idea behind putting a diagram of the Pythagorean theorem on Voyager 1 in the hope that the truths of Euclidean geometry were universal and would be recognized by any intelligent being. Is the Newtonian theory of gravity objectively true? Most people would say that it is even if that truth is only an approximate one. Is the relativistic theory of gravity objectively true? Is it an eternal truth? Maybe it is but the jury is still out. At any rate a scientific theory is objective in this sense if the theory exists independent of the scientist who formulates or studies it. So some believe that the scientist *discovers* what is already there, that the rules of the universe and of mathematics are built-in, so to speak. Let us call this *strong objectivity.*

I believe that mathematics is objective in the weak but not the strong sense (This is what I take away from the writing of Reuben Hersh in books like (1980) *What Is Mathematics, Really?*, Oxford University Press, and is, I believe, the essence of the point of view that is celebrated in this volume.). This implies that the truths of mathematics are human truths not Platonic truths. So the truths of Euclidean geometry are objective in the weak but not the strong sense. In exactly the same way one could subscribe to the hypothesis of "weak AI" and not "strong AI." Notice that the "weak" position depends on an essential connection between human thought and science; the "strong" position on the other hand holds that once the theory (or technology) is established, human beings are redundant. One holds that the process of rationality can operate independent of human beings, the other that human beings are the essential measure of rationality for rationality is the process by which human being understand the world and themselves.

14 Contradiction and Ambiguity

Because we conflate reason with logic and are oblivious to the role of paradigm change in mathematics, we tend to be hostile to the presence of contradiction, paradox, and ambiguity within mathematics much less to be open to their constructive uses. Most people believe that mathematics and science are free of ambiguity, contradiction, and paradox and yet I wrote a long and well-received book about the constructive uses of these things. To give an elementary example, the number zero was unacceptable to the Greeks because of its contradictory nature – something that represents nothing as one author put it. The contradictory nature of zero remains but this has not stopped us from using it productively in science and technology. The same thing could be said for *infinity* (remember mathematics has been called the science of infinity), randomness, and many other basic concepts of math not to speak of the idea of complementarity in Quantum mechanics which makes what appears to be a paradoxical situation into a foundational principle.

The human mind has discovered creative ways to make use of contradictory elements. Indeed the legendary physicist Niels Bohr distinguished between two kinds of truth. "An ordinary truth is one whose opposite is a falsehood. A profound truth is one whose opposite is also a profound truth." Not only is he saying that something can be contradictory and true at the same time but he is further asserting that containing an inner contradiction may be an indication of its profundity.

Thus there is a vast difference between what I have been calling human rationality and logical, rule-based argumentation. The human mind is extraordinary! It can create something profound out of nothingness or chance. It can make constructive sense out of what appears to be meaningless, create something out of nothing. It can create regularity out of the very lack of regularity, that is, randomness.

This gaping hole in the prevailing philosophy of mathematics arises because of the view that identifies mathematics with its formal structure and ignores informal mathematics and process in general. In order to accomplish this it has been necessary to elevate rule-based logical argumentation to a kind of foundational principle, which is, of course, what formalism does. In this way process, creativity, and learning not to say history have been banished from the way the subject is described and understood.

A key element in all of this is a misunderstanding of the meaning and the role of ambiguity. The conductor and composer Leonard Bernstein once said that the richness of music is proportional to its ambiguity and the same can be said of mathematics. A measure of how deep a result is can be given by the number of ways it can be understood, the number of different contexts that it fits into. We continually look at particular mathematical concepts from multiple points of view. For example a matrix can be thought of as an object, a rectangular array of numbers, or a process, a linear transformation. Notice that a reified structure is ambiguous since it is both process and object. And you choose to look at it from one or the other perspective depending on your needs at any given time. In fact we think of the matrix in both ways at the same time and move freely from one context to the other. The point here is that the ambiguity, thought of as the ability to look at one situation from

a multidimensional point of view, contributes to the power and creative potential of an idea. When you start to think about things in this way one quickly sees that ambiguity crops up everywhere in mathematics and moreover that it is a generator of the process of deepening our understanding of a particular subject. Logic may have no place for ambiguity but rational thought uses it productively all the time.

15 Logic, Algorithms, and the Human Mind

Rationality is best understood as the cognitive function of the human mind. One isolates and abstracts some particular aspect of cognition, logic for example, and then the case is made that one can reconstruct all of human cognition from that one property. There was the famous attempt by Russell and Whitehead to reconstruct all of mathematics starting with the rules of logic. Of course that attempt failed but it was interesting and was used for later developments like the Incompleteness Theorem of Gödel. In the long run extreme reductionism never works. You cannot capture the full potential of the human mind using one particular model of mind. The mind is just too large and free for that to be possible. Logic and algorithms are one way the mind functions but not the only way.

The way to see this conclusively is to look at the history of science and mathematics. This history consists of the development of paradigms, conceptual systems, or ways of understanding aspects of the world. Inevitably, as the subject progresses, anomalies or inconsistencies develop. The response of the scientist is ultimately to develop a new theory that contains the old theory but expands it so as to explain the anomaly. Thus the nonrationality of the root of two is resolved by moving from the rational numbers to the geometric numbers, the problem of trisecting the angle or squaring the circle ultimately leads to the real number system. In just his way the problem of subtraction (2–5 is undefined in the counting numbers) leads to the development of the (positive and negative integers). The problem of division (is 123 divided by 47 a number?) leads to the rationals.

The process of (human) rationality has room for, and even thrives upon, the problematic in the form of ambiguities, contradictions, and even paradoxes. It would be a great shame and a loss to our collective humanity to reduce human rationality to a process that eliminates the problematic and focuses exclusively on one particular form of argumentation.

16 Conclusion

This volume commemorates a change in the philosophy of mathematics that is long overdue, a fundamental change in the way we think about the nature of mathematics. This change involves elevating the role we ascribe to process. Process, doing mathematics, then becomes part and parcel of what we consider mathematics to be. The mathematician cannot be separated definitively from the mathematics he or she produces. Far from living in some remote Platonic realm, mathematics is seen to

be a vital part of human culture and society. Thus we can and should study the interactions between the mathematical culture and the larger scientific and general cultures. One consequence of this change is that the world of mathematics becomes less isolated from the larger culture. In particular mathematicians would stop thinking of mathematics as a self-contained world of its own. At this time when we are both benefitting and suffering from the effects of digitalization, it is imperative that people who know something about mathematics take the time to think about the assumptions behind the technological revolution that we are living though today and, moreover, speak out about their conclusions.

These considerations lead me to emphasize once more the essential role of human creativity in mathematics and elsewhere. I will repeat that learning properly understood and the creation of new mathematics are both based on what is essentially the same mechanism – human beings infinite capacity for creativity and innovation. Reason is our birthright as human beings and reason is no more and no less than the means by which human creativity finds expression. Our job as mathematicians is to develop our own creativity, nurture it in our students, and resist the temptation to reduce the adventure of creativity to something mechanical and ultimately sterile.

References

Byers W (2007) How mathematicians think: using ambiguity, contradiction, and paradox to create mathematics. Princeton University Press, Princeton

Byers W (2015) Deep thinking: What mathematics can teach us about the mind, World Scientific, Singapore

Byers W (2017) Can you say what mathematics is? In: Srirman B (ed) Humanizing mathematics and its philosophy. Birkhäuser, Cham

Carey S (2009) The origin of concepts. Oxford University Press, New York

Davis PJ, Hersh R (1981) The mathematical experience. Birkäuser, Boston

Grosholz E (2007) Representation and productive ambiguity in mathematics and the sciences. Oxford University Press, New York

Hersh R (1998) What is mathematics really? Oxford University Press, Oxford

Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, New York

Oliver Sacks (2019) Everything in its place, New York Review of Books. LXVI(10)

Thurston W (1994) On proof and Progress in mathematics. Bull AMS 30(2):161–177

Platonism, *De Re*, and (Philosophy of) Mathematical Practice

Marco Panza

Contents

1 Preamble 2308
2 A Misconception 2310
3 An Ill-Posed Question 2313
4 Plato's Problem 2316
5 On the Platonist Side 2319
6 Philosophical Purposes 2320
7 Objects by Reification 2322
8 Object by *De Re* Epistemic Access 2327
9 History, Cognition, and Logic 2331
9.1 By History 2331
9.2 By Cognition 2332
9.3 By Logic (or Formalization) 2333
10 A Few Words, in Guise of Conclusion 2334
References 2334

Abstract

The chapter advances a reformulation of the classical problem of the nature of mathematical objects (if any), here called "Plato's problem," in line with the program of a philosophy of mathematical practice. It then provides a sketch of a platonist solution, following the same perspective. This solution disregards as nonsensical the question of the existence of abstract, and specifically

Thanks to Jeevan Acharya, Annalisa Coliva, Silvia De Toffoli, Michele Friend, Giovanna Giardina, Peter Jipsen, Alex Kurz, Abel Lassalle Casanave, Fabio Minazzi, Andrew Moshier, Michael Otte, Andrea Sereni, Stewart Shapiro, Daniele Struppa, Ed Zalta, two anonymous referees, and countless other friends and colleagues.

M. Panza (✉)
CNRS, IHPST (CNRS and University of Paris 1, Panthéon Sorbonne), Paris, France

Chapman University, Orange, CA, USA
e-mail: marco.panza@univ-paris1.fr; panza@chapman.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_56

mathematical, objects, by rather focusing on the modalities of our access to them: objects (in general, both concrete and abstract) are regarded as individual contents that we have (or can have) a *de re* epistemic access to. The question of the existence of mathematical objects is then replaced by that of the modalities of our *de re* epistemic access to individual mathematical contents.

Keywords

Platonism · Nominalism · Mathematical objects · Plato's problem · *De re* vs *de dictum* epistemic access

1 Preamble

Mentioning platonism and *de re* (attitude) in the title of a chapter of a handbook about philosophy of mathematical practice might surprise some adepts of the latter. While platonism might be seen as a metaphysical option, the distinction between *de re* and *de dicto* attitudes might appear as proper to the analytical philosophy of language. These two facts might be considered enough for making the philosophy of mathematical practice principally alien to these notions. In echoing this view, while reporting on a previous version of my chapter an anonymous referee has qualified it as "an ultra-conservative piece in the philosophy of mathematical practice." The reason for this is that it would have told "nothing about mathematical practice, divorcing itself starkly from the aspirations of some philosophy-of-mathematical-practice pioneers such as the PhiMSAMP (Löwe and Müller 2010) members, nor does it agree with Mancosu's (2008) conservative understanding of PMP as exploration of new philosophical questions about mathematics."

This report made me decide to write the present preamble, in whose absence some readers might have been tempted to agree with this judgment, by misconstruing the very purpose of my chapter.

It is certainly true that I hereby explore a traditional (or "ultra"-traditional) question for the philosophy of mathematics: so traditional to date back to ancient Greek philosophy. However, being new is a required value neither for good philosophical questions nor for fruitful philosophical approaches. A philosophical investigation is often made valuable and innovative because of the way traditional questions are restated and tackled and of the tentative answers that stem from time-honored approaches. It is certainly not for me to say whether the investigation I present here is a valuable and innovative one. What I can say is that it aspires to be so because of this reason: since it aims at using a well-established distinction, as that between *de re* and *de dicto* attitudes, to answer a traditional question in a purportedly original way.

It is also certainly true that I hereby explore no specific aspect or episode of mathematical practice. Promoting and arising similar explorations is not, however, the only purpose of philosophy of mathematical practice (as I understand it, at least). Another equally important purpose is accounting for some essential general feature

of this practice, by so insisting that the term 'mathematics' in 'philosophy of mathematics' should stand for it: for the human activity of doing mathematics. This activity has social and material aspects, but it is not because of them that it is mathematical. What makes it so are the modalities of its being an intellectual activity. This makes it necessary for the philosophy of mathematical practice to engage with the purpose of accounting for the specific feature of that intellectual activity that we perform by doing mathematics, namely, with the question of wondering what makes an intellectual activity mathematical in nature. This is just the question I explore in this chapter (even if only partially, of course). And it is just because of such that I take this chapter to be, in the full sense of the term, a piece of philosophy of mathematical practice – whether (ultra-) conservative or not, I leave to the readers to decide.

More than having a history, mathematical activity is (or generates) a history. Since I in no way think that the philosophy of mathematics should have any normative duty concerning mathematics itself, I take this history to be, properly speaking, what the term 'mathematics' should stand for in 'philosophy of mathematics.' Establishing what and how mathematics should be is a question for professional mathematicians, not for philosophers. Even less, it is a duty of philosophers to establish whether something is there or not. In particular, philosophers of mathematics are entitled to decide neither what is actually there (or has happened and happens) that one would consider as mathematical, nor what is genuinely mathematical and what is not so, among all that is there (or has happened and happens) that mathematicians and professional historians of mathematics call by this name.

In my view, philosophy essentially aims at accounting for independent realities (unless it is some sort of metaphilosophy, which it might of course legitimately be, in some cases). This is even obvious, it seems to me, when it is the philosophy of something, as the philosophy of mathematics clearly is, since, in order for philosophy to be the philosophy of something, this something should be available to its inquiry. This inquiry can (or even cannot avoid to) shape this something, just as science does for the reality it studies. But this shaping is just an essential part of accounting for it, of explaining how it is, and of spelling it out, not of establishing how this should be, or what is there to form this reality, and what is not there.

Arguing from these theses is, in my view, just the same as promoting a practical turn in philosophy, if there is one that one might promote. Hence, if philosophy of mathematical practice is to be part of this turn, it has to tell us of which something that is there and is called 'mathematics' it is intended to be philosophy of. When I say that the term 'mathematics' in 'philosophy of mathematics' should stand for the historical fact of mathematical activity, I simply mean that this something should be nothing but this historical fact: what this activity has been and is today. This fact makes the intersection of history and philosophy of mathematical practice not only quite strict, but also essentially constitutive: The philosophy of mathematical practice is, to my mind, the philosophy of a historical fact.

Still, there is no need, I hope, to refer to Croce (the historicist philosopher *ex excellentia*) for remarking that history (*res gestæ*) is not the same as historiography (*historia rerum gestarum*). This constitutive connection does not make philosophy

of mathematical practice the same thing as the history of mathematics, in the usual sense where 'history' wrongly stands for 'historiography' (and, then, philosophers of mathematics the same thing as historians of it). History (in this sense, i.e., historiography) reconstructs relevant facts and accounts for their specific and detailed aspects. Philosophy accounts for general features, the way of being, and the peculiarity that makes them mathematical facts. The former can be done without the latter (though possibly not very lucidly). The latter cannot be done without the former, since the former provides to the latter with the reality which it is the philosophy of (since history, properly called, only appears to us through the lens of historiography: We have no direct access to it). This is another way to say, in the language of Lakatos's *dictum*, that historiography of mathematics without philosophy of mathematics is certainly blind, but can still subsist, while the philosophy of mathematics without the history of mathematics (and therefore its historiography that provides it to us) is so empty that it cannot subsist unless as a caricature of itself.

Another important clarification is suggested by the same referee of a previous version of my chapter. What I am arguing for is not at all that the philosophy of mathematics cannot but consist of a collection (or even an organic system) or case studies. My point is precisely the contrary. A collection (or system) of case studies cannot but be what it is, namely, a more or less ordered configuration of particular inquiries devoted to single neighborhoods in the space of history. Philosophy is much more and much less than this: It is an account of the space itself. While the former accounts for some elements in a quite large set, the latter defines a structure based on this set and studies it as such. This is why the reader will find no case study in what follows (while some of them can be easily found in other, more properly historical essays I have written), but rather a reflection of what makes the historical fact of mathematical activity mathematical in nature.

I conclude my preamble with a proviso. Though I will be naturally brought to mention a few philosophers of mathematics and some of their works, my purpose here is not that of discussing the opinion of this or that colleague or master. Rather it is that of promoting a general philosophical option, and rejecting some others, by presenting the former in my personal (and I hope, original) guise. Any reader that would like to have more bibliographical information about the present discussion in the philosophy of mathematics that I shall deal with should look for it elsewhere. I contributed to offering a survey of some of these discussions (in Panza and Sereni 2013), and (apart fort the most recent developments, elapsed after its publication) I can certainly refer the reader to this work.

2 A Misconception

Philosophy goes with many temptations. And the philosophy of mathematics is not without them, too. The most dangerous one is taking the latter as a mere extension of the former, due to the possibility of using mathematics as a testbench for philosophical theses, as a storehouse of examples and problems on which to measure the mutual merits of opposite philosophical views. There are two kinds of philosophers: those that have problems and look for solutions for them, and those that have

solutions and look for problems to solve them with. The philosophers of the latter kind are victim to a misconception, I contend. The dangerous temptation I am alluding to is a form of this misconception. One of the ways it takes place in the philosophy of mathematics is using mathematics and/or mathematical examples to measure the mutual merits of realism and nominalism, understood as general metaphysical theses, the former asserting and the latter denying the existence of abstract objects.

In the course of this chapter, realism, or, more precisely, the particular version of it that is often called 'platonism' – which I write, as is often done (following Parsons 1964, p. 275), with a lowercase 'p,' for disguising it from other theses more faithful, also in their intention, to Plato's philosophy – is presented as the thesis that "there exist mathematical objects" (Balaguer 1998, p. 1). More precisely, it is the thesis that, independently of our course of action and thought, there exist abstract objects in their own right, which are intrinsically mathematical in nature, and that mathematics deals with them, and aims at describing them, by so revealing how they are and how they are related to each other. Mathematics would, then, be successful when what it asserts of these objects is true, that is, when it truly describes them and their world as they are, in fact. It follows that a mathematical theory is good if and only if its theorems are true of the relevant objects. Among these objects, there are sets, numbers, and geometrical figures, or, possibly, only sets, starting with the empty one, according to the most reductionist ones among these realist self-styled philosophers of mathematics.

These, or other, less radical, partisans of ontological reductionism are certainly ready to admit that good mathematical theories can include theorems that are not literally true, or are so only if the relevant terms are properly understood according to appropriate reductionist clauses. Still, once the reduction is finally complete, or these clauses are fully operative, there is no more need for rephrasing: The theorems of these theories plainly mean what they appear to mean *prima facie*, and what makes them true of the relevant objects is that these objects just are and stand to each other like these theorems literally say that they are.

Nominalism is presented as the opposite thesis that denies the existence of abstract objects, and, *a fortiori*, of mathematical ones. It denies, for instance, "that numbers, functions, sets, or any similar entities exist" and, by consequence, "that it is legitimate to use terms that purport to refer" to them (Field 1980, p. 1). To avoid the most classical of paradoxes, this cannot but be understood as the thesis that the concept of an abstract object is empty, which immediately entails that this is also so for the concept of a mathematical object (provided it be admitted that mathematical objects could not be but abstract). These two concepts cannot be taken as ill-formed of course, since, if they were so, it would be hard to sensibly claim that they are empty. The same is even more clearly the case for more specific mathematical concepts, as that of a number, a set, or a function. These concepts are so defined to make some particular mathematical items purportedly fall under them. It follows that either a nominalist is prepared to also argue that these items are not objects (but, then what?), or cannot avoid to conclude that the usual mathematical language is both well-formed and empty, making mathematical statements either literally false, or only vacuously true.

Supposing that the doubtful distinction between mathematical items and objects is dropped out, it follows that appropriately understanding a mathematical statement depends on two possible moves, locally alternative to each another, but possibly operating together within a more general program aiming at a global understanding of mathematical discourse. On the one side, one might require, in a Quinlan vein (inaugurated in Quine 1939, but pursued in many subsequent works), that mathematical statements be paraphrased as statements on nonabstract, and then nonmathematical, objects. Once the paraphrase is complete, these statements would become, finally, literally true or false. This would allow one to say that a mathematical theory is good if its theorems admit a paraphrase making them literally true, but no more mathematical, statements. One might also require, on the other side, as suggested in Field (1980), that mathematical theories be added to empirical ones, which only deal with concrete objects to generate conservative extensions of them. Though remaining literally untrue and not necessarily admitting any paraphrase making them true, the theorems of a good mathematical theorems would then have a very welcome virtue: that of making it easier (or even practically possible) to establish empirical truths.

Both the partisans of platonism and nominalism so conceived may easily find, within the mathematical storehouse, many examples appearing, at first sight, favorable to one or the other party.

A nominalist might easily observe, for example, that a sentence like 'Mars has two moons' can be faithfully rephrased as 'There are x and y such that they are distinct to each other and both moons of Mars, and for any z, either it is not a moon of Mars or it such that either $z = x$ or $z = y$', where no reference is made to any number. The use of the term 'two' might, then, be taken as a *façon de parler*, simply used for short.

A platonist might concede this, after all, but observe that the former sentence does not express, in fact, a (purely) mathematical statement, and that the simplicity of such an example vanishes, when we speak of numbers as such, rather than of the numbers of the moons of Mars, or of pears or apples. When we say, for instance, that 2 plus 3 is 5, we do not speak at all of two particular things and three other particular things that, when put together, result in five things. At most we speak, in agreement with def. VII.2 of Euclid's *Elements* (Euclid, TBEH), of two units and three units, that, when put together, result in five units. But units are certainly not, as such, in all their generality, concrete objects. This makes it impossible – the platonist might conclude – to follow the same strategy used above to restate the sentence to avoid any reference to abstract objects. Indeed, this would make it loose its generality, which is a salient feature of mathematical statements.

The nominalist can try to replay by suggesting a new paraphrase that would be too long to write here. It will be enough to say that it might try to render the generality of the original sentence by replacing the reference to units either by appropriate universal quantifiers ranging on the totality of concrete objects, or by a talk of concrete objects whatsoever. At this point, the platonist might either present another simple example, as that of the sentence there are infinite prime numbers, by challenging the nominalist to suggest a faithful nominalist paraphrases, or, go for a much more general strategy and remember Plato's distinction between vulgar and

knowledgeable (or philosophical) arithmetic, respectively concerned with "unequal units [μόνᾳ ἄνισος]," such as "two armies" or "two oxen," and with "countless units [μόνᾳ μυρίαι]," none of which differs from each other (*Philebus*, 56d-e; see *Theaetetus*, 195e–196a). The nominalist might retort, in turn, under Quine's authority (Quine 1986, p. 400), that the latter arithmetic is merely "recreational" and should in fact be ignored by naturalized (and, then, good) philosophy, to the effect that an example as the last one just mentioned should be taken as perfectly immaterial for such a (good) philosophy.

This last move might also open the road for the second nominalist strategy, in which the recourse to paraphrase is replaced by the talk of conservative extensions of scientific, and, then, nominalistic, theories. Faced with this strategy, the platonist would have no reason to surrender, since there would be still room to argue: (i) that scientific mathematized theories are far from nominalist, in fact, since mathematics is constitutive of its objects; and (ii) that pure mathematic is far from recreational and should constitute the appropriate text-bench for deciding the realism versus nominalism dispute.

Who is right? It might appear that there are good reasons for both sides.

For instance, there is no need to invoke the number two, as an abstract object, to count the moons of Mars if counting objects (or transitive counting: Benacerraf 1965, pp. 49–50) is understood as such, that is as an intellectual operation on these objects (rather than as the establishment of a bijection between them and the numbers), which is all that is needed for all scientific purposes. There is no more any need to invoke numbers to decide whether there are more planets or moons in the solar system. Using them is surely quite convenient; it makes things simpler. However, it is not indispensable. If one had to work on much bigger collections, one would probably need to pass through numbers, but this would be only a practical need. In principle one could avoid them. This might appear to be a strong enough reason to think that number words enter our statements concerned with these empirical practices only as useful, but dispensable, *façons de parler*, as the nominalist suggests.

But is it the use in such empirical practice that decides the nature of mathematics? For making the realist win over the nominalist on the testbench of mathematics, it does not seem at all necessary that no occurrence of mathematical words might be explained as the nominalist suggests. It seems enough that some occurrence cannot be so explained. But then, the case of the infinity of prime numbers might become decisive: Either the nominalist can explain it nominalistically, or the platonist has a point, after all, unless, in a last and decisive gasp, the nominalist succeeded in excluding this case from the relevant testbench.

3 An Ill-Posed Question

But, then, again: Who is right? Neither one, I contend. Since the question the two parties offer opposite answers to is simply ill-posed.

Even supposing that the realist versus nominalist dispute has any sense, in general, it does not extend to mathematics in these terms, at least if such an extension has to take the form of a philosophy of mathematics. Since, if philosophy of mathematics is to be about mathematics, it has to start from mathematics itself, and reflect on it, rather than begin elsewhere and come to mathematics by accident, by reproducing on it the same general questions and answers already looked over outside it, with the mere replacement of schematic place holders with mathematical ingredients. In other terms: It is not enough to replace 'abstract' with 'mathematical' in the general question 'do abstract objects exist?' to get a sensible question for the philosophy of mathematics.

Let us be a little bit more precise. Let us assume that the word 'abstract' is the only schematic word in this general question. This means that the replacement of this word with any other word of the same type within this question is not to go together with a corresponding replacement of any other words such as 'objects' or 'exist,' or, at least, with a corresponding particular understanding of them. This being granted, suppose, for the sake of the argument (and only for that), that, when it is so understood, such a general question has a clear sense. The point is then the following: Also under these provisos, it would still remain that the mathematical instance of the question – 'do mathematical objects exist?' – is not a genuine question for philosophy of mathematics, or, at least, it is not so if it is also admitted that philosophy of mathematics is actually to be about mathematics.

Of course, philosophers are perfectly allowed to have no particular interest for mathematics, as well as for any other field of knowledge. But we, philosophers of mathematics, are also allowed, and to my mind, even required, to regard the philosophy we practice as a genuine enterprise only if it reasons on mathematics as an external material that it takes as given as such, before beginning its work.

If this is admitted, there is room to argue that the very question 'do mathematical objects exist?' is nonsensical. Or, at least, it is so, if mathematics is supposed to be given, and it is granted that the words 'objects' and 'exist' designate primitive notions, asking for no appropriate elucidation depending on mathematics itself. Consequently, the connected question 'is mathematics about its own objects?' also becomes nonsensical, under similar conditions.

This makes any putative answer to these questions, advanced in agreement to such conditions, not only doubtful, but inevitably misleading: a claim that is not so far, after all, at least for its form, from Kant's claim about dogmatic metaphysics.

A symptom of this nonsensicalness is the difficulty of answering (and in agreeing on the good answer) to a connected question: 'what could be conceived of as an evidence for the existence, or nonexistence, of mathematical objects?' If mathematics is taken as given, it should include such evidence. Of course, this might be very well hidden, quite far from the surface that immediately appears to an inexperienced observer. This would make discovering and exhibiting such evidence quite hard. But it should be, in any case, possible to clearly say how it might be like. Since the question 'do mathematical objects exist?' presents itself as a factual question, deciding which is the right answer to it should be a matter of fact. Moreover,

assigning a clear meaning to this question, by making it sensible, should require clearifying how this matter of fact could become manifest to our understanding. We should, in other terms, be able to say what would make us unquestionably aware that mathematical objects do or do not exist.

I know only two attitudes face to this last crucial question, which might appear plausible, at last at first glance.

The first is only available to nominalists: The burden of providing evidence for deciding about the existence of mathematical objects is only on the platonist side, since platonists are the only ones that make a positive claim. Here is how the point is made by C. Cheyne (2001, p. 108), for instance,

> [. . .] it is a contingent matter whether numbers exist with mathematical necessity in the actual world. The claim that this could be decided by purely a priori methods becomes dubious. Discovering whether or not this world is favored by the existence of entities that might not exist would seem to be a matter of observation. But how could we observe such entities if they lack causal powers? The burden of proof is [. . .] on platonists [. . .].

The nominalist's burden is then nothing but that of accounting for the available evidence in a way that does not suppose the existence of abstract, and, then, mathematical objects. This is not properly an answer to the question; it is instead a move for turning the absence of any answer in favor of the nominalist side. But the move is weak enough. For it turns this absence in favor of an agnostic attitude, at most, by so making clear that the nominalist thesis is, in fact, a *petitio principi*. The nominalist might retort by appealing to the Ockham razor, and claim that, in the absence of any positive evidence, all that can decide whether mathematical objects exists or not is whether it is necessary to postulate their existence in order to account for the available evidence. However, this inverses the burden of proof, since it now requires the nominalist shows that this is not necessary. In J. Burgess' words (Burgess 1983, pp. 95–96):

> Actually, the burden of proof is on such enemies of numbers [. . .], to show either: (*a*) that science, properly interpreted, already does dispense with mathematical objects, or (*b*) that there are scientific reasons why current scientific theories should be replaced by alternatives dispensing with mathematical objects.

The problem becomes, then, that the nominalist account, both in terms of paraphrases and in terms of conservative extensions, is partial, at most, if not largely unsatisfactory. Possibly, it accounts for some scant evidence. But it does it at the price of declaring this evidence untrustworthy, that is, of tacking it as a fake appearance of a hidden reality that our language is often incapable to describe faithfully, or can only describe in a quite entangled and odd way.

The second attitude is that of some platonists: The required (positive) evidence is provided by the truth of some appropriate mathematical statements involving terms that purportedly denote mathematical objects. In B. Hale and C. Wright's words (Hale and Wright 2001, p. 8):

> [. . .] no more is to be required in order for there be a strong prima-facie case that a class of apparent singular term have reference, than they occur in true statements free of all epistemic, modal, quotational, and other forms of vocabulary standardly recognized to compromise straightforward referential function.

The point here is that a statement about certain objects is true only if these objects exist. Fair enough. But this leaves other questions open: What evidence do we have for the truths of the relevant mathematical statements? What evidence do we have that these statements are actually about the objects they seem *prima facie* to be about? Even more crucially: Are we really sure that mathematical truths require existence?

The second of these three questions brings grist to the mill of nominalists. Considering the third one would bring us too far from the purpose of the present chapter. So, I limit myself to the first. A possible answer to it is that the relevant truths are logical truths, or, at least, logical consequences of so innocent assumptions that we could not reject them as untrue. This was essentially Frege's strategy, and, in a different version, is still the neologicist one (as defended in Hale and Wright 2001, for instance). But it has problems, in turn, and not only because of the difficulty of deriving suitable mathematical statements from consistent systems of innocent enough assumptions. But also, among other things, because of the doubt one can raise about the fact that the innocence of these assumptions entails their truth, or, better, an appropriate kind of truth for supporting existence. One might, of course, redefine existence in terms of (the relevant kind of) truth. But this would violate the condition mentioned above, about the presence of a single schematic word in the question about the existence of abstract objects, making the platonist thesis essentially different from a simply application to mathematics of the realist one. Under this guise, the thesis would rather involve a claim about a peculiar notion of existence, proper to mathematics. It should, then, be supported, not by a general metaphysical attitude, but by a specific inquiry about the nature of mathematics, which is exactly what I am encouraging.

4 Plato's Problem

I can stop the *pars destruens* of this chapter here. It was aimed at critically describing an attitude that no sort of philosophy of mathematical practice might share, to my mind. Still, arguing that the philosophy of mathematical practice might not share this attitude is not the same as arguing that the question it deals with, though nonsensical, as such, is not rooted in a relevant and perfectly sensible genuine problem about mathematics (which philosophy of mathematical practice is certainly not only entitled, but also required to tackle). What I wanted to suggest is that the trouble with this attitude is much less due to the opposite answers that it suggests for this question, than to the question itself, better to the way this question renders the crucial problem that stays at the roots of it. The point at issue is that this way makes this question essentially external to mathematics, by so making hard to decide which evidence, coming from mathematics itself, might count as a justification for an answer to it, whatever this answer might be.

In Panza and Sereni (2013), A. Sereni and I designated the problem I am referring to as Plato's one. A way to address it is by quoting a famous passage by the *Republic* (527a–b), in an appropriate translation:

> This <. . .> will not be disputed, at least, by those that <are> even a bit acquainted with geometry, that this science [αὕτη ἡ ἐπιστήμη] is in full opposition [πᾶν τοὐναντίον] with the language spoken within it by its adepts <. . .>. Since <the way they> speak <is> as much ludicrous as necessary [μάλα γελοίως τε καὶ ἀναγκαίως]: they shape all their discourses as if they were operating in practice and in the purpose of practice [ὡς γὰρ πράττοντές τε καὶ πράξεως ἕνεκα πάν-τας τοὺς λόγους ποιούμενοι]; [they] speak of squaring, applying and adding, and they all say this way, whereas all <their> learning is somehow suited for the purpose of knowledge [τὸ δ' ἔστι που πᾶν τὸ μάθημα γνώσεως ἕνεκα ἐπιτηδευόμενον]. <. . . > [This is] knowledge about that which always is, and not that which now comes into being and now perishes.

Contrary to the way this passage has been often understood (and consequently translated), this is in no way a criticism of the language and practice of geometers (the point has been firstly made *mutatis mutandis* by M. F. Burnyeat 1987). Plato notices that this language seems to be shaped "πράξεως ἕνεκα [in the purpose of practice]," whereas geometrical learning is "γνώσεως ἕνεκα [for the purpose of (theoretical) knowledge]" of eternal (and abstract) matters. He takes this language to be not only "ludicrous [γελοίως]," but also "necessary [ἀναγκαίως]." He seems then to acknowledge that there is no other option than using such a practical, ludicrous language for theorizing about such an abstract matter.

This is exactly the opposite of the nominalist claim: Mathematics does not use an abstract language for speaking of concrete objects, by calling for appropriate nominalist paraphrases; it rather speaks of *abstracta* as if they were material artifacts, by so calling for an appropriate understanding. The ludicrous (or metaphorical) aspect of its language does not consist in its surreptitiously dealing with concrete objects by *prima facie* speaking of abstract ones, but in its (having no other option than) using a language originally fashioned to speak of the former objects for reasoning about its abstract matter.

But this is no more the same as plainly asserting that there are abstract objects that mathematics is about and call with their proper names. If it were so, there would be no opposition, no necessary ludicrousness. At most, the problem would be that of inventing a new language, just fashioned for speaking of these abstract objects, just as the vulgar one had been invented to speak of concrete ones.

According to Plato, mathematics is certainly not about "ideas [εἴδη]." It cannot depend on the exercise of νοῦς but rather depends on that of διάνοια, or, possibly φαντασία, as Proclus will say later. The soul can contemplate, at most, the idea of a triangle, or that of a number, possibly those of even and odd, but mathematics requires more, or less, than that: It requires triangles, and numbers; not single ideas, but pluralities of items. What do these pluralities consist of? How can we conceive of these items? More in general: How can we account for the intellectual exercise that mathematics pertains to?

Such is Plato's problem. And this should be the ontological problem of the philosophy of mathematics. Not whether there are μαθηματικά or not. But how

can we account for our reasoning on them? Since it is of them that mathematics is reasoning, unless one wanted to argue, *contra* Plato, and because of a biased and preestablished misconception, that its form is fictitious and is to be amended by true philosophy (which would just be the opposite of philosophy of mathematics, being rather philosophy against mathematics).

Take this statement 'Any vector space has a basis'. What is it speaking of? The most obvious answer is that it speaks of vector spaces and its bases. Does it not? Why not? Possibly since vector spaces must be abstract objects and abstract objects do not exist? I suggest a nominalist supporting this answer to explain it to any working mathematician, and consider the reaction. I suspect that it would be better not to describe it. But, does this mean that the platonist described above has finally won? Does this prove that vector spaces exist and have bases? And that there is nothing more to say on that matter, except, possibly, for adding that this makes the statement true, and its proof a justification for its truth?

Suppose to work within ZF, i.e., whiteout the axiom of choice. The proof that any vector space has a basis does no more hold. But what does the axiom of choice say? Let us consider one of its most usual formulations (all equivalent to each other in appropriate frameworks): "For any set of pairwise disjoint non-empty sets, there exists at least one set that contains one and only one element from each set in the former set." What is it speaking of? Again, the most obvious answer is that it speaks of (nonempty) sets, or, more specifically, of sets of (nonempty) sets. Our platonist would say that these sets exist (let us avoid asking him/her whether this makes the set exist formed by all sets). But, then, is this statement true of them? If it were not, I suppose the same platonist should maintain that it is false. Then its negation should be true. But, based on this negation, we might now prove that there are vector spaces with no bases. Hence our previous statement would be false. The former statement ('any vector space has a basis') is, then, true only if the latter (the axiom of choice) is so. But how might we prove the latter? We might certainly prove it, in ZF, by grounding on another equivalent or even stronger statement, for example, on the former one (which is provably equivalent to the latter over ZF). But this would only displace the question. What is it if it is not a proof that justifies the axiom of choice? And if we have no justification, how can we be sure that it is true? And if we cannot be sure that it is so, how can we be sure of it for the former statement, about vector spaces? So, possibly the two statements are not true: Possibly vector spaces exist, but some of them have no bases. All that we can know, then, according to our platonist, is that vector spaces exist and either all of them have a basis, or some do not. Or, possibly, this is not so, since, in fact, we have, after all, a clear justification of the axiom of choice: We see that sets are as it says; we see it with our insightful eyes of mathematicians, even, if we cannot make other eyes see it, even eyes of other (nonclassical or constructivist) mathematicians.

The problem with both conclusions is that they do not follow from any evidence coming from mathematics itself. They do not really provide a solution to Plato's problem (relatively to the particular fragment of mathematics that it is concerned with) but merely dissolve it. If these conclusions hold, the problem would simply not arise; it would be solved before being posed, since, what causes the problem is just

that mathematics speaks of vector spaces just like geography speaks of lakes, mountains, or towns, whereas vector spaces are not like lakes, mountains, or towns. So, how can mathematics speak of them this way? The previous platonist conclusions simply deny the root of the problem: They merely presume that vector spaces are just like lakes, mountains, or towns, at least at the insightful eyes of (classical) mathematicians.

5 On the Platonist Side

Let us restate the problem, then, in a more general setting. Mathematics results from a quite peculiar intellectual activity, or it is such an activity when it is considered as an enterprise rather than as a corpus. This activity appears to be twofold. It seems to consist both in the creation from noting of eternal and immutable items, and in the attribution of (some) properties and relations to them. Though these items appear to be eternal and immutable, their creation is the work of human beings (not of Gods or God), and it takes place in time, or even in history (not before any time). These human beings have, moreover, no other way for justifying their attributions of properties and relations to these items than proving statements that assert that these items are and stand to each another accordingly to these attributions. The proofs of these statements start from other statements of the same sort, some of which have necessarily to be admitted both without grounding on a similar justification, and without having acquaintance with these same items. How can we explain it?

This is not at all a problem about existence, about what exists or does not exist. What exists is clear: It is mathematics, and this existence is made manifest by history and actual experience (of teachers, students, and professionals), just as it happens for the existence of other intellectual enterprises and social phenomena, as well as for their products and, *mutatis mutandis* (but not *mutandis* so much), for empirical middle-size objects. The problem is rather about how this existence should be better accounted for. As such, this is a genuine, even crucial problem for a philosophy of mathematical practice, since this existence is just that of mathematical practice itself and of its outcomes.

When one looks at the problem this way, the theses that the platonist and nominalist I have described above try to test (or, better, confirm) on the testbench of mathematics are invested in another light, and they appear, then, dramatically naive and implausible. The platonist thesis becomes the thesis that this evident reality is to be explained by postulating the existence of another reality, inaccessible and unverifiable, but somehow (either causally or not) operative enough so as to make it possible to take the former existence as a more or less deformed imagination of the latter. The nominalist thesis becomes the thesis that this evident reality is not as it looks, since what appears is a theater performance disguising a much more trivial reality, not dissimilar in nature from our everyday life. When a serious philosophy of mathematical practice tackles the problem, it cannot but reject both these preposterous answers.

To my mind, the rejection should not be symmetric, however, since these theses are not so: whereas the former aims at explaining the available evidence through a

far-fetched assumption, the latter aims at denying this evidence in the name of a preconception. If I had to take a side in the dispute, though judging it misleading, I would have, then, no doubt in taking the platonist side. Since this is the side which aims at accounting for mathematics as it appears to be, rather than as it is required to be in agreement to an independent preconception. The purpose of the following considerations is to suggest a plausible way to stay on this side for a philosophy of mathematical practice: a way of being a platonist without recourse to the far-fetched assumption of the existence of mathematical objects.

One might argue that staying on the platonist side because of the reason I just advanced is still not the same as being a platonist. If it were so, the preceding pronouncement would be incongruous, since what comes after the colon would go far beyond what comes before it. Anyone that agreed with this would also probably find the position I will defend in what follows far from platonist in nature. Though I do not care so much about -ist and -ism-words, I regard this point as relevant, and I want to clarify why.

To my mind, this is not a purely terminological point, indeed. It rather depends on a judgment on what is essential in the platonist option, in all the different forms it has taken along the history of philosophy of mathematics. More than that, it also depends, in a sense, on what is considered to be the main task of philosophy. For many, this task is that of establishing what there is, and platonism is an important part of philosophy since it clearly claims that there are abstract objects. Not for me! I do not take this to be at all a task for philosophy. At most, it is one for positive sciences (whether they be physical, natural, biological, economical, or historical). I neither take the task of philosophy to say how what is there should be (if it could be in different ways). This is certainly a task for politics, but philosophy is not politics at all (though it could be useful for it). Philosophy should rather aim at providing the most appropriate conceptual categories to be used to account for what there is, or, better, what appears to be there.

This makes me think that platonism is a worthwhile option in philosophy of mathematics since it suggests that mathematics deals with the very contents that the language of mathematics speaks of, *prima facie*. In other terms, I take platonism, in the philosophy of mathematics, to be the quite general thesis that mathematical talk is to be literally understood as a talk of individual contents that provide the subjects of different forms of predication – in a word, as a talk of objects – and that this talk, so understood, consists of statements that hold as such, and as such have to be admitted and accounted for. My purpose, in what follows, is to articulate and better detail this general thesis in an appropriate way, along the line of a philosophy of mathematical practice.

6 Philosophical Purposes

Before doing that, an *excursus* is in order, however. It aims at shortly answering this further question: What should the purpose be, in general, of a philosophy of mathematical practice? Or even, more simply: What should such a philosophy be

like? I have already tackled the problem in my preamble, but it seems in order to come back to it now, in the light of the previous considerations.

Let us put the question this way: Should philosophy of mathematical practice be conceived as a reaction against a metaphysical, analytical, or foundational trend in the philosophy of mathematics? Or rather as an attempt to tackle new, previously unnoticed problems that, though strictly connected with the everyday work of mathematicians, do not strictly have a mathematical nature, being rather concerned with methodological choices, value judgments, or alike? Or, again, as an effort to rethink classical problems of philosophy of mathematics, restate them, and tackle them from a different perspective?

I do not like the first, purely antagonistic, option, and consider that the second one is simply not enough, since I think that classical philosophical problems are so, for they are good and difficult problems, though they are often conceived and restated in an ill-posed way. Plato's problem is an example of that: What makes it be, since Plato himself, a traditional problem for philosophy of mathematics is that it is a crucial problem that no serious philosophy of mathematics can dismiss, and even less if the dismissal is simply motivated by its being an old problem. This makes me think that the second option might be truly fruitful only if pursued as a complement of the third. It is, then, in the frame of this third option that I reason here.

When this option is joined with what I have said at the end of Sect. 5 above, it brings me to consider that a crucial task for the philosophy of mathematical practice is that of suggesting good conceptual categories to be used to account for the structural form of mathematics (intended as an existent reality), as it appears to be.

I say 'account for the structural form' for the sake of brevity, of course, in order to make clear that the relevant conceptual categories are not those required to describe the specific contents of mathematical theories, but rather some general features of these theories and contents, which are often characterized, by using a technical philosophical language, as logic, epistemic, ontological, etc. The specific contents of mathematical theories are also to be described, of course, using appropriate conceptual categories, but shaping these categories, and providing such a description, is a task for mathematicians rather than philosophers (though it would be wise for philosophers to try to understand this description, without falling into the opposite temptation to that mentioned in Sect. 2: the temptation of replacing mathematicians in this task, to provide a supposedly deeper description of these contents, which inevitably results much less profound than confused and defective; a temptation, it should be said, that some alleged philosophers of mathematical practice are not unfamiliar to).

In this framework, it seems to me that nothing might forbid philosophy of mathematical practice from embracing a platonist perspective, under the form of the thesis that the peculiar nature of mathematical activity is to (or, at least, might suitably) be accounted for as an intellectual activity dealing with (the) abstract objects (denoted by the usual terms figuring in the mathematical language).

This is still a very broad view, however, that can be declined in very different ways. One might wonder, for example, what one should mean in this context by 'intellectual activity,' 'dealing with,' 'abstract objects,' and, even, by 'objects' *tout*

court. I cannot enter all these matters, here, of course. I will, then, take for granted that my audience understands the term 'intellectual activity' more or less as I do. I only add that, according to the version of platonism that I would like to illustrate (and defend) here, 'dealing with x' means something like 'supplying (shaping, giving birth to) x, then studying it in order to describe it.' I also suppose, that this is enough to restrict my ongoing task to (quite programmatically) answer to the following single question: What should one mean, in the present context, by 'objects,' in general, and 'abstract objects,' in particular. The attention to the context is essential, since I am far from pretending to explain what can make an object abstract, in general. I shall limit myself to the case of abstract objects that one might take as mathematical ones.

My purpose, in what follows, will then be nothing but that of trying to make clear as the notion of an abstract (mathematical) object should be understood, in order to provide a good conceptual category to be used while describing mathematics as an intellectual activity dealing with abstract objects.

7 Objects by Reification

Objects are usually taken to be self-existing individuals, or, better, individual self-existing contents. An object is, then, taken to be abstract if this content is so, that is, as it is usually said, it is not spatiotemporally identifiable or determinable. An abstract object is, then, taken to be mathematical if (pure) mathematics involves putative reference to it (without receiving it from outside, for instance, from logic). If we understand an object this way, it becomes unavoidable to wonder how an individual content can exist (or have existed, or will possibly exist) without being spatiotemporally identifiable or determinable. It is just by taking this possibility for granted, without any further explanation, or, worse, in trying to provide quite improbable, and, often, even (quasi-)mystical answers to this question that platonism becomes generally implausible.

A way to avoid the problem is by ceasing to take existence, at least if intended as a primitive notion, as a necessary condition for objecthood. This is in no way an original perspective, and it is rather quite common in philosophy, also outside philosophy of mathematics. It goes, however, quite often together with some troubles, which is neither necessary nor possible to consider here in any detail. It will be enough to observe that a common move made to escape (some of) these troubles is taking abstract objects to be individual contents coming from reification of (first-level) concepts. This could go together with making the notion of an abstract object essentially independent of the more general notion of an object, namely, with renouncing at defining abstract objects as a particular sort of objects. What is important here is, however, not how this last task could be achieved, but rather how one can make the notion of reification of a (first-level) concept both precise enough, and such to avoid that the claim that mathematics deals with abstract objects becomes empty.

An example will make the point clearer. Let us take for granted that ordinary arithmetic deals with natural numbers. A very low-cost way to also grant that this

makes it deal with (abstract) objects is by providing a clear enough definition of the concept of a natural number, by means, for instance, of some appropriate version of Peano axioms, then taking natural numbers to be abstract objects just insofar as they are reifications of the concepts of the distinct items falling under this concept, reiteratively defined in turn, on the base of the definition of this last concept. The notion of reification of a concept is not very well defined here, but it does not seem *prima facie* implausible to consider, for instance, that taking the number one to be the successor of zero, by also admitting that the concepts of zero and successor are implicitly defined through the relevant version of the Peano axioms, be a way to take this number as the reification of a concept (that of successor of zero), and that the same also happens, in principle, for all other positive natural numbers.

Though certainly not perfectly crystal clear, this reasoning does not seem to me preposterous. But it would not help up to make the claim it supports – that arithmetic deals with abstract objects, understood as reification of concepts – anything more than an empty (though comfortable) claim, a contrived *façon de parler* to merely say that mathematics uses singular terms with a referential intent. If we want to make platonism an interesting philosophical thesis, we need to do much better than this.

Today's philosophy of mathematics presents at least two ways of doing it, each of which has unquestionable merits.

The first is provided by *Ante Rem* Structuralism (ARS), mainly advocated by Steward Shapiro (1997; a similar view that differs, however, from Shapiro's in many important details is advanced in Resnik 1997; other forms of structuralism are also available, which are not necessary to discuss here, since they do not openly adhere to the idea that mathematical objects result from a form of reification of concepts). To say it in general, the basic idea behind ARS is characterizing the appropriate form of reification by identifying a particular way, typical of mathematics, for defining concepts and organize them with the aim of fixing individual contents. This modality pertains to the identification of a categorical structure: The relevant abstract objects are identified with the poles (or places) in such a structure. They are, then, taken to exist insofar as the categorical structure they are poles of exists, and this exists, in turn, if and only if it is coherently defined, and obeys a general theory of structures, which (*mutatis mutandis*) mimics set theory.

Though much more might be said on ARS, this short account should be enough for making clear the way it answers to Platos' problem. What is important, indeed, is that it leaves to mathematics itself the task of defining categorical structures, and then insuring the existence of mathematical objects, though taking on it the task of indicating the appropriate way in which that has to be done (namely, by defining categorical structures). Leaving apart the usual, more technical objection about indiscernibles (originally advanced by Keränen 2001, to which Shapiro has replied, for instance, in Shapiro 2008), its main shortcoming comes, to my mind, from its being unapt to account for the historical evolution of mathematics, and, more in general, for large portions of past and present mathematics which are certainly not structuralist in nature, as well as for the current pervasive practice of transcendental proofs (proofs of theorems of a certain theory, involving tools from other theories).

The second way requires a more detailed explanation, since it appeals to less commonplace technicalities. It also adopts a much more descriptive (rather than normative) stance. It is provided by E. Zalta's Object Theory (Zalta 1983, 2021). The basic idea of this theory (or OT from now on) is providing a logical rendering of the very idea of reification of concepts in the shape of a particular form of predication, called 'encoding,' essentially distinct from that customarily in use in predicate logic. Formally, this depends on admitting two distinct forms for an atomic predicate formula, namely, 'Fa,' and 'aF,' where 'a' is a term and 'F' a monadic predicate, expressing the case where *a,* respectively, exemplifies (or has) *F*, and encodes it. Both forms of predication are embodied in a system of modal higher-order logic admitting types, but also working without them for simplicity. In its nontyped version, this system includes the following axiom

$$\Diamond[aF] \Rightarrow \Box[aF]$$

(where 'a' and 'F' are schematic letters, and '$\Diamond$' and '$\Box$' are the usual modal operators), asserting that *a* can encode *F* only if it encodes it necessarily. As the reciprocal implication also holds, of course, because of the usual properties of the modal operators, this renders that idea that *a* is the reification of *F* if and only if it is necessarily so.

This idea goes together with two further crucial ingredients of the system: the clause that an object encodes a (monadic) property only if it is abstract, and a form of comprehension asserting that, for any condition on monadic properties expressible in the formal language at issue, there is an abstract object that just encodes the properties that meet this condition (which allows an abstract objet to be conceived, in fact, as the reification of a bunch of properties).

The former ingredient is implemented by the introduction of the nonlogical monadic predicate constant '$E!$' denoting the property of being a concrete object, through which the two other monadic primitive predicate constants '$A!$' and '$O!$' respectively, denoting the properties of being an abstract and an ordinary one, are defined by

$$A! =_{df} [\lambda x \neg \Diamond \mathsf{E!x}] \quad \text{and} \quad O! =_{df} [\lambda x \Diamond \mathsf{E!x}].$$

It follows that any object considered in the system is either abstract or ordinary, and it is abstract if and only if it cannot be concrete, that is, it is concrete in no possible world, while it is ordinary if and only if it can be concrete, namely, it is concrete in some possible world. To these definitions, the following axiom-schema is added:

$$\forall \ldots \ [O!x \Rightarrow \neg \exists P[xP]],$$

(where '$\forall \ldots \ [\varphi]$' stands for the universal closure of 'φ'), asserting that an object is ordinary (and, then, not abstract) only if it fails to encode (is not the reification of) any (monadic) property. By contraposition, it immediately follows that an object

encodes (is the reification of) of a (monadic) property only if it abstract, as announced above.

The latter ingredient completes the picture. It consists of the following comprehension axiom-schema:

$$\exists x[A!x \wedge \forall P \ [xP \Leftrightarrow \varphi]]$$

where 'φ' is any formula in the formal language at issue in which 'x' does not occur free. Let, then, $\boldsymbol{P}$ be whatever (monadic) property already defined in the system. From this axiom schema, together with the first axiom mentioned above and a principle of identity for abstract objects making x the same abstract objects as y if and only if x necessarily encodes the same properties as y, we immediately get that

$$\exists!x\,[A!x \wedge \forall P \ [xP \Leftrightarrow \forall z \ [Pz \Leftrightarrow \boldsymbol{P}z]]],$$

which ensures that there exists a single abstract object that encodes (is the reification of) any (monadic) property coextensive with $\boldsymbol{P}$. Extensionally speaking, we can then say that for whatever (monadic) property $\boldsymbol{P}$, there is a single abstract objet that encodes (reifies) it and only it. But this is not all, since the formula 'φ' can be more complex than that. It can express any condition on (monadic) properties, which makes the schema ensure that for any such condition there is an abstract object that just encodes the (monadic) properties that meet it, as also announced.

With this formal apparatus at hand, and with the help of some further enrichments of the language of his theory – including the introduction of the definite descriptions operator '$\imath$' (such that '$\imath x\,[\varphi]$' is a term denoting the single object, if any, that meets the condition expressed by 'φ') – B. Linsky and Zalta himself have been able to provide a rigorous explicit characterization of an intratheoretic mathematical object (Linsky and Zalta 1995, 2006, 2019; the reason for I speak of an intratheoretic mathematical object, rather than of a mathematical object *tout court,* will become clear below). This is an object κ_{T} such that there is a mathematical theory T (identified, in turn, with an abstract object appropriately fixed within OT, so as to render an actual mathematical theory), such that this object encodes all and only the properties it has in T, that is:

$$\kappa_{\mathrm{T}} = \imath x \ [A!x \wedge \forall F \ [xF \Leftrightarrow \mathrm{T} \ [\lambda z F \kappa_{\mathrm{T}}]]],$$

where $[\lambda z F \kappa_{\mathrm{T}}]$ is the 0-place (or propositional) property that any object has just in case the world it belongs to is such that $F\kappa_{\mathrm{T}}$ (in Zalta's terminology, this is the propositional property "constructed out" of $F\kappa_{\mathrm{T}}$).

This is properly not a(n explicit) definition, of course. Its very logical form prevents it from being so, since the term 'κ_{T}' figures in it both in the left- and the right-hand sides. It is rather a "theoretical description," offered in OT's (rigorous) language. Whether there is a theory like T, if it is actually a mathematical one, how it its done, which individual contents (denoted by singular terms) it deals with, which of them are taken to have in it this or that property, and what makes they have them

are questions that OT neither can decide nor pretend to decide. It simply provides a formal setting making possible to rigorously claim that these contents are intratheoretical mathematical objects, and that nothing else is such an object. This entails that intratheoretical mathematical objects are all abstract, and are just the individual contents (possibly denoted by appropriate singular terms) that a mathematical theory deals with. More generally, each of them is the reifications of the concept of having all the properties that this theory assigns to it.

Just as Shapiro's structuralist definition, Zalta's characterization or "theoretical description" of intratheoretic mathematical objects is intrinsically unsuited for accounting transcendent proofs and extratheoretical contents, that is, for mathematical objects possibly studied by different theories and then independent of them. The problem is serious, since denying this possibility makes it impossible to maintain, for instance, if not quite informally, that different versions of arithmetic all deal with natural numbers, as a platonist account should be able to do, it seems to me. OT has however the resources for also characterizing extratheoretical mathematical objects.

On the one side, it accounts for the reference of singular mathematical terms in statements that make no (explicit) mention of a particular mathematical theory, as a reference to abstract objects that encode a relevant concept. Supposing we are able to deal with the concept of a triangle, however it might be fixed, it would be enough, for instance, to identify this concept with a particular concept, call it '$\boldsymbol{T}$', dealt with it within this theory, to easily render the statement of a natural or informal language 'a is a triangle' by the statement of such a theory '$a\boldsymbol{T}$'. This is the same as interpreting the 'is' in the former statement as an informal occurrence of the encoding operator. If it happens that the concept $\boldsymbol{T}$ is defined within a certain particular mathematical theory, be it formal or not (and this theory is, in turn, identified with an abstract object appropriately fixed within OT), we can prove (within this theory) that a encodes $\boldsymbol{T}$ if and only if a exemplifies $\boldsymbol{T}$. The statement '$a\boldsymbol{T}$' is then a statement about an intratheoretic mathematical object, and this makes it suitable for rendering 'a is a triangle' only if the same also happens for this latter statement. Still nothing forbids considering $\boldsymbol{T}$ (informally) fixed otherwise (by a way that OT is not required to specify), and still rendering 'a is a triangle' by '$a\boldsymbol{T}$'.

On the other side, OT also allows to define within it a quite weak, pseudofreegan set theory, and use it to define natural number *à la* Frege, as extensions of concepts (Zalta 1999). To this purpose, it is enough to take the extension $\varepsilon\boldsymbol{F}$ of a (sortal) concept $\boldsymbol{F}$ as the object

$$\iota x \ [A!x \wedge \forall P \ [xP \Leftrightarrow \forall z \ [Pz \Leftrightarrow \boldsymbol{F}z]]]$$

that just encodes $\boldsymbol{F}$, to define the appurtenance relation $\in$ (or its Fregean *alias* $\frown$) by

$$x \in y \Leftrightarrow \exists P \ [x = \varepsilon P \wedge Py]$$

and to adopt an appropriate modal principle to get second-order Peano axioms as Frege did (*mutatis mutandis*) in *Grungesetze*. According to Zalta, extratheoretical natural numbers can, then, be taken as the objects meeting these axioms (on which

no order relation, no addition, and no multiplication are defined yet), and the same can be done for other mathematical objects definable within the same weak set theory.

All this should be enough to make clear, in the same time, the great expressive (or descriptive and elucidatory) strength of OT, and its intrinsic and purposeful presumptive weakness. This weakness consists in its deliberately avoiding any endorsement in favor of a specific view about the nature of mathematics. The theory succeeds in providing a clear logical account of the reification process and the corresponding constitution of abstract objects. However, it purposely refrains from accounting for what distinguishes the mathematical form of reification and makes the intellectual exercise that mathematics pertains to result in such a powerful and stable corpus. This makes OT particularly suited as a theoretical context for advancing a pluralistic view about mathematics (Zalta forthcoming). But it makes it also unsuited to provide, alone, an answer to Plato's problem. This does not mean that it cannot contribute to provide this answer. It simply means that some other independent insight is required to achieve this task.

8 Object by *De Re* Epistemic Access

This is why I want to suggest another perspective here, essentially independent of ARS and OT. In particular, I suggest looking at an object, both concrete or abstract, as an individual content we could have (had) a *de re* epistemic access to.

If I say that a has the property $\boldsymbol{P}$, I can mean two quite different things, whose difference depends on the modality of my epistemic access to a. I can mean, of a, identified as such in advance, that it has the property $\boldsymbol{P}$, or I can mean that a, which has not been necessarily identified as such in advance, has the property $\boldsymbol{P}$.

I was born in Varese, a nice little town in the north of Italy, and lived permanently there the first 27 years of my life, and even now, though I live abroad, I try to come back to Varese anytime I can, spending time in the house where I grew up and lived for so many years. From the balcony of this house, one can see the Varese Lake, on the shores of which I went to walk and bike myriads of times. I also went boating many times on this lake, and, when I was a kid, pollution did not advise against swimming in it, and so I swam in it many times. Now, it just happened that I read on Wikipedia that the Varese Lake has a surface area of 14.5 km^2. Despite my familiarity with the Varese Lake, this is a new piece of information for me, and it is, literally speaking, a piece of information about the Varese Lake. It is about it that I got this information, which I did not have before (simply since I never thought before to look for it). Not having this information did in no way make me unable to perfectly know what the Varese Lake is, to understand what one was speaking about when speaking of it, and to locate it on the surface of the Earth, and the image of it in my memory: in few words, to have a perfect epistemic access to it. Because of this access, I now know what the new information that I got is about. So, if I now say, that the Varese Lake has a surface area of 14.5 km^2, I definitely mean, of the Varese Lake, that it has a surface area of 14.5 km^2.

After having got this new information about the Varese Lake, I opened another Wikipedia page and did read there that the "Eridania Lake is a theorized ancient lake [on Mars] with a surface area of roughly 1.1 million square kilometers [. . .][, a] maximum depth of 2400 meters [and a] volume[. . .][of] 562,000 km^3" I had never heard anything concerning the Eridania Lake, and never heard of the theory that there could have been lakes on Mars in the far past. I am, then, very much surprised that such a huge mass of water on Mars could have existed. This new piece of information is not about something I knew before having it and independently of it, that I was able to locate in the universe, and I had an image of in my memory. In few words, it is not about something that I had an epistemic access to, before getting it. This piece of information is instead, for me, constitutive of the Eridania Lake: It is what makes me have an epistemic access to it. So, if I now say that the Eridania Lake had a surface area of about 1,100,000 km^2, I cannot but mean, literally speaking, that the Eridania Lake had such a surface area, and this is what makes me now able to think of it somehow.

This illustrates the distinction I want to make between two radically different modalities of epistemic access to a content that works as a reference of a proper name. In the former example, I take my epistemic access to be *de re*. In the latter, I take it to be purely *de dicto*.

This having been said, let us come back to the latter example (the former seems to me clear enough for not requiring other considerations). Though no human being (and also no nonhuman animal presumably) has certainly never had any direct empirical experience of the Eridania Lake, it seems to me clear that it counts as an object, and even more so as a concrete one. What does make it so clear (at least to me)? I advance that the reason is this: The fact that nobody has ever had any direct empirical experience of the Eridania Lake does in no way make it logically impossible for someone to have had such a direct empirical experience. Moreover, even if (for some reasons that I am not frankly able to imagine) we were urged to consider such an experience as logically impossible, it would still remain that this would hardly depend on the intrinsic features of the Eridania Lake, but rather of the limitations imposed to our (or other animal's) capacity of empirical experience. As such, the Eridania Lake would have allowed anyone to have a direct experience of it, and this experience would have, in turn, allowed a human being to have a *de re* epistemic access to it. The (logical) possibility of such an access is, in my view, what makes the Eridania Lake count for us as an object. The fact that this is the possibility of an empirical experience is, instead, what makes this object count for us as a concrete one.

There is an important aspect of this argument that deserves clarification. What I just described are in no way conditions that make the Varese and the Eridania Lakes concrete objects *tout court*, but rather conditions that make them count for us as concrete objects. I am, in fact, not interested in the question of establishing what makes a concrete or an abstract object be so. I even consider this question perfectly meaningless. In agreement with the approach to philosophy that I described above, I am rather interested in the question of establishing what makes us consider an object

to be so. It is just to this question that I suggest answering by saying that this is nothing but the possibility of having (had) a *de re* epistemic access to it.

This having been made clear, let us come back to the main point. The two examples I advanced are both such to make the possibility of having (had) a *de re* epistemic access to a certain individual content depending on the possibility of a direct empirical experience of this content. This is just what makes these examples concern concrete objects. The next question is, then, this: Are there other sorts of epistemic relations with some individual contents that can make us able to have a *de re* epistemic access to them. The possibility of (sensibly) taking some individual content as an abstract object – rather than, merely, as the putative referential target of a proper name used in a statement that is admitted to hold – depends, in my view, on the answer to this question. Moreover, if the answer is positive, it is just the nature of these alternative relations that decides the nature we can ascribe to the object at issue, and whether it should count for us as an abstract object or a concrete one. This makes me replace the question of whether there are mathematical (abstract) objects with the question of whether there are forms of epistemic relations with individual (abstract) contents that are proper to mathematics (and differ, then, from direct empirical experience) that make it possible for us to have a *de re* epistemic access to these contents.

To better understand the question, let me try to better clarify it.

What makes my epistemic access to the Varese Lake *de re* is my capacity of identifying this content as such while ascribing a certain property to it, but fully independently of my doing it. This is, in other terms, my capacity of fixing this content in my epistemic horizon independently of recognizing it as the content having this property.

Now, it seems plain to me that this also happens very often with individual contents that we deal with in mathematics.

We can certainly identify the number 31 independently of ascribing to it the property of being the 11th prime number (in the usual order on natural numbers). Hence, properly speaking, we do not say that 31 is the 11th prime number; we rather say, of 31, that it is the 11th prime number.

In the same way, we do not properly say that the ratio of two successive terms of the Fibonacci sequence tends to the golden ratio, or even that this ratio is (equal to) $\frac{1+\sqrt{5}}{2}$; we say of the ratio of two successive terms of the Fibonacci sequence that it tends to the golden ratio, and, of this ratio, that it is equal to $\frac{1+\sqrt{5}}{2}$.

Despite naming it after Euclid, we do not properly say that the Euclidean closure of the rational numbers is the field of the numbers constructible by ruler and compass; we say, of the Euclidean closure of the rational numbers, that it is the field of the numbers constructible by ruler and compass.

Again, we do not properly say that 4 is the minimal number of colors required to color a map by avoiding that two adjacent regions be colored with the same color; we say, of 4, that it is the minimal number of colors required for so coloring a map.

We no more properly say that 18 is the Ramsey number $R(4, 4)$ (i.e., the smallest number of vertices of a graph containing either a four-cycle or an independent set

with four vertices), or that $43 \leq R(5, 5) \leq 48$; we say, of 18, that $R(4, 4)$ is equal to it, and, of 43 and 48, that they are the limits of the closed interval including $R(5, 5)$.

Finally, we do not properly say that $\frac{\pi^2}{6}$ is the limit of the series $\sum_{i=0}^{i=\infty} \frac{1}{i^2}$; we say of $\frac{\pi^2}{6}$, that it is the limit of this series.

Examples might be multiplied. Some of the previous ones might, by the way, also be inverted. We might perfectly say, for instance, of $\frac{1+\sqrt{5}}{2}$, that it is (equal to) the golden ratio, or of the limit of the series $\sum_{i=0}^{i=\infty} \frac{1}{i^2}$, that it is (equal to) $\frac{\pi^2}{6}$. The point is not which examples we consider, or what we consider to be, in each of them, the given content, and the property that we attribute to it. What is relevant is that these examples illustrate an important fact about mathematics: that in it, we can exhaustively fix individual contents in our epistemic horizon, independently of ascribing to them relevant properties that are rather ascribed to them *post festum*. In other terms: We can identify such contents without specifying all the properties they have in the very setting in which they are identified or in other settings. This fact is just what allows us to export individual contents from setting to setting, or, better, to deal with the same individual contents in different settings. Again, we can fix such contents in a certain setting, then extend the setting by defining appropriate relations or functions on these same contents, which allows us to ascribe new properties to them, without making them different contents. We do both things pervasively. And this is just what makes, to my mind, a sort of spontaneous (though often irreflexive) platonism so popular among working mathematicians.

This is, indeed, my own version of platonism for a philosophy of mathematical practice: Mathematics can be (is) suitably accounted for as an intellectual activity dealing with *sui generis* individual abstract contents, so fixed as to make us able to have a *de re* epistemic access to them. As it should be clear, this claim has strictly nothing to do with the existence of these contents (and the truth of some statements on them). Not only does it not assert the existence of mathematical objects, it also does not deny it. It rather assumes that the notion of existence of abstract objects is, when conceived as a primitive notion, simply nonsensical, by leaving open the possibility of defining this notion based on other ones, by so ending up recognizing that, yes, mathematical objects exist (and mathematical theorems are true of them), but in a new, derived sense. Among other things – particularly the admission that mathematical statements are to be understood *prima facie* – this assumption of nonsensicality essentially distinguishes the claim from any version of nominalism. It is, rather a claim for genuine platonism, though this is, as it were, a platonism without existence (and truth), at least if intended as primitive notion(s).

I say 'sui generis,' since the way the relevant individual contents are fixed is peculiar to mathematics. And it is just this peculiarity that makes this thesis not only compatible with the program of the philosophy of mathematical practice, but also dependent on this very program for its clarification and advocacy. Indeed, it is only by parsing mathematical practice and its history that we can identify the different ways these contents are fixed, so as to make such an epistemic access possible, and we can try to recognize what makes them mathematical in nature.

Such a project is structurally similar, in nature, to that of any empirical science: It is, so to say, the project of an empirical science of mathematical reality, this reality being not, however, that of self-standing objects, as a traditional platonist would argue for, but rather that of mathematical activity and production.

9 History, Cognition, and Logic

Of course, realizing this project is not something I can do here. Both the limits of space and the very purpose of this chapter force it to be essentially programmatic, in nature. What I can and want to do is delineating a philosophical perspective, a research program for a philosophy of mathematical practice able to answer Plato's problem. This only requires, for the time being, to briefly suggest three general sources (certainly not alternative, but rather complementary) for our ability to fix abstract individual contents so as to make possible for us to have *de re* epistemic access to them (and, then, to take them to be mathematical objects): history, cognition, and logic (or, more generally, formalization).

9.1 By History

Mathematics largely evolves by (conceptual) recasting (I borrow this term from several talks of Ken Manders). Theories are firstly (mainly informally) advanced, then variously restated, reorganized, and formalized. This important part of mathematical activity explicitly and essentially depends on the admission that these different forms of recasting conserve contents, that some contents (not only individual, of course, but crucially also individual ones) remain the same under the transformation of the setting in which they are defined and studied. This suggests to inverse the model-theoretic perspective.

According to this last perspective, mathematical objects are conceived of as the elements of the first-order domain(s) involved in the model(s) of a (formal) mathematical theory, namely, with the individuals forming the range(s) of the (different sorts of) first-order variables. What I mean by inverting this perspective is conceiving most mathematical theories as codifications – or, possibly, in some cases, formalizations – of a portion of available knowledge inherited from the history of mathematics (in the proper sense of 'history,' i.e., as *res gestæ*, rather than *historia rerum gestarum*), in which contents (not only individual, but crucially also individual ones) are crystallized, so as to make their transmission possible. What many mathematical theories would, then, do is taking some contents as given, and suggest a new conceptual configuration for them, possibly connecting them with other contents, historically crystallized in turn, but dissociated from the former ones, in the original piece of knowledge. This would make a mathematical theory be not about its models but about the contents it captures by recasting.

The availability of a single model, under isomorphism (and within a given semantic setting), would merely show that the theory captures these contents making them part of a univocal structure (within this setting). The presence of

nonisomorphic models would rather show that these contents are not so captured, leaving, then, space for alternative interpretations. This lack of univocality (or relative categoricity) can be certainly beneficial, since it can open new horizons of knowledge (the case of nonstandard analysis is symptomatic), but it says nothing about the original contents. It rather says something about new contents that, in this way, we become able to distinguish from the former, and study together with, or separately from them.

Here, the basic idea is that mathematical activity is intrinsically historical (as argued for in the preamble), though this characteristic and peculiar nature of it remains hidden to superficial regards. The reason of the concealment is that this historical nature manifests itself by a sort of (temporary) negation of itself, just due to recasting.

9.2 By Cognition

Mathematics requires (both common and specific) cognitive abilities. These are not the usual abilities to realize practical performances, often considered by alleged studies in mathematical (typically numerical) cognition, in order to display widespread forms of cognitive access to mathematical objects (typically small integer positive numbers), some of which would also be proper to human infants and some species of primates. These merely appear to me as mundane abilities to play some linguistic or prelinguistic games, which either involve mathematical words, or admit a retrospective description appealing to these words or the corresponding concepts. Subtilizing is only one among many other possible examples of these practical performances, which have, to my mind, no attested relation with mathematical theories and the contents they actually deal with.

What I am instead referring to are cognitive abilities that are required to crystallize intellectual contents involved in abstract reasoning. These are typically (but not limited to) abilities of recognizing an invariant structure within different (sorts) of exemplification of it, for example, the ω-sequence structure within the succession of Arabic numerals, or that of the reciprocal of the natural numbers, or the structure of a regular polygon within equilateral triangles, squares, pentagons, etc.

More generally, these are abilities of detecting abstract invariants, either individual or not (for instance, operational ones) and of both considering them as such, and working on particular instances of them, in order to identify general structural properties. In Kant's well-known terminology, this is the capacity of dealing with the "universal *in concreto.*" However flawed his conclusions may have been, both by a dramatically mistaken experimental practice and by a misleading confusion between ontogenetic and historical evolution, in his study of the cognitive development of children, J. Piaget famously tried to locate and isolate abilities like those.

Here, the basic idea is that performing mathematics activity depends of the (acquisition and) possession of cognitive abilities of different levels and sorts. These are the abilities that make us able to identify stable intellectual contents (either individual or not), which are both phylogenetically forged, and ontogenetically refined, also through cognitively complex and specific exercises (even if certainly not only through them).

9.3 By Logic (or Formalization)

Mathematical (conceptual) recasting mostly depends on a codification of language and intellectual acts, such as (stable or provisional) suppositions, inferences, definitions, constructions, etc. This codification is essentially logical, insofar as it depends on restrictions of the forms of expression and of the possibility of actions, in agreement with a number of general clauses to be applied in any specific occasion. These clauses might be seen as norms granting permissions, to be appealed to yield an argumentative setting that excludes all that is not allowed by them. A given setting is, of course, often extended, either occasionally, by unenvisaged admissions or acts, or systematically, by the addition of clauses granting further permissions. Moreover, different settings are frequently merged with each another. If Cantor was certainly right in claiming that "the essence of mathematics just lies in its freedom" (Cantor 1883, p. 564; my translation), this is essentially the freedom of fixing appropriate settings or, at most, of locally contradicting them by merely foretelling new permissions, rather than that of leaving anything go.

So broadly described, codification can result in either formal or informal theories. The border between the two cases is far from rigid and largely depends on historical or cultural conditions, branches of study, and methodological attitudes and sensibilities. What can be taken as formal within some mathematical communities cannot be taken this way within other communities. Even the criterion of (the possibility of a) computer implementation is much vaguer than it might appear at first glance, since it depends on the machine languages used and the interfaces with higher-level languages. Still, be it as it may, the crucial point is that any form of codification, as broad as it might be, goes together with identifying a delimited space of possibilities, involving some poles that can count as fixed individual contents.

Conceiving a particular codification as an act of recasting makes these individual contents count as particular occurrences of a more general content that they render within a certain setting. This is just what allows the last content to be taken, as such, as an available target for attributions of properties or the definition of relations or functions, which are in no way constitutive of it, either in general or within the particular setting at issue.

Fixing natural numbers as poles of the (relatively) categorical structure defined by Peano second-order axioms allows, for instance, both taking these poles as targets of a fist-order (uncategorical) axiomatization – which count, then, as an act of recasting within a different setting – and defining on them, within their own constitutive setting, a linear order, and an additive and a multiplicative law of composition, without changing their nature, but rather attributing to them, as such, a new bunch of properties. Moreover, making codification meet some appropriate conditions also allows its outcome to become, as such, the subject matter of a new mathematical inquiry, depending, in turn, on a metacodification. Such is the very well-known phenomenon of metamathematics, which is also a distinctive aspect of our capacity of having, in mathematics, a *de re* epistemic access to individual contents.

Here, the basic idea is that mathematics uses codified languages within delimited argumentative spaces. These spaces are not only suitable for fixing some appropriate

structures that are intended to render previously available pieces of knowledge, by bringing them to a new life. They also have specific structural properties that can be studied as such by new mathematical theories.

10 A Few Words, in Guise of Conclusion

When they are so conceived, as sources of our capacity to fix individual contents, so as to make possible having a *de re* epistemic access to them, history, cognition, and logic become crucial ingredients of a form of platonism. Its being independent of any talk of existence of abstract objects (and of truth of statements on them), if not appropriately redefined as derivative notions, makes the theses advanced by it perfectly factual claims.

Whether mathematics complies or not with such a form of platonism becomes, then, a pure matter of fact: a matter of fact that the philosophy of mathematical practice is not required to abstractly argue for, but rather to verify by an inquiry of this practice and its history. This makes this philosophy work, in turn, as a sort of high-level empirical science.

References

Balaguer M (1998) Platonism and anti-Platonism in mathematics. Oxford University Press, New York/Oxford

Benacerraf P (1965) What numbers could not be. Philos Rev 74:47–73

Burgess JP (1983) Why I am not a nominalist. Notre Dame J Formal Logic 24(1):93–105.

Burgess JP (2008) Mathematics, models and modality. Cambridge University Press, Cambridge/New York, etc., pp 31–45

Burnyeat MF (1987) Platonism and mathematics: a prelude to discussion. In: Graeser A. (ed) Mathematics and metaphysics in Aristotle. P. Haupt, Bern/Stuttgart, pp 213–240

Cantor G (1883) Ueber unendliche, lineare punktmannichfaltigkeiten, 5. Math Ann 21:545–586

Cheyne C (2001) Knowledge, cause, and abstract objects. Causal objections to platonism. Springer, Dordrecht

Euclid (1926) The thirteen books of the Elements, 2nd edn. Cambridge University Press, Cambridge. Translated with introduction and commentary by Sir Thomas L. Heath; 3 vols

Field H (1980) Science without numbers. Princeton University Press, Princeton. Second edition: Oxford University Press, New York, 2016

Hale B, Wright C (2001) The Reason's proper study. Clarendon Press, Oxford

Keränen J (2001) The identity problem of realist structuralism. Philos Math 3(9):308–330

Linsky B, Zalta E (1995) Naturalized platonism versus platonized naturalism. J Philos 92(10): 525–555

Linsky B, Zalta E (2006) What is neologicisms? Bull Symb Log 12(1):60–99

Linsky B, Zalta E (2019) Mathematical descriptions. Philos Stud 176(2):473–481

Löwe B, Müller T (eds) (2010) PhiMSAMP. Philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London

Mancosu P (2008) The philosophy of mathematical practice. Oxford University Press, Oxford/New York

Panza M, Sereni A (2013) Plato's problem. An introduction to mathematical platonism. Palgrave Macmillan, Basingstoke

Parsons C (1964) On platonism in mathematics. Translation of P Bernays (1935) Sur le platonisme dans les mathématiques. L'enseignement mathématique 34: 52–69.
Quine WVO (1939) Designation and existence. J Philos 36:701–709.
Quine WVO (1986) Reply to Charles Parsons. In: Hahn L, Schilpp P (eds) The philosophy of W.V. Quine. Open Court, La Salle, pp 396–403
Resnik MD (1997) Mathematics as a science of patterns. Clarendon, Oxford
Shapiro S (1997) Philosophy of mathematics: Structure and ontology. Oxford University Press, Oxford/New York
Shapiro S (2008) Identity, indiscernibility, and ante rem structuralims: the tale of i and $-i$. Philos Math 3(16):285–309
Zalta E (1983) Abstract objects. Reidel P. C., Dordrecht/Boston/Lancaster
Zalta E (1999) Natural numbers and natural cardinals as abstract objects: a partial reconstruction of Frege's *Grundgesetze* in object theory. J Philos Log 28(6):619–660
Zalta E (2021) Principia Logico-Metaphysica. Most recent version of a work in progress, regularly updated and posted at https://mally.stanford.edu/principia.pdf
Zalta E (forthcoming) Mathematical pluralism. Available online at https://mally.stanford.edu/Papers/math-pluralism.pdf

Section X

Pluralism

Michele Friend

One Mathematic(s) or Many? Foundations of Mathematics in Twentieth-Century Mathematical Practice

Andrei Rodin

Contents

1 "Practical" Foundations of Mathematics and Foundations of Mathematics *Simpliciter* 2340
2 The Received FOM and Metamathematics 2341
3 Is the Received FOM Adequate? 2343
4 Foundations at Work 2347
4.1 Bourbaki 2348
4.2 Category Theory as a Foundation 2351
4.3 Univalent Foundations 2353
5 Renewing Foundations and Progress in Mathematics 2357
6 Tolerance and Pluralism 2360
7 Cross-References 2362
References 2362

Abstract

The received Hilbert-style axiomatic foundations of mathematics has been designed by Hilbert and his followers as a tool for metatheoretical research. Foundations of mathematics of this type fail to satisfactory perform more basic and more practically oriented functions of theoretical foundations such as verification of mathematical constructions and proofs. Using alternative foundations of mathematics such as the univalent foundations is compatible with using the received set-theoretic foundations for metamathematical purposes provided the two foundations are mutually interpretable. Changes in foundations of mathematics do not, generally, disqualify mathematical theories based on older foundations but allow for reconstruction of these theories on new foundations. Mathematics is one but its foundations are many.

A. Rodin (✉)
Institute of Philosophy, Russian Academy of Sciences, Moscow, Russia

Saint-Petersburg State University, St. Petersburg, Russia
e-mail: andrei@philomatica.org

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_28

Keywords

Metamathematics · Foundations of mathematics · Univalent foundations

1 "Practical" Foundations of Mathematics and Foundations of Mathematics *Simpliciter*

Foundations of a theoretical discipline are traditionally conceived as a join between this discipline and its philosophy; accordingly, the philosophy of mathematics typically focuses on the foundations of mathematics (FOM) rather than on other areas of mathematical research. The philosophy of mathematical practice attempts, among other things, to enlarge the scope of philosophical analysis of mathematics by paying more attention to past and present mathematical developments beyond FOM and, more specifically, beyond the *Grundlagenkrise* that occurred in mathematics at the turn of the nineteenth and twentieth centuries and the related subsequent developments (Mancosu et al. 2008). Without trying to diminish the importance and philosophical relevance of mathematical practices and contents beyond FOM, I would like to stress that the mismatch between theory and practice in mathematics, which provides a rationale for distinguishing the philosophy of mathematical *practice* into a separate area of research, concerns FOM in the first place.

Described in few words, the current situation in FOM is as follows. The mainstream FOM is a relatively isolated and very specialized area of today's mathematical research. The majority of working mathematicians and mathematical educators know very little about FOM and cannot see its relevance to their works. This is certainly not how FOM, traditionally conceived, was expected to function. This unusual situation produces a range of different reactions in the mathematical community.

Some mathematicians including many FOM researchers explain and justify the gap between FOM and the rest of mathematics by a natural division of labor. They assume (i) that an ordinary working mathematician does not need to master details of the foundations of her discipline and (ii) that the research in FOM just like research in any other special area of mathematics requires a long professional training that cannot possibly be provided for all working mathematicians. A version of this view may go along with the conviction that FOM is of little or no significance for other mathematical disciplines. This view, however, is also compatible with the belief that all mathematical disciplines crucially depend on FOM whether or not working mathematicians are aware of it.

A number of mathematicians working outside the mainstream FOM research believe that mathematics needs foundations in a more traditional sense of the term, according to which FOM should represent some core mathematical contents and be known to and used by all mathematicians. On this basis Lawvere and Rosebrugh boldly disqualify the received FOM as a valid foundation:

> A foundation makes explicit the essential general features, ingredients, and operations of a science, as well as its origins and general laws of development. The purpose of making these explicit is to provide a guide to the learning, use, and further development of the science. A "pure" foundation that forgets this purpose and pursues a speculative "foundations" for its own sake is clearly a nonfoundation. (Lawvere and Rosebrugh 2003, p. 235)

Some other mathematicians try to make a peaceful agreement between the two concepts of FOM. They reserve the name of *practical foundations* for the traditional notion without trying to challenge the received FOM explicitly (Taylor 1999). On a parallel – albeit chronologically more recent – development in the philosophical camp a group of philosophers constituted the *philosophy of mathematical practice* as a subfield of philosophy of mathematics outside the mainstream FOM-oriented philosophy of mathematics (Mancosu et al. 2008). Since pure mathematics is a theoretical discipline *par excellence*, the grounds of the relevant distinction between mathematical theory and mathematical practice are far from clear, moreover so that philosophers of mathematical practice in their work systematically refer to mathematical theories and other mathematical content, but not only to practical issues related to mathematics research and education. The popular wisdom according to which "in theory there is no difference between theory and practice, while in practice, there is" perfectly applies in this case. Practically speaking, it is indeed the case that philosophers who associate themselves with the mainstream FOM research and philosophers of mathematical practice are doing different research and form different communities. But from a theoretical point of view it is not clear to me why these lines of philosophical inquiry into mathematics should be so different.

In order to clarify the situation I will not go deeply into philosophical generalities about theory and practice. Instead, I will analyze the nature and the genesis of the aforementioned gap between the received FOM and the rest of mathematics. I will use the above concise nonorthodox description of FOM coined by Lawvere and Rosenburgh as my working definition of this concept.

2 The Received FOM and Metamathematics

Today's received concept of FOM stems from Hilbert's seminal works in axiomatic geometry (David Hilbert 1899) and his later works on what is known today under the name of *Hilbert's Program*, which includes a modern form of proof theory (Hilbert and Bernays 1934, vol. 2). A special feature of this conception of FOM, which distinguishes it from more traditional conceptions, is its *metamathematical* character. The idea of metamathematics, as Hilbert and Bernays explain it in (Hilbert and Bernays 1934, Chap. 2, Sect. 5), is to provide *mathematical* answers to certain general mathematically related questions, which have been earlier seen as purely logical and philosophical (Ferreiros 2005). (Such as Hilbert's *Entscheidungsproblem*.)

The idea of metamathematics can be traced back to Descartes and Leibniz but it is not constitutive in the traditional FOMs. In particular, Euclid's *Elements* (which

qualify as FOM in our sense) has no metamathematical capacity. Let me use this classical example for explaining where the idea of empowering FOM with metamathematics comes from.

Euclid's *Elements* comprise a number of impossibility results such as a proof of impossibility to find a common measure for a pair of arbitrary straight segments. (See the *Elements*, Book 10, Problem 10.) Such proofs show that certain geometrical constructions can be specified in Euclid's theoretical language but, on pain of contradiction, cannot be performed within Euclid's theory. Such impossibility proofs solve certain mathematical problems in the *negative* way by showing that certain desired constructions are impossible.

However, there are also well-known mathematical problems that are are well posed in Euclid's mathematics but cannot be solved using only Euclid's principles either positively (by performing the required construction) or negatively (by showing that the required construction is impossible). The angle trisection problem is an example. A trisected angle is not an impossible geometrical construction. One can easily perform an instance of such construction by tripling a given angle (as usual by ruler and compass). Yet, as we know it today thanks to Galois' theory, the trisecting of arbitrary given angle by ruler and compass is not possible. This latter result is metamathematical (relative to Euclid's mathematics) because it belongs to a different theory, to wit Galois' theory, built on different principles. In what follows I call a mathematical theory M_E *meta*mathematical with respect to mathematical theory E just in case theorems of M_E describe relevant general features of E such as its consistency, limits of its methods, etc. (This informal description of the concept of metamathematics is broader and less exact than formal conceptions found in standard textbooks (Kleene 2009) but it better elucidates the epistemic role of metamathematics than more technical definitions.)

Let us call a theory *epistemically complete* when it allows one to solve – positively or negatively – any well-posed problem within this very theory (see (Pastin 1979) for a formal treatment). The concept of epistemic completeness raises a bunch of questions such as: Are epistemically *in*complete theories (epistemically) acceptable?; Is the epistemic completeness epistemically valuable?; How does epistemic completeness relate to the semantic completeness and other forms of theoretical completeness? and, last but not least, Are epistemically complete theories possible? David Hilbert was occupied with issues of epistemic completeness during all of his long-term work in FOM. His view on this matter was formed in the context of the so-called *Ignoramibusstreit* (*ignoramibus* debate; the debate about unknowable) that was started back in 1872 by Emil du Bois-Reymond who claimed that certain scientific questions and problems are unsolvable in principle. Responding to du Bois-Reymond Hilbert developed an opposite view according to which every well-established theory (and every mathematical theory at the first place) must be epistemically complete. Hilbert's belief remained remarkably stable during his long academic career. Hilbert expressed it publicly on many occasions, most famously in his 1900 address delivered in the Sorbonne at the International Congress of Mathematicians where he pronounced his famous "no ignoramibus" (in mathematics),

and also in his last major public lecture delivered in 1930 in Königsberg, which Hilbert concluded with the words "We must know. We will know" (McCarty 2005).

Hilbert believed that the epistemic completeness of mathematics (appropriately reformed) can, and should, be irrefutably proved by rigorous mathematical means, and not only be supported by general philosophical arguments. Whence the idea to identify FOM with a metatheory of all mathematics that could rigorously prove the epistemic validity of every mathematical theory deserving the name. Hilbert's metamathematical conception of FOM was based on the assumption that the wanted foundational theory could be kept minimal and form a reliable, trusted core of contemporary mathematics without imposing restrictions onto its conceptual contents. In 1927 Hilbert speculated that since "a formalised proof, like a numeral, is a concrete and surveyable object" proving the formal consistency of a given formalized theory "is a task that fundamentally lies within the province of intuition just as much as does in contentual number theory the task, say, of proving the irrationality of $\sqrt{2}$" (Hilbert 1967, p. 471). Retrospectively, it is clear that at this point Hilbert seriously underestimated the complexity of metamathematical tasks related to his foundational goals.

The question of how much (if any) of Hilbert's aspiration remains today sound in view of later FOM-related metamathematical and logical research has been widely discussed since 1930s (Zach 2003; Detlefsen 1990; Artemov 2020). Leaving aside this and other special questions about the Hilbert-style FOM I would like to stress once again that the very idea Hilbert had of providing FOM with a metamathematical capacity is highly original, and should not be simply taken for granted. From the traditional perspective on FOM, which has been outlined above, a metamathematical capacity of FOM appears as useful additional feature. It goes without saying that this new feature of FOM cannot serve its intended epistemic purpose unless the given FOM also performs its more basic functions, which Lawvere and Rosenburgh describe as "making explicit the essential general features, ingredients, and operations" of mathematics. Unless a given FOM performs these basic functions satisfactorily, its metamathematical power remains useless and irrelevant.

3 Is the Received FOM Adequate?

Any FOM has two sides. One is normative, logical, and epistemological. It deals with questions like: What is a correct mathematical proof? and provides tools for building sound mathematical theories and means for ruling out erroneous theories. The other side is "practical" and descriptive. It has to do with applications of the given FOM in mathematical practice including mathematics education and research, and, in particular, with the practical enforcement of epistemic norms associated with this given FOM.

An epistemological analysis of FOM needs to take both these sides into account. On the one hand, stripping away the normative dimension would reduce the epistemological analysis to historical, sociological, and psychological accounts and thus

destroy its specific character and purpose. On the other hand, ignoring the past and present mathematical practice deprives the epistemological analysis of mathematics of its subject matter. This concerns FOM just as much as any other part of mathematics. A normative epistemological judgement, which is grounded solely on general philosophical principles and does not accurately take into account what falls under its scope (in our case this is FOM as they are practiced), can hardly be justified.

The adequacy of the received Hilbert-style FOM is questionable on both these accounts. The most obvious problem concerns its *feasibility*: it turns out that attempts to represent standard mathematical contents in symbolic languages reserved by this FOM for this purpose require manipulations with chains of symbols, which cannot be represented in printed form and surveyed by a human. A notorious example is Bourbaki's definition of cardinal number 1 that comprises about 10^{15} symbols (Grimm 2010). At first glance the unfeasibility problem appears to be merely "practical," and can be successfully dealt with by using symbolic shortcuts such as "1." What is problematic here from an epistemological point of view is the relationships between mathematical theories represented "informally" or "semi-formally" with colloquial means and their formal counterparts represented with formal languages that make part of Hilbert-style axiomatic FOMs. The situation when theory T is represented formally as $\lfloor T \rfloor$ and the situation when T is represented by different means but can *in principle* be translated into form $\lfloor T \rfloor$ are very different from an epistemological point of view. In the latter case the given FOM have only a weak and indirect control over T, which works counter-positively. If T is shown to be *not* representable in the form $\lfloor T \rfloor$ then one concludes that the given FOM does *not* justify T. But the mere fact that T *is* representable in the form $\lfloor T \rfloor$ does not, by itself, allow one to judge whether T is valid by the given FOM's standards or not. Such a judgement is possible only when the formal representation $\lfloor T \rfloor$ is built effectively. Since $\lfloor T \rfloor$ is usually unfeasible, such a judgement is usually not available. This feature of the received Hilbert-style axiomatic FOM is at odds with the traditional expectation that mathematical theories can be justified by the very fact of being "based on" an appropriate FOM. Justification of a given mathematical theory is a basic FOM-related task, which is, generally, not performed via metamathematical considerations concerning the epistemic (in-)completeness of the given theory. Thus the feasibility problem has an epistemological dimension, which does not allow one to call this problem "merely practical."

Another problem concerns basic epistemological assumptions about mathematics that underpin the received axiomatic conception of FOM. Hilbert assumes that (ready-made) mathematical knowledge is representable in the form of axiomatic theories, and that such theories are systems of sentences, where each sentence is deducible from a set of distinguished sentences called axioms. This conception assumes that all established mathematical knowledge has a propositional character. This strong epistemic assumption about the mathematical knowledge typically extends to all scientific knowledge and is not specific for Hilbert's axiomatic approach to FOM. As Bertrand Russell remarked in 1900 "That all sound

philosophy should begin with an analysis of propositions, is a truth too evident, perhaps, to demand a proof" (for references and further discussion see (Gaskin 2009, p. v)).

However the assumption according to which all mathematical knowledge deserving the name has a propositional character does not square well with the existing mathematical practice. This can be clearly seen with the historical example of Euclid's theory of elementary geometry presented in Books 1–4 of his *Elements*. The first principles of Euclid's theory are of two types: geometrical *postulates* and *common notions* aka "axioms," which apply in the geometrical and in the arithmetical Books of the *Elements*. Euclid's postulates 1–3 are not propositions but schemes of elementary geometrical operations that intake some geometrical objects and output some other objects. (I assume here that a proposition, generally, admits for a truth-value. Euclid's postulates 1–3 are not propositions because they do not admit for truth-values.) For example, Postulate 1 is a scheme of operation that intakes two different points A, B and outputs a straight segment AB. Unlike the postulates the common notions 1–3 do operate with propositions, namely with propositions of the form $A = B$ (equalities). However, as Lewis Carrol rightly stressed back in 1895, these common notions should be read, on the pain of infinite regress, not as propositions but rather as schematic rules of inference, i.e., schemes of operations that allow one to obtain new equalities from some given equalities (Carroll 1895). Under this reading Common Notion 1 intakes two equalities of the form $A = C$ and $B = C$ and outputs equality of the form $A = B$.

Accordingly, the main body of Euclid's geometrical theory comprises items called *problems* along with theorems. Euclid's problems are recipes for making geometrical constructions provided with proofs that these constructions have their desired properties; these recipes are not expressed in a propositional form. For example, Problem 1 that opens Book 1 of the *Elements* tells us how to construct an equilateral triangle given its side; it provides a construction using Postulates 1,3, and a proof (using Common Notion 1) that the output of this construction qualifies as an equilateral triangle. Most of Euclid's geometrical theorems also involve constructions, which in this case are colloquially called "auxiliary." So problems and theorems in Euclid's *Elements* are intertwined and form a single deductive order, where earlier solved problems and earlier established theorems are used for solving further problems and establishing further theorems. Prima facie there is no reason for thinking that the operational *knowledge-how* expressed in Euclid's *Elements* in the form of problems is epistemically less significant than the propositional *knowledge-that* expressed in Euclid's geometrical theorems. (Logical reconstruction of Euclid-style geometrical reasoning remains problematic and cannot be more systematically explored here. For a more detailed analysis of Euclid's geometrical reasoning see [55, Ch. 1.2]. Back in 1974 Ian Müller claimed that no contemporary system of modern logic could adequately account for Euclid's form of geometrical reasoning (Müller 1974), see also (Müller 1969). In my opinion this situation did not significantly change during the following decades.)

Hilbert's work in the axiomatic foundations of Euclidean geometry (David Hilbert 1899) demonstrated that a rigorous propositional rendering of Euclid-style

traditional geometry is possible. However, it also showed that such a rendering is not innocent from a logical point of view, and requires a fundamental revision of Euclid's style of reasoning. In the introductory part of their *Foundations of Mathematics* published in 1934 Hilbert and Bernays stress the difference between their propositional axiomatic approach in theory building and Euclid's "genetic" aka "constructive" approach, which involves nonpropositional operations such as construction of new geometrical objects from given objects. The authors do not dismiss the genetic approach as an outdated version of their axiomatic approach but recognize its specific character, and conceive of a possibility to combine the two approaches (Hilbert and Bernays 1934, v. 1, Introduction).

The traditional Euclidean pattern of geometrical reasoning presented above remains very much alive in today's mathematics; various patterns of "genetic" reasoning coexist in today's mathematics with the axiomatic reasoning (Rodin 2014, p. 35 ff). As we shall shortly see (Sect. 4.3 below), the univalent foundations of mathematics proposed in the mid-2000s provides a formal framework where the two patterns of mathematical reasoning work jointly. Thus the assumption according to which all ready-made mathematical knowledge is essentially propositional is at odds with the current mathematical practice where various forms of knowing-how continue to play a major role.

Let us now consider the normative dimension of this assumption. A theoretical epistemological objection to this assumption can be formulated as follows. *Models* are at least as important constituents of mathematical theories as the sentences that these models make true, and arguably are even more important (as will be explained shortly). Today's mathematics is model based with very little exception. Mathematical models (but not only mathematical theories construed as systems of sentences) play a key role in today's science. The thesis that mathematical knowledge reduces to propositional knowledge does not do justice to the epistemic role of mathematics in science and technology.

Back in the 1950s Patrick Suppes put forward an influential epistemological thesis according to which scientific theories should be identified with classes of models rather than with certain syntactic structures (like $\lfloor T \rfloor$) that are interpreted as systems of sentences. Since one and the same class of models can, generally, be determined by different theories identified syntactically, theories in Suppes' sense are invariant or "robust" with respect to a range of possible syntactic choices used for axiomatic presentation of these theories (Suppes 2002). Suppes' view of theories is known in the literature under the name of *semantic* aka *nonstatement* view (Halvorson 2016). Suppes and his followers developed this view of theories attempting to apply Bourbaki-style axiomatic techniques in formal presentations of physical and other scientific theories. In spite of the fact that they referred to Bourbaki's project explicitly (Suppes, 2002, p. 33), (Stegmüller 1979) the discussion of the concept of theory for some reason remained limited to philosophy of science and never extended to philosophical logic and philosophy of mathematics. Similar views in philosophy of mathematics, however, were in different words expressed by proponents of *mathematical structuralism* who were equally motivated by Bourbaki's project but preferred to use Bourbaki's term "structure" instead of the

term "model" avoiding in this way a terminological interference with model theorists (Hellman 2006).

The standard semantic approach in FOM does not do full justice to the model-based character of mathematical reasoning because it offers a new way of thinking about existing formal techniques (to wit, about the Hilbert-style axiomatic method empowered with Tarkski-style formal semantics) without developing new techniques appropriate for the purpose. True, building models (structures) and experimenting with them is a key element of today's mathematics. But the Hilbert-style axiomatic FOMs provide no means for the formal encoding such operations and giving them a stable theoretical form. This is because such FOMs comprise only formal rules, which are logical rules including rules of truth-preserving logical inferences. FOMs of this type do not include rules similar to Euclid's postulates that regulate manipulations with nonpropositional objects. Hilbert-style FOM may justify the *existence* of a given mathematical structure with certain properties but is indifferent as to how this structure is obtained or "found." The idea that one may stipulate an object or structure with certain desired properties by fiat without asking how to build such an object from some familiar elements makes the axiomatic method in mathematics very useful and powerful. But the constructive aka *genetic* reasoning that does take such issues into account is at least equally important in mathematics (Term "constructive" in logic and foundations of mathematics is highly overloaded. Following Hilbert and Bernays, I call here a mathematical reasoning constructive when it involves nonpropositional operations such as building complex mathematical objects (aka constructions) from some given simple constituents, see (Rodin 2018) for further discussion.). This constructive aspect of mathematics is particularly important in the applied mathematics where mathematical methods (applied outside of pure mathematics) have a major epistemic role. As we shall see in Sect. 4.3 below, the univalent FOM provides a technique for encoding mathematical knowledge-how and implementing this knowledge computationally.

4 Foundations at Work

In this section I provide a very brief overview of three different "practical" FOMs that emerged in the twentieth and early twenty-first centuries: Nicolas Bourbaki's *Elements of Mathematics*, the category-theoretic foundations first proposed by William Lawvere and the univalent foundations more recently proposed by Vladimir Voevodsky. This list is not supposed to be complete: there is a number of other interesting FOM-related developments in the twentieth-century mathematical practice that I do not consider here. (This concerns, in particular, early attempts to adopt the Hilbert-style axiomatic presentation of elementary geometry in school education (Halsted 1904) and current attempts to introduce computer-assisted proofs into mathematical practice, which are independent of univalent foundations (Avigad 2018).) It is essential to keep in mind that we talk here not about some accomplished foundational theories per se but about continuing efforts to organize some large fragments of mathematics in a particular way with certain open-ended foundational

theories: Set theory in the case of Bourbaki, Category theory in case of categorical foundations, and Homotopy Type theory (HoTT) in case of univalent foundations. The three projects started at different times: Bourbaki in the mid-1930s, category-theoretic foundations in early 1960s, and univalent foundations in the mid-2000s. (More exact dates are, of course, conventional. Bourbaki group has been created in 1934 (Baulieu 1994). The idea of category-theoretic FOM has been first put forward by Lawvere in 1963 in his Ph.D. thesis (Lawvere 1963). The idea of univalent foundations (at this point without the name) has been first publicly presented by Voevodsky during his talk in the Institute of Advanced Studies (Princeton) on March 22, 2006 (Voevodsky 2006).)

Since the development of each later project involves a critique of earlier foundational projects and purports to fix their shortcomings, it is not unreasonable to regard these three projects as consecutive stages of the same overall process. However, the real history of these projects does not fit into such a linear Hegelian scheme. Bourbaki-style set-theoretic FOM, category-theoretic FOM, and univalent foundations (UF) all play roles in current mathematical practice – typically not in their "pure" forms but combined with some older foundational schemes and patterns. These roles are very different in the three cases. The set-theoretic FOM still maintain their "official" status (at least in eyes of the majority of logicians) and elements of Bourbaki-style presentation of mathematical reasoning are still widely used in practice (including mathematics education) in spite of the general skepticism about Bourbaki's foundational project shared by the majority of working mathematicians. Category theory is presently widely used as a convenient abstract theoretical language that extends and in some cases wholly replaces the Bourbaki-style set-theoretic language in many areas of mathematics. Situations when the category theory has a role of *foundation* are rarer and more specific; some such examples are referred to below. Finally, UF is a relatively new project (work in progress) aiming, in Voevodsky's words, at "a computerized version of Bourbaki."

4.1 Bourbaki

The continuing multivolumed *Elements of Mathematics* produced by a group of (mainly French) mathematicians of several generations that uses the pseudo-name of *Nicolas Bourbaki* is a unique long-lasted attempt to build Hilbert-style axiomatic foundations of mathematics and implement them into a broad mathematical practice in research and education. (The first volume of Bourbaki's *Elements* was (first) published in 1939 (Bourbaki 1939). The latest original volume appeared in 2016 (Bourbaki 2016); in 2019 appeared a new edition of an earlier published volume. Further references and updates concerning the work of Bourbaki's seminar and recent publications of the group can be found on the group's webpage https://www.bourbaki.fr) The title "Elements of Mathematics" refers to Euclid's *Elements*; following Euclid's example Bourbaki aims at providing a self-contained introduction into the contemporary mathematics, which systematically presents not only its principles but also its basic contents. The original French title is *Eléments de*

mathématique, which uses the unusual singular form "mathématique" that expresses Bourbaki's goal to unify mathematics. A concise general description of this project, which makes explicit some grounding ideas behind it, has been published in 1950 as a separate programmatic article titled the "Architecture of Mathematics" (Bourbaki 1950). (The article has been written by Jean Dieudonné and signed by the name of Nicolas Bourbaki, see [Corry 2004, Chap. 7].)

After a recognition of the unifying power of the Hilbert-style axiomatic method Bourbaki makes an interesting move by distinguishing between the *logical* aspect of the axiomatic method from another aspect, which can be called *structural*; in Bourbaki's view this is the latter rather than former aspect that makes the axiomatic method a powerful instrument of the unification:

> [E]very mathematical theory is a concatenation of propositions, each one derived from the preceding ones in conformity with the rules of a logical system [...] It is therefore a meaningless truism to say that this "deductive reasoning" is a unifying principle for mathematics. [...] [I]t is the external form which the mathematician gives to his thought, the vehicle which makes it accessible to others, in short, the language suited to mathematicians; this is all, no further significance should be attached to it. What the axiomatic method sets as its essential aim, is exactly that which logical formalism by itself cannot supply, namely the profound intelligibility of mathematics. (Bourbaki, 1950, p. 223)

In order to understand Bourbaki's misgivings about the "logical formalism" it is essential to take into account how Bourbaki's axiomatic method differs from Hilbert's. Using the term borrowed from the philosophy of science I have characterized in the last section Bourbaki's method as *semantic*. Bourbaki's "axiomatic" theories are not Hilbert-style theories provided with some default interpretations but theories of set-theoretic models of such Hilbert-style theories. For example, each particular group (identified either up to set-theoretic equality or up to isomorphism) qualifies as a model of Bourbaki's axioms for group theory (nowadays standard). But group theory does not reduce to proving general theorems that follow from these axioms alone and thus apply to all groups indiscriminately. For example, the elementary Lagrange theorem does not belong to this kind because it involves two different groups, i.e., two models of the axioms. Since all such models are set theoretic, Lagrange theorem as well as any other group-theoretic theorem proved by Bourbaki is deducible from axioms of set theory. This is how Bourbaki's semantic approach makes set theory into a core of their FOM.

Interestingly enough Bourbaki expresses some misgivings about set theory along with the aforementioned misgivings about the "logical formalism":

> [T]he axiomatic studies of the nineteenth and twentieth centuries have gradually replaced the initial pluralism of the mental [mathematical] representation [...] first to the concept of the natural number and then, in a second stage, to the notion of set. This latter concept, considered for a long time as "primitive" and "undefinable", has been the object of endless polemics [...]; the difficulties did not disappear until the notion of set itself disappeared [...] in the light of the recent work on logical formalism. From this new point of view, *mathematical structures* [my emphasis - A.R.] become, properly speaking, the only "objects" of mathematics. (Bourbaki 1950, pp. 225–226)

The provided references do not allow one to identify the precise sense in which Bourbaki talks here about the "disappearance of sets." But the above quotation makes it clear that already at an early stage of their project Bourbaki were motivated by the idea to represent mathematical structures without using set theory. Yet, officially, all Bourbaki's structures are construed set theoretically, namely, as set-theoretical models of certain systems of Hilbert-style axioms. The intended notion of "self-standing structure" (without an underlying set) is nowhere rigorously defined in Bourbaki's *Elements*. But as a matter of practical implementation of this intended notion Bourbaki systematically disregard details of their own set theory and its associated formal logical machinery.

A reason why Bourbaki feel so uneasy about their set-theoretic representation of mathematical structures concerns the idea (that historically dates back at least to Hilbert) that *isomorphic* structures should count as identical (MacLane 1996). The problem is that no (Hilbert-style) axiomatic set theory that comes with its internal identity relation provides a rigorous sense in which isomorphic structures can be said to be identical. We shall shortly see how this problem drives the category-theoretic approach in FOM and how it is settled in the univalent FOM (Ahrens and North 2019). In Bourbaki's *Elements* the notion of self-standing mathematical structure independent of any set-theoretic background remains a wishful thinking or, on a more charitable interpretation, a *regulative idea* in the Kantian sense of the word.

Outside a narrow group of enthusiasts who continue today to push further forward Bourbaki's project there are very few research mathematicians and mathematical educators who are ready to use Bourbaki's volumes in their teaching and build their research on them. Nevertheless it is clear that Bourbaki's *Elements* indeed "make explicit the essential general features, ingredients, and operations" of certain style and pattern of mathematical thinking that has played a major role in twentieth-century mathematics. Even if a number of important achievements of twentieth-century mathematics have been wholly unrelated to Bourbaki's edifice, Bourbaki provided a certain vision of mathematics as a whole and contributed to its unification and connectedness by developing the semi-formal "set-theoretic language" that is still routinely used in various fields and areas of mathematical research and helps their cross-fertilization. Even if there are serious reasons today to regard Bourbaki's set-theoretical FOM as outdated it still serves mathematicians and students of mathematics for many everyday purposes.

The story of the rise and fall of the Bourbaki-based approach in the elementary mathematics education, which took place worldwide in the course of twentieth century, deserves special attention. It starts with an attempt at reforming school mathematical education on the national level according to the new standards known as the *New Math* in the USA at the end of 1950s as a key element of the American response to the Soviet launch of the first Sputnik in 1957 (Phillips 2015). This educational reform aimed at the reinforcement of national technological and intellectual power and promoted the Bourbaki-style approach at the elementary school level. It was led by a number of prominent American mathematicians including Marshall Stone and was predictably opposed by the majority of school teachers and pupils' parents. More remarkably, this reform was also severely criticized by some

leading mathematicians and scientists, including such a prominent figure of the time as Richard Feynman (Feynman 1965). The opinion that the *New Math* was a pedagogical error prevailed already in the early 1970s (Kline 1973), and this controversial reform was soon largely abandoned.

Ironically or not, in the Soviet Union a similar radical reform of mathematical education was started almost simultaneously with the American *New Math*. In 1959 Boltyansky, Vilenkin, and Yaglom published a programmatic paper (Vilenkin NYa et al. 1959) where they urged for renewal of standard school mathematical curriculum. In the late 1960s the reform of mathematical education was supported by Soviet authorities and soon developed into a project of national scale; a Bourbaki-style version of school mathematical curriculum was chosen for implementation in school textbooks, developed by a group of leading mathematicians and mathematical educators under the general supervision of Andrey N. Kolmogorov. The fate of this Soviet reform was similar to that of the American *New Math*: in the late 1970s under the pressure of complaining school teachers, pupils' parents, and harsh critique of some mathematicians, this reform was abandoned and the new mathematical textbooks were again replaced by more traditional ones.

Similar reforms were undertaken in France and some other countries, and their fate was also similar. These national reforms of school mathematics did not proceed in isolation; they were parts of an international trend in mathematics education where the leading role belonged to French mathematicians including André Lichnerowitcz and some other members of the Bourbaki group.

The failure of Bourbaki-style approach in the elementary mathematics education cannot be fully analyzed here. Let me only mention that proponents and critics of the Bourbaki-style approach in mathematics and mathematics education agreed that this approach was not helpful for pointing to, illuminating, and supporting applications of mathematics in science and technology (Stone 1961; Feynman 1965; Arnold 1998). However they evaluated this fact differently: while the opponents argued that linking mathematics with science and technology was essential, the proponents of Bourbaki's approach argued that it was not.

4.2 Category Theory as a Foundation

Explaining the legacy of set theory in the twentieth-century mathematics, Yuri Manin remarks that in the second half of this century "sets gave way to categories" (Manin 2002, p. 7). The mathematical concept of *category* emerged in 1940s, its official birth is marked by 1945 paper by S. Eilenberg and S. MacLane (Eilenberg and MacLane 1945). Category theory (CT) soon proved to be a useful mathematical language that helped to organize a number of emerging concepts and new results into a stable theoretical form. Among early examples that justified the usefulness of CT in eyes of many mathematicians were (Eilenberg and Steenrod 1952; Eilenberg and Cartan 1956; Quillen 1967). Some new mathematical concepts and mathematical theories developed during the same period would not even emerge without CT. A number of today's mathematical disciplines including algebraic geometry,

homological algebra, and functional analysis use CT as their basic language (Kromer 2007; Marquis 2009).

The success of CT as an alternative language that allowed for effective theoretical accounting for mathematical structures in terms of their maps bypassing the redundant set-theoretic details appeared to many mathematicians, among them some Bourbaki members (including H. Cartan, S. Eilenberg, and A. Grothendieck) as a way to realize Bourbaki's dream about the "disappearance of sets" in their role of universal foundational background. Remarkably, Samuel Eilenberg and Norman Steenrod in the *Preface* to their (Eilenberg and Steenrod 1952) as well as Daniel Quillen in (Quillen 1967) boldly claim that using the language of CT they build "axiomatic" theories. Their approach was indeed axiomatic in its spirit in the sense that these authors stipulated the existence of categories with some special properties and on this basis developed abstract theories, which generalized upon many earlier known examples. However, these theories did not meet standards of formal and logical rigor used by mathematicians working in FOM, and in fact did not pursue this goal.

The idea to develop CT into a genuine FOM is due to William Lawvere. In his 1963 Ph.D. thesis (Lawvere 1963) he proposed Hilbert-style axioms for CT that are sometimes referred to in the current literature as Eilenberg–MacLane axioms and the corresponding theory is called EM; in the same thesis Lawvere built the elementary theory of category of sets (ETCS), where the word "elementary" refers to the first-order character of this theory, see also (Lawvere 2005) for a more detailed exposition of this work. Unlike the aforementioned theories of homological algebra and homotopy (Eilenberg and Steenrod 1952; Quillen 1967) that also use the language of categories and are called by their authors "axiomatic," ETCS is an axiomatic first-order theory in the standard sense of the term familiar to logicians and philosophers. Later Lawvere developed a more general project of category-theoretic FOM, namely, an axiomatic theory of category of categories (category of categories as a foundation or CCAF for short) (Lawvere 1966). In these early works Lawvere challenged the notion according to which set theory in some form should be a core part of modern FOM. In particular, he showed the possibility to develop an axiomatic set theory on the basis of axiomatic CT rather than the other way round. However at least at this stage of his career he did not challenge the notion according to which a modern FOM should involve the Hilbert-style axiomatic approach in some form.

The most remarkable of Lawvere's achievement in the category-theoretic axiomatization of mathematics is his work in topos theory. The concept of *topos* emerged in the circle of Alexandre Grothendieck in the early 1960s as a far-reaching generalization of the concept of topological space. Grothendieck's concept of topos had no special relevance to logic; the discovery of logical structure of topos is wholly due to Lawvere and his foundational approach. In the beginning of his 1970 paper (Lawvere 1970) Lawvere provides his definition of topos usually called today the definition of *elementary* topos; the title "elementary" reflects the fact that Lawvere's definition unlike Grothendieck's original construction almost straightforwardly translates into the standard first-order formal language (McLarty 1992).

A necessary and sufficient condition under which a given elementary topos is a Grothendieck topos was found in 1972 by Grothendieck's student Jean Giraud (Giraud 1972).

Even if Lawvere's theory of elementary topos qualifies as a Hilbert-style axiomatic theory it has a feature, which makes it quite unlike ZFC and other similar theories. The standard approach applied by Hilbert and later described and justified by Tarski from a metodological viewpoint assumes that an axiomatic theory is based on a certain symbolic calculus with fixed logical semantics, which allows one to formulate some specific axioms and then interpret nonlogical terms of this syntactic structure in certain intended models. The logical part of such theories, i.e., their "underlying logic," is supposed here to be rigidly fixed in advance (Tarski 1941). Lawvere's approach is different in this respect. His insight behind the axiomatization of topos theory is that every topos has an internal logical structure, which is analogous to that present in the category of sets. As Lawvere puts it "logic is a special case of geometry" (Lawvere 1970, p. 329). This understanding of relationships between logic and geometry is very unlike Hilbert's and Tarski's.

Studies in the category-theoretic FOM not only provided a new grounding for CT independent of set theory but also helped to simplify many category-theoretic concepts and apply them in various areas of mathematics as well as outside of pure mathematics, in particular, in computer science. Lawvere's axioms for elementary topos made topos theory more accessible and helped many people outside the community of experts in algebraic geometry to enter this field and conduct fruitful research in it. Although the category-theoretic FOM so far did not have a general impact on mathematics comparable with Bourbaki's impact it also had a significant effect on mathematics education (at the university level) and research.

4.3 Univalent Foundations

The univalent foundations (UF) is a novel foundation of mathematics first proposed by Vladimir Voevodsky in 2006 (Voevodsky 2006); UF uses homotopy type theory (HoTT) as a language and purports to express in this language various mathematical contents. HoTT in its original form is an interpretation of Martin-Löf's type theory (MLTT) in terms of homotopy theory where types are interpreted as homotopy spaces and terms of these types are interpreted as points of those spaces. This interpretation was discovered by Vladimir Voevodsky and, independently, by Steven Awodey and Michael Warren in the mid-2000s (Univalent Foundations Group 2013).

For Voevodsky a major motivation for developing UF was pragmatic and concerned the possibility of computer-assisted formal verification of long mathematical proofs. As we have already stressed, the received FOM does not support a formal checking and verification of mathematical proofs beyond trivial cases. However the practical need for such a reliable verification becomes more and more pressing. Even if the use of computer-assisted proofs remains today uncommon in the mainstream mathematical research, there is a growing corpus of

mathematical results proved with computer; in many such cases nonassisted "purely human" proofs are not known and there are indications that such proofs may not exist. A recent overview of existing proof assistants and presentation of the state of the art in the interactive computer-assisted theorem proving – which however does not cover the *UniMath* and other UF-motivated projects – is found in (Avigad 2018) and references therein.

It is remarkable that no existing proof assistants straightforwardly implement Hilbert-style formal representations of mathematical theories but use instead some form of type theory in their theoretical background. Per Martin-Löf designed and developed his constructive type theory with dependent types (MLTT) back in 1980s already in view of its possible implementation on computer; its fragments were effectively implemented in Coq and some other proof assistants. He provided this formal system with a *proof-theoretic* logical semantics, which essentially differs from Tarskian model-theoretic logical semantics used in the received Hilbert-style FOM (Piecha and Shroeder-Heister 2015); syntactically, MLTT is a Gentzen-style (rule-based) but not Hilbert-style (axiom-based) formal system. The discovery of homotopical interpretation of MLTT in the mid-2000s, which gave rise to HoTT, convinced Voevodsky that this theory can serve as a basis for a new FOM that supports an automated proof checking. Voevodsky called this FOM *univalent* because of the *univalence axiom* (UA) explained in what follows. The origin of term "univalence" is somewhat arbitrary: it comes from the expression "faithful functor" translated into Russian on some occasion as (in the backward translation due to Voevodsky) "univalent functor" (Grayson 2018, footnote 4). See (Grayson 2018) for a concise introduction into UF, and (Univalent Foundations Group 2013) for a more systematic exposition.

The homotopical semantics of MLTT allows one to distinguish between types of different *homotopy levels* and interpret types of different levels differently: types of level (-1) are interpreted as propositions, types of zero level are interpreted as sets, and types of level 1 are interpreted as groupoids (namely, fundamental groupoids of homotopy spaces) and so on up to higher homotopy groupoids. Formal rules of the underlying type theory are interpreted, at the propositional level, as rules regulating logical inferences; at higher homotopy levels the same formal rules regulate geometrical (to wit homotopical) constructions, which make certain propositions true. Thus HoTT justifies the aforementioned view of Lawvere on logic as "a special case of geometry." This feature of HoTT makes the UF-based mathematical reasoning in some respects similar to the traditional Euclid-style geometrical reasoning where geometrical constructions help to prove theorems. Since homotopy-theoretic concepts are spatial and intuitive in low dimensions, UF helps to visualize and understand many abstract mathematical concepts and proofs. In particular, it helps to have an intuitive grasp of formal proofs represented with a computer code and ready for an automated verification. This feature of UF addresses the issue of "opaqueness" (i.e., lack of surveyability) of computer-assisted proofs widely discussed in the literature (Bradley Bassler 2006).

Like in case of the traditional geometrical reasoning the epistemic role of UF-based constructions does not, generally, reduce to witnessing certain

mathematical sentences but have other important epistemic roles. Notice that by constructions I mean here rule-based mathematical procedures but not only their outcomes. Such constructions represent mathematical know-how in an explicit form. For a historical example think of Euclid's construction of five regular polyhedra aka platonic solids in Problems 13.13-17 of *Elements*, Book 13. This is the crowing achievement of Euclid's edifice; when one traces these constructions back to their first principles, i.e., to postulates and common notions of Book 1, then one can see that the necessary theoretical background of these five problems comprises a significant portion of the mathematical knowledge presented in the *Elements*. This large bulk of mathematical knowledge is packed in the five aforementioned problems in the operational but not propositional form. For another important historical example think of algebra in its early form presented by al-Khwarizmi in the book, which gave the discipline its current name (Rashed 2010). Al-Khwarizmi's algebra is a systematized collection of recipes (provided with an Euclide-style geometrical background) for solving problems of a certain sort. The current mathematical practice does not give one any indication that the role of the operational knowledge-how in mathematics is less significant today than it used to be in the past. The fact that UF unlike the standard Hilbert-style axiomatic FOMs allow for a formal and meticulous representation of operational knowledge in the form of UF-based homotopical constructions is one more reason in favor of the claim that UF are more adequate.

The significance of knowing-how is particularly evident in the applied mathematics where mathematics is seen as an epistemic tool rather than an independent discipline. Mathematical models of natural phenomena and technological artifacts always verify certain mathematical sentences; to put it colloquially, such models have certain mathematical properties. However, the main epistemic role of such models is not to instantiate and witness these mathematical properties but to account for empirical data. In such contexts knowing *how* to build an appropriate mathematical model is a key issue.

Let me briefly illustrate this point with two examples: the historical example of Ptolemean astronomy and the up-to-date example of topological data analysis (TDA). Ptolemean astronomy involves the traditional Euclid-style geometry as its mathematical foundation (but goes beyond the geometrical content of Euclid's *Elements* since it applies spherical geometry). Visible motions of celestial bodies are reconstructed as superpositions of circular motions. In case of planets this requires using a number of supplementary circular motions, which are not observable. The trajectories of such motions Ptolemy calls *epicycles*. The resulting geometrical model, by Ptolemy's famous word, "saves the phenomena," i.e., is empirically adequate (with a certain degree of approximation). Importantly, Ptolemy's geometrical models of astronomical phenomena are built step by step and thus adjusted to observational data; such genetic procedures are based on constructive geometrical principles closely related to Euclid's postulates (think of epicycles). Such a systematic way of model building is made possible by the constructive (genetic) character of Euclid-style geometry explained in Sect. 3 above.

In TDA, which is a rapidly developing area of today's applied mathematical research (Patania et al. 2017), large datasets of various nature are accounted for in

terms of topological spaces and their invariants. This allows one to extract from the given data some essential information, which turns out to be relevant and useful in many cases. All relevant characteristics of those topological spaces – and, in a sense, also these very spaces – are computed on the basis of the given data. Even if such a procedure can, in principle, be fully specified in terms of truth conditions within an appropriate external theoretical framework, UF offers a possibility to interpret such computations internally (and yet formally and rigorously) as genetic topological constructions. Such a direct internal interpretation better squares with the colloquial language and intuition of working mathematicians and computer scientists working in the field. This specific feature of UF has a potential to narrow the existing gap between today's abstract pure mathematics and applied mathematics (Rodin 2021).

Let us now turn to the univalence axiom (UA), which gives the univalent foundations its name. Let A, B be types, in symbols $A, B\ TYPE$. Consider the identity type $A =_{TYPE} B$ and the type of *equivalences* $A \simeq_{TYPE} B$ where by equivalence one understands a function $f : A \to B$, which is in an appropriate sense invertible; under the homotopical interpretation such equivalences are homotopy equivalences. The rules of MLTT/HoTT allow one to construct a canonical map of the form

$$e\,(A = B) \to (A \simeq B)$$

which witnesses the fact that identity is a special case of equivalence. The *univalence axiom* states that this map e has an inverse and thus is itself an equivalence. In other words, UA says that the type

$$(A = B) \simeq (A \simeq B)$$

is inhabited. More recent versions of UF based on cubical type theory (CTT) do not use UA as an axiom but prove it as a theorem. The aforementioned inverse map, which is not constructible by using the rules of MLTT, is constructible by using the rules of CTT. Thus in CTT-based UF there is no need to postulate univalence with a special axiom. This fact supports the claim that CTT is a better formal carrier for UF than MLTT (Bezem et al. 2019).

Univalence is important for a variety of reasons. A reason which has a particularly strong philosophical appeal is that the univalence property apparently justifies the structuralist idea of mathematical reasoning "up to isomorphism" (or up to another type of equivalence), which reflects the existing practice in the Bourbaki-style set-based mathematics and in the more recent category-theoretic mathematics but lacks a firm logical grounding in these foundational frameworks. A detailed analysis of this claim, which includes further qualifications and certain reservations, is found in (Ahrens and North 2019).

It should be born in mind that the homotopical semantics of MLTT is an extension (not fully conservative) of its intended logical semantics but is not a model of this theory in homotopy theory. *Models* of HoTT are further interpretations of its syntactic structures, which respect the homotopical semantics in an appropriate sense. When formal rules of HoTT are interpreted in external structures they can

be understood as rules for building such structures from their elements. This feature of UF has no analogue in the standard Bourbaki-style semantic approach to theory building where all structures are supposed to be "static." From a historical perspective this aspect of UF can be described as a synthesis of axiomatic and *genetic* approach, which has been considered by Hilbert and Bernays back in 1934 as a theoretical possibility (Rodin 2018).

Since UF is still a work in progress it is too early to evaluate its success and practical impact on mathematics against the impacts of earlier FOMs. However, already today one can see that UF is designed as a practical tool for learning and doing mathematics, which unlike the received Hilbert-style axiomatic FOMs does not suffer from the "meta-mathematical bias" (see Sect. 2 above). Thus UF presently appears as a strong candidate for bridging the existing wide gap between FOM and the rest of mathematics, between mathematical foundational theories and mathematical practices.

5 Renewing Foundations and Progress in Mathematics

In the remaining part of this chapter I would like to describe my general views about mathematics and its foundations in terms proposed by Michele Friend (Friend 2014). First of all, my position is *foundationalist* insofar as I believe that FOM is an important part of mathematics both practically and theoretically; I also believe that FOM is an important (albeit not unique) point of contact between mathematics and its philosophy. Second of all, I am a *pluralist* about FOMs. This should be clear from the earlier sections where I described several different FOMs without saying which is the right one, and which are wrong.

Since my working general concept of FOM borrowed from Lawvere and Rosenburgh is broader than the one used by Friend, my position is foundationalist in a weaker sense of this term than in Friend's sense. Friend defines FOM as "an axiomatically presented mathematical theory to which all or most of successful existing mathematics can be reduced" (Friend 2014, p. 8). I assume that Friend uses here the term "axiomatic" in its usual Hilbertian sense, and I qualify an axiomatic presentation of mathematics in *this* sense as a special feature of the received Hilbert-style FOMs rather than as an essential feature of FOM in general. Univalent foundation is an example of FOM that is not axiomatic in this usual sense because the UF applies a Gentzen-style (i.e., rule-based) but not a Hilbert-style formal architecture. (It makes sense to broaden the received notion of being axiomatic, so it may include also Gentzen-style formal systems. This terminological convention, however, creates a high risk of confusion since systems like MLTT that comprise no axiom in the received sense of this term would also qualify as axiomatic. For this reason I stick in this chapter to the received meaning of "axiomatic." For discussion of a broader conception of being axiomatic see my (Rodin 2018).)

Further, I do not assume together with Friend that mathematics can be "reduced" to its foundations in any epistemically relevant sense of this term. When such a

reduction is understood in a strong epistemic sense, so the reducibility condition reads:

In order to learn all existing mathematical knowledge it is sufficient to learn FOM.

this condition is obviously too strong and even absurd. When the reduction is understood as formal provability or in some other sense that has no definite epistemic content, the same condition is too week or irrelevant. The fact that all texts written in English can be encoded into 0-1 sequences does not make 0 and 1 into building blocks of all English literature. Similarly, the fact that a lot of mathematics can be encoded into fragments of ZFC does not, by itself, make ZFC into a foundation of mathematics. (I am not trying here to dismiss the set-theoretic FOM in one stroke and realize that set-theoretic FOMs involve more than formal provability. I would like to stress, however, that formal provability of theorem *T* in ZFC in the usual sense of existence of truth-preserving inference from the axioms of ZFC to *T* does not, by itself, shed a light on how *T* is (or is not) *known*, see (Prawitz 1986) for a discussion.) Instead of looking further for an appropriate notion of reduction I shall describe the relationships between mathematics and its foundations by looking at it from the opposite end. The following epistemological condition on FOM is, in my view, appropriate:

In order to acquire any fragment of the existing mathematical knowledge it is necessary to learn (a relevant fragment of) FOM.

The latter condition implies that FOM or their relevant fragment are necessary in justification of every mathematical result but it does not imply that FOM are sufficient for that purpose.

In spite of these differences between my working concept of FOM and Friend's concept, they share many essential features in common, in particular, a normative character, a capacity to unify mathematics, and a justificatory function. So my notion of being a foundationalist is not wholly at odds with Friend's, after all.

Let me now explain how I think about multiple FOMs. My first point is historical. I believe that the development of FOM is an important part of the overall historical development of mathematics. I assume that this overall development is progressive, and understand this progress as a growth of accumulated mathematical knowledge (Bird 2007).This notion of progress is compatible with a picture of mathematics that develops on some fixed foundations (say, Euclid's), so the growth of mathematical knowledge amounts to proving more and more geometrical theorems and solving more and more geometrical problems by ruler and compass. This process is potentially infinite, at least if we abstract away from issues of feasibility. However mathematics as we know it is more interesting, it simply does not develop in this simplistic way, at least not in the long term. Our real mathematics involves a permanent *revision* of foundations, which contributes to the overall mathematical progress.

Recall that the trisection of angle and some other geometrical problems are unsolvable on Euclid's foundations. These old geometrical problems were successfully solved (in the negative way) in the nineteenth century using the then new algebraic foundations of geometry laid out during the course of seventeenth to eighteenth centuries by Descartes and others. These new foundations (Arnauld 1683) were not a conservative extension of Euclid's foundations with certain additional principles but involved a very deep conceptual revision of these old foundations. However, this revision did not disqualify Euclid's foundations as being wholly erroneous either. Euclid is still with us, and we see nothing fundamentally wrong about his geometrical constructions by ruler and compass (except for some fixable logical gaps). This situation can be compared with the foundational change in physics that was brought about with the development of quantum mechanics: this foundational theory did not wholly disqualify the earlier accepted foundational theory (classical mechanics) but explained the success of this older theory while showing the limits of its applicability. The principles according to which new foundational theories in physics should relate to older foundational theories in this way is known as Bohr's *correspondence principle*. The same principle applies to foundational changes in mathematics.

In order to use the new foundations for solving geometrical problems formulated in a different foundational framework these problems had to be reformulated in a new way. What we learn today under the name of Pythagorean theorem is not quite the same statement that we typically identify as this same theorem in Euclid's *Elements* (Book 2, Theorem 47); the corresponding proofs of these statements are moreover very different. How and in which exact sense mathematical contents preserve their identities through such foundational changes is a tricky and little explored question, which I should leave aside. However the very fact that the acquired mathematical knowledge is capable to survive through foundational changes is remarkable.

The above example of foundational change in geometry is typical rather than exceptional. Sometimes new foundations help to identify old errors, but I do not know about examples of foundational revisions that have invalidated large bodies of mathematical knowledge established earlier on different foundations. Thus I observe that a revision of FOM typically has these three effects: (i) it preserves earlier acquired mathematical knowledge but expresses it in a new form, (ii) it helps to solve open problems, and (iii) it triggers new developments that were not possible with only the resources of the old foundations. So the continuing revision of FOM contributes to the overall progress of mathematics.

The above historical point does not commit me to a historical relativism about epistemic norms provided by different FOMs used in different historical times. It is quite justified to judge older FOMs by epistemic standards of more recent FOMs. One should bear in mind, however, that different FOMs often have very different designs, so some such judgements can be easily ill-formed. To repeat, new FOM, generally, does not disqualify the older FOM that they replace but put them in a new perspective. I can see no reason for thinking that such a renewal of FOM should ultimately converge to a certain "true" FOM. The continuing revision of FOMs has a

character of expansion (in its depths and scope) rather than convergence. I believe that this specific form of progress is worth encouraging rather than being stopped by choosing the current FOM as ultimate.

6 Tolerance and Pluralism

My final point concerns what Friend calls the principle of *tolerance* to conflicting philosophical views on mathematics associated with different FOMs (Friend 2014). What has been said above about the continuing revision of FOMs during the historical development of mathematics should not be understood in the sense that at any given moment of historical time mathematics relies or should rely on a unique FOM. In practice, mathematicians of the same generation often use different foundations, including more traditional and more "progressive" ones. The failure of the *New Math* and akin reforms in mathematics education in the 1960–1970s suggests that such a diversity is justified at least in education. I have described above how new foundations replace and "absorb" older foundations without disqualifying these older foundations completely. A similar scheme applies to contemporary FOMs. There exists a large body of mathematical knowledge that can be reconstructed on different alternative FOMs, say, on the set-theoretic and the category-theoretic FOMs. Let us consider a more problematic case when certain mathematical proof (construction) P is justified by one FOM (F) and not justified by another FOM (G). It appears that such cases make a tolerant attitude to conflicting FOMs impossible: in order to judge whether or not P is valid one needs to take a side and reject either F or G. This contradiction, however, can be avoided. There is a powerful mechanism that allows one to avoid the choice between F and G without becoming an epistemological relativist or a skeptic.

A more precise description of the above conflict between F and G involves two different versions of P, namely, P_F that construes and justifies P according to standards of F, and P_G which construes P according to standards of G and invalidates it. Now suppose that G admits an interpretation g of F that extends to P_F and provides an explanation in terms of G of how and why proof P_F is valid by the standards of F. Reciprocally, F admits interpretation f of G that extends to P_G and explains in terms of F why P_G is invalid by the standards of G. When such interpretations f, g exist, F, G are called *bi-interpretable*. These mutual interpretations can be pictured with the following diagram:

$$F/P_F \underset{f}{\overset{g}{\rightleftarrows}} G/P_G$$

So a user of G gets another reconstruction of P in terms of G, namely, $P'_G = g(P_F)$ which unlike P_G is valid by the standards of G. Reciprocally, a user of F gets another reconstruction of P, namely, $P'_F = f(P_G)$ which unlike P_F is invalid by the standards of F. The users of F and G now share a mutual understanding of principles needed for justifying or rejecting P. The user of F and the user G will think about P in

different ways and disagree about the epistemic status of *P*. But at this point their disagreements belong to the philosophy of mathematics rather than mathematics itself.

The above scheme does not apply universally but it applies to bi-interpretable FOMs, which include many cases of interest. For a trivial example consider the case when *F*, *G* are two different Hilbert-style axiomatic theories which share some axioms in common. One can use ZF as a foundation and use the axiom of choice (AC) in certain proofs as a separate hypothesis. Alternatively, one can take ZFC as a foundation, and this time view the AC as an axiom. From an epistemological and foundational point of view the two approaches differ. There is, however, a large body of mathematical knowledge that is wholly neutral with respect to choice between accepting AC as an axiom and using it as a hypothesis. In this case the interpretations $ZF \rightarrow ZFC$ and $ZF \leftarrow ZFC$ are trivial and amount to changing the epistemic status of AC from hypothesis to axiom and back. In cases when different FOMs use different formal architectures (as in the case of ZF and UF) and differ conceptually, such mutual interpretations can be very involved and nontrivial. The long-lasted continuing history of rebuilding Euclidean geometry on different foundations provides a lot of suggestive examples of this sort.

Some conditions stronger than the mere bi-interpretability can also be considered. Let *f*, *g* be composable in a associative way. If $g \circ f = 1_F$ and $g \circ f = 1_G$, i.e., the image of image of *F* in *G* is *F* itself and, reciprocally, the image of image of *G* in *F* is *G* itself, the two FOMs are equivalent – in a way which is specified by the character of these interpretations. The elementary category theory, this time in its role of a metatheoretical tool, suggests many other possible configurations between different FOMs, which I leave for another study.

The possibility of the coexistence of different FOMs in mathematical practice does not undermine the significance of philosophical arguments in favor of certain FOMs and against certain other FOMs. Such philosophical discussions have multiple impacts on mathematics including shaping its new trends. The tolerant attitude to multiple FOMs and their philosophical underpinnings, which I defend here, does not imply a neutrality and lack of interest to FOM-related philosophical matters. I understand tolerance as a methodological principle according to which a foundational debate in mathematics needs an environment, which is both concurrent and cooperative. As far as one accepts that conflicting philosophical ideas about mathematics deserve a common stage where they are systematically presented, criticized, and further developed, one should also accept the notion that a part of this stage belongs to mathematics and accommodates multiple FOMs. As I have explained above, the multiplicity of FOMs does not break mathematics down into disconnected pieces. Different FOMs may serve different special purposes and be useful in different ways. We showed that the received set-theoretic axiomatic FOM was designed for metatheoretical purposes and that this FOM cannot be effectively used for the purpose of formal proof verification. Using UF for this latter purpose does not exclude a possibility to use the received set-theoretic FOM for metamathematical purposes. I do not believe that these different epistemic tasks – to verify proofs and to provide a metatheoretical view on a given theory – can be adequately

distinguished as being practical and being theoretical. Both tasks have theoretical aspects that are analyzed in normative logical and epistemological terms and practical aspects that are described in terms of past experiences and recommendations for novices.

As many authors have earlier remarked, the traditional architectural metaphor behind the concept of theoretical foundation can be not only useful and suggestive but also very misleading. Unlike traditional buildings mathematical theories do not rise and fall with their foundations but have a capacity to survive through foundational changes. Martin Krieger compared mathematics of the nineteenth century with an industrial city (Krieger 2015, Chap. 6). In line of Krieger's metaphor FOM should be thought of as a city infrastructure, which uses elements inherited from earlier historical times. We may extend Krieger's metaphor and think about today's FOMs as even larger, more flexible, and more robust infrastructures, which are more suitable for our diverse postindustrial economies. Pluralism in Friend's sense is a sine qua non of FOM so conceived. My short answer to the question asked in the title of this chapter is this: one mathematic(s) and multiple foundations.

7 Cross-References

▶ Grothendieck: A Short Guide to His Mathematical and Philosophical Work (1949–1991)

References

Ahrens B, North P (2019) Univalent foundations and equivalence principle. In: Centrone S, Kant D, Sarikaya D (eds) Reflections on the foundations of mathematics: univalent foundations, set theory and general thoughts, vol 407. Springer, Synthese Library, pp 137–150

Arnauld A (1683) Nouveaux Eléments de Géométrie. Guillaume Desprez, Paris

Arnold VI (1998) On mathematical teaching (in Russian). Russ Math Surv 53(1):229–236

Artemov SN (2020) The provability of consistency. https://arxiv.org/abs/1902.07404

Avigad J (2018) al. Introduction to 'Milestones in interactive theorem proving'. J Autom Reason 61(1):1–8

Baulieu L (1994) Dispelling a myth: questions and answers about Bourbaki's early work, 1934–1944. In: Chikara S, Mitsuo S, Dauben JW (eds) The intersection of history and mathematics. Birkhäuser, Basel, pp 241–252

Bezem M, Coquand T, Huber S (2019) The univalence axiom in cubical sets. J Autom Reasoning 63:159–171. https://doi.org/10.1007/s10817-018-9472-6

Bird A (2007) What is scientific progress? Noûs 41(1):64–89

Bourbaki N (1939) Eléments de Mathématique. Théorie des Ensembles. Hermann, Paris

Bourbaki N (1950) Architecture of mathematics. Am Math Mon 57(4):221–232

Bourbaki N (2016) Eléments de Mathématique: Topologie Algébrique, ch. 1–4. Springer

Bradley Bassler O (2006) The surveyability of mathematical proof: a historical perspective. Synthese 148(1):99–133

Carroll L (1895) What the Tortoise Said to Achilles. Mind 4(14):278–280

Corry L (2004) Modern algebra and the rise of mathematical structures. Birkhäusert

Detlefsen M (1990) On an alleged refutation of Hilbert's program using Gödel's first incompleteness theorem. J Philos Log 19(4):343–377
Eilenberg S, Cartan H (1956) Homological algebra. Princeton University Press, Princeton
Eilenberg S, MacLane S (1945) General theory of natural equivalences. Trans Am Math Soc 58(2): 231–294
Eilenberg S, Steenrod N (1952) Foundations of algebraic topology. Princeton University Press, Princeton
Ferreiros J (2005) Dogmas and the changing images of foundations. Philos Sci CS(5):27–42
Feynman R (1965) New textbooks for the 'new' mathematics. Eng Sci 28(6):9–15
Friend M (2014) Pluralism in mathematics: a new position in philosophy of mathematics. Springer, Dordrecht
Gaskin R (2009) The unity of proposition. Oxford University Press, Oxford
Giraud J (1972) Classifying topos. In: Lawvere FW (ed) Toposes, algebraic geometry and logic (volume 274 of Lecture notes in mathematics). Springer, 2, pp 43–56
Grayson D (2018) An introduction of univalent foundations for mathematicians. Bull Am Math Soc (New Series) 55(4):427–450
Grimm J (2010) Implementation of Bourbaki's elements of mathematics in coq: part one, theory of sets. J Formalized Reason 3(1):79–126
Halsted GB (1904) Rational geometry: a text-book for the science of space; based on Hilbert's foundations. Wiley, New York
Halvorson H (2016) Scientific Theories. In: Humphreys P (ed) The Oxford Handbook of philosophy of science. Oxford University Press, Oxford
Hellman G (2006) Structuralism, mathematical. In: Borchert DM (ed) Encyclopedia of philosophy, vol 9, 2nd edn. Macmillan, New York, NY, pp 270–273
Hilbert D (1899) Grundlagen der Geometrie. Teubner, Leipzig
Hilbert D(1967) Foundations of mathematics. In: van Heijenoort J (ed) From Frege to Gödel: a source book in the mathematical logic, vol 2. Harvard University Press, Cambridge, Mass. pp 464–480
Hilbert D, Bernays P (1934–1939) Grundlagen der Mathematik. Springer, Berlin
Kleene SC (2009) Introduction to Metamathematics. Ishi Press International, San Jose, California
Kline M (1973) Why Johnny can't add: the failure of new Maths. St James Press, New York/ London
Krieger MH (2015) Doing Mathematics: convention, subject, calculation, analogy. World Scientific Publishing Company, Singapore
Kromer R (2007) Tool and object: a history and philosophy of category theory. Birkhäuser, Berlin
Lawvere FW (1963) Functorial semantics of algebraic theories. PhD thesis, Columbia University
Lawvere FW (1966) The category of categories as a foundation for mathematics. In: Proceedings of the La Jolla conference on categorical algebra, pp 1–21
Lawvere FW (1970) Quantifiers and sheaves. In: Berger M, Dieudonne J, et al (eds) Actes du congres international des mathematiciens, Nice, pp 329–334
Lawvere FW (2005) An elementary theory of the category of sets (long version) with the author's commentary. Reprints in Theory and Applications of Categories 11:1–35
Lawvere FW, Rosebrugh R (2003) Sets for mathematics. Cambridge University Press, Cambridge
MacLane S (1996) Structure in mathematics. Philos Math 4(2):174–183
Mancosu P et al (2008) Philosophy of mathematical practice. Oxford University Press
Manin Y (2002) Georg Cantor and his heritage (talk at the meeting of the German mathematical society and the Cantor medal award ceremony). arXiv:math.AG/02009244 v1/
Marquis J-P (2009) From a geometrical point of view: a study of the history and philosophy of category theory. Springer, Dordrecht
McCarty DC (2005) Problems and riddles: Hilbert and the Du Bois-Reymonds. Synthese 147:63–79
McLarty C (1992) Elementary categories, elementary Toposes. Clarendon Press, Oxford
Müller I (1969) Euclid's elements and the axiomatic method. Br J Philos Sci 20:289–309

Müller I (1974) Greek mathematics and Greek logic. Ancient logic and its modern interpretations. Synthese Historical Library, vol 9. pp 35–70
Pastin M (1979) The need for epistemology: problematic realism defended. In: Pappas GS (ed) Justification and knowledge. New studies in epistemology. D. Reidel Publishing Company, Dordrecht, pp 151–168
Patania A, Vaccarino F, Petri G (2017) Topological analysis of data. EPJ Data Sci 6(7). https://doi.org/10.1140/epjds/s13688-017-0104-x
Phillips CJ (2015) New math: a political history. The University of Chicago Press, Chicago
Piecha T, Shroeder-Heister P (eds) (2015) Advances in proof-theoretic semantics. Trends in logic, vol 43. Springer
Prawitz D (1986) Philosophical aspects of proof theory. G. Fløistad, La philosophie contemporaine. Chroniques nouvelles, Philosophie du langage 1:235–278
Quillen D (1967) Homotopical algebra. Lecture notes in mathematics, no. 43. Springer, New York, NY
Rashed R (2010) Al-Kwarizmi: The Beginning of algebra: History of Science and Philosophy in Classical Islam; London: Saqi Books
Rodin A (2014) Axiomatic method and category theory. Synthese library, vol 364. Springer, Dordrecht
Rodin A (2018) On constructive axiomatic method. Logique et Analyse 61(242):201–231
Rodin A (2021) Voevodsky's unfinished project: Filling the gap between pure and applied mathematics. Biosystems vol. 204, 104391
Stegmüller W (1979) The Structuralist view of Theories: a possible analogue of the Bourbaki Programme in physical science. Springer, Berlin
Stone M (1961) Revolution in mathematics. Am Math Mon 68(8):715–734
Suppes P (2002) Representation and invariance of scientific structures. CSLI Publications, Stanford
Tarski A (1941) Introduction to logic and to the methodology of deductive sciences. Oxford University Press, New York
Taylor P (1999) Practical foundations of mathematics. Cambridge University Press, Cambridge
Univalent Foundations Group (2013) Homotopy type theory: univalent foundations of mathematics. Institute for Advanced Study (Princeton). Available at http://homotopytypetheory.org/book/
Vilenkin NYa, Boltyansky VG, Yaglom IM (1959) On the content of school mathematical curriculum (in Russian). Matematicheskoe Prosveschenie (Mathematical enlightement) 4:131–143
Voevodsky VA (2006) Foundations of mathematics and homotopy theory (lecture delivered at the Princeton Institute of Advanced Studies on march 22, 2006). https://www.math.ias.edu/vladimir/Lectures
Zach R (2003) Hilbert's program. Stanford encyclopedia of philosophy

Multiverse and the Society

Mirna Džamonja

Contents

1 Introduction 2366
2 Notation and Assumptions 2367
3 On the Methodology of Mathematical Modeling 2368
3.1 Set Theory as a Modeling Tool 2368
4 Formulating the Societal Interpretation 2370
4.1 ZF in the Societal Interpretation 2370
4.2 The Axiom of Infinity and Relativization 2371
4.3 The Remaining Axioms 2374
5 Plurality of Models of Set Theory 2374
6 Universe and Models in the Societal Interpretation 2376
6.1 The Silent Majority 2376
6.2 Universal Democracy Is Undefined 2377
6.3 Back to Kant 2378
7 Conclusion 2379
8 Appendix: Axioms of ZF 2380
References 2381

Abstract

This chapter follows the philosophical line of mathematical pluralism to make two points: that if there is a view that mathematics is a basis for sciences including social sciences and if there is a pluralism in mathematics, then this implies that the

This research was supported by the European Union Horizon 2020 research and innovation programme under the Maria Skłodowska-Curie grant agreement No. 1010232. The author would also like to thank the Institut d'Histoire et de Philosophie des Sciences et des Techniques CNRS-Université Paris 1 Panthéon-Sorbonne, Paris, where she is an Associate Member and the School of Mathematics at the University of East Anglia in Norwich, where she holds the title of a Visiting Professor.

M. Džamonja (✉)
IRIF(CNRS & Université de Paris-Cité), Paris, France
e-mail: mdzamonja@irif.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_29

only way to model society is a pluralist one. Along the way to this conclusion, we argue why mathematical logic is a good part of mathematics to use in modeling and, in particular, explain the use of set theory in this context.

Keywords

Pluralism · Set theory · Society

1 Introduction

Mathematicians of the eighteenth century, starting with the celebrated Carl Friedrich Gauss, were the first in the Western world to question the uniqueness of mathematical interpretation. Thinking of geometry on the surface of the globe, Gauss realized that it exhibits a non-Euclidean geometry, consistent with all but the fifth postulate of Euclid. Due to its departure from the then established view that the whole of mathematics is based on the axioms of Euclid, the core of mathematics, this discovery was quite shocking to Gauss. So shocking that he did not publish it.[1] Yet, there is no way to stop history and the fact was re-discovered independently by János Bolyai in 1832 (Bolyai 1832) and Nikolaï Lobatchewsky in 1837 (Lobatschewsky 1837). Speaking of this and the similarly dramatic discovery of non-commutative algebra, Paul Dirac has said in (Dirac 1931) "Non-euclidean geometry and non-commutative algebra, which were at one time considered to be purely fictions of the mind and pastimes for logical thinkers, have now been found to be very necessary for the description of the general facts of the physical world."

Following this line of reasoning, this chapter demonstrates the usefulness as a modeling tool of mathematical logic, a part of mathematics that some contemporaries may consider as "purely fictions of mind and pastimes for logical thinkers," while in fact it has "often been found to be very necessary for the description of the general facts of the physical world." We concentrate on an aspect of mathematical logic that is considered even more abstract than the others, set theory, and on its application as a model of what is even more physical than the physical world studied by Dirac: the society. Set theory is considered so abstract for reasons similar to the non-Euclidean geometry being considered too abstract in its own time: it shakes the conviction that there is only one universe of mathematics, as much as geometry shook the conviction in the universality of the Euclidean thought. Both geometry and set theory are examples of the phenomenon of mathematical pluralism, where mathematicians have to find ways to grasp the situation in which absolute truths are a rarity, rather than a norm that they were once thought to be. The reader is probably ready to accept that plurality, given that mathematicians have been

[1] In fact, Gauss was not the only one. In 1733, Giovanni Girolamo Saccheri (1733) and in 1766 Jean-Henri Lambert, separately found counter-examples to the parallel postulate (see Kagan 1974). Saccheri's also did not quite know how to treat this discovery and it only appears implicitly in his final work, while Lambert discretely added it to the second edition of his book (Lambert 1759).

studying it for several centuries. We would like to use that good will of the reader to explain a logical consequence: if we accept that mathematics is a basis for sciences, including social sciences, and if there is a pluralism in mathematics, then this implies that the only way to model the society is a pluralist one.

The chapter will lead to accept that pluralist model of the society on the scientific basis, while using set theory as foundations. It will not add any knowledge to set theory, nor to the discussion of the place of set theory within science. It will rather try to use our experience in set theory and the multitude of its universes to understand something about the order of the human society.[2]

History and heuristics seem to show that the only chance for a society to stand united in the moments of crisis is the acceptance of the plurality of opinion. We shall defend the view that this conclusion follows from a very fundamental logical model. That conclusion feeds into our philosophical position which is that the acceptance of the plurality of opinion is crucial both in mathematics and in the society.

2 Notation and Assumptions

In this chapter, we understand the standard Zermelo-Fraenkel axiomatization of set theory given in the first-order classical logic. This is known as ZF. The Axiom of Choice could be added to this to obtain ZFC, as it is known since Kurt Gödel's work in (Gödel 1938) that if ZF is consistent, so is ZFC. However, we have no use of the Axiom of Choice in what are interested in here and hence we shall skip it.

There are several other worthy formulations of set theory that have been studied for philosophical or mathematical reasons (see (Fraenkel et al. 1973) for more on this matter), but ZF(C) is the mainstream one and the one to which all the results cited in the paper refer to. It is part of the standard toolbox of a mathematical logician to know that nothing essential would change in the philosophical conclusion of this material by either changing the axioms or even by changing the foundations from set theory to some other ones. This is because a trace of Gödel's Incompleteness (Gödel 1931) would catch up with us one way or another (see, e.g., Fraenkel et al. 1973).

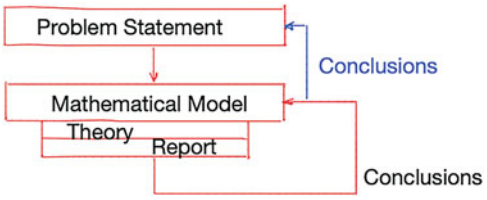
The reader not familiar with mathematical logic may simply think of ZF(C) as a significant representative of the method that mathematicians use to obtain mathematical foundations. The reader not familiar with the concept of mathematical foundations may simplify further and remember that any theory in human experience is based on some foundational truths, some unquestionable statements that are a base for all reasoning within the theory. Mathematical foundations are what mathematicians would like to declare their foundational truths. The question of what exactly mathematical foundations are is not a simple one, and no simple answer is possible; taking set theory as foundations is a very good and very common partial answer.

[2] In this text the noun *human* is used for members of any gender identity and the pronoun *he* is used as a generic human pronoun as well as the masculine pronoun.

3 On the Methodology of Mathematical Modeling

As stated, we wish to develop a model of the human society based on our knowledge of mathematical logic and then use the conclusions about the model to make some conclusions about the society. In other words, we develop a mathematical model. At this point, it is worthwhile to discuss what this means exactly, addressing both the advantages and the limitations of this process.

The steps involved in mathematical modeling are rather standard, let us, for example, use the description from Neumaier (2004). We start with a Problem Statement and provide a mathematical model, including elements of Theory and their report with the problem. We draw conclusions about the mathematical model and hope that the fit with the problem is sufficient to be able to carry the conclusions back to the Problem Statement. Then, we make conclusions about the Problem Statement.



While some parts of this diagram are purely logical, we point out that the fit of the model with the problem is a leap of faith. This in some cases can be defended by heuristic information and experiments, for example, in biology, and in others, as is the case here, by a philosophical position. In all honesty, the art of mathematical modeling is a bit of a mixture of experiments, and the philosophy and believing a model is never more than that: believing a model.[3]

3.1 Set Theory as a Modeling Tool

There are many examples where mathematical logic has made considerable advances in modeling social situations. Perhaps, the most striking recent example is Jouko Väänänen's Dependence Logic (Väänänen 2007). This logic, although of great independent interest in pure abstract logic, for its original treatment of dependence, independence, and teams, has also found numerous applications including some in social sciences, such as an analysis of Arrow's theorem (Arrow 1950) (of which more will be said in Sect. 6) done by Eric Pacuit and Fan Yang (2016). Game theory, which is in close connection with logic through game-theoretic

[3] Those who claim otherwise, waging equations and statistics in their own defense, need only to consider the remarkable lack of success that such a dogmatic approach has had in various concrete contemporary situations. Modelling is hard and nobody has a monopoly.

definitions of semantics, was in part developed as a theory of social organization in the tumult of the Second World War, as is documented in the book (Leonard 2010) by the well-known historian of economic thought, Robert Leonard. In fact, much earlier in mathematics, we can mention Condorcet's Jury Theorem and Condorcet's Paradox by Nicolas de Condorcet in (de Condorcet 1785), as one of the first attempts to model democratic processes by mathematics.

We shall stay within the area of mathematical logic, specifically set theory. Set theory has proved its worth as a modeler in many mathematical contexts, making contributions to a diversity of fields, within logic, such as model theory, and in a broader way in mathematics, such as in topology, measure theory, analysis, algebra, and combinatorics, and further in computer sciences. There are whole fields of mathematics whose name starts with the words "set-theoretic," such as the set-theoretic topology, and there are innumerable books on various aspects of mathematics whose title is "Set-Theoretic Methods in . . .," for example (Blanchini and Miani 2015), in control theory, (Michał Skrzypczak. Descriptive set theoretic methods in automata theory, volume 9802 of Lecture Notes in Computer Science. Springer-Verlag, Berlin 2016) in the theory of automata, and (Eklof and Mekler 2002) in algebra, to cite just a few. Moreover, set theory has played a major foundational role in mathematics, in that it is understood by many mathematicians as *the* foundations of mathematics and by all as *a* possible foundation of mathematics. This success speaks to the modeling power of set theory. Simply, a great variety of mathematical concepts are sets, with some additional properties. This is indeed what made the reputation of set theory in the foundations of mathematics and incited David Hilbert to say, perhaps somewhat too optimistically, that "nobody will take us out of the paradise created by Cantor" (Hilbert 1926).

We shall test the modeling power of set theory by going one step further, to social sciences. Set theory has famously entered into humanities through the work of Alain Badiou, including the book (Badiou 1988) which gives a mathematical ontology through set theory and uses the axioms of ZFC to identify the relationship of a being to the Nature, to the God, to the State. As the name of the book suggests, Badiou discusses the greatest philosophical questions using set theory, which he continues to do in further volumes of work (Badiou 2006, 2019). Of course, the four pillars of Badiou's work being Art, Love, Politics, and Science, he certainly uses his set-theoretical ontology to a great variety of social ideas. Our ambition is more modest and rather concrete: it will evolve around the idea of membership as a founding concept in set theory, but also in the society. Therefore, if we accept to define the society using the idea of membership as a defining feature, as we argue in the following, then it is reasonable to model that same society as a model of some sort of set theory. As mentioned above, we shall work with the axioms of ZF. We shall show, firstly, that the axioms of ZF(C) set theory have a natural interpretation in a model of the society and that, hence, the society can be modeled using the ideas we have about the universes of set theory. We shall call the resulting model of the society *the societal interpretation*.

4 Formulating the Societal Interpretation

The basic object in set theory is, evidently, a set. It is not a concept that we define, rather we take it as given and we build set theory by studying the relation of membership between sets. This relation is guided by the axioms of ZF(C) and deductions from the axioms, which are done using the rules of deduction in first-order logic. The resulting theory is a rich domain of mathematics, and it is made even richer by the fact that many objects in mathematics can be modeled using sets and the membership relation: the set of natural numbers, the set of real numbers, the set of all lines in a plane, and so on. As much as many mathematical concepts are sets, even more *social* concepts are sets: the set of members of a family, the set of students in a class, the set of members of a political party, followers of a religion, or the set of citizens of a country. Indeed, when introducing the concept of a set to a new audience, one would usually first go through this type of example, before even entering into the examples coming from mathematics. Furthermore, the basic relation of set theory is that of belonging to a set and we support the idea that this is also a basic relation in the society. While we might find that the idea proposed by Robert Park (1921) that "the individual is a product of group life," that is, the identity of a member of a society is determined by which sets he belongs to, is somewhat extreme, it would be difficult to disagree with the converse, that is *the identity of any society is determined by its members*.

This is exactly one of the basic axioms of set theory, the *Axiom of Extensionality*. Next, in Sect. 4.1 we explain how the rest of the axioms of ZF can be interpreted in the society. The reader not familiar with the axioms of ZF can find them listed in the Appendix. We shall not attempt to duplicate the work already done by Badiou (Badiou 1988) in interpreting ZF(C) set-theoretic axioms as ontology. Rather, we shall give a self-contained and very concrete interpretation of the axioms in the context that interests us.

4.1 ZF in the Societal Interpretation

To interpret a list of axioms in a logical context, we need to understand the domain where the variables appearing in the axioms are allowed to range. This domain for us will be the society. Therefore, the societal interpretation will have one society, the society of all humans on the planet and many societies, which represent smaller groups of humans. Humans individually, in families, in their work place, in their street, in their city, in their political party, on their Facebook page. All of these examples of societies will be represented as sets.

In the introduction to this section, we have already discussed the Axiom of Extensionality in the societal interpretation. Several following axioms in the list of ZF have a very natural interpretation in the model, as we now discuss.

The *Axiom of Pairing* says that for every two sets, there is a set which contains exactly the two of them. This would simply mean that for any two societies, one can form a third which consists of exactly the two we started with, which conforms with

observation. The *Union Set Axiom* says that if we have a set Z of sets, there is one set which contains exactly the elements that belong to some element of Z. Therefore, if we have a society of societies, say the Russell group of universities in the UK, each university consisting of students in that university, then we can form the union of all students of all Russell group universities. The *Power Set Axiom* says that if we have a set A, then we can form a set which contains exactly the subsets of A. For example, we can form a set which consists of all possible social subgroups of the students of all Russell group universities. The *Foundation Axiom*, which is given a very prominent role in (Badiou 1988), has a rather natural interpretation, saying that in any society, one can go back to the root to break the society into its constituent members which have no societies as members.

4.2 The Axiom of Infinity and Relativization

Now we arrive to the Axiom of Infinity, which might pose a problem. Namely, if we put ourselves at the external point of view, clearly the number of all inhabitants of the planet is finite, and so the number of all societies is finite. The *Axiom of Infinity* states that there is an infinite set. Do we have a contradiction?

No, we do not have a contradiction, and this is the main point. To understand this, we need to introduce the idea of the relativization, an idea which plays a major role in set theory since the work of Gödel on the constructible universe, in (Gödel 1938), but which, stated in different terms has already played a role in the work of Bolyai (1832) and Lobatchewsky (1837), on the models of geometry. Let us explain that idea abstractly.

Suppose that we have two sets M_0 and M_1 in which we have interpreted some mathematical language $\mathcal{L}$. An important observation to make is that the sentences of $\mathcal{L}$ might have different truth values in M_0 and M_1. For concreteness, let us consider as M_0 the set of $\mathbb{N}$ of natural numbers, as M_1 the set of $\mathbb{R}$ of the real numbers and, for the language in question, the language $\mathcal{L} = \{\leq\}$, which we interpret as the usual orderings on the sets $\mathbb{N}$ and $\mathbb{R}$. Then to every abstractly formed sentence in $\mathcal{L}$, say

$$(\exists x)(\forall y)\, x \leq y$$

we may associate the truth value that this sentence has in the models M_0 and M_1, and this truth value will not only depend on the sentence but also on the model in which the sentence is interpreted. For example, if we interpret the above formula in $(\mathbb{N}, \leq)$, then it is true, and if we interpret it in $(\mathbb{R}, \leq)$, then it is false. This is so because $(\mathbb{N}, \leq)$ has the least element and $(\mathbb{R}, \leq)$ does not. To formalize this idea, we shall introduce some notation and a formal definition, following the presentation in our book (Džamonja 2020), where the reader may find more details on the topic.

Let us recall how formulas of first-order logic are constructed, concentrating simply on the language $\mathcal{L} = \{\in\}$ of set theory. The basic, *atomic*, formulas, are of the form "$x = y$" or '$x \in y$'. The short definition of the set of first-order formulas in this language is that this is the collection of well-formed formulas that we obtain

from the atomic formulas by applying negation $\neg$, the conjunction $\wedge$, and the existential quantifier $\exists$ any finite number of times and in any order desired.[4] Well-formed formulas are those strings of symbols that obey the grammar rules for the formation of formulas, for example, $(\exists x)\varphi(x)$ is well formed, but $(\exists x)$ is not. Sentences are those formulas in which all variables are bound by some quantifier. For example, $(\exists x)(\forall y)\ x \leq y$ is a sentence, but $(\exists x)\ x \leq y$ is not.

We need the following notation:

Notation 4.1 *We write* $(\exists x \in M)\varphi(x)$ *as a shorthand for* $(\exists x)(\varphi(x) \wedge x \in M)$, *and similarly, we use the shorthand* $(\forall x \in M)\varphi(x)$.

Definition 4.2 *Suppose that* $(M, \in)$ *is a model of the language* $\mathcal{L}$ *of set theory and* φ *a formula of* $\mathcal{L}$. *We define the* relativization φ^M *of* φ *to* M *by induction on the complexity of* φ:

- *If* φ *is* $x = y$ *or* $x \in y$, *then* $\varphi^M = \varphi$.
- $(\varphi \wedge \psi)^M = (\varphi^M \wedge \psi^M)$, *and similarly for* $\vee$ *and* $\neg$.
- *If* φ *is* $(\exists x)\psi(x)$, *then* φ^M *is* $(\exists x \in M)\psi^M(x)$, *and similarly for* $\forall$.

For example, the relativization of the formula $(\exists x)(x^2 = -1)$ to $\mathbb{C}$ is $(\exists x \in \mathbb{C})(x^2 = -1)$, which is true, while its relativization to $\mathbb{R}$ is $(\exists x \in \mathbb{R})(x^2 = -1)$, which is not true. In general, note that the definition of $\models$ is set up so that $M \models \varphi$ iff φ^M is true.

The example above shows that there can be a large difference between a formula and its relativization. If we imagine two models of the same theory, say $M \subseteq N$ such that M is a part of N, then we can consider the formulas relativized to M and N as two different points of view on the truth in M. For a given sentence φ, we may think of the *internal* point of view of M as being represented by φ^M and of the *external* point of view, known to N and represented by φ^N. The two statements φ^M and φ^N come from the same formula, but are not the same, and they do not necessarily have the same truth value.

Let us go back to the point that we started this discussion, the Axiom of Infinity. The formalization of this axiom in ZF (see the Appendix) basically posits the existence of a non-empty set that is closed under the successor operation, defined by $S(a) = a \cup \{ a \}$. This operation is not particularly interesting in the society, and it would be rather futile to take this as a definition of what infinite means in everyday life. That is, we are not interested in that external view of an infinite society. Rather, we are interested to model the society as it appears to an internal observer, an inhabitant of the planet. In this we take a line of thought that is presented in applied mathematics, where "infinite" stands for a finite object of such large finite

[4] Note that the other connectives $\vee$, $\Rightarrow$, and $\Leftrightarrow$, as well as the quantifier 8, can be derived from the ones given above.

size that for all practical purposes this size may be infinite. For example, in modeling waves in deep water, Jean-Marc Vanden-Broeck (2014) uses the expression "infinite depth" to describe that the depth being very large, the effect of the bottom of the sea can be neglected. The well-known set-theorist Đuro Kurepa, in one of his popular presentations of his subject, used to describe his favorite object of study: a Souslin tree (an object of size $\aleph_1$), as a tree of life of a microorganism which starts dividing and at regular intervals has to face a choice of which units will be allowed to continue (the limit levels of the tree), raising the question of any unit surviving to the end (Kurepa 1987). From the point of view of the microorganism, going to the next limit level might very likely represent a life span, or even a life span of a whole generation. This microorganism cannot imagine what comes after, all the relevant information is already encapsulated in the several steps ahead. We may interpret "infinity" as being closed under a very large number of repetitions of a given operation, like measuring 1 mm of the depth of water in (Vanden-Broeck 2014) or a division of a microorganism in Kurepa's example.

Let us then try to interpret infinity in the societal interpretation. We shall do this by the notion of *distance to the influence*. To start, observe that there are some small societies, such as the nuclear family of an individual, or him and his two best friends. In both of these societies, this individual has a very direct input and engagement. Then, there are larger and larger societies he might belong to, and we might observe that the direct influence of him as an individual decreases. We could in fact model the relation of an individual a to the society or to a society, by considering the membership graph generated by the individual: a belongs to a society A, and it directly influences the decisions in A. However, A belongs to the society $\mathcal{A}$, where a also has an influence, but only an indirect one, passing by A. The membership distance of a to A is 1, while the membership distance of a to $\mathcal{A}$ is 2 and so on. We may notice that the influence that the individual as a unit has in a society decreases with the membership distance. Moreover, although from the external point of view there is no truly infinite membership distance on the planet and from the internal point of view there are membership distances which are so large that for all practical purposes, they are infinite. The distance of a typical Greek citizen λ to the Eurogroup, which in 2015 is deciding if Greece will be forced to exit the Euro or will have to increase the austerity measures which would perhaps cost the citizen his livelihood, is infinite for all practical purposes.[5]

We posit that large societies and legal entities such as countries but even smaller entities such as large cities or regions are best modeled as infinite sets. In fact, we claim that this is the point of view taken in everyday practice. For example, in what is the most popular understanding of democracy at the moment, it consists of asking a large number of citizens their opinion on a multiple choice question (which usually

[5] Of course, this example is somewhat subjective, which is why we emphasise that like in any model where large finite numbers are treated as infinite, there is an element of fine-tuning the model to the purpose. Usually, the rule followed in applied mathematics is that a quantity that is being treated as infinite is indeed infinite in some reasonably taken limit. In our case that limit would be the limit taken over the growing number of inhabitants of the planet.

only has two choices) in elections and then deciding by simple majority. Obviously, this idea is much more fruitful in elections in which there is a large number of participants-elections in which only a small number of citizens vote are usually cancelled. This point that was expressed mathematically by Condorcet in his Jury Theorem (de Condorcet 1785) and it basically says that the probability that a jury will make a decision that best represents the majority increases as the size of the jury goes to infinity. Believing in democratic elections by simple majority in large societies is basically confirming that everyday practice treats these societies as infinite, as we shall do in the societal interpretation.

4.3 The Remaining Axioms

It remains to interpret the Axiom Schemas of Comprehension and of Replacement.

Comprehension states that if we have a set y and a property φ of its elements, then the collection of all members of y that satisfy φ is also a set. In the society, this would correspond to forming subsocieties that is isolating those members of a given society that have a certain property. This, of course, conforms with the everyday practice, where we might want to address the set of all citizens of Paris that live on the left side of the river Seine, or the set of all of Mr. X's children that are older than 13. Comprehension is actually a scheme, since it gives a separate axiom for every formula φ.

Similarly, Replacement is a scheme. It says that if we have a formula φ which acts as a function, in the sense that to every argument it assigns exactly one value, then we can collect all these values in one set. In everyday life, this would correspond to the practice of choosing something from every member of a society. When every adult Estonian creates an income tax return declaring the flat income tax of 21% of their income, the resulting set of all the returns collected is guaranteed to exist by the Axiom of Replacement. So is the set of all exact moments of birth of all inhabitants of the Earth, or the set of all first books read of the citizens who have ever read a book.

5 Plurality of Models of Set Theory

A careful reader might have already noticed a curios property of the societal interpretation. It is not unique. There are many societies! This is a phenomenon well studied in set theory. There are many models. In this part we shall describe how set theorists view the models or universes of set theory, with the plan to use the knowledge of these universes in Sect. 6 to make conclusions about the societal interpretation.

The present state of knowledge in set theory has fully acknowledged that there are some simply stated statements about sets which are not decided by ZF or ZFC. A canonical such statement is the Continuum Hypothesis (CH), which posits in ZFC that every infinite subset of the real line is either bijective with the set of natural numbers or with the set of real numbers, the two being known not to admit a

bijection. If there is a model of ZF at all, then there is one in which ZFC+CH holds (Gödel 1938) and there is another one in which ZFC+¬ CH holds (Paul Cohen 1963). We say that CH is *independent* of the axioms of set theory. In fact, there are many other such independent statements, and even if we were to axiomatize a truth value to CH, so accepting that it is true or false, there still would be independent statements, concrete examples of which are well known. One such statement is the Souslin Hypothesis which is independent of ZFC+CH and of ZFC+¬CH (see books (Kunen 1980; Shelah 1998), or (Džamonja 2020) for expositions).

The state of affairs with CH is the tip of the iceberg. Modern set theory abounds with ways to produce new models of set theory from the existing ones by using forcing or large cardinals or both. This leads to rather opposing views of the structure of the part of mathematics described by the axioms of set theory. These are, roughly, the universe view and the multiverse view. The societal interpretation has elements of both and in fact is the best described by an elegant combination of them in Jouko Väänänen's view from (Väänänen 2014). Let us describe the universe and the multiverse views.

The universe view holds that there exists the true Universe of set theory, axiomatized categorically by some axioms, leading to a unique model. These axioms, evidently, are still to be found, given the situation with CH and other independent statements. Gödel's Incompleteness theorems (Gödel 1931) tell us that without a radical solution such as changing the logic, there is no way out of this conundrum. There is a solution if we change the logic to the second order logic, which does not lend itself to the arguments used in the Incompleteness Theorems. Moreover, second-order ZFC set theory enjoys what is known as quasi-categoricity. Namely, Ernst Zermelo (1930) proved that the second-order version of ZFC axioms, known as ZFC^2, has one single type of models, namely the sets V_κ for a strongly inaccessible cardinal κ. There are unfortunately many disadvantages of second-order set logic for the practical purposes in mathematics, since exactly by being abundantly expressive, it fails to satisfy any properties that make mathematical logic useful in practice. For example, the completeness or compactness theorems or Downward-Skolem theorems fail for second-order logic. In fact, a theorem of Lindström (1969) characterizes the first-order logic as the maximal satisfying both the compactness theorem and the downward to $\aleph_0$ Löwenheim-Skolem theorem. So, if we want to have a useful, rather than merely expressive, set theory out of ZF(C), we better stuck with the classical first-order version, and then the universe view needs to be defended or defeated.

This leads to the multiverse view, which defends the idea there is no canonical Universe of set theory; rather, there are many models, and it is their plurality that gives the real information on the subject. Many authors defend this view, which is well known in set theory, but got known to a wider philosophical audience through the much-cited paper (Hamkins 2012) by Joel Hamkins. This beautifully written paper centered on forcing extensions is sometimes over-simplified in applications, which do not always account very carefully for the seeming contradiction that in order to take forcing extensions of a model in the multiverse, this model already needs to be in a larger universe. To understand this, one has to distinguish the

internal and the external views. This important point is underlined in Väänänen's view from (Väänänen 2014), where the *multiverse* consists of a multitude of universes. We are going to take this view, except that for us a model of set theory will not be called a universe but simply a model, and the multiverse of Hamkins and Väänänen will be called Universe. This terminology aligns us with Shelah's view from (Shelah 2003), where he explains a fine line that differentiates a pure Platonist from the one that believes in a universe but also in undecidable problems.

6 Universe and Models in the Societal Interpretation

For the application that we are interested in here, the societal interpretation, there is an obvious Universe, namely the society as a whole. However, like in any other context where set theory appears, there are also many smaller models of set theory, sometimes in contradiction with each other, sometimes defending the opposite axioms, but it is their multitude that gives the real information of the Universe. By applying the same logic that we have used in constructing the societal interpretation, we may interpret as models of set theory the large societal units, concretely the states-countries. These countries obey the basic laws of the society, but they differ greatly in their interpretation of further laws. Some are communist, some are liberalist, some are religious, some are western-democratic, and so on. Each defines its own model. We shall interpret the set-theoretic knowledge about the societal interpretation to the society to defend the idea that it is the multitude of societies that describes the truths in the society and that is only through their multitude that the society can advance.

The strength of set theory is to recognize the fallacy of modelling an absolute truth. A set theorist can look at the Universe and see two models, M and N such that a model M sees certain statements as true, while N sees them as false. However, both M and N being models of set theory, they agree on the basics and can live in collaboration in the Universe of models. By no means, the vision of M will negate the vision of N or vice versa; there is simply a plurality of visions. This is the main point we wish to retain regarding the societal interpretation. There are many societies in the society, and various societies may have different opinions on some statements, but the plurality of their opinions is what determines the society.

6.1 The Silent Majority

The view that it is only the plurality of voices that determines the society is in contradiction with the reality of the lives of most inhabitants of the planet, whose opinion is not heard. They pass under the various qualifications, from the "silent majority" to some more negative descriptions. The fault for this silence is often put on the specific society that they inhabit, the lack of something or other in their country, and the lack of freedom or education, for example. While these are undeniably aspects that the society suffers from, we would like to point out another

aspect that passes like a taboo under the radars of most but the most outspoken critics.

What we claim is that the silent majority exists because there is a particular idea that has permeated the political life of this planet and which silences and marginalizes everybody who is not adhering to it. Under the umbrella of making the political choices more just, we have started believing in the existence of a universal system which will make everybody's voice heard. However, we have not defined what this would mean exactly. Those who think that as soon as everybody has the right to vote we will have reached the ideal society, we have still to define what is actually understood by this ideal society and then defend its existence. However, in today's society, a particular idea, these days called Democracy, pretends to being an absolute truth. Everyday experience shows that the latter is clearly not the case. What we claim here is that there is a mathematical proof that it cannot be the case. This Universal Democracy, as pushed on us as it may be by certain big powers, is not the absolute truth simply because the absolute truth does not exist.

Even the simple belief that such an absolute truth exists has stopped the society from accepting the plurality of its constituents and advancing towards a more fruitful international cooperation.

6.2 Universal Democracy Is Undefined

Universal Democracy, which has been presented as the only true way, has not ever been correctly defined. In fact, universal democracy has a lot to do with the Continuum Hypothesis. We do not seem anywhere close to having a decision on the value of the Continuum Hypothesis, and we are not anywhere close to having a good definition of what the word "democracy" might actually mean and yet less close to what "universal" might mean, apart from the simplistic and unsatisfactory answer of everybody having the right to vote. The question of how the votes are distributed, influenced, and used, and moreover, what it is that we are voting for, is very variable.

In fact, there is a theorem which states that the question is much more difficult than just giving the votes to everybody, namely a theorem that illustrates the impossibility of having an ideal voting structure.

Theorem 6.1: Arrow's Impossibility Theorem (Arrow 1950) *In all cases where preferences are ranked, it is impossible to formulate a social ordering without violating one of the following conditions:*

- *Nondictatorship: The wishes of multiple voters should be taken into consideration.*
- *Pareto Efficiency: Unanimous individual preferences must be respected: If every voter prefers candidate A over candidate B, candidate A should win.*
- *Independence of Irrelevant Alternatives: If a choice is removed, then the others' order should not change: If candidate A ranks ahead of candidate B, candidate A*

should still be ahead of candidate B, even if a third candidate, candidate C, is removed from participation.
- *Unrestricted Domain: Voting must account for all individual preferences.*
- *Social Ordering: Each individual should be able to order the choices in any way and indicate ties.*

Arrow's Impossibility Theorem forms a basis of social choice theory. There are many theorems which indicate that simple voting systems are impossible or at least very difficult to implement; for an early example, one can see Condorcet's Theorem, (de Condorcet 1785). The experience including the two world wars and the Cold War and many other wars since has taught us that universal democracy is an ideal which might not be easy to reach. Some think that universal democracy is to have the rule of the market, others think that it is to have the pure rule of the majority, yet others find themselves embracing various compromises between the visions of democracy. Democracy seems to live on a spectrum. At the moment of writing, there are at least two big powers in the world who are located on the two opposite parts of that spectrum, both entirely convinced in their own model, and many others who remain unconvinced by any of the two. Therefore, due to a still rather vague and discorded understanding of what the word "democracy" actually means, searching for a universal meaning of this word might not even be an aim that one would want to reach, or at least not an aim that seems to be anywhere in the remit of the society of today.

Like CH, the Universal Democracy is a vision which does not admit discord. Like CH, for the moment we do not know how to obtain it and moreover we have proofs that we cannot do it. So, in our opinion, the universal democracy is an Absolutely Undecidable Proposition, in the language of (Väänänen 2014).

6.3 Back to Kant

The proponents of the Universal Democracy find that their vision is so clear that it does not even have to be defined. In their minds, it goes without saying that every great philosopher has started with the Universal Democracy as given. For example, Kant is reputed to have claimed that perpetual peace could be secured by universal democracy and international cooperation in (Kant 1795). However, closer inspection of (Kant 1795) shows that there is no discussion of universal democracy, not even of what today is a given, a vote for everybody. Instead, there is a lot of discussion of international cooperation and very concrete suggestions of how such a collaboration could be achieved. Our view is completely in line with Kant, in believing that international cooperation is a definable and achievable goal.

As for the democracy, we still can have a measure of understanding it, by observing how it behaves in various models. In our paper with Marco Panza, (Džamonja and Panza 2018), we have argued that concentrating on resolving or not resolving CH and hitting their head on the wall has made set theorists blind to what they actually can resolve. It has made them, for a long time, neglect the parts of the cardinal hierarchy where their methods are very powerful. In (Džamonja and Panza 2018), we have discussed the situation of singular cardinals as an example of where ZFC has much to

say, in spite of it not having anything to say on CH. It is the same with the universal democracy. Once we accept that for the moment we do not really have the means to give a just definition of this concept, then, rather than burning the planet in yet another war of ideological supremacy to show that one or another possible definition of the concept is correct, we should concentrate on what we *can* do.

International cooperation has received much less attention than the universal democracy, and when it did, it has often been overwhelmed by the latter. Many wars repeatedly came to the point of destroying the planet, run by ideologies. Bringing "democracy" or "liberating" was always somewhere in the background, and it almost always went wrong. This is because even when genuine, those messengers of the universal democracy have acted in a biased way that has been based on their belief that they have all the answers. It is painfully clear that this is not a way to proceed. Moreover, we feel that the moment is propitious to point out that no society is more worth than another, as the crisis we live in has brought up nationalist and selfish behavior on the parts of many governments who feel that the "natural supremacy" of their system gives them the right to make decisions and choices that badly influence livelihoods and even lives of those living in other systems. Yet, if we unite in an international effort to resolve the crisis, we might succeed.

To conclude, we suggest that the ideal of international cooperation is possible even without the fuzzy notion of universal democracy. International cooperation is exactly the lesson we can learn from believing that the society is a multiverse. Nobody has all the answers.

7 Conclusion

As much as science and mathematics can explain certain societal questions, societal questions do influence science. This point of view complements what is done in this chapter and is material for another article. However, we can openly admit that our views are not divorced from the society in which we live. Mathematicians and logicians are engaged in the society even when they believe not to be. Their research is in phase with that society. This is an honor and a responsibility, and it does not apply only to the ones who do an obviously applied research. We have advanced one way in which research in mathematical logic can be engaged in this process. To put it in context, many mathematical competences have been used in modeling possible consequences of economic and political decisions. Sometimes they obtained surprising results which show that ill-defined moves may have consequences which are completely opposed to what the superficial definition might indicate. The inherent error in the theory leads to disastrous consequences (see, for example, the book (Varoufakis et al. 2009) by Yanis Varoufakis, Joseph Halevi and Nicholas J. Theorcarakis who, among other competences, use game theory, to make the above conclusion in the concrete example of the economic reflexivity). Game theory is a staple of economic reflection and the inherent study of the society. In this chapter, we suggest a complementary competence coming from set theory to make the same conclusion, which we study on the concrete examples of universal democracy and international cooperation.

In simple words, the conclusion of this chapter is short. Unify in respect of each other or perish while searching for dominance, which is nothing but an elusive and ill defined concept.

8 Appendix: Axioms of ZF

The following is the standard axiomatization of the ZF system, as found, for example, in our book (Džamonja 2020). We express our axioms in first-order logic, where the variables are taken to range over sets. It is common to assume that the domain of discourse is non-empty, so we do not need Axiom 0 from (Kunen 1980). We shall use the following notation.

Notation 8.1

(1) *For any choice of* $y, x_1, \ldots, x_n$ *the notation*

$$(\exists! x)\ \varphi(x, y, x_1, \ldots, x_n)$$

is a replacement for

$$(\exists x)(\varphi(x, y, x_1, \ldots, x_n) \wedge (\forall z)(\varphi(z, y, x_1, \ldots, x_n) \Rightarrow z = x)).$$

(2) *When writing* $\varphi(x, y, x_1, \ldots, x_n)$*, we mean that the free variables of* φ *are among* $\{x, y, x_1, \ldots, x_n\}$*, but do not necessarily include them all.*
(3) *For any set* x*, the operation* $S(x)$ *is defined as* $x \cup \{x\}$.[6]

Axiom of Extensionality Two sets have exactly the same elements if and only if they are the same set or

$$\forall x \forall y\ \ (\forall z(z \in x \Leftrightarrow z \in y) \Rightarrow x = y).$$

Axiom of Pairing For every two sets x, y there is a set whose elements are exactly x, y or

$$\forall x \forall y \exists z(x \in z \wedge y \in z \wedge \forall w(w \in z \Rightarrow w = x \vee w = y)).$$

Union Set Axiom

$$(\forall A)(\exists Z)(\forall b)(b \in Z \Leftrightarrow (\exists a \in A)(b \in a)).$$

[6] One needs the axioms of Union and Pairing to verify that for any set x, the resulting set $S(x)$ is a set.

Replacement Axiom Scheme For each formula $\varphi(x, y, x_1, \ldots, x_n)$ we have

$$(\forall x_1, \ldots, x_n, I)((\forall x \in I)(\exists! y\, \varphi(x, y, x_1, \ldots, x_n) \Rightarrow (\exists A)(\forall z)(z \in A \Leftrightarrow (\exists x \in I)\varphi(x, z, x_1, \ldots, x_n))).$$

Comprehension Axiom Scheme For each formula $\varphi(x, y, x_1, \ldots, x_n)$ we have

$$(\forall x_1, \ldots, x_n, y)(\exists z)(\forall x)(x \in z \Leftrightarrow x \in y \wedge \varphi(x, y, x_1, \ldots, xn)).$$

Axiom of Infinity There exists an infinite set, which we formalise as.[7]

$$(\exists x)(\emptyset \in x \wedge (\forall y \in x)\, S(y) \in x).$$

Power Set Axiom For every x there is y whose elements are exactly those z which are a subset of x or

$$(\forall x)(\exists y)(\forall z)\ (z \subseteq x \Leftrightarrow z \in y).$$

Axiom of Foundation Any non-empty set is disjoint from at least one of its members or

$$\forall x(\neg(x = \emptyset) \Leftrightarrow \exists y(y \in x \wedge y \cap x = \emptyset)).$$

References

Arrow KJ (1950) A difficulty in the concept of social welfare. J Polit Econ 58(4):328–346
Badiou A (1988) L'Être et l'événement. Seuil
Badiou A (2006) Logiques des mondes. L'Etre et l'Evénement, 2. Seuil
Badiou A (2019) L'immanence des vérités. L'Être et l'evenement 3. Fayard
Blanchini F, Miani S (2015) Set-theoretic methods in control. Systems & control: foundations & applications, 2nd edn. Birkhäuser/Springer, Cham
Bolyai J(J) (1832) Appendix. In: Tentamen, Bolyai, Farkas (Wolfgang) (ed) Conspectus Arithmeticæ Generalis. Collegeium Reformare, vol 1. Maros Vásárhelyini (Târgu Mures)
Cohen P (1963) The independence of the continuum hypothesis. Proc Natl Acad Sci U S A 50(6): 1143–1148
de Condorcet M(N) (1785) Essai sur l'application de l'analyse á la probabilité des dècisions rendues á la pluralité des voix. Imprimerie Royale, Paris
Dirac PAM (1931) Quantised singularities in the electromagnetic field. Proc Royal Soc Lond Ser A 133:60–72

[7] The existence of the unique empty set $\emptyset$ used in the formulation of the Axiom is justifiable by the axioms above.

Džamonja M (2020) Fast track to forcing. Cambridge University Press
Džamonja M, Panza M (2018) Asymptotic quasi-completeness and ZFC. In: Carnielli W, Malinowski J (eds) Contradictions, from consistency to inconsistency, volume 47 of trends in logic. Springer, pp 159–182. To appear in the volume "Between Consistency and Inconsistency" of trends in logic (eds. W. Carnielli and J. Malinowski)
Eklof PC, Mekler AH (2002) Almost free modules, volume 65 of North-Holland mathematical library, revised edition. North-Holland Publishing Co., Amsterdam. (Set-theoretic methods)
Fraenkel AA, Bar-Hillel Y, Lévy A (1973) Foundations of set theory (2nd revised edition). North-Holland, Amsterdam
Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. Monatsh Math Phys 38:173–198
Gödel K (1938) The consistency of the axiom of choice and of the generalized continuum hypothesis. Proc Nat Acad Sci 24:556–557
Hamkins JD (2012) The set-theoretic multiverse. Rev Symb Logic 5:416–449
Hilbert D (1926) Über das unendliche. Math Ann 95:161–190
Kagan V (1974) Lobatchevski. Mir, Moscow
Kant I (1795) Zum ewigen Frieden. Ein philosophischer Entwurf. Friedrich Nicolovious, Königsberg
Kunen K (1980) Set theory, volume 102 of Studies in logic and the foundations of mathematics. North-Holland
Kurepa Đ (1987) Ramification methods in set theory. In: A presentation for the "6. Konferencija Algebra i Logika". Sarajevo
Lambert J-H (1759) La perspective affranchie de l'embaras du Plan géometral. Heidegguer et comp., Zurich
Leonard R (2010) Von Neumann, Morgenstern and the creation of game theory. Cambridge University Press, Cambridge
Lindström P (1969) On extensions of elementary logic. Theoria 35(1):1–11
Lobatschewsky N (1837) Géométrie imaginaire. J Reine Angew Math 17:295–320
Michał Skrzypczak. Descriptive set theoretic methods in automata theory, volume 9802 of Lecture Notes in Computer Science. Springer-Verlag, Berlin (2016) Decidability and topological complexity. University of Warsaw, Warsaw, Dissertation
Neumaier A (2004) Modelling languages in mathematical optimization. In: Kallrath J (ed) Applied optimization, chapter mathematical model building, vol 88. Springer, Boston, pp 37–43
Pacuit E, Yang F (2016) Dependence and independence in social choice: Arrow's theorem. In: Väänänen J, Abramsky S, Kontinen J, Vollmer H (eds) Dependence logic. Birkhäuser, Cham, pp 235–260
Park RE (1921) Sociology and the social sciences, the group concept and social research. Am J Sociol 27(2):169–183
Saccheri GG (1733) *Euclides ab omni naevo vindicatus*. Pauli Antonii Montani, Mediolani (Milan)
Shelah S (1998) Proper and improper forcing, 2nd edn. Perspectives in Mathematical Logic. Springer-Verlag, Berlin
Shelah S (2003) Logical dreams. Bull Amer Math Soc (NS) 40(2):203–228
Väänänen J (2007) Dependence Logic. Cambridge University Press, London Math. Soc. Student Texts
Väänänen J (2014) Multiverse set theory and absolutely undecidable propositions. In: Kennedy J (ed) Interpreting Gödel. Cambridge University Press, Cambridge
Vanden-Broeck J-M (2014) On periodic and solitary pure gravity waves in water of infinite depth. J Eng Math 84:173–180
Varoufakis Y, Halevi J, Theocarakis NJ (2009) Modern political economics: making sense of the post-2008 world. Routledge
Zermelo E (1930) Über Grenzzahlen und Mengenbereiche. Fundam Math 16:29–47

Implicitly Defining Mathematical Terms

A Path Toward Pluralism

Demetra Christopoulou

Contents

1 Introduction 2384
2 Implicit Definitions and Availability of Meanings 2385
3 The Uniqueness Problem in Case of Mathematical Implicit Definitions 2388
4 Pluralism in Foundations 2390
5 Structuralism, Implicit Definitions, and Pluralism 2394
6 Conclusion 2397
References 2398

Abstract

This chapter expounds the view that mathematical implicit definitions, i.e., systems of axioms as well as abstraction principles, underpin some aspects of pluralism in mathematics. Firstly, it takes in account Horwich's "uniqueness problem" concerning implicit definitions and examines it particularly in the case of mathematical systems of axioms as well as abstraction principles. Meanings accrued to the *definienda* are not uniquely determined, and multiple systems of objects may satisfy the system of axioms or the abstraction principle in question. Multiplicity in ontology is one of the aspects of pluralism which characterize the ways in which we implicitly define mathematical terms. Systems of axioms as well as abstraction principles are taken to play a foundational role in building up mathematical theories. However, the chapter turns the focus of interest to a structural view of mathematical implicit definitions. A structural account of mathematical implicit definitions supports a view of modest pluralism.

D. Christopoulou (✉)
Department of Mathematics, National and Kapodistrian University of Athens, Athens, Greece
e-mail: christopdem@math.uoa.gr; chrdemet@yahoo.gr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_30

A structuralist is not interested in foundations. In this view, systems of axioms as well as abstraction principles implicitly define structures. Shapiro's *ante rem structuralism* involves implicit definitions in procedures of *disclosure* of certain mathematical structures. Structures have an ontological status independent of their exemplary cases. In the penultimate section, the chapter takes into account the case of *satisfiable* abstraction principles. Such principles are taken to implicitly define certain structures independently of the "virtues" or "defects" that characterize the principles themselves. In particular, satisfiable abstraction principles that preclude each other, e.g., *Hume's Principle* and *Nuisance Principle*, describe alternative structures with different exemplary systems of objects.

Keywords

Axiomatic system · Abstraction principle · Implicit definition · Pluralism · Foundation · Structuralism · Model

1 Introduction

This chapter presents the way that implicitly defining mathematical terms implies some connections with certain versions of pluralism in mathematics. Systems of axioms as well as abstraction principles are regarded as implicit definitions of mathematical concepts. In Sect. 2, the chapter describes how implicit definitions work and discusses Horwich's "existence" and "uniqueness problem." Then it takes under consideration those problems particularly in case of mathematical implicit definitions. In regard to the "uniqueness problem," meanings which are accrued to the *definienda (: the mathematical terms)* are not uniquely determined, so multiple systems of objects may satisfy the postulates (Sect. 3). After a short outline of some obvious instances of pluralism in foundations, the chapter describes the way some philosophers (M. Balaguer, Ø. Linnebø) endorse *pluriverse* accounts in order to rescue realistic positions by reconciliating them with pluralism (Sect. 4). In the sequel, the chapter turns the focus of interest to a structural view of mathematical implicit definitions (Sect. 5). Structuralism is not interested in foundations, but it appeals to model theory. Shapiro's account of *ante rem structuralism* takes implicit definitions to be involved in procedures of disclosure and description of certain structures. According to it, structures have an ontological status independent of their exemplary cases, since they are one level of generality up. The structures provide the place for certain examples. A basic presupposition for a mathematical implicit definition (either a system of axioms or an abstraction principle) to work successfully is satisfiability. Besides, truth is dependent on structure. In particular, the chapter takes in account satisfiable abstraction principles which preclude each other, and it maintains that they describe alternative structures. This account presupposes a notion of truth in a structure (or according to the structure). The chapter describes a connection between the structural account of mathematical implicit definitions and some modest version of mathematical pluralism.

2 Implicit Definitions and Availability of Meanings

An explicit definition regularly substitutes the *definiendum* (e.g., a scientific term) by some already known expressions. In contrast, an implicit definition does not work in that way. It settles certain conditions under which scientific terms can acquire meanings. If L is a language and "t" a new term, then we take the extension of the initial language L+{t}. An implicit definition is a sentence (or a set of sentences) in the extension L+{t} that includes (include) the *definiendum* and accrues (accrue) meaning(s) to it. There may be more than one terms (*definienda*) which are assigned meanings on the basis of their mutual relations stated by a system of postulates. According to a commonly held view about implicit definitions, a sentence that contains the *definiendum* is *stipulated* as *true*. B. Hale and C. Wright (2001, p. 117) describe this (traditional) view as the following: "Implicit definition, taken as the complement of explicit, embraces a variety of subtypes. What all have in common is the idea that we may fix the meaning of an expression by imposing some form of *constraint* on the use of longer expressions -typically, whole sentences containing it. On one traditionally influential conception, this constraint is imposed by the free *stipulation of the truth* of a certain sentence, or range of sentences, embedding the *definiendum* and composed of otherwise previously understood vocabulary." On this account, we take a sentence (or a range of sentences) that contains (contain) the *definiendum* (*definienda*) and some already well understood expressions, and we stipulate that sentence (those sentences) *as true*. In similar words, "the meaning accrued to an expression *F* (or expressions *G, H* etc.) is such that the postulate(s) in which it (they) occur(s) is (are) true" (Psillos and Christopoulou 2009). Accordingly, the *definiendum* should have *the meaning it needs to have* for the sentence in question (the vehicle of the implicit definition) to express a truth. The "traditional" (as being characterized in the above passage) model of implicitly defining theoretical terms in science is discussed by philosophers like P. Horwich (1998) and the advocates of Neofregeanism Hale and Wright (2001). The view in question expects that appropriate meanings for the *definienda* should be *available* in order that the outcome of the stipulation in question is an a priori true proposition. A priori true propositions are independent of sensory experience. Rationalists standardly regard basic mathematical truths, e.g., axioms, as a priori. Implicit definitions are connected with the issue of non-inferential a priori knowledge. In the case of mathematics, discussion about implicit definitions includes mainly systems of axioms and Fregean abstraction principles as indicated below.

Horwich (1998, pp. 134–138) has exposed some serious difficulties the above account faces. When one assumes that the term "*f*" should have *the meaning it needs to have* for the stipulated sentence to express a true proposition, she presupposes that (a) *there is* a meaning such that if "*f*" has that meaning then the sentence which works as an implicit definition *is* true and (b) there is at most one such meaning. Those points made by Horwich have produced a discussion about the "existence problem" and the "uniqueness problem," respectively. Let us make use of the symbol "#*f*" to present an implicit definition of the expression "*f*." The sentence bearing the implicit definition says that *f* satisfies *such* and *such* conditions (: #). Hence, if an appropriate

meaning is available for "*f*," then the stipulation of the sentence "#*f*" gives a *true proposition*. However, there may be no such meaning for "*f*" or there may be more than one. Horwich argued that it is possible, for example, for a physical theory to be falsified (1998, pp. 136–138). In those cases, the sentences that bear implicit definitions are questioned. So Horwich preferred to say that a meaning is accrued to the *definiendum* "*f*" so that the sentence in question "#f" *is held to be true* or it is *accepted as true*. Furthermore, he defended an account according to which "a meaning is accrued to a term" means that a certain "pattern of use" is established for that term. In other words, Horwich maintained that a *use theoretic account of meaning* is more appropriate than an account requiring the preexistence and availability of meanings. A pattern of use of the theoretical term "*f*" is regularly established by the postulation of a sentence as true which contains this term.

Horwich's suggestions appear to concern implicit definitions of physical terms rather than implicit definitions of mathematical terms. His proposal aims to provide a solution for a difficulty that emerges in case of scientific terms like "phlogiston," "ether," etc. Such terms occurred in scientific theories which eventually were falsified and thereby abandoned. Those facts have raised certain questions concerning the truth of the postulate, i.e., the sentence "#*f*" which bears the implicit definition of a theoretical term like "f" = "phlogiston." This is because that term is denotationless as D. Lewis (1970, p. 429) remarks. In general, worries arise if we endorse certain semantics according to which if there is no object of reference of a term "f" then the sentence "f is *so* and *so*" cannot be true. Nevertheless, there are differences between implicit definitions of physical terms and implicit definitions of mathematical terms since mathematical propositions are not falsifiable for empirical reasons. Philosophical discussion concerning implicit definitions takes systems of axioms to be implicit definitions of mathematical terms (e.g., in case of arithmetic: the terms "*number*," "*zero*," "*successor*"). Meanings are obtained collectively on the basis of mutual relations they are taken to satisfy by means of a series of postulates. D. Hilbert elaborated the idea of treating systems of axioms as implicitly defining mathematical terms (Resnik 1974, pp. 390–391). An example of such implicit definition is the system of Euclidean axioms, let be "#(*f*, *g*, *h*)." That system is taken to define collectively the geometrical terms "*f*," "*g*," "*h*" (: "*point*," "*line*," "*plane*") by laying down certain mutual relations ("#") that points, lines, and planes should satisfy. Hilbert stated that his total system of axioms defined his primitive terms; however, his axioms did not uniquely determine any single concept. He thought that every theory is a schema of concepts together with their necessary relationships with each other (Resnik 1974, p. 390). So, the first question that arises is whether the *existence problem* pointed out by Horwich emerges also in the case of systems of mathematical axioms working as implicit definitions. The second question concerns Horwich's *uniqueness problem*.

Perhaps, the case of definitions of mathematical terms is clearer and more tractable than the case of physical terms. In mathematics, a "phlogiston case" cannot appear, and the stipulations in question do not bear empirically falsifiable propositions. Mathematical implicit definitions are expected to work successfully as the vehicles of true propositions in which the *definienda* appear. Horwich's *existence*

problem can be dealt with in the mathematical case if only certain conditions are met. A condition that should be required is *satisfiability*. If a system of axioms is satisfiable, then the system has models, and certain meanings can be accrued to the mathematical terms under definition.

There is a second kind of implicit definitions in mathematics, i.e., abstraction principles. An abstraction principle coming from the philosophical work of G. Frege (1884/1953) is formulated as a universally quantified biconditional which includes an identity statement on the left-hand side and an equivalence relation on the right-hand side. The general form is as the following:

$$(\forall a)(\forall b)\ [(\Sigma(a) = \Sigma(b)) \leftrightarrow (a \approx b)],$$

($\approx$: an equivalence relation between a, b, Σ : an operator that forms new singular terms, "Σ(a)": a singular term that is purported to refer to an object).

A famous example of Frege, presented in *Grundlagen* §63, §69, is a second-order quantified abstraction principle called "Hume's principle" which introduces the concept of natural number. The name is due to a comment of Frege (in *Grundlagen* §63) that his inspiration concerning 1-1 correspondence had come from D. Hume.

$$(\mathrm{HP})\ (\forall F)(\forall G)\ [(Nx : Fx = Nx : Gx) \leftrightarrow (F\ 1-1\ G)]$$

(HP) says that for every two concepts F and G, the number of Fs is identical to the number of Gs if and only if there is a 1-1 correspondence between the instances of F and the instances of G.

Abstraction principles have been characterized as implicit definitions in recent philosophical discussion (Hale and Wright 2001, pp. 117–150; MacBride 2003, pp. 103–163). They assign certain meanings to the singular terms appearing on the left-hand side, e.g., "Nx:Fx," "Nx:Gx" in the above principle. Abstraction principles settle the truth conditions of identity statements in which the *definienda* appear. For example, in the case of (HP), the *definienda* "the number of the concept F," "the number of the concept G" appear in the identity statement on the left-hand side ("the number of the concept F = the number of the concept G"). The truth conditions of that identity statement are settled by the abstraction principle so that if "F 1-1 G" is true, then "the number of the concept F = the number of the concept G" is true and vice versa. An abstraction principle is a kind of implicit definition with a *conditional* form as B. Hale and C. Wright (2001, pp. 142–145) note. Moreover, the advocates of Neofregeanism have suggested that a requirement that implicit definition should fulfill is to be conditional in form (Hale and Wright 2001, pp. 129–131).

In fact, abstraction principles are universally quantified formulas. (HP) introduces a pattern of *use* for the operator N so that anyone who understands that there is a 1-1 correlation between the instances of the concept F and the instances of the concept G acquires the ability to grasp a meaning for the expression "the number of the concept F" and ascertain that "the number of the concept F = the number of the concept G" is true. As F. Macbride (2003, p. 110) remarks, in the general case of an abstraction

principle, a pattern of *use* for the operator Σ is introduced so that anyone who is familiar with the equivalence relation ≈ becomes also familiar with identity statements of the form "Σ(a) = Σ(b)" and gets acquainted with the meanings of the terms of the form "Σ(a)." Hence, an abstraction principle like (HP) implicitly defines mathematical terms.

Since abstraction principles work as implicit definitions, the *existence* problem pointed out by Horwich arises here too. The suggestion made above in the case of systems of axioms will be stated here, that is, the least requirement is satisfiability (Christopoulou 2013). An abstraction principle which is satisfiable has models. Then certain meanings are accrued to the *definienda*. For example, Hume's principle presented above is satisfiable. By contrast, there is another abstraction principle, Frege's Basic Law V, which is not satisfiable so it fails.

$$(\text{Law V})\forall F\forall G\,[(\{x : Fx\} = \{x : Gx\}) \leftrightarrow \forall x(Fx \leftrightarrow Gx)]$$

Nevertheless, although in mathematics there is no empirical test, there is a consistency test of implicit definitions' operation.

3 The Uniqueness Problem in Case of Mathematical Implicit Definitions

With regard to the second problem, that is, the *uniqueness* problem stated by Horwich, both kinds of mathematical implicit definitions are affected. According to Horwich's "uniqueness problem," meanings assigned to implicitly defined terms are not uniquely determined. Systems of axioms as implicit definitions suffer from this disease, and it appears that Hilbert was aware of that semantic characteristic of axioms. His system of axioms did not uniquely determine any single geometrical term. The axioms determine many systems of concepts "each of which will satisfy the axioms so they fail to determine a unique model" (Resnik 1974, p. 393). Hence, Hilbert claimed that *whatever* entities satisfy the interconnections and mutual relations settled by the Euclidean axioms can stand for a *point*, a *line,* and a *plane*. *Whatever* entities satisfy the postulates (provided that the system is consistent) are such that the implicitly defined concepts can apply appropriately to them.

Similarly, in case of Peano arithmetic, it may be remarked that *whatever* entities satisfy the postulates in question can stand for the natural numbers. For example, the set theoretical surrogates of natural numbers according to Zermelo and Von Neumann:

$$\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \ldots$$

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

Of course, those models are isomorphic; however, they provide different ontologies. In the first model, every set contains just the previous set in the hierarchy,

whereas in the second model, every set contains all the sets that precede it. The dilemma "which of the two sets {{Ø}} or {Ø, {Ø}} is the number 2 ?" was raised by P. Benacerraf (1965), and it is often called a "semantical" dilemma. A similar question concerns the reference of the singular term "2": which of the two sets is the object of reference of the term "2"?

M. Balaguer's (1998) answer to the issue was that arithmetical terms might possess multiple reference, for example, both {{Ø}} and {Ø, {Ø}} can be taken to be objects of reference of the term "2" in regard to different mathematical universes (: models). In fact, a *pluriverse* account can deal with such problems. Models are associated with domains (universes) of mathematical objects which bear certain properties and satisfy relations. Nevertheless, standard models of first-order arithmetic grasp the intended interpretation of the axioms of the theory. Since those models are isomorphic, one can think of just *one* model *up to isomorphism*. However, this is an extensional approach to the issue. From an intensional point of view, meanings of the defined terms may differ, as indicated in the above example concerning two different hierarchies of sets which stand for natural numbers. Moreover, first-order arithmetic allows for nonstandard models which have quite different characteristics. Intensional differences cannot been seen from an extensional point of view which punctuates uniqueness up to isomorphism. Nonetheless, mathematicians usually appeal to extensionality.

Does Horwich's *uniqueness* problem impinge on abstraction principles too, working as implicit definitions? The answer to that question is positive. For example, *Hume's principle* (HP) is the abstraction principle that implicitly defines natural numbers. The operator of it is N: "the number of the concept ( )" which applies to concepts and thereby forms singular terms of the form "Nx:Fx": "the number of the concept F.. Hence, every singular term of the form "Nx:Fx" arises from the completion of the functional expression "the number of the concept ( )" by an expression of a concept (e.g., "F").

An interesting remark which settles the issue of *plural meanings* here comes from H. Hodes (1984, pp. 134–135). Hodes claims that the operator N in question is not uniquely fixed. In general, it is a function expression (an *enumerator*) which correlates concepts to objects of reference of arithmetical singular terms. Particularly, it correlates the concept F to the object of reference of the singular term "Nx:Fx" ("the number of the concept F"). Hodes believes that the enumerator N may receive various semantic interpretations. Hodes's view is important since (HP) itself working as an implicit definition cannot fix the interpretation of the operator N as unique. Yet one of the interpretations N can receive implies the *standard enumerator* n. This works in accordance with the *intended* interpretation concerning familiar natural numbers. According to the *intended* interpretation, the following biconditional holds: "nx:Fx $=$ n $\leftrightarrow$ $\exists$nx Fx," that is, the number of the concept F is the number n if and only if there are exactly n objects which fall under the concept F. In detail, the *standard enumerator* correlates the concept "x$\neq$x" to the number *zero* (i.e., the object of reference of the singular term "nx:[x$\neq$x]"). It also correlates the concept "x=0" to the number *one* (i.e., the object of reference of the singular term "nx:[x=0]"). It correlates the concept "x=0 $\vee$ x=1" to the number *two* (i.e., the

object of reference of the singular term "nx: [x=0 ∨ x=1]") and so on. The *standard enumerator* produces an interpretation of numbers in accordance with *familiar* arithmetic. However, the very possibility of more than one interpretation of the operator N shows that multiple meanings can be accrued to the terms of the form "Nx:Fx," i.e., the *definienda* of (HP). Nevertheless, the standard enumerator is quite privileged since it draws the intended meanings of the arithmetical terms "Nx:Fx."

In fact, Horwich's *uniqueness problem* applies to both kinds of mathematical implicit definition. Many alternative kinds of mathematical "things" satisfy systems of axioms. The case of abstraction principles is similar. This situation shows that some signs of pluralism come up in the procedure of implicitly defining mathematical terms. Implicit definitions appear to be associated to pluralism in mathematical ontology.

4 Pluralism in Foundations

Although systems of axioms working as implicit definitions of mathematical concepts play a foundational role, it appears that they do not safeguard uniqueness of foundations beyond isomorphism. This chapter will present some clear cases of the position called "pluralism" in foundations. According to M. Friend (2014, p. 24), "the pluralist holds that there is insufficient evidence to think that there is a unique foundation for mathematics." Further, she argues, for example, that "the word 'set' is not fully determined by ZF" (2014, p. 22) and that the implicit definition of "set" is different in alternative theories. She also sets out alternative theories like Aczel's non well-founded set theory, Vopênka's set theory, and Hrbacek's set theory. Those theories include different axioms and describe different characteristics of the notion of set; however, they use ZF as a starting point. Nevertheless, ZF set theory is "a standard of measure" for mathematicians, while various mathematical theories are reduced to it. However, it does not constitute a unique foundation for mathematics. A position of a pluralist is that there is no guarantee that ZF is "an absolute guide" to mathematical ontology or epistemology, yet it can be taken to be "a point of reference" for mathematical practice (Friend 2014, p. 23).

In the case of arithmetic, one could mention alternative theories, for example, Robinson arithmetic which is Peano arithmetic without the axiomatic schema of mathematical induction. Instead, it includes an axiom which holds that every (natural) number has a successor (Boolos and Jeffrey 1980). Another alternative theory is Presburger arithmetic that proves every proposition which concerns only the addition of natural numbers. Yet although there are alternative arithmetics, Peano arithmetic is still taken to be the point of reference for mathematicians.

Besides, in case of analysis, aside from standard Analysis, there is Robinson's nonstandard analysis which endorses infinitesimals. Nonstandard analysis is an alternative theory with a different ontology of numbers. Apart from real numbers, it includes infinitesimals as well as infinitely great numbers. The axiom of Archimedes-Eudoxos does not hold in that theory.

To return to mathematical implicit definitions, systems of axioms are supposed to play a foundational role in mathematical theories. Whole theories are based on them. However, the advocates of Neofregeanism Hale and Wright have defended the view that at least in case of Peano arithmetic, an abstraction principle has a prior epistemological status than the axioms themselves. In fact, *Hume's principle* + second-order logic is a consistent system that provides foundation for arithmetic since Peano axioms are derived by that system. Wright (1983, pp. 158–169) proved that Peano axioms are deduced by *Hume's principle* in second-order logic. So Hume's principle plays the role of foundation of arithmetic. The question whether other abstraction principles can play similar roles for other mathematical theories has not been answered yet, and such expectations do not seem to have been fulfilled.

In most cases, mathematical axioms are regularly supposed to have a special epistemological priority and formulate basic mathematical beliefs which are regarded as self-evident and not reducible to other primitive beliefs. This is due to a kind of intrinsic persuasiveness they hold themselves. Further, certain inferential correlations (deductive relations) connect axioms with theorems. Foundationalists insist on the supposed evidence of axioms, and some of them appeal to mathematical intuition. Gödel's platonistic inclinations are widespread. He regarded sets as abstract entities, and he thought that the elementary axioms of set theory are "forced upon human mind" (Benacerraf and Putnam 1983, p. 484). Much discussion has taken place either in favor or against the notion of intuition. For example, J. Brown (1990) and J. Katz (1995) defended the case of intuition, whereas Benacerraf (1973) and Ch. Chihara (1982) criticized it. Recently, J. Bengson (2015) made an attempt to clarify the function of intuition as a representational skill about which an adequate epistemic account was never given. A presentational mental state not only represents world, but moreover, it brings and makes the objects and states of the world present to human mind. Furthermore, he holds that mental states that originate from mathematical intuition, e.g., the intuitive grasp of the case that every natural number has a successor, *make* this fact *present* to our mind. So, intuition is a process of epistemic access to basic mathematical beliefs. Axiomatization of set theory makes those beliefs clear and strict, whereas it corrects any various ill-formed primitive principles. Besides, in concern to the *intuitive* approach to axioms, M. Potter (2004, p. 36) asserts that it invites us to clarify our conceptions of sets to such an extent as to determine the axioms those conceptions satisfy. The notion of mathematical intuition is always a tempting issue in philosophy of mathematics as far as it concerns our access to *basic* mathematical beliefs like axiomatic principles. Further, if one works inside a certain mathematical theory, she gets familiar with the intellectual environment of this theory, so her familiarity might count as intuition too.

However, another interesting point is that there are axioms without self-obviousness for the acceptance of which other criteria are necessary. Decisions concerning the choice of those axioms which lack self-evidence are not based on any intuition skill. Instead, we have to consider their fruitfulness, the way they shed light on certain areas of the theory, etc. The choice of such axioms in mathematical practice is usually based on certain pragmatic criteria according to a set of research targets and theoretical perspectives. So, Potter (2004, p. 35) has taken into account an alternative

strategy of justification of axioms which he calls "regressive" that is not foundational. The *regressive strategy* regards the axiomatic base as successful in case it can produce theorems whose results are accepted on other mathematical grounds. The justification of axioms is not based on any intrinsic evidence or intuitive force, but it comes from their participation in a consistent theoretical context. The "regressive" approach described by Potter is pragmatic in principle and connected to a kind of justificatory coherentism.

Further, a more interesting fact is that different systems of axioms prove to support alternative theories. Moreover, extensions of certain theories imply different directions of building up alternative ontologies. In cases of controversial axioms like the axiom of choice, a set theorist has two paths to follow: either to (a) accept the axiom of choice and make an extension of ZF theory by adding this axiom (taking ZFC theory) or (b) reject the axiom of choice and make an extension of ZF theory adding the denial of the axiom (ZF + ¬AC). Both theories are consistent if ZF is consistent. The decision in favor of one of the two options or the other is made on more or less pragmatic grounds. Many mathematicians accepted the axiom of choice on the basis of its fruitfulness inside set theory and its interconnections with other mathematical areas (Maddy 1997, pp. 54–55). However, a difficulty arises particularly for the proponents of set theoretical realism since they should be able to offer an account of *which* of the two universes is *the real* one. Yet they have no such option since both "universes" are acceptable on different grounds, and none is a clear winner. Just one solution is available for the mathematical realist: to accept a kind of pluralistic account, endorsing alternative mathematical universes.

A similar case is that of CH (the *continuum hypothesis*), the hypothesis that $2^{\aleph_0} = \aleph_1$ (the first infinite cardinal which follows $\aleph_0$ is the cardinal number of the continuum). CH is undecidable. Undecidable propositions are not provable, while their opposites are not provable either. In this case, one can extend ZFC toward two alternative directions: either to (a) add CH to ZFC and form ZFC + CH or (b) add the denial of CH to ZFC and form ZFC + ¬ CH. Both extensions of the original theory ZFC are consistent if ZFC is consistent. Then one might make up one's mind and make use of that theory which is more fruitful than the other. Yet, again, realists are not able to answer the dilemma of *which* of the two universes described by the two alternative theories is *the* real one. Gödel believed that an independent mathematical reality exists, and the question about the truth or falsity of the continuum hypothesis should have a definite answer. Further, he thought that the axiomatic system of Zermelo-Fraenkel set theory perhaps does not provide an adequate picture of the set theoretical realm. Nonetheless, CH remains today an *undecidable proposition* in the context of ZFC theory. The different extensions of ZFC (plus CH or plus ¬CH) are incompatible with each other. Again, Friend maintains: "The pluralist is exactly someone who will at least in principle entertain the possibility that there are several extensions of ZF. The extensions are not equivalent to each other, and some pairs of extensions preclude each other – because if added together they would make set theory trivial" (2014, p. 21).

Since one might choose ZFC+CH and another might choose ZFC+¬CH for various reasons, a pluralistic account can be endorsed here. In a first view, it gives

the impression that it works against mathematical realism. However, this appears to be just an incorrect impression since one can go on by accepting a kind of mathematical realism that endorses more than one mathematical universe. For example, M. Balaguer (1998), the proponent of *full-blooded platonism*, has defended the existence of a plenitudinous mathematical reality. Besides, Balaguer (2017) argues that platonism and mathematical pluralism are perfectly compatible with one another. In that paper, he defends the view that mathematical reality is plenitudinous; hence, he calls his version of platonism "plenitudinous platonism." It seems that a contemporary mathematical realist should accept more than one set theoretical universe with different ontologies of sets in order to retain her realistic inclinations. Additionally, Linnebø (2017, p. 177) notes: "Today, set-theoretic pluralism often takes the form of a "pluriverse" view, which holds that there are many universes of sets, each compatible with our full conception of sets. CH is true in some of these universes and false in others. On this view, CH has a truth-value only relative to some particular universe, ..." The trick is to be a Platonist one level of generality up. They are Platonist about ontologies of "all" legitimate theories. The quantifier is over ontologies of theories, not ontologies within a theory. It's a higher-order quantifier.

In regard to *undecidability* (like the case of CH), D. Anapolitanos (2016, p. 208) writes: "There are models of the theory that make the sentence in question (i.e. CH) true and models that make it false. . .The truth of mathematical sentences *depends* on the specific universe of a certain model of the theory. The true sentence *in regard to* a model describes the characteristics of the relative universe and the characteristics of all universes of the models of the theory in which the sentence is true." In the above passages, the notion of truth is taken to be *relative* to certain semantical interpretations of the language of the theory in question. A standard *semantical interpretation* includes a universe (domain) of objects, whereas certain symbols (constants) are purported to refer to such objects, and others (predicate symbols) denote properties or relations ascribed to those objects. Models can exactly be regarded as those semantical interpretations which make the axioms of a theory true. So, this account takes truth to be *relative* to models. Such an account allows for multiple systems of mathematical objects. Of course, undecidable sentences allow for alternative extensions of the initial theory.

Again, Linnebø ascertains that also in case of geometry, we come across a pluralistic situation, since "it makes no sense to ask, without further qualification, whether the parallel postulate is true or false: it is true of some spaces and false of others" (2017, p. 176). For example, Euclidean geometry and Hyperbolic geometry make a pair of mutually contradictory mathematical systems. They differ over the parallel postulate so a number of true propositions of Euclidean geometry are not accepted in the context of Hyperbolic geometry. In similar words, Euclidean geometry and non-Euclidean geometries describe different possible spaces, and they are each acceptable.

In conclusion, it appears that there are alternative consistent systems of axioms in each one of the most well-known mathematical areas. Those systems of axioms support alternative mathematical theories. Besides, the existence of independent mathematical sentences like CH shows that we possess different options of building

up different extensions of certain theories. The very notion of truth appears to be *relative* to certain semantical universes associated with the axiomatic systems. Besides, an abstraction principle, that is, (HP) plays a foundational role in arithmetic; however, it allows for multiple systems of arithmetical objects. Pluralism in foundations amounts to accepting this.

5 Structuralism, Implicit Definitions, and Pluralism

Both kinds of mathematical implicit definitions (systems of axioms and abstraction principles) can also be considered in the light of some version of structuralism. According to Friend, "Structuralism consists in the view that within the discipline of mathematics, we find a number of theories. Theories can be thought of as structures or as a set of formulas closed under some operations. . .. Objects are only ever objects in a structure. The truth of a theory can only be judged from a meta-perspective: a meta-theory or meta-structure. The standard meta-perspective is provided by model theory. . ." (2014, p. 52). Structuralism is not purported to be foundational. In particular, Friend holds that S. Shapiro's structuralism expresses a kind of anti-foundationalism "which is inspiring for the pluralist" (2014, p. 51). Shapiro (2000) examines how implicit definitions work in the context of *ante rem* structuralism. A successful implicit definition is taken to characterize a structure, that is, a system of mutual relations among certain elements.

A structuralist focuses her attention especially on the relations, not the objects. If certain objects fill in the positions of a structure, then the system of objects in question satisfies the structure and constitutes an exemplary case of it. For example, the axioms of Peano arithmetic determine a structure. The system of natural numbers is regarded as an exemplary case of the structure in question. What concerns a structuralist is not the numbers themselves as objects but their positions in the structure and the mutual relations among them. Such relations might be, for example, the relation between a natural number and its successor, the relation of a as greater than b, the relation between a and b in case a divides b (without remainder), etc. S. Hedman notes: "Structures naturally arise in many branches of mathematics. For example, a vector space is a structure. The groups, rings, and fields of abstract algebra also provide examples of structures. . .. The real numbers provide examples of structures that should be familiar to all readers" (Hedman 2004, p. 58).

Shapiro is an advocate of "ante rem structuralism" according to which structures have an existence independent of all their exemplary cases. Structures are ontologically prior to the systems of objects which satisfy them. Appropriate singular terms are purported to refer to positions in a structure. As Friend (2014, p. 69) remarks, Shapiro develops his structuralism in accordance with model theory.

Besides, Shapiro holds that an implicit definition might define a structure even though no exemplary case of it is known. The Euclidean system of axioms defines a structure (the Euclidean structure of space). As noted earlier, Hilbert believed that *whatever* entities satisfy the relations in question can stand for a *point*, a *line*, and a

plane. Though not a structuralist, Hilbert focused his attention on the system of relations laid down by the axioms. He did not focus on the objects playing the role of a *point*, a *line*, and a *plane*. Additionally, he showed that there is a model for the system of axioms in question.

Shapiro (2000) states two conditions for successful implicit definitions. The first requires that at least one structure is characterized by an implicit definition. According to the second condition, at most, one structure is described by the implicit definition up to isomorphism. What the second condition means is that all systems of objects that satisfy the implicit definition in question share a common structure. Further, Shapiro has chosen second-order logic as the underlying logic of his structures, so he guarantees a unique characterization of those structures up to isomorphism. For example, in the case of arithmetic, the logical background of second-order logic makes possible avoiding all structures which are related to the Löwenheim-Skolem theorems. Unique characterization could not be achieved if the logical background were that of first-order logic. Nonstandard models would appear which satisfy alternative structures. As Shapiro (2000, p. 285) states, in first-order logic, the Löwenheim-Skolem theorems are developed so "no theory that is true of an infinite system characterizes a unique structure."

Another crucial point in Shapiro's account is that an implicit definition needs to be consistent with the rest of the theory in order that it determines a structure. Particularly, Shapiro makes use of the notion of *satisfiability* instead of the notion of (deductive) consistency. Yet, he explains that his choice is quite appropriate: "In the second order logic that I favour there are sentences that are deductively consistent but have no models" (2000, p. 287). As stated above, second-order logic has different characteristics from first-order logic. For example, in second-order logic, the downward Löwenheim-Skolem theorem is not true. Besides, second-order logic is not complete; hence, it does not guarantee that every universally valid second-order formula, under standard semantics, is provable (Hedman 2004, p. 391). Further, what Shapiro stresses in the above passage is that in second-order logic, deductive consistency and satisfiability are not equivalent notions. So, he has chosen satisfiability which implies consistency. In this setting, both kinds of mathematical implicit definitions should be satisfiable: a satisfiable system of axioms defines a structure, and similarly, a satisfiable abstraction principle defines a structure.

The abstraction principle (HP) is satisfiable. In the context of ante rem structuralism, that principle is regarded as an implicit definition that characterizes an arithmetical structure:

$$(\forall F)(\forall G)\,[(Nx : Fx = Nx : Gx) \leftrightarrow (F\ 1-1\ G)]$$

Of course, this would be a structural view of (HP), not the Fregean view (which is foundational). Frege's arithmetic might be regarded just as an exemplary case of that structure. In fact, more than one nonequivalent exemplary case of that structure is possible (see Hodes' considerations in Sect. 3).

An interesting characteristic of implicit definitions is called "non-eliminability" of the *definienda* (Shapiro and Weir 1999). If L is the original language and S is a set of new terms (the *definienda*), then the implicit definition is formulated in the extension of the original language L, i.e., L+S. The expression "non-elimination of the *definienda*" means that there is at least one formula Φ lying in the extension of the original language for which there is no formula Φ′ in the original language L, so that Φ can be translated in terms of Φ′. That is, the *definienda* cannot be translated in terms of the original vocabulary. Moreover, a consequence of that characteristic is no *unique extendability*, i.e., if M is a model of the original language, then there is more than one way to extend that model to a model of the extension of the original language. Hence, this account can explain why the meanings which are accrued to the terms introduced by an implicit definition are not uniquely determined up to isomorphism. This situation holds equally for systems of axioms as well as abstraction principles functioning as implicit definitions.

In general, multiple *meanings* are accrued to the *definienda* of an implicit definition (as Horwich's *uniqueness* problem indicates). Yet, the issue is now addressed by means of a structural perspective: implicit definitions describe structures with *various* exemplary systems of objects. As Hilbert pointed out, *whatever* objects satisfy the supposed system of mutual relations might be regarded as objects falling under the defined notions.

Apart from the pluralistic aspects of structuralism which were just discussed, there are more options to take into account. In fact, different satisfiable abstraction principles characterize alternative structures. For example, the *Nuisance principle* (NP) describes a structure quite different from that of *Hume's principle* (HP). The Nuisance principle is formulated as the following:

$$(\mathrm{NP})\ (\forall F)(\forall G)\ [\upsilon(F) = \upsilon(G) \leftrightarrow F \approx G]$$

$F \approx G$: *a finite number of objects fall under F but not G or under G but not F*

The *Nuisance principle* has only finite models. It is true only in finite domains.

The two abstraction principles (HP) and (NP) are incompatible, that is, they preclude one another. The exemplary cases of the former exclude the exemplary cases of the latter. However, both can be accepted from a pluralistic point of view.

The advocates of *Neofregeanism* Hale and Wright maintain that the Nuisance principle (NP) is a *nonconservative* abstraction principle so it has a serious defect. A principle is *conservative* in case, if when it is added to a theory T, it does not provide new consequences concerning the preexisting ontology of the initial theory T. In fact, (NP) is nonconservative, since it imposes restrictions on the size of the extensions of the concepts F, G, etc. Besides it is usually taken to be a threat for (HP) since the two abstraction principles cannot be together true. Yet, the fact that (NP) has a serious defect entitles the *Neofregeans* to dismiss it.

However, if one is not interested in the motivations of the *Neofregeans*' account (which works as a foundational program), one is not obliged to endorse certain virtues or defects *Neofregeanism* often ascribes to abstraction principles in order to vindicate or dismiss them. Moreover, a pluralist does not need to dismiss any

nonconservative abstraction principle. After all, the above abstraction principles are both satisfiable and describe alternative structures with different kinds of exemplary cases. None of them might be a threat for the truth of the other, because the truth of every such principle should be regarded as relative just to the structure it defines and its own models. Nonetheless, the situation concerning the abstraction principles (HP) and (NP) gives rise to a pluralistic option. All *satisfiable* abstraction principles can be accepted. They define various structures with possible exemplary systems of objects. Of course, unsatisfiable abstraction principles should be rejected as unsuccessful (unless one is interested in exploring an inconsistent situation). This is the case of Frege's Basic Law V which was presented earlier in Sect. 2 and defines the extensions of concepts:

$$\forall F \forall G\, [(\{x : Fx\} = \{x : Gx\}) \leftrightarrow \forall x(Fx \leftrightarrow Gx)].$$

Shapiro's structuralism is not interested in providing mathematical foundations; it appeals just to models. As Friend (2014, p. 156) puts it, model theory is used in order to compare structures. Accordingly, implicit definitions do not play a foundational role in the context of structuralism, but they are supposed to define structures. Theories are about structures, and truth is not independent of structure. Particularly, truth of the bearers of implicit definitions is relative only to the models of the structures so defined. For example, a satisfiable abstraction principle describes a structure, and it is regarded as *true* in respect to the structure. So, there is no purely *mathematical* reason to favor one structure. There are, of course, other reasons: logicality, aesthetics, simplicity, and fruitfulness. In fact, no satisfiable implicit definition (axiomatic system or abstraction principle) describing a structure is favored over other satisfiable implicit definitions on purely mathematical grounds. Especially, in Shapiro's view, model theory individuates mathematical structures. Isomorphic mathematical structures count as *one* structure, i.e., structures are individuated up to isomorphism.

6 Conclusion

Mathematical systems of axioms as well as abstraction principles work as implicit definitions of mathematical terms. If they are satisfiable, then certain meanings are accrued to the *definienda* (: mathematical singular terms or predicates). However, Horwich's *uniqueness* problem indicates that multiple systems of meanings are assigned to the *definienda*. This situation reminds us of Hilbert's remark that *whatever* objects satisfy the Euclidean system of axioms can stand for a *point*, a *line,* and a *plane*. Nonetheless, in Sect. 3, a similar situation was presented for abstraction principles working as implicit definitions. Horwich's *uniqueness* problem affects (HP) too. Except for pluralism in mathematical ontology, there are other aspects of pluralism which are also detected in regard to the issue of implicit definitions. Systems of axioms and abstraction principles are regularly taken to play a foundational role. Systems of axioms support mathematical theories, and, as

Sect. 4 reminds us, the abstraction principle (HP) has been taken to play a foundational role for arithmetic in terms of Wright (1983). However, in various mathematical areas, there are alternative mathematical theories and alternative systems of axioms which support them. Moreover, the case of independent axioms (like the axiom of choice in regard to ZF set theory) shows alternative directions of development of mathematical theories. The case of undecidable sentences indicates similar situations. So *pluralism* in foundations appears to characterize mathematical implicit definitions, particularly systems of axioms as foundations of certain theories. From a different point of view, structuralism offers an interesting account of both kinds of mathematical implicit definitions (systems of axioms and abstraction principles). That account is not foundational. Systems of axioms and abstraction principles are regarded as implicitly defining mathematical structures. In the case of ante rem structuralism, they are taken to *disclose* structures. Satisfiability is a necessary condition for implicit definitions to successfully define structures. From a structuralist view, both kinds of mathematical implicit definitions describe structures with multiple exemplary systems of objects. Besides, Sect. 5 of this chapter deals with satisfiable abstraction principles which preclude each other. Those principles describe alternative structures with different exemplary systems of objects. This is again a situation supporting pluralism to the extent that incompatible but satisfiable abstraction principles describe alternative structures which are equally accepted.

In the context of a structural view of mathematical implicit definitions, it appears that a *modest* form of *pluralism* is brought out. This is a kind of pluralism with some limitations or "a very conservative form of pluralism" as Friend (2014, p. 69) notes. Limitations come from the fact that structuralism appeals to models and model theory. It is also wedded to set theory and classical logic so it gives account of those mathematics which are recognized by model theory. By contrast, a version of *maximal mathematical pluralism* might take in account also cases of intensional logics, paraconsistency, mathematics which are in the stage of development, etc. However, in this chapter, mathematical implicit definitions were discussed in connection with a modest version of pluralism that has been detected both in case (a) implicit definitions are taken to have a *foundational* role and (b) they are considered in the context of *ante rem structuralism*.

References

Anapolitanos D (2016) Labyrinths, epistemological cracks, philosophical fragmentations and consolations. Athens Ekdotiki, Athens

Balaguer M (1998) Platonism and anti-platonism in mathematics. Oxford University Press, New York

Balaguer M (2017) Mathematical pluralism and platonism. J Indian Council Philos Res 34:379–398

Benacerraf P (1965) What numbers could not be. Philos Rev 74:47–73

Benacerraf P (1973) Mathematical truth. J Philos 70:661–679

Benacerraf P, Putnam H (eds) (1983) Philosophy of mathematics, selected readings, 2nd edn. Cambridge, Cambridge University Press

Bengson J (2015) The intellectual given. Mind 124(495):707–760
Boolos GS, Jeffrey RC (1980) Computability and logic. Cambridge University Press, Cambridge
Brown JR (1990) π in the sky. In: Irvine A (ed) Physicalism in mathematics. Kluwer, London, pp 95–120
Chihara CS (1982) A Gödelian thesis regarding mathematical objects. Do they exist? And can we perceive them? Philos Rev 91:211–227
Christopoulou D (2013) On the synthetic content of implicit definitions. Logic Log Philos 22:75–88
Frege G (1884/1953) The foundations of arithmetic. Blackwell, Oxford
Friend M (2014) Pluralism in mathematics: a new position in philosophy of mathematics. Springer, New York/London
Hale B, Wright C (2001) The reason's proper study. Clarendon Press, Oxford
Hedman S (2004) A first course in logic. Oxford University Press, New York
Hodes H (1984) Logicism and the ontological commitments of arithmetic. J Philos 81:126–135
Horwich P (1998) Meaning. Oxford University Press, New York
Katz J (1995) What mathematical knowledge could be. Mind 104:491–522
Lewis D (1970) How to define theoretical terms. J Philos 67(13):427–446
Linnebø Ø (2017) Philosophy of mathematics. Princeton University Press, Princeton/Oxford
MacBride F (2003) Speaking with shadows: a study of Neo-Logicism. Br J Philos Sci 554:120–121
Maddy P (1997) Naturalism in mathematics. Clarendon Press, Oxford
Potter M (2004) Set theory and its philosophy. Oxford University Press, New York
Psillos S, Christopoulou D (2009) The a priori: between conventions and implicit definitions. In: Kompa N, Nimtz C, Suhm C (eds) The a priori and its role in philosophy. Mentis, Paderborn, pp 205–220
Resnik D (1974) The Frege-Hilbert controversy. Philos Phenomenol Res 34:386–403
Shapiro S (2000) Thinking about mathematics. Oxford University Press, Oxford
Shapiro S, Weir A (1999) New V, ZF and abstraction. Philos Math 7:293–321
Wright C (1983) Frege's conception of numbers as objects. Aberdeen University Press, Aberdeen

Non-deductive Justification in Mathematics

A. C. Paseau

Contents

1 Introduction 2402
2 A Case Study: Goldbach's Conjecture 2403
3 Skepticism About Enumerative Induction 2406
4 The Last Bastion of the Euclidean Program 2408
5 Justifying the Consistency of Mathematics as a Whole 2410
6 An Unorthodox Claim: Non-deductive Knowledge of Mathematics 2412
7 Conclusion 2415
References 2415

Abstract

In mathematics, the deductive method reigns. Without proof, a claim remains unsolved, a mere conjecture, not something that can be simply assumed; when a proof is found, the problem is solved, it turns into a "result," something that can be relied on. So mathematicians think. But is there more to mathematical justification than proof?

The answer is an emphatic yes, as I explain in this chapter. I argue that non-deductive justification is in fact pervasive in mathematics, and that it is in good epistemic standing.

Keywords

Proof · Justification · Mathematical knowledge · Mathematical justification · Non-deductive evidence · Empiricism · Non-deductive knowledge · Deductive knowledge · Induction · Foundationalism · Euclidean program

I am grateful to Michèle Friend for the invitation to contribute to this handbook and to Jim Franklin for helpful comments on a previous draft.

A. C. Paseau (✉)
Oxford University, Oxford, UK
e-mail: alexander.paseau@philosophy.ox.ac.uk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_116

1 Introduction

The focus of this survey article is on non-deductive justification in mathematics, by which I mean any kind of justification for p other than a proof of p. (Unless otherwise stated, p will be a mathematical proposition.) I shall argue that this sort of justification is pervasive, and that it is in good epistemic standing. My chapter is complementary to that by James Franklin in the present handbook (Franklin 2021a). His is more focused on the practice side of things, mine more on the philosophical side. The two articles may profitably be read side by side. (It is worth noting, however, that the present chapter is not committed to an objective Bayesian analysis in the way Franklin's is. For all I say here, justificatory relationships between mathematical beliefs or truths may or may not be cashed out in Bayesian terms, objective or otherwise.)

We begin in Sect. 2 with a case study: the non-deductive evidence behind Goldbach's conjecture (GC). A famous conjecture in number theory, it was first put forward by Christian Goldbach in a 1742 letter to Euler. Although GC has much non-deductive evidence behind it, it remains unsolved. Section 3 briefly looks at skepticism about the value of enumerative inductive evidence in arithmetic. What is the justificatory value of, say, verifying the first trillion instances of GC, all of which are in some sense small? Section 4 ties the rise of non-deductive methods to the decline of the Euclidean ideal in mathematical epistemology. Section 5 considers how one might justify the consistency of mathematics as a whole. The natural, and perhaps only, way to do so is by means of non-deductive evidence. We conclude in Sect. 6 by examining a radical suggestion: perhaps non-deductive evidence is enough for *knowledge* of a mathematical proposition. We expound one of the arguments for this suggestion.

Right at the outset, we need to be clear that most mathematical justification, even knowledge, is testimonial, because it is acquired from other sources (people, books, journals, websites, social media, etc.). For example, I am justified in believing that Fermat's Last theorem is true because I have heard of its proof from various reliable sources. (Or at least from a preponderance of reliable sources. Fermat's Last Theorem states that if $a^n + b^n = c^n$, where a, b, c and n are all positive integers, then $n = 1$ or 2.) Testimonial justification is non-deductive because it is not based on a mathematical proof of your source's reliability. Any evidence that Andrew Wiles is a reliable mathematician is partly but not strictly mathematical: it compares his mathematical pronouncements (an empirical matter) to the mathematical facts. And evidence that the media and personal channels through which I have heard of Wiles's proof of Fermat Last Theorem's are accurate is not strictly mathematical either. (Calling it 'Wiles's proof' is a simplification; as is well-known, Wiles's original 1993 proof of Fermat's Last Theorem contained a flaw. The 1995 patch-up is owed to Wiles and Richard Taylor.) So most justification of a mathematical proposition p is non-deductive, even if the justification chain ends in a proof. That much is agreed on all hands. A more contested question concerns the extent and significance of justification that does *not* end in a proof because there is currently no proof to be had. It is this sort of justification that is our topic here.

2 A Case Study: Goldbach's Conjecture

Franklin (2021a) illustrates the notion of non-deductive evidence in mathematics with several examples. Section 5 of his article notably presents some of the evidence for the presently unsolved Riemann hypothesis. The present section does the same for GC. (The material here draws significantly on p. 779 of my (2015). My survey of the non-deductive evidence for GC is far from exhaustive.)

Before we get to brass tacks, two notes on terminology. I prefer to talk about "the justification of mathematical statements" rather than "mathematical justification" as it might be insisted that properly mathematical justification must be deductive. As a matter of definition, one might insist, any justification for p that does not take the form of a proof of p is not *mathematical*, sensu stricto. I am not sure that's right; but in any case, it's a terminological fight not worth having, why is why I called my 2015 article on the subject "Knowledge of Mathematics Without Proof" rather than "Mathematical Knowledge Without Proof."

The second point is that the word "induction" is multiply ambiguous. It has a procedural or ceremonial meaning (e.g., "induction into the Rock and Roll Hall of Fame") and a physical meaning ("electromagnetic induction"). Even setting these aside, it can mean (at least) three more things. (1) The first is an inference from particular instances to a generalization. An induction is something of the form "A_1 is F, …, A_n is F; therefore all As are F." (Or even, in some cases, an inference to this conclusion from infinitely many instances.) I shall reserve the term *enumerative induction* for this. (2) The second is a broader sense: any non-deductive inference or any form of non-deductive evidence. Bertrand Russell consistently used the word in this way, and I followed him in my (2015). But upon reflection, confusion is avoided and clarity promoted if we simply write "non-deductive" for "inductive" in this sense. That will be my policy here. (This is the sense of induction that is relevant to Hume's problem of induction, i.e., the problem of justifying an inference from observed to unobserved cases. Some recent mathematical results are highly pertinent to this problem, as discussed in my (2011) and (2008).) (3) In its third sense, it forms part of the expression "proof by mathematical induction," i.e., a proof that all numbers have P by showing that 0 has P and that if n has P then $n + 1$ also does. I shall not be concerned with induction in this sense.

On to GC, then, which states that every even number greater than 2 is the sum of two primes. Number theorists are highly confident of GC's truth on the basis of non-deductive evidence. What form does this evidence take?

First, there is enumeratively inductive evidence for GC. Specifically, GC has been checked for every even number up to about 4×10^{18} and double-checked up to a number not much smaller. (The latest results can be found on Tomás Oliveira e Silva's website http://www.ieeta.pt/~tos/.)

Second, various slightly weaker claims than GC have been proved. An example is the ternary Goldbach conjecture, that every odd number ≥ 7 is the sum of three primes. The Soviet mathematician Ivan Vinogradov proved in the 1930s that every sufficiently large odd number is the sum of three primes; in 2013, Harald Helfgott proved the ternary conjecture outright. Two more results in the same vein are:

(a) every sufficiently large number is the sum of a prime and either a prime or the product of two primes and (b) every even number is the sum of no more than six primes. Both of these have been proved, so are bona fide theorems. The idea here is that if a slightly weaker claim than GC is proved, that makes GC more likely to be true. Note in passing that this illustrates our broad use of "non-deductive evidence for p" to include *proofs* of statements distinct from p but related to it in some interesting way.

This sort of justification exhibits the following pattern (compare Mazur (2014, p. 25); Pólya (1956, 1968) is a classic early discussion of non-deductive reasoning in mathematics):

$$\begin{gathered} A \Rightarrow B_1, B_2, \ldots, B_n \\ B_1, B_2, \ldots, B_n \text{ all hold} \\ A \text{ gains in justification} \end{gathered}$$

Of course, "$A \Rightarrow B$, B; therefore A" is a formal fallacy, known as *affirming the consequent*. (Or "modus morons," as Haack (1976, p. 115) playfully calls it.) This is another way of saying that the bare formal bones of this sort of argument are non-deductive.

The schema can be further elaborated. For example, A gains *more* in justification the more consequences B_1, B_2, …, B_n are verified (i.e., the greater n is); the more varied and independent of one another B_1, B_2, …, B_n are; the "closer" each of B_1, B_2, …, B_n is to A; the more of the B_1, B_2, …, B_n were discovered after conjecturing A and independently of that conjecture; and so on. (On closeness: the claim that every even number >2 is the sum of six primes is for instance closer to GC than the claim that every even number >2 is the sum of 10 primes. Needless to say, how to make precise sense of this notion of closeness is tricky.)

Third, let G(n), the Goldbach number of n, be the number of different ways in which n can be written as the sum of two primes. GC can then be expressed as the claim for all even n greater than 2, G(n) $\geq$ 1. As Echeverría (1996) points out, computer evidence shows that the function G(n) broadly increases for even n as n increases (with oscillation, but with an increasing trend), so that for instance for even $n \approx 10^5$, G(n) $\geq$ 500. In light of this evidence, that G(n) will suddenly drop to 0 is regarded as deeply unlikely.

This third sort of evidence combines neatly with enumerative inductive evidence, because it suggests that smaller numbers are more likely to be counterexamples to be GC than larger ones. (This example is also discussed by Baker (2007, pp. 69–70).) In the case of GC, there are therefore conjecture-specific reasons for thinking that the earliest cases are the "hardest." GC, incidentally, is by no means unique in this respect. Another example is Legendre's conjecture, also currently unproved, which states that for every positive integer N there is a prime between N^2 and $(N+1)^2$. As in the case of GC, the enumerative inductive evidence suggests that not only is Legendre's conjecture true, but also that the number of primes between N^2 and $(N+1)^2$ non-strictly increases with N. Incidentally, there are heuristic arguments for this conclusion too, of

the sort we will shortly mention for GC. The prime number theorem implies that the number of primes between N^2 and $(N+1)^2$ is asymptotic to a quantity which increases with N. So for Legendre's conjecture, as for GC, smaller numbers are "hard" cases – they are more likely to yield counterexamples than larger numbers. (There are other conjectures for which we have specific reason to believe that early cases are *easy* cases. An example is the conjecture that all perfect numbers are even; a natural number N is perfect iff its factors sum to 2 N. Since Ochem and Rao (2012), we know that the smallest odd perfect number, if it exists, must be greater than $10^{1,500}$. In this respect, such conjectures are diametrically opposed to GC and Legendre's conjecture.)

Fourth, the ratio $R_N = \frac{1}{N}$. (numbers $1 < k \leq N$ such that $G(2\ k) = 0$) has been proved to tend to 0 as N tends to infinity. In other words, the density of counterexamples to GC is zero. Of course, if GC is true then R_N is simply equal to 0 for all $N > 1$.

Fifth, Hardy and Littlewood's formula for the asymptotic number of representations of

$$N = p_1 + \ldots + p_m,$$

where $p_1 \leq \ldots \leq p_m$ are m primes, has been proved for $m \geq 3$. If true for $m = 2$, it implies GC for sufficiently large even numbers.

Sixth, a well-known heuristic probabilistic argument suggests the same conclusion. The prime number theorem states that the number of primes up to N tends to $N/\ln N$ asymptotically, i.e., the ratio of these two quantities tends to 1 as N tends to infinity (here ln N is the natural logarithm of N). Using the prime number theorem, one can estimate the number of ways in which an even number N can be written as the sum of two primes. One simple way of doing so based on the theorem estimates that N can be written as the sum of two primes in $\frac{1}{2}.N/\ln^2 N$ ways, a quantity that increases with N. (The argument seems to be mathematical folklore.)

This last argument is very rough and ready, and, to stress, heuristic rather than demonstrative. But it can be improved to yield much better estimates that point to the same conclusion: the greater N is, the greater $G(N)$ is likely to be. Such arguments remain heuristic – GC has not been proved – but they make it plausible that the first counterexample to GC, if one exists, will be a "small" number. These sorts of arguments are much more convincing than one might initially think: if the numerical facts did not follow the probabilistic expectation, the thought goes, there would be some totally unknown mathematical phenomenon that would cause the deviation – and there is no reason to expect this.

On the basis of this and other evidence – our list is illustrative rather than exhaustive – mathematicians are close to certain of GC's truth. The non-deductive evidence behind GC is justification for its truth, even in the absence of proof. One might say that the truth of GC best explains this wide range of evidence, and therefore that we should infer its truth on this basis. (Lange (2022a) discusses inference to the best explanation in mathematics.)

In illustrating some of the non-deductive evidence for GC, we have taken into account the strictly mathematical evidence for it. This is, if you like, the first-order

evidence for GC. But for any given p, such as GC, there is also higher-order evidence for p: what other mathematicians make of this first-order evidence for p. To determine how much the evidence supports p, mathematicians also take each other's judgements into account. In any real-life situation, a mathematician's judgement of how much the mathematical evidence supports p will also depend on what other mathematicians make of the same question. When you think about how likely p is to be true, your peers' judgements also matter.

This last point interestingly complicates the picture of mathematical justification, and applies to proof-based and non-proof-based justification alike. We will not dwell on it further here, but simply note that the sum total of the non-deductive evidence for p includes other experts' judgements.

3 Skepticism About Enumerative Induction

The first sort of evidence for GC mentioned in the previous section consists of the first 4×10^{18} verified instances of GC. In the absence of supporting reasons, mathematicians may mistrust such evidence for arithmetical generalizations, more so than most other forms of non-deductive evidence. Some philosophers have also expressed skepticism about the value of enumerative inductive evidence in arithmetic.

Why? The reason usually given is that known instances of an arithmetical conjecture are almost always small. (We say "almost always" because, for example, one could have an argument that all *odd* numbers satisfy some arithmetical property. That would constitute enumerative inductive evidence for the claim that *all* numbers have that property. But the evidence would consist of infinitely many cases, which could not all be said to be small.) For example, in the case of GC, the evidence is potentially biased, as it consists only of the *first* 4×10^{18} natural numbers. Since the size of a natural number significantly affects its properties, our enumerative inductive evidence seems biased with respect to size.

Following Frege (see §10 of his 1884), Alan Baker has given voice to this sort of skepticism. In an article devoted to the subject, he concludes with the following normative and descriptive point (about arithmetic): mathematicians ought not and in general do not "give weight to enumerative induction *per se* in the justification of mathematical claims" (2007, p. 72). But following other writers, Baker allows that circumstantial reasons can come to the rescue of an enumerative induction, in arithmetic as well as elsewhere in mathematics.

We limit ourselves to making three quick points against the skeptic, which a more expansive treatment would develop. The first point, already made in Sect. 2, is well appreciated by Baker and is worth stressing. In several cases, we have good reason to think that early cases are hard cases. We mentioned GC and Legendre's conjecture as examples. For these conjectures, small cases may be biased, but they are *favorably* biased: they are the cases most likely to yield a counterexample. If no such counterexample is found among them, the conjecture has passed an important test, like a climber who has got past the steepest part of the mountain face. As a result, confidence that all remaining cases fall in line can reasonably increase.

In the case of GC, the evidence for thinking that early cases are hard cases is, of course, partly based on enumerative induction. However, unless one is an out-and-out skeptic about enumerative induction, the circle here is virtuous. Enumerative inductive evidence is deployed to show that, so far as GC is concerned (say), the size bias works in favor of someone deploying enumerative inductive evidence to confirm it. In any case, as we also saw in Sect. 2, a heuristic argument also points to the same conclusion.

A second and related point, also acknowledged by Baker, was implicit in Sect. 2. The situation in which *all* we have is enumerative inductive evidence is a rare one. Almost always, this evidence is accompanied by other sorts of non-deductive evidence. Section 2 detailed some of that accompanying evidence in the case of GC.

The third point is that the viability of this sort of "size-skepticism" depends on what's motivating it. In a recent article (Paseau 2021), I distinguished three sorts of size-skeptics and pointed out that some are better motivated than others. In particular, some are better able to respond to the following *frontloading argument*.

Let E, a finite subset of the natural numbers, consist of our enumerative inductive evidence for a particular arithmetical conjecture. In other words, E is the set of known instances of a generalization over the natural numbers. Let the function v be our evidential function, with domain all finite subsets of the natural numbers. We assume only that v's codomain is the closed unit interval [0, 1] with the usual order; the higher v's value in [0, 1], the stronger the evidence. Evidential values may be thought of as measuring the subject's rational degree of confidence in the generalization in question, though without commitment to the whole panoply of probabilistic ideas.

Consider next the following, presumably uncontroversial, evidential principle:

More is Better
If n is not in E then $v(E \cup \{n\}) > v(E)$.

As its name indicates, *More is Better* simply captures the idea that more evidence is better than less; so the evidential value of more evidence is greater than that of less. Next, define $l = \lim_{n \to \infty} I_n$, where $I_n = v(\{0, 1, \cdots n\})$. By *More is Better*, if $m < n$ then $I_m < I_n$; and since 1 is an upper bound for the I_n, the limit l exists. The real number l itself, of course, may be 1 or smaller than 1, but it has to be greater than 0 (by *More is Better*). So we deduce that $0 < l \leq 1$. Now by the definition of a limit, for any $\varepsilon > 0$, however small, there is an N_ε such that for any $N^* \geq N_\varepsilon$, I_{N^*} is to within ε of l.

Here's another way of putting it: for ε much smaller than $l - \varepsilon$, almost all the evidential value stems from the first N_ε instances of the enumerative induction. The remaining instances add very little evidential value. The evidential value of any finite amount of numerical instances is therefore concentrated almost entirely in an initial segment. Whatever arithmetical conjecture you wish to test, the value of further instances beyond some finite bound (that depends on the conjecture) will be vanishingly small. An initial segment provides the lion's share of the confirmation.

The conclusion of this remarkably simple argument appears to contradict size-skepticism. Paseau (2021) discusses in detail which forms of size-skepticism are genuinely affected by it. The conclusion there is that some but not all are.

4 The Last Bastion of the Euclidean Program

What is the structure of mathematical justification? The traditional picture is foundationalist. More specifically, it is a form of foundationalism largely inspired by Euclid's geometrical method in *The Elements* (c. 300 BC). So what was Euclid's method? Starting from some definitions, postulates, and common notions, Euclid derives the geometry of his day theorem by theorem, in a cumulative manner over the course of 13 books. A concise summary of what he calls the *Euclidean program* is given by Imre Lakatos in the following passage, where he contrasts it with the *Empiricist program*:

> The Euclidean programme proposes to build up Euclidean theories with foundations in meaning and truth-value at the top, lit by the *natural light of Reason*, specifically by arithmetical, geometrical, metaphysical, moral, etc. intuition. The Empiricist programme proposes to build up Empiricist theories with foundations in meaning and truth-value at the bottom, lit by the *natural light of Experience*. Both programmes however rely on Reason (specifically on logical intuition) for the safe transmission of meaning and truth-value. (Lakatos 1962, p. 5)

The most obvious way to spell out the Euclidean program would be to base it on what Euclid himself has to say about it in *The Elements*. But that would give us very little to go on, because Euclid offers us no philosophical gloss on his method, as many commentators down the ages have noted.

Lakatos offers us more. He characterizes the Euclidean program as follows:

> I call a deductive system a 'Euclidean theory' if the propositions at the top (axioms) consist of perfectly well-known terms (primitive terms), and if there are infallible truth-value-injections at this top of the truth-value True, which flows downwards through the deductive channels of truth-transmission (proofs) and inundates the whole system. (If the truth-value at the top was False, there would of course be no current of truth-value in the system.) Since the Euclidean programme implies that all knowledge can be deduced from a finite set of trivially true propositions consisting only of terms with a trivial meaning-load, I shall call it also the Programme of Trivialization of Knowledge. Since a Euclidean theory contains only indubitably true propositions, it operates neither with conjectures nor with refutations. In a fully-fledged Euclidean theory meaning, like truth, is injected at the top and it flows down safely through meaning-preserving channels of nominal definitions from the primitive terms to the (abbreviatory and therefore theoretically superfluous) defined terms. A Euclidean theory is eo ipso consistent, for all the propositions occurring in it are true, and a set of true propositions is certainly consistent. (Lakatos 1962, pp. 4–5)

Now in this passage Lakatos speaks of truth (and meaning) injection; but this is somewhat misleading. The Euclidean program represents an epistemological conception, and the hierarchical path from axioms to theorems is a path the subject, as

opposed to reified truth, follows. The flow-of-truth metaphor is better construed as transmission of an epistemic good of some sort, such as justification say. The picture is then a foundationalist one in which one gains justification for axioms first and thence for theorems by inferring them from the axioms.

Historical proponents of mathematical epistemologies that, to one or degree or another, approximate the Euclidean conception are many and varied. The conception's high point came in the seventeenth century; see in particular Pascal's posthumously published *On the Geometric Mind* (written in the 1650s) or even Descartes' *Discourse on Method* (1637). For various reasons, the Euclidean program is no longer tenable as a mathematical epistemology for *all* mathematics. Paseau and Wrigley (2024) explains the reasons why for the case of set theory. In brief, the standard axioms of set theory are no longer generally regarded as self-evident; at least, not all of them are. And deductive rules that take us from theorems to axioms are also not thought to be certainty-preserving, or even rational-credence-preserving. On top of that, the Euclidean program's ideal of completeness is also, post Gödel, not realizable. Any reasonable axiomatic organization of set-theoretic truths in a deductive system will omit some of them.

Arithmetic much more closely approximates the Euclidean picture than set theory does. (As discussed in Sect. 5 of Franklin (2021b), for example.) The justification of its axioms seems more intrinsic than extrinsic. Extrinsic evidence for a principle consists in its instrumental value, in drawing consequences, forging connections between different areas, making for better explanations, and the like. This is the kind of evidence on which theoretical principles in science are mostly, if not exclusively, based. Intrinsic evidence we may take to be non-extrinsic evidence: the sheer obviousness or plausibility of a principle, as well as how it fits with the broader conception of the subject matter. And many would see the axioms of arithmetic as intrinsically evident.

One of the Euclidean program's tenets that's still very much standing is that deduction is prized as the highest form of justification available for a mathematical proposition and regarded as necessary for knowledge. In fact, in some quarters, deductive justification is thought to be the *only* available form of justification; and for many, non-deductive justification pales in comparison to the real McCoy: proof. To challenge these sorts of proof-centric ideas, or at least weaken their hold, is thus to challenge one of the last bastions of Euclideanism.

Friends of non-deductive reasoning in mathematics will find what Lakatos called the Empiricist program much more congenial than the Euclidean one. This post-Euclidean outlook was perhaps first articulated by Russell, who drew a "close analogy between the methods of pure mathematics and the methods of the sciences of observation" (1907, p. 272). On a thumbnail, the idea is this. We commonly conceive of natural-scientific propositions as being divided into two broad kinds: data and more theoretical principles, ultimately laws. On this (simplified) conception, the data are empirical propositions that we take to be the facts, and the principles/laws are propositions formulated in order to predict the facts. Indeed, the prediction of the data is the primary means of verifying these principles/laws. Whether or not the principles/laws are intrinsically plausible, to a first

approximation we take them to be true if they predict all the data and don't predict anything false.

By analogy, mathematical axioms are supposed to be verified by "predicting" mathematical propositions of some privileged kind identified as the data. Prediction here is simply deductive implication. The justification of these principles or laws would then be extrinsic. This is broadly Russell's view of the matter, which he calls the "regressive method." It plays a major role in his own foundational system; for example, in the introduction to the first edition of *Principia Mathematica*, he (with Whitehead) had this to say about the controversial so-called axiom of reducibility:

> That the axiom of reducibility is self-evident is a proposition which can hardly be maintained. But in fact, self-evidence is never more than a part of the reason for accepting an axiom, and is never indispensable. The reason for accepting an axiom, as for accepting any other proposition, is always largely inductive, namely that many propositions which are nearly indubitable can be deduced from it, and that no equally plausible way is known by which these propositions could be true if the axiom were false, and nothing which is probably false can be deduced from it. (Whitehead and Russell 1910/1962, pp. 59–60)

It is essential to the Empiricist program that certain propositions are identified as being data, and that these have a special epistemological status which explains their role in the program. Various manifestations of the Empiricist program will have different ideas about which propositions are properly classified as data, or which propositions are lit by the natural light of experience. Plainly, not just any mathematical truth can be considered a datum, otherwise any true axiom would be self-certifying in an unacceptable way. For Russell and Whitehead, notably, this moral applies to arithmetic as well: the axioms are justified because they entail propositions such as "$1 + 2 = 3$," not the other way round.

The use of non-deductive methods in mathematics therefore tallies better with the Empiricist than the Euclidean program. To put it very roughly, mathematical justification is much more like scientific justification than traditionally imagined. At the very least, it has an important extrinsic component. It is then a short step to the idea that non-deductive evidence has an important role to play in mathematical justification, just as it does in science. Coming at it from the other direction, to recognize the important role this sort of evidence plays in mathematics is to chip away at the epistemological dimension of the empirical science/mathematics divide. Mathematics is much more like science than our philosophical forebears imagined. Just as in science, non-deductive justification has an important role to play.

5 Justifying the Consistency of Mathematics as a Whole

Let's now turn to an apparently unrelated question, whose connection to our main topic will emerge shortly. Set theory is regarded by many as a foundation for mathematics – though in what sense exactly remains a source of controversy. An incontrovertible fact is that almost all mathematics can be carried out in set theory. And we can prove the consistency of virtually all mathematical theories – arithmetic,

analysis, geometry, and so on – in set theory. (The consistency of a theory may here be understood as its not implying every sentence.) If set theory were inconsistent, a proof of, say, the consistency of arithmetic in set theory would be cold comfort, since set theory would also prove that arithmetic is inconsistent – indeed, it would prove anything statable in its language. A set theoretic proof of arithmetic's consistency is thus best understood as a relative consistency proof: it establishes the consistency of one theory (arithmetic) on the assumption that another (set theory) is consistent. To interpret such a proof as telling us that arithmetic *is* consistent, we need reason to think that set theory itself is consistent.

But now a difficulty looms: we have no proof of this latter fact, that is, of set theory's consistency. And this for a principled reason: by Gödel's second incompleteness theorem, if set theory is consistent then it cannot prove its own consistency. Of course, we can take our preferred system of set theory S and extend it to S^+ and then proceed to prove the consistency of S within S^+. But that only pushes the question back one stage: how do we convince ourselves that S^+, our new ultimate theory, is consistent? By Gödel's theorem again, we cannot do so in S^+ (assuming S^+ is consistent).

So far we have assumed that set theory is the foundation of mathematics. But actually this is inessential. The question of consistency can be raised for any putative foundation for mathematics, e.g., category theory as opposed to set theory. And even if you think mathematics has no foundation, you still face the question of why we are justified in thinking mathematics as a whole is consistent. Everything said here can be easily reformulated to accommodate either of these alternative views.

The question of the foundation's consistency is particularly pressing if you take all justification in mathematics to be deductive, i.e., to be given by proof and exhausted by proof. For, on that view, how on earth are we justified in believing set theory's fundamental principles? The reply that an axiom has a zero-step proof from the axioms, though true, is hardly consoling. In any axiomatic system, the axioms are trivially provable, for systems consisting of true axioms and of false axioms alike. Any crank can put forward a hare-brained mathematical "system" and give a zero-step proof that each of its axioms is true. So how do we know that the axioms of our *not* hare-brained but trusted (we think) system are true? Or at least why are we justified in so thinking?

Indispensabilists have an answer. Indispensabilism is the philosophy of mathematics inspired by Quine and Putnam; see Paseau and Baker (2023) for a more recent account. In a succinct formulation, the idea is that we are justified in believing the axioms of mathematics as a whole – whatever exactly these are – because they are successfully and indispensably applied in science. Mathematical justification ultimately rests on scientific justification. This is the indispensabilist's broad answer to the question of how axioms are justified. Plainly, though, the answer is indirect and holistic: to justify the thought that set theory is true, or even consistent, requires no less than a detour through the whole of science, or at least large parts of it. It would be better to combine this holistic justification with more direct evidence for the consistency of mathematics. As we have seen, this evidence apparently cannot be deductive, so has to be non-deductive. In sum, we have here an important role for

non-deductive evidence to play. Such evidence, if it exists, can be used to show that set theory (and hence mathematics as a whole) is consistent.

Does such evidence exist? Most definitely. Maddy (1988) is a then state-of-the-art discussion of non-deductive evidence for the axioms of set theory and their possible extensions. As such, it offers plenty of reasons for believing in set theory's consistency. Chapter 8 of Paseau (forthcoming) is a shorter and less technical survey of some non-deductive evidence for the same.

6 An Unorthodox Claim: Non-deductive Knowledge of Mathematics

We have so far been concerned with non-deductive *justification* and *evidence* in mathematics. We end this chapter by considering a much more radical idea: that there can be non-deductive (non-testimonial) *knowledge* of mathematical propositions. This claim goes squarely against the way mathematicians speak, since mathematicians typically equate p's being known with there being a proof of p. (For mathematical p, obviously. In my (2015), I argue that gainsaying mathematicians is much less problematic when they are talking not about mathematics but about its epistemology, as in this case.) This is the quasi-universal, orthodox view. I am among the very few who dissent from it, because I believe that, in the best cases, non-deductive evidence can yield knowledge of a mathematical proposition (Paseau 2015). In this section, I shall present an argument for the unorthodox view, adapted from Sect. 5 of Paseau (2015); my article contains several other such arguments.

Epistemologists have given a good deal of general thought to knowledge. Not just mathematical, but empirical, scientific and moral knowledge, self-knowledge, knowledge of the past and the future, a priori and a posteriori knowledge, and so on. Ever since the publication of Gettier (1963), one particularly prominent concern has been to provide necessary and sufficient conditions for knowledge. Notoriously, none of the myriad proposed conditions has achieved consensus (for general reasons lucidly discussed in Zagzebski (1994)), although many have been thought to be along the right lines, or at least to cover an important range of cases. For brevity, call any analysis that has gained at least some traction in the literature a "right-track analysis." (Some examples will shortly follow.)

The argument for non-deductive knowledge of mathematics is that any right-track analysis falls into one of two categories. *Either* it allows that knowledge of mathematics may be obtained by non-deductive means; call this category *Type A*. *Or* it does not apply to knowledge of mathematics, so *a fortiori* does not privilege deductive, as opposed to non-deductive knowledge of mathematics; call this *Type B*. The underlying thought here is that if (non-testimonial) knowledge of a mathematical proposition could only be deductively acquired, at least *some* right-track analyses would have that implication, when supplemented with some generally accepted principles.

So let's take a look at a few accounts of knowledge. The first and most venerable one is that knowledge is justified true belief. As Gettier (1963) notes, Plato in the

Theaetetus and the *Meno* may have respectively considered and proposed such an account. This account is of Type A: it allows for non-deductive knowledge of mathematics, since a mathematical proposition may be justified non-deductively. For example, the evidence behind GC (some of which we encountered in Sect. 2) justifies our belief in it.

An influential revised account was offered in Goldman (1967). Goldman suggested that a subject knows that p just when her true belief that p is causally connected to the fact that p. Whatever the merits of this account for other domains, it does not apply to mathematics, since the subject matter of mathematics does not have causal powers: the number 53 itself does not causally impinge on our senses any more than a Banach space or an ordered field do. Goldman's "causal analysis" of knowledge is therefore of Type B. Goldman himself could not have been clearer on this point, declaring in the first paragraph of his famous article that "[m]y concern will be with knowledge of empirical propositions only, since I think that the traditional analysis is adequate for knowledge of nonempirical truths [including those of mathematics]" (Goldman 1967, p. 357).

Now that we have seen how this sort of argument goes, it can be extended to other post-Gettier right-track analyses more swiftly. These typically supplement the first two clauses, that p is true and S believes that p, with a third condition that aims to improve upon the justification clause. Here is a sample, along with their classification:

- The belief that p is not inferred from a false lemma (Clark 1963). Of Type A, since non-deductive evidence for p need not contain a false lemma.
- The justification that p must not essentially rest on a false assumption (Harman 1973). Similar to the previous: of Type A, since non-deductive justification for p need not, and in many usual cases will not, essentially rest on a false assumption.
- There is a law-like connection between the fact that p and the belief that p (Armstrong 1973). Armstrong intended it to cover empirical cases only, so this analysis is of Type B.
- The belief that p is produced by a reliable process not undermined by the subject's cognitive state (Goldman 1976). Of Type A, since non-deductive evidence can be the product of a reliable process.
- If it were the case that not-p then the subject wouldn't believe p (Nozick 1981). (Ignoring Nozick's other condition, which is problematic and does not affect the moral. More generally, in presenting these accounts, I have omitted qualifications, refinements, and extra clauses that do not affect the general point.) To apply this, note that mathematical propositions are usually thought of as necessary. The standard account of counternecessaries in the literature takes them to all be true, in which case Nozick's analysis is of Type A – it allows for non-deductive knowledge of mathematics. (The standard accounts are derived from Stalnaker (1968) and Lewis (1973). As Baker (2021) notes, following an observation by Ralph Wedgwood, counternecessaries are usually known as "counterpossibles." If the analogy with "counterfactual" is to be exact, however, "counterpossibles"

should be known as counternecessaries. See Baker (2021) for more on these types of conditionals, whatever they should be called.)

Even if we reject the standard account of counternecessaries, it seems that Nozick's account is of Type A. For example, we are liable to think that if the Riemann Hypothesis were false, we would lack the evidence we in fact possess for it, and consequently that we wouldn't believe it. Similarly, if the perpendicular bisectors of Euclidean triangles did not meet in a point then trying to confirm this fact empirically—by drawing the lines with the utmost care on a plane surface—would not lead us to conclude that they do. If a particular number were composite rather than prime, the primality test evidence would be different. And so on.

- The subject could not easily have falsely believed that p (this is the safety condition discussed in, e.g., Williamson 2000). This condition can easily be failed by true mathematical beliefs, e.g., if they are the product of happenstance or of an unreliable method that happens to be right in this instance. For instance, suppose I believe that $1^3 + 5^3 + 3^3 = 153$ because I recognize it as an instance of the generalization $a^3 + b^3 + c^3 = abc$, where abc is written in decimal notation, which I believe to be true. The generalization is patently false, but this instance of it happens, quite fortuitously, to be correct. The condition is of Type A: the best forms of non-deductive evidence satisfy it.

Observe in passing that the best inductive mathematical cases are quite unlike lottery cases. Say that p is the proposition that the subject won't win the prize in a large lottery, with N tickets, that she enters. In a lottery case, the subject's evidence for p is insensitive to whether p is true. She has the same evidence and belief in the scenario in which she holds the winning ticket as in the $N - 1$ nearby scenarios in which she holds a losing ticket. Not so for the best cases of non-deductive evidence in mathematics.

Evidently, this is just a sample from a vast post-Gettier literature. But it includes most of the leading candidates; and none of them was chosen with a view to confirming the "either Type A or Type B" moral. In any case, broadening the range of examples would not alter the moral. General epistemology has not brought to light any reasonably popular condition that excludes a non-deductive (non-testimonial) route to knowledge of mathematical propositions. No right-track analyses allow deductive routes to mathematics while ruling out non-deductive ones. As a prominent field of inquiry, mathematics is, and should be, a test case for general epistemology. If the only (non-testimonial) route to knowledge of a mathematical proposition were deductive, you would expect some prominent general accounts of knowledge to have this consequence. That they don't supports the idea that such knowledge can be acquired non-deductively.

Of course, this is not a knockdown argument for the conclusion that we can know a mathematical proposition without proof. (Which, to repeat, is to be understood in the strong sense that no one has a proof of it. As stressed, we often know p testimonially without being able to prove p ourselves.) It is one of many arguments and considerations supporting that idea. To further strengthen it, one ought to consider those as well. A fuller treatment would examine not just these but the conservative backlash as well, if we may call it that. Lange (2022b), for example, is

an interesting recent article that tries to support the standard view by proposing a necessary condition on knowledge that Lange believes (i) is independently motivated and (ii) rules out non-deductive knowledge of mathematics.

A closely related point worth stressing is that, at least in some cases, there is a gap between strong justification for one's true belief that *p* and knowledge that *p*. Gettier's refutation of the "JTB" analysis of knowledge showed that justification is not enough for knowledge, even when the belief is true. (Nagel (2014, p. 58) observes that discussion of so-called Gettier cases may be found in Indo-Tibetan philosophy that predates Gettier by centuries.) In fact, one can go further: one of the lessons of so-called lottery cases is that even justification for a true belief that falls short of complete certainty by a tiny but positive margin *may* not be enough for knowing it. In a fair lottery in which there are *N* tickets, my chances of winning are 1/*N*, which for large *N* is very small, and I may truly believe that I don't hold the winning ticket; but – and this is the key point – I don't know that my ticket won't win.

What do the lottery cases teach us? The fact that a mathematician's non-deductive justification for her belief in the true mathematical proposition *p* is extremely strong does not, in itself, show that she knows that *p*. Something more is needed, at least in general. Clearly, though, not all epistemic situations are similar to lottery cases, and one should not overgeneralize from them. Their moral is not that extremely strong justification can *never* transmute true belief into knowledge; only that it *sometimes* fails to do so.

Let's take stock. Suppose you think, like I do, that non-deductive justification can yield knowledge in mathematics. Then you have a job to do. It's not enough for you to simply point out that in many such cases, the evidence for *p* is overwhelmingly strong. You must produce reasons for thinking that, at least in some cases, it's strong enough for knowledge. The argument in this section that nothing in general epistemology rules it out is just one example of an argument you could give. There are many others.

7 Conclusion

The present chapter has looked at some of the more philosophical aspects of the role non-deductive evidence plays in mathematics. The role is an important one, and its contours have only started to be investigated in detail in recent decades. There is more philosophical work to be done to understand it better.

References

Armstrong DM (1973) Belief, truth, and knowledge. Cambridge University Press

Baker A (2007) Is there a problem of induction for mathematics? In: Leng M, Paseau A, Potter M (eds) Mathematical knowledge. Oxford University Press, pp 59–73

Baker A (2021) Counterpossibles in mathematical practice: the case of spoof perfect numbers. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer. online

Clark M (1963) Knowledge and grounds: a comment on Mr. Gettier's paper. Analysis 24:46–48
Echeverría J (1996) Empirical methods in mathematics. A case-study: Goldbach's Conjecture. In: G. Munévar (ed) Spanish Studies in the Philosophy of Science, pp. 19–55
Franklin J (2021a) Bayesian perspectives on mathematical practice. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer. online
Franklin J (2021b) 'Let no-one ignorant of geometry...': mathematical parallels for understanding the objectivity of ethics. J Value Inq 55:1–20
Frege G (1884/1974) Die Grundlagen der Arithmetik, transl. as Foundations of Arithmetic by J.L. Austin. Blackwell
Gettier E (1963) Is justified true belief knowledge? Analysis 23:121–123
Goldman A (1967) A causal theory of knowing. J Philos 64:357–372
Goldman A (1976) Discrimination and perceptual knowledge. J Philos 73:771–791
Haack S (1976) The justification of deduction. Mind 85:113–119
Harman G (1973) Thought. Princeton University Press
Lakatos I (1962) Infinite regress and foundations of mathematics. In: Worrall J, Currie G (eds) Mathematics, science and epistemology. Cambridge University Press, pp 3–23
Lange M (2022a) Inference to the best explanation is an important form of reasoning in mathematics. Math Intell 44:32–38
Lange M (2022b) Why is proof the only way to acquire mathematical knowledge? Australas J Philos. (forthcoming)
Lewis DK (1973) Counterfactuals. Blackwell
Maddy P (1988) Believing the axioms. J Symb Log 53(481–511):736–764
Mazur B (2014) Is it plausible? Math Intell 36:24–33
Nagel J (2014) Knowledge: a very short introduction. Oxford University Press
Nozick R (1981) Philosophical Explanations. Harvard University Press
Ochem P, Rao M (2012) Odd perfect numbers are greater than 10^{1500}. Math Comput 81:1869–1877
Paseau AC (2008) Justifying induction mathematically: strategies and functions. Logique et Analyse 203(2008):263–269
Paseau AC (2011) Proving induction. Australas J Log 10:1–17
Paseau AC (2015) Knowledge of mathematics without proof. Br J Philos Sci 66:775–799
Paseau AC (2021) Arithmetic, Enumerative Induction and Size Bias. Synthese 199: 9161–9184
Paseau AC (forthcoming) What is mathematics about? Oxford University Press
Paseau AC, Baker AR (2023) Indispensability. Cambridge University Press
Paseau AC, Wrigley W (2024) The Euclidean programme. Cambridge University Press
Pólya G (1956) Mathematics and plausible reasoning, volume 1: induction and analogy. Princeton University Press
Pólya G (1968) Mathematics and plausible reasoning, volume 2: patterns of plausible inference. Princeton University Press
Russell B (1907) The regressive method of discovering the premises of mathematics. In: Lackey D (ed) Essays in analysis. Braziller, pp 272–283
Stalnaker R (1968) A theory of conditionals. In: Rescher N (ed) Studies in logical theory. Blackwell, pp 98–112
Whitehead AN, Russell B (1910/1962) Principia mathematica to *56. Cambridge University Press
Williamson T (2000) Knowledge and its limits. Oxford University Press
Zagzebski L (1994) The inescapability of Gettier problems. Philos Q 44:65–73

Section XI

"Experimental" Mathematics

Jessica Carter

Experimental Mathematics in Mathematical Practice

Jessica Carter

Contents

1 Introduction 2420
2 Experimental Methods in Contemporary Analysis 2423
2.1 The Structure of Graph-Algebras 2424
2.2 The Order of λ_α 2425
2.3 Thompson Group and Calculating the Bounds of an Expression 2426
3 The Contributions 2427
References 2430

Abstract

This chapter presents an overview of the contributions to the section on Experimental Mathematics by focusing in particular on how they characterize the phenomenon of "experimental mathematics" and its origins. The second part presents two case studies illustrating how experimental mathematics is understood in contemporary analysis. The third section offers a systematic presentation of the contributions to the section.

Keywords

Experimental mathematics · Heuristics · Verification · Experimental methods in contemporary mathematics · Graph algebras · Thompson group

J. Carter (✉)
Department of Mathematics, Aarhus University, Aarhus C, Denmark
e-mail: jessica@css.au.dk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_121

1 Introduction

The label "experimental mathematics" is today associated with the recent trend of using computer-based methods in mathematical research. As such, it can still be understood broadly, ranging from the use of computational tools for heuristic purposes, e.g., pointing to new results or to visualize a situation for obtaining a better understanding, to a more radical interpretation where computations are regarded as evidence for results (see, for example, Zeilberger 1994). In this light, it is natural to ask how it is possible to characterize the phenomenon of "experimental mathematics." In this connection, one might wonder whether the so-called experiments done in mathematics bear any likeness to the experiments of the natural sciences. Both of these questions are addressed by the chapters in this section. Furthermore, we also find accounts of some of the recent events that led to the establishment of "experimental mathematics."

"Experimental mathematics" is interpreted in a variety of ways by the contributors to this section. A unifying theme is that the phenomenon is understood in a broad sense. For one thing, experiments are not only done by computational tools – they might include, for example, pencil and paper as well as visualizations. This is also noted in the aims of the journal *Experimental Mathematics*: "Many mathematical experiments these days are carried out on computers, but others are still the result of pencil-and-paper work, and there are other experimental techniques, like building physical models."[1]

In addition, the use of experiments in mathematics (in some interpretations of the term) is not a new phenomenon. Mathematicians have since long performed "experiments," for example, in the form of computations, trying out and rejecting ideas. Tradition, however, prescribes that mathematics is presented in the form of definitions, propositions, and proofs that hides the work that came before. To use a famous quote by Reuben Hersh: "the "front" of mathematics is mathematics in "finished" form, as it is presented to the public in classrooms, textbooks, and journals. The "back" would be mathematics as it appears among working mathematicians, in informal settings, told to another in an office behind closed door" (Hersh 1991, p. 128).

In the literature are a few attempts to characterize "experimental mathematics" inspired by how experiment is understood in science. van Bendegem (1998) assumes that "experiment" can be understood in the naïve sense of manipulating objects and "setting up processes in the 'real' world and with observing possible outcomes of these processes" (p. 172). He examines possible descriptions of what these processes could be in the context of mathematics such as "experiment is computation" or "experiment is real-world-experiments" but concludes that they fail for different reasons. The second description fails since it is only rarely used in contemporary mathematical practice. The second occurs more often (and seems to be what

[1] Taken from the webpage of *Experimental Mathematics*, see https://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=uexm20

mathematicians are mainly referring to), but it is not "experimental" as understood in the sciences. Baker (2008) partly disagrees with this conclusion and formulates a tentative description of the genre, as he refers to it. He first notes that "experiment" can be interpreted in two distinct ways. In the first sense, one might focus on the method of formulating and testing hypotheses as experimental. As noted above, this is frequently done in mathematics and is (usually) considered as part of the context of discovery. In the second sense, we might understand "experiment" as a form of inductive evidence for a hypothesis or a claim. Based on an analysis of different cases, Baker combines these two and suggests that experimental mathematics "involves calculating instances of some general hypothesis" (Baker 2008, p. 340).

Among the contributions are both broad interpretations, as well as specific characterizations of experimental mathematics. de Villiers and Jahnke find that experimental mathematics and its aims should be understood in a variety of ways in order to demonstrate how they can fruitfully be used when teaching mathematics. For this purpose, they take inspiration from the long list of definitions of experimental mathematics found in Borwein's "Exploratory Experimentation: Digitally-Assisted Discovery and Proof" (Borwein 2012, pp. 73–74):

- Discovering new relationships and conjecturing (looking for inductive patterns, generalizations, etc.)
- Visualizing math principles (viewing graphical or other visual representations of analytically derived results)
- Verification (obtaining certainty about the truth or validity of a statement or conjecture)
- Computing replacing lengthy hand derivations (use of computer algebra, data handling, and/or graphical representations)
- Global and heuristic refutation (disproving a false statement by generating a counterexample, as well as reformulating, refining, or polishing a true statement by means of local counterexamples)
- Understanding (gaining insight/intuition, and grasping the meaning of a proposition, concept, or definition or assisting in the discovery of a proof)

In contrast, Schlimm and Gonzàlez present a simple framework that is useful for their purpose: to show how experiments using a variety of tools such as doodling and visualization have played a role in the development of a result on how to divide a line segment in equal parts. They characterize an experiment in terms of a setup, a process, and an evaluation. Although their framework is simple, it is still flexible enough to contain aspects of all items in Borwein's list. In the chapter by Ferraro, analyzing the role of examples in the mathematics of Euler, we find experiments in the sense of verification. In this case, we learn how examples can be used as evidence for both results and proofs.

A second question in Baker's (2008) paper is whether the use of experimental methods in mathematics is philosophically interesting. It appears that it is. Sørensen's contribution lists four – partially overlapping – topics that have been discussed in recent writings in the philosophy of mathematics. They include

the question whether it is possible to derive formal results from inductive evidence and the problem of opacity of computer calculations. There is also a focus on experimental methods in mathematics in connection with heuristics, using in particular computers as explorative tools.

The computer proof of the four-color theorem (4CT) by Appel and Haken in the mid-1970s is commonly mentioned as a turning point for experimental mathematics. The proof of the 4CT was followed by discussions among mathematicians whether it could really be considered a valid proof. A later debate, after the proof had been accepted in mainstream mathematics, concerns whether the proof of the 4CT changes the conception of a proof. This debate was spaked by the influential paper "The Four-Color Theorem and Its Philosophical Significance" (Tymoczko 1979). Proofs are normally considered to be a priori (or our knowledge of proof is a priori), certain, and surveyable. Tymoczko finds that the proof of the 4CT apparently is convincing, but not surveyable. Furthermore, since computers are used to check a (large) number of instances, there is an element of empirical data that is drawn upon and so the proof becomes a posteriori.

The chapter by McEvoy presents Tymoczko's claims and takes us through the main issues and line of arguments of the ensuing debate. Sørensen's contribution further informs us about some of the debates on the role of "experimental" methods among mathematicians during the 1990s: Jaffe and Quinn's proposal to distinguish between theoretical and rigorous mathematics and responses published in the Bulletin of the AMS as well as the heated discussion between Zeilberger and Andrews in the Mathematical Intelligencer. Furthermore, Sørensen writes about the formation and consolidation of the journal of the field, *Experimental Mathematics*.

The story of the use of computational tools in mathematics is, of course, much older. The first known computational device is the abacus that was used for calculations in ancient Mesopotamia. Somewhat later, we find the idea of using mechanical tools to aid our reasoning, forming what Leibniz called a *calculus ratiocinator*. Avigad (2018) traces Leibniz's ideas to the thirteenth-century Franciscan monk from Mallorca, Ramon Llull. One of his great achievements was to invent various devices that are based on the three important assumptions (Avigad 2018, p. 683):

- We can represent concepts, assertions, or objects of thought with symbolic tokens.
- Compound concepts (or assertions or thoughts) can be obtained by forming combinations of more basic ones.
- Mechanical devices, even as simple as a series of concentric wheels, can be helpful in constructing and reasoning about such combinations.

The usefulness of using representations of various types is noted in several of the contributions. Starikova and Schlimm & Gonzàlez discuss the use of visual representations, or diagrams in the process of generating new results. Discussing the case of thought-experiments, Starikova's diagrams are both in the form of external and mental representations of mathematical situations. Schlimm & Gonzàlez refer to external representations as cognitive aids and explain some of their advantages.

These advantages are not limited to the use of diagrams but hold also for other types of representations such as calculations. One mentioned advantage is that external representations "allow us to keep a durable and shareable record of the reasoning process and they can be used to determine the boundaries of accepted inferences, such as the inferential steps within a regimented language or the restriction to a set of tools, e.g., compass and straightedge. Moreover, rather than just representing a mental reasoning process, external representations can themselves be the vehicles of epistemic inferences" (p. 6). One is reminded of the logician C.S. Peirce's characterization of mathematical reasoning as constructing, observing, and *experimenting upon* mathematical representations (which he refers to as diagrams)[2]:

> The truth, however, appears to be that all deductive reasoning, even simple syllogism, involves an element of observation; namely, deduction consists in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of *experimenting upon this image in the imagination*, and of observing the result so as to discover unnoticed and hidden relations among the parts. . . . As for algebra, the very idea of the art is that it presents formulae which can be manipulated, and that by observing the effects of such manipulation we find properties not to be otherwise discerned. (CP 3.363, italics added)

In the last part of the introduction, I present examples of experimental mathematics (as understood by the authors) illustrating different ways that computer-based methods play a role in contemporary analysis. At the end is a systematic summary of the six contributions to the section on Experimental Mathematics.

2 Experimental Methods in Contemporary Analysis

We have noted that "experimental" is understood in two distinct ways. Mathematics – as the sciences – uses a variety of methods and tools to "experiment" when to discover new relations. In the second sense, "experimental" methods, such as computations, sometimes play a justificatory role. In this section, I present results from contemporary mathematics illustrating the varied use of experiments using computational tools. The first example considers so-called graph-algebras and illustrates two different interpretations of experimental methods. The second example concerns the Thompson group and the question whether its subgroup F is amenable. The first example is a case of what McEvoy refers to as a computer-assisted proof. The considered result is verified by having a computer program compute all instances of a large number of cases. The second and third example are experimental in the sense that hypotheses are formulated and tested with the help of computational tools. The two last cases fit the characterization given by Eilers and Johansen in their *Introduction to Experimental Mathematics* (2017, p. 1):

[2] The quote is from Peirce's Collected Papers 1931–1960, volume III, paragraph 363. I refer to (Marietti 2010) and (Carter 2020) for explanations of Peirce's diagrammatic reasoning.

"the systematic investigation of concrete examples of a mathematical structure in the search for conjectures about its properties."

2.1 The Structure of Graph-Algebras

Graph-algebras are C^*-algebras generated from so-called directed graphs. The structure of graph-algebras can be investigated by studying particular maps, i.e., endo- and automorphisms, that can be defined on them. Two different experimental methods are illustrated in this case. The first constructs an algorithm that can be turned into a computer program and so performed successfully by a computer. A diagrammatic representation of an endomorphism leads to an algorithm that gives further information about the map. In simple cases, the calculations can be done by hand. If not, the algorithm can be performed by a computer. The second type of use is closer to an experiment in its original sense – as they are done in the experimental sciences: A problem concerning the automorphisms of a graph-algebra – assumed to be undecidable – is investigated by the use of computers. In the first stage, certain calculations give rise to a pattern. The pattern is further explored – and *experimentally* verified – by further calculations or by mathematical arguments.[3] I give some more details about the case below. More details can be found in "Visualizing Automorphisms of Graph-algebras" (Avery et al. 2018) and in some articles referred to there.[4]

A directed graph is formally defined as the four tuple, $E = (E^0, E^1, r, s)$ where E^0 consists of vertices, E^1 of edges, and r, s: $E^1 \rightarrow E^0$ are range and source functions of the edges. This means they tell for each edge which vertices it ends and starts at. When dealing with a particular directed graph, it is most often easier to present it as a diagram. Figure 1 shows an example.

Whenever there is given a directed graph, E, a construction can be performed so that it generates a C^*-algebra. I refer to (Avery et al. 2018) for details.

Given a particular graph, it is possible to define the length of paths between two vertices (counting the number of edges). See Fig. 1 for an example. Given two vertices u and v, one further considers the collection, $E^k_{u,v}$, of paths of a fixed length k between them. Finally, a permutation on the paths of $E^k_{u,v}$ is referred to as an end point fixing permutation. An end point fixing permutation induces an endomorphism on the generated C^*-algebra. If the permutation is named α, the endomorphism that it induces is denoted by λ_α.

Avery et al. deal only with finite graphs (i.e., finite number of vertices and edges).[5] This means that they are discrete and finite combinatorial objects. As a

[3] The fact that the result is most likely undecidable contradicts Baker's (2008) statement that mathematics need not rely on inductive methods since it is always possible that a formal, or deductive, proof can be found of a given proposition (p. 337).

[4] I thank W. Szymanski for conversations about these cases.

[5] There are further technical restrictions imposed on the graphs that are not relevant for this brief presentation.

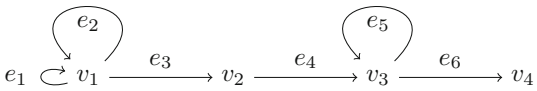


Fig. 1 A directed graph with four vertices and six edges. A path from v_1 to v_3 of length 3 is, for example, $v_1 \to v_1 \to v_2 \to v_3$.

contrast, when constructing a graph-algebra from a particular graph, one obtains an infinite, analytic object, whose study requires analysis. The idea in the study presented here is to study symmetries of the graph-algebras captured by actions of groups that can be expressed via the combinatorial properties of the directed graphs. In other words, certain "accessible and visible" symmetries of graphs are generated and studied.

The induced map, λ_α, can be shown to always be $1-1$, but it need not be onto. The considered results concern the properties of this map. The two questions that are investigated are:

1. Is λ_α onto (which would entail that it is an automorphism)?
2. If so, what is the order of λ_α?

For each particular permutation α, there exists an algorithm that will tell in finite time whether (1) holds or not. That is, thegraph-algebras solution of (1) can be found computationally.

2.2 The Order of λ_α

The answer λ_α to question (2) is likely to be undecidable, and so in this case there is no algorithmic solution. The problem has similarities to problems in automata theory that are undecidable, and it is assumed that the same holds for this problem. Therefore, an experimental approach has been taken.

The basic idea is to translate the endomorphism (which is an element of a group) into a formal language. The complexity of this language can be calculated, and this gives a particular natural number. In this way, each order of λ_α gives a natural number. If this number turns out to be zero for a particular order, k, then λ_α^k is the identity and it follows that λ_α has order k. Below is an illustration of this process when the sequence terminates quickly:

Order of λ_α	λ_α	λ_α^2	λ_α^3	...	$\lambda_\alpha^k = Id$
Corresponding number	c_1	c_1	c_3	...	0

The interesting, or challenging, part is when this process does not terminate quickly. (It should be noted that the calculations have been done by parallel computing and using Super Computers – they are huge.)

By making a number of calculations of orders of lambda for different permutations, a pattern arose. Either the sequence of natural numbers oscillated or it grew in an exponential way. Avery, Johansen, and Szymanski therefore made the following hypothesis or prediction: If the sequence oscillates, then it will eventually become zero – and so λ_α will have a finite order. If the sequence looks exponential, then the order is infinite.

These two hypotheses have been tested – and experimentally confirmed. The first case, when the sequence oscillates, has been verified by testing a number of cases – calculating even higher powers of λ_α. In each of the chosen examples, it turned out that the number eventually became zero. In the second case – where it is conjectured that the order is infinite – this procedure does not work. Instead proofs using analysis have been carried out. More precisely, a number of cases among the automorphisms that displayed an exponential pattern were selected and analyzed one by one using tools from analysis. For each of these cases, it was possible to prove that the order λ_α was infinite.

2.3 Thompson Group and Calculating the Bounds of an Expression

The Thompson group T is a famous example of a countable group that has a finite presentation. It can be generated by two elements C and D, which is reproduced in Fig. 2 below. A long-standing problem is whether one of its subgroups, the subgroup F, is amenable.[6] We consider a study by S. Haagerup, U. Haagerup, and M. Ramirez-Solano published in *Experimental Mathematics* where calculations performed by computer programs indicate that the group might not be amenable (Haagerup et al.

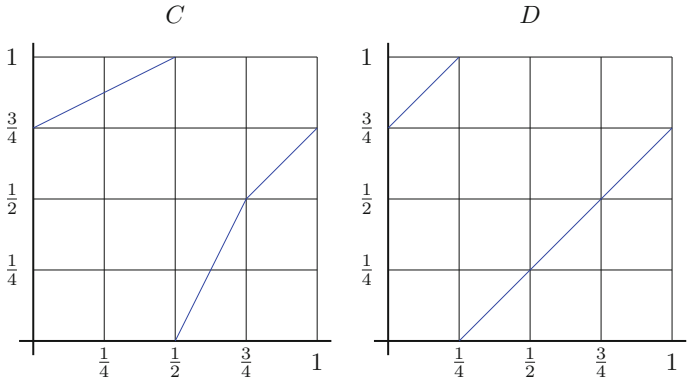


Fig. 2 Graphs of the generators C and D of the Thompson group

[6] The concept of an amenable group has been introduced by von Neumann in connection with his work on the Banach-Tarski paradox.

2021). In this case, the value of a certain expression composed by generators of the Thompson group is computed. The conclusion is based on the result of this computation as well as other indications. Before showing the expression, I present the definition of the Thompson group.

I refer to (Haagerup et al. 2021) for further details on the Thompson group and the subgroup F.

The Thompson group T is the group of cyclic order preserving homeomorphisms $f\colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ for which the following hold:

1. f and f^{-1} are piecewise linear with finitely many breakpoints.
2. All breakpoints lie in $\mathbb{Z}\left[\frac{1}{2}\right]/\mathbb{Z}$.
3. All slopes of f are in the set $2^{\mathbb{N}} = \{2^n | n \in \mathbb{Z}\}$.

The group T can be generated by the elements C and D shown in Fig. 2 and the following finite collection of relations[7]:

$$\begin{aligned} T = \langle C, D \mid \quad & C^3 = D^4 = (CD)^5 = e, \\ & \left[CDC, D^2(CDC)D^2\right] = \left[D^2(CDC)D^2, C\left(D^2CDCD^2\right)C^{-1}\right] = e\rangle \end{aligned} \tag{1}$$

S. Haagerup, U. Haagerup, and M. Ramirez-Solano use numerical methods to determine an approximate value of the expression $\frac{1}{12}\left\|\left(I + C + C^2\right) \times \left(I + D + D^2 + D^3\right)\right\|$. Formal results have shown that if the value of this expression is $2 + \sqrt{2}$ then F is amenable. The value that they find is

$$\frac{1}{12}\left\|\left(I + C + C^2\right)\left(I + D + D^2 + D^3\right)\right\| \approx 3.28$$

This result is combined with formal proofs that

$$\sqrt{2} + \sqrt{3} < \frac{1}{12}\left\|\left(I + C + C^2\right)\left(I + D + D^2 + D^3\right)\right\| \leq 2 + \sqrt{2}$$

Since the numerical calculation indicates that the value is less than $2 + \sqrt{2}$, they conclude (referring also to other results) that the group F *might* not be amenable.

3 The Contributions

The contribution by Henrik Kragh Sørensen, "The role of experiments in *Experimental Mathematics,*" presents the background for contemporary experimental mathematics. It takes us through some of the debates during the 1990s and considers the formation and consolidation of the journal *Experimental Mathematics*. The second part of the chapter considers in more detail various aspects of the journal

[7] The group commutator $[g, h]$ is given by the expression $[g, h] = ghg^{-1}h^{-1}$.

obtained using data-mining techniques. It finds that the journal publishes contributions by a large number of authors in a broad range of topics. The conclusion is that experimental mathematics (whatever it means) is not of marginal interest in the mathematical community. Finally, Sørensen examines whether there is a detectable difference in how mathematicians use the methods of "observation" and "experiment" in the articles published in *Experimental Mathematics*.

Mark McEvoy's "The New Four-Color Problem Revisited" runs through the main lines of the discussions that followed Tymoczko's (1979) seminal paper formulating the "new four-colour problem." The four-color theorem states that it is possible to color any map with four different colors so that no neighboring countries have the same color. This problem was originally formulated in the middle of the nineteenth century but only solved in the 1970s by Appel and Haken partially using a computer. Whether or not to accept this as a proof is the old four-color problem. The new problem is to decide whether the proof changes our conception of the properties of a proof.

McEvoy's contribution gives us a clear presentation of the main claims and lines of argument in the discussions that followed Tymoczko's article. The main topic concerns the a priori versus a posteriori nature of the proof (also taking into account the more recent developments in computer-assisted proofs, the development of proof checkers, and formal proof assistants) in the light of the claimed unsurveyability of the proof. Despite Tymoczko's claim (and others agreeing with him), McEvoy defends the view that computer-assisted proofs such as the proof of the four-color theorem can be a priori even though they may not be surveyable.

Giovanni Ferraro's "Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice" offers a glimpse of Euler's use of examples in his analysis of the infinite. In contemporary usage, examples are offered as illustrations of, e.g., concepts and theorems. In contrast, Ferraro shows that Euler uses examples for verification. Ferraro illustrates that examples in this sense play two different roles in Euler's mathematical writings. One role is as verification of a result by showing that the result holds for a particular, but generic example. This type of use is referred to as arbitrary exemplification. It is further possible to find cases where Euler uses what is referred to as inductive generalization, in the sense that a number of examples are used as evidence for a general result. Euler warns of using this procedure in general, however, giving examples of sums of infinite sequences that do not generalize.

In the case of arbitrary exemplification, a result is obtained without depending on essential properties of the considered example. In such cases, it is considered safe to generalize the obtained result: The procedure can be repeated in every instance. It occurs, for example, when Euler determines the differential of polynomials: There is no main difference between determining the differential of $y = x^3$ or the general case of $y = x^n$.

In the second sense, Euler used examples to verify proofs, to demonstrate that they are indeed valid and to convince others of the correctness of the procedures they involve. Ferraro argues that this should be seen as a way to demonstrate the soundness of the new calculus which at the time did not yet rely on a solid foundation. Euler demonstrates in multiple cases that different techniques give the same results to convince of the correctness and internal coherence of the calculus.

"Mathematical experiments on paper and computer" by Dirk Schlimm and Juan Fernàndez Gonzàlez takes us to the "back" of mathematics, discussing the different processes that have led to a result obtained by Gonzàlez. We learn about the different stages, from the first vague ideas to the generation of new concepts and finally the formulation of the final theorem and proof. Their chapter further illustrates how contemporary "experiments" can be done without computers – simply by a pencil and a notebook, or even on the frozen window of a bus. The considered result is a procedure to divide a line segment in any number of equal parts (and so also by a rational number) by only bisecting line segments.[8] Schlimm and Gonzàlez understand an experiment to consist of a setup, a process with an outcome, and an interpretation. The setup is defined as a "specification of an initial mathematical situation or configuration, such as a formula or algorithm with some initial parameters." It may, but need not, include certain expectations about the outcome of the experiment formulated in the form of a hypothesis. The hypothesis is evaluated by the outcome of the experiment. The process consists of a calculation or construction that can be done on paper, by a computer, or in the mind. Schlimm and Gonzàlez illustrate how their description of an experiment applies to the development of the considered theorem focusing on the variety of tools used, ranging from the initial doodles that sparked the first ideas and the ensuing explorations giving rise to different conjectures to the later systematic use of a computer program.

Irina Starikova's "Thought experiments in mathematics: from fiction to facts" consider the particular case of thought-experiments in mathematics. Starikova notes that philosophers have argued that thought experiments cannot exist in mathematics since experiments are a form of a posteriori reasoning whereas mathematics is a priori. Others have claimed that thought experiments are just a variation of mathematical reasoning that obviously depend on thought. In contrast, Starikova maintains that thought experiments do exist in mathematics and argues that they should be considered as an independent form of mathematical reasoning.

Starikova characterizes a mathematical thought-experiment as "a method that allows epistemic progress through experimental thinking, which extends the rules prescribed by a given practice." As such, thought-experiments play a role not just in (generating) proofs but also in the formation of new ideas and conjectures. While thought-experiments may draw on both sensory and physical experience, the mathematical ideas remain a priori.

The chapter is rich in examples illustrating their varied use of thought-experiments in mathematics. They range from simple examples such as the mutilated check-board to examples of how thought-examples play a role in contemporary mathematics. The latter include how to determine the diameter of a graph using sensory experience and the "zooming out" of Cayley graphs using our imagination.

Thought-experiments are sometimes – but need not be – accompanied by visual representations. Starikova further explains that thought-experiment is not just

[8] A previous well-known method to divide a line segment in equal parts depends on the stronger assumption that it is possible to draw parallel lines.

"experimenting in the mind"; the flexibility of our imagination is an important advantage of mathematical thought experiments. In this way, they extend both our deductive and visual intuition by introducing fictional elements. She points to the example of Hilbert's hotel where physical properties no longer hold, and that it is possible to remove physical properties or combine different mathematical properties, or even mathematical with physical properties.

Finally, having noticed that experimental methods permeate mathematics, Michael de Villiers & Niels Jahnke find it natural that they should be present also when teaching mathematics. As suggested by their title "The Intimate Interplay Between Experimentation and Deduction: Some Classroom Implications," they further emphasize the dual role between experiment and proof. By including experiments in the curriculum, they argue that it is possible to stimulate the natural curiosity of children. After all, mathematics is not just about reproducing solutions to known problems, but also discovering new relations and establishing by proofs that they hold. Experiments play a role in both of these contexts.

The chapter is a great source of inspiration for teachers who wish to include experimental practices in their teaching as it contains a large number of examples that can be (and have been) used for teaching purposes. It includes different aspects of experiments in mathematics, from the simple of guessing a result or checking the outcome of a calculation to the use of both physical and mathematical models.

References

Avery J, Johansen R, Szymanski W (2018) Visualizing automorphisms of graph algebras. Proc Edinb Math Soc 61(1):215–249

Avigad J (2018) Opinion: the mechanization of mathematics. Not Am Math Soc 65(6):681–690

Baker A (2008) Experimental mathematics. Erkenntnis 68:331–344

Borwein JM (2012) Exploratory experimentation: digitally-assisted discovery and proof. In: Hanna G, de Villiers M (eds) Proof and proving in mathematics education, New ICMI study series 15. Springer, New York, pp 69–96. https://doi.org/10.1007/978-94-007-2129-6_4

Carter J (2020) Logic of relations and diagrammatic reasoning: structuralist elements in the work of Charles Sanders Peirce. In: The prehistory of mathematical structuralism. Oxford University Press, New York, pp 241–272

Eilers S, Johansen R (2017) Introduction to experimental mathematics. Cambridge University Press, Cambridge

Haagerup S, Haagerup U, Ramirez-Solano M (2021) Computational explorations of the Thompson Group T for the amenability problem of F. Exp Math 30(1):105–126

Hersh R (1991) Mathematics has a front and a back. Synthese (Dordrecht) 88(2):127–133

Marietti S (2010) Observing signs. In: Moore ME (ed) New essays on Peirce's mathematical philosophy. Open Court, Chicago/La Salle, pp 147–167

Peirce CS (1931–1960) Collected Papers of Charles Sanders Peirce, Vol I–IV, Hartshorne C, Weiss P (eds), The Belknap Press of Harvard University Press, Cambridge

Tymoczko T (1979) The four-color problem and its philosophical significance. J Philos 76(2):57–83

Van Bendegem J-P (1998) What, if anything, is an experiment in mathematics? In: Anapolitanos D, Baltas A, Tsinorema S (eds) Philosophy and the many faces of science. Rowman & Littlefiel, London, pp 172–182

Zeilberger D (1994) Theorems for a price: tomorrow's semi-rigorous mathematical culture. Math Intell 16(4):11–18

The Role of Experiments in *Experimental Mathematics*

Henrik Kragh Sørensen

Contents

1 Introduction: Experiment, Observation, and Induction 2432
2 Background Discussions Among Practitioners in the Early 1990s 2435
3 Tasks of Computers and Experiments in Mathematical Practice 2436
4 Experimental Mathematics Characterized by Practice 2438
4.1 Zeilberger and Andrews on Warrant from Unlikely Circumstances 2439
4.2 Epstein and Levy on Warrant from Systematic Experimentation 2440
4.3 Borwein Et al. on Warrant from Crucial Experiments 2440
5 Empirical Studies of *Experimental Mathematics* 2442
5.1 The Journal and Its Authors 2444
5.2 A Cross-Field Enterprise 2445
5.3 Experiments and Inductions: Quantitative Analyses 2447
5.4 Experiment and Observation: Linguistic Analyses 2448
6 Conclusions and Perspectives 2452
7 Cross-References 2455
References 2455

Abstract

This chapter is devoted to discussions about the nature of experiments in mathematics during the formative period of *experimental mathematics* and to a quantitative and qualitative analysis of the journal *Experimental Mathematics* founded in 1992. Particular attention is paid to interpret the roles of experiment and observation in the mathematical practices surrounding *Experimental Mathematics*.

H. K. Sørensen (✉)
Section for History and Philosophy of Science, Department of Science Education, University of Copenhagen, København Ø, Denmark
e-mail: henrik.kragh@ind.ku.dk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_113

Keywords

Experimental mathematics · Observation · Numerical induction · The journal *Experimental Mathematics* · Mathematical debates in 1990s · Borwein · Zeiberger · Epstein · Digital philosophy of mathematical practice

1 Introduction: Experiment, Observation, and Induction

In the early 1990s, debates over new practices arose in the mathematical research community, sparked by enthusiasm and skepticism about the role of computers in research and the potential turn toward a more *experimental* epistemology of mathematics. Indeed, the theme was not entirely new (for the history of experiments in mathematics, see, e.g., Hazewinkel 1985; Goldstein 2007; De Mol 2014), but the focus was shifting during the 1990s.

These debates roared in many of the broad mathematical journals that serve as professional discussion fora, like the *Notices of the AMS*, *the Mathematical Intelligencer*, and the *Bulletin of the AMS*; but they also had more direct consequences for publishing research mathematicians. Before the end of the century, a new mathematical journal devoted to *experimental mathematics* had been established and consolidated, a number of *schools* (or perhaps better, *cultures*) had formed which might disagree on the philosophical issues but still professed themselves to be *experimental*.

A large proportion of the philosophical interest in *experimental mathematics* has centered on the possibility of deriving formal (deductive) knowledge from inductive evidence (Baker 2008, 2007; Van Bendegem 1998), or whether the computer and the ensuing unsurveyability (opacity) of computer-assisted proofs might make mathematical knowledge dependent on non-a priori knowledge (Van Bendegem 1993; Johansen and Misfeldt 2016; McEvoy 2008, 2013). These analyses have produced important corrections to standard views of mathematical proofs, such that today demands of *formalizability* or *explanation* of proofs require consideration of the role of externalizing epistemic steps of huge proofs. These are some philosophical results derived from the debates, but when practitioners disputed the roles of computers and experiments, they often had more concrete philosophical goals in mind.

In philosophically analyzing the emergence of this experimental approach to mathematical research, practitioners and philosophers have turned to many different philosophical frameworks, from the quasi-empiricism suggested by Imre Lakatos (Lakatos 1976), over broadly generalized accounts of experiments in science (Medawar 1979), to viewing the mathematical experiment as neither heuristic nor justificatory but *exploratory* in the tradition of philosophers of science (Bailey and Borwein 2011; Franklin 2005; Sørensen 2010; Steinle 2016), or discussing whether *experimental mathematics* is a field of its own or a *culture* which can transcend the ordinary mathematical subject fields (Sørensen 2016).

As a means of mapping these debates from perspectives in the philosophy of mathematical practice, Fig. 1 provides a conceptual mapping of discussions. The

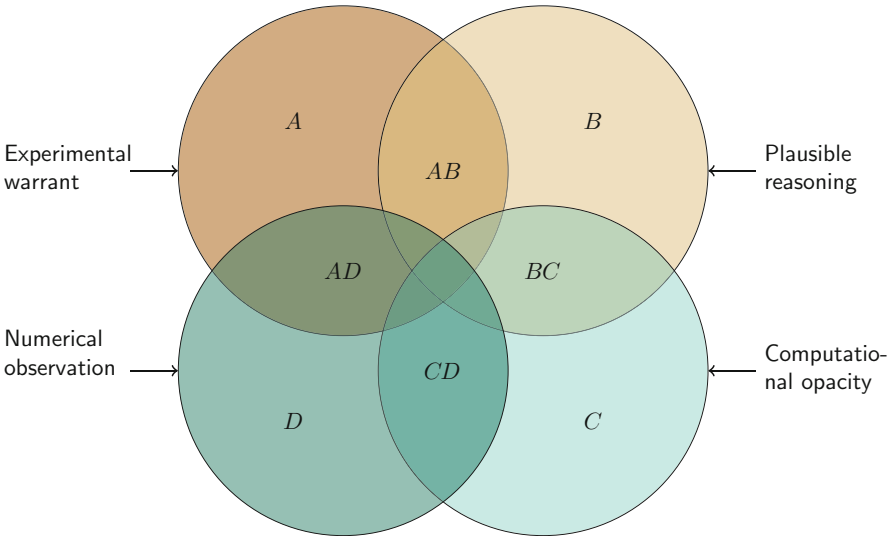


Fig. 1 Overview of the discussions neighboring *experimental mathematics* in the late twentieth century. In the following, attention will be focused on the intersection (*AD*) of *Experimental warrant* (*A*) and *Numerical observation* (*D*)

circles each signify a major topic of debate and philosophical analysis, which are introduced below. But particular attention should be paid to the overlaps between the circles, where discussions overlap and influence each other. Two of these overlaps (*AB* and *AD*) will be important in analyzing the emergence of a flagship publication of *experimental mathematics*: the journal *Experimental Mathematics* (other journals such as the *ACM Journal of Experimental Algorithms* also played important roles in shaping *experimental mathematics*, especially from a computer science perspective, but here the focus will be on *Experimental Mathematics*).

- The overlap *AD* between *Experimental warrant* and *Numerical observation* contains discussions about the extent to which experiment, observation, and induction might provide epistemic warrants in mathematics (for more on epistemic warrants of experiments, see also Lupacchini and Corsi (2008), Lolli (2008), and Sørensen et al. (2024)). One such discussion was the aptly titled paper "Strange Series and High Precision Fraud" in which Jon Borwein and his brother, Peter, show that even surprising numerical agreement for the first n instances of a complex formula studied experimentally and with the aid of computers does not guarantee the general validity of a result (Borwein and Borwein 1992): even high-precision verification is no substitute for formal proof. Consequently, the inductive support that calculating more and more digits of π or searching for bigger and bigger Mersenne primes has led mathematicians and philosophers to reaffirm that *numerical observation does not* belong in the epistemic toolbox of mathematicians.

- The overlap *AB* between *Experimental warrant* and *Plausible reasoning* points to the many examples where experimental results have been useful in the process of *conjecturing* and exploring plausible connections, often in search of counterexamples. Going back to Polya (1968, Vol. 1, Chap. 4–7), number theory had been one of the core examples of noncomputerized plausible reasoning and heuristics in mathematics. The discussion about drawing epistemic warrants from plausible reasoning is thus another recurring one, which typically is settled by pointing to a fundamental difference between heuristics and justification.
- The final circle, *C* indicating *Computational opacity*, is among the typical reasons against computers in gaining mathematical knowledge as *unsurveyable* computer-assisted proofs challenge the epistemic autonomy of the mathematician. This idea claims that mathematical knowledge depends on an autonomous epistemic agent who, alone and without external cognitive resources, comes to know a mathematical result by verifying it step by step. The view was succinctly expressed by Thomas Tymoczko as the *Simon Says Analogy* where a Martian (Simon) comes to pose a deep epistemic challenge to mathematics since he is correct in all he says, but humans have no means of inspecting his reasoning. Arguments about the necessity of epistemic autonomy is currently being challenged by work into the social epistemology of mathematics (Andersen et al. 2021a).

The relationship between induction through observation and experimentation has been a major recurring theme in the philosophy of science. In the nonformal sciences, empirical support is the sine non qua, and the distinction is often drawn from the stipulation that experimental knowledge is obtained through intervention, whereas observations leave the observed system untouched. If this distinction is transferred to mathematics, it requires expounding and interpreting the key notions:

- In mathematics, observation could be interpreted as "empirical support gained from (numerical) data" (Baker 2007, 2009, 2010).
- In mathematics, experiment could be interpreted as "in silico exploratory manipulation of mathematical objects" where the effects of changes can quickly be seen (Giardino 2022; Misfeldt et al. 2015; Sørensen 2010).

With these interpretations, the observations would connect with *D* (Numerical observation), and exploratory experimentation would influence the kinds of warrant to be drawn from experiments (*A*, *Experimental warrant*).

Therefore, the following sections are designed as follows: First, I provide some background on the scene for discussions about experiments and computers in mathematics around the turn of the century. Then I go into more detail about the different views about the experimental mathematics during that period, drawing out instances of both observations and experiments. Then I turn the focus to the founding and consolidation of *Experimental Mathematics* from its inception in 1992 to recent issues, combining quantitative analyses with short qualitative examples. Finally, before summarizing and add perspectives, I again use the combination

of quantitative and qualitative approaches to address the relations between observation and experiment in mathematical practice.

2 Background Discussions Among Practitioners in the Early 1990s

In 1992 an important step was taken in the consolidation of a new movement with mathematical research, when the journal *Experimental Mathematics* was founded. At least since the new proof of the *four-color conjecture* in the mid-1970s, mathematicians had to confront the introduction of a new epistemic tool in their arsenal: that of heavily computer-dependent practices of investigation and proof.

Although debate over the proof of *four-color theorem* – as it had then become – was initially engaging many in the community, the ensuing philosophical discussion focused mainly on one point, centered around the *unsurveyability* of large computer calculations (Tymoczko 1979). This in turn led to discussions whether mathematics was becoming a *quasi-empirical* and *a posteriori* science. Some of the concerns about quasi-empiricism went back to the controversy of Lakatos' philosophy of mathematics, indirectly drawing on Lakatos (Lakatos 1986) advancing a claim that is now a cornerstone of much philosophy of mathematical practice:

> The history of quasi-empirical theories is a history of daring speculations and dramatic refutations. (Lakatos 1986, p. 43)

However, to a large extent, these discussions shifted from the mathematical practice into the philosophical domain of epistemology.

Thus, when the discussion about the role of computers reemerged in the mathematical community in the later 1980s, it took a slightly different form. Desktop computers supplemented the mainframes which had proved to be both a bit inaccessible for the noncomputer expert mathematician and did not allow for the quick interaction that would be useful in *exploring* unknown territory.

In the first issue of *Experimental Mathematics*, the editors David Epstein and Silvio Levy published a "letter to the author," which remains on the article website in almost the same form and content. In that letter, the editors presented their "belief that theory and experiment feed on each other" (Epstein et al. 1992, p. 1). As a consequence of this belief, they explained it was deplorable that the means by which mathematicians *discovered* new results were mostly not shared.

Thereby, Epstein and Levy touched upon ideas that were also being presented in Davis and Hersh (1981), most vividly when Reuben Hersh introduced the idea that mathematics has a back and a front stage (Hersh 1991). The front stage consists of the logical exposition of proofs, but the back, he suggested, was full of the *dirty practice* that mathematicians engage with to arrive at the results to be presented – exactly the part of the practice that the mathematical community was lacking an outlet for, according to Epstein and Levy.

When Epstein and Levy introduced the new journal, they also suggested that mathematics might stand to gain from sharing their ideas at an earlier stage in the research process (Epstein et al. 1992, p. 1). This idea was, at the time, increasingly being addressed by the newly established `arxiv` (https://www.arxiv.org), where the first few mathematics papers were uploaded in 1992. And the same idea was addressed from a more skeptical point on the pages of the *Notices of the AMS* soon thereafter (Quinn 1995), expressing concern about the standards of mathematical articles if access to publication or (online) sharing became too easy.

Frank Quinn had already been debating the dual nature of mathematics – proof and exploration – in the years before, when his paper with Arthur Jaffe sparked another debate, this time in the *Bulletin of the AMS*. Jaffe, who would become president of the American Mathematical Society (AMS) later in the decade, and Quinn suggested a "synthesis of mathematics and theoretical physics" (Jaffe and Quinn 1993), in which mathematics would stand to gain from closer interaction between its "theoretical" and "experimental" sides (see Stöltzner 2005). To them, the theoretical part was responsible for exploring or conjecturing, which one could see as the backstage of Hersh. Thus, confusingly in the present context but in line with their view of physics, Jaffe and Quinn meant the "experimental" part of mathematics to be providing critical experiments and proofs. This entire suggestion was taken up by other prominent mathematicians, including Michael Atiyah who – after agreeing to much of the argument – addressed how the status of conjectures should not be diminished but recognized as a creditable part of mathematics (Atiyah et al. 1994).

Thus, around the time of the creation of the journal *Experimental Mathematics*, discussions were again running among prominent mathematicians about the relevance of conjectures and (more) open exploration and experimentation in mathematics.

3 Tasks of Computers and Experiments in Mathematical Practice

Having thus established a venue for publication, a next keystone in the consolidation of experimental mathematics can be said to be the exemplary books that came out of a group centered at Simon Fraser University in Vancouver, Canada. At Rutgers University in the USA, Doron Zeilberger was not shying away from provocation and controversy, publishing with his computer as a coauthor (e.g., Ekhad and Zeilberger 1996) – or even having Ekhad appear as the single author in papers on his webpage – and suggesting that mathematical knowledge comes with a price tag of the computational power required for it (Zeilberger 1994). Yet, the group forming around Jon Borwein and David Bailey set out another approach, showing the relevance of computers and experimentation by providing both results from their approach, as Zeilberger was doing, and philosophizing about it, as Zeilberger was also doing, but coating it in a form more amenable to their colleagues, and students and future mathematicians.

Around 1992, as the field and the journal were still finding footings, the precise meaning and expectations of *experimental mathematics* were still clouded. It seems mostly to have been a juxtaposition of words rather than a clearly defined concept. Part of this uncertainty had to do with how to classify the new content: was it a field or even a discipline of its own, were it a tool for other fields, was it comparable to parts of physics, or was it rather a community and a culture of mathematical practice such as has been suggested more recently (Sørensen 2016).

Indeed, setting out the scope of the journal in the very first issue, Epstein and Levy had given a first kind of definition of *experimental mathematics* – at least as it pertained to the journal under their editorship (and beyond). Highlighted in an introduction to the author, Epstein and Levy called attention to the following topics which might be publishable in the journal (Epstein et al. 1992):

1. Experiments
2. New theorems
3. Algorithms
4. Practical issues
5. Computer programs and software
6. Surveys
7. Essays, humor, polemic, and correspondence

Soon, however, the journal became "more professional" in the sense that the first three points came to vastly dominate its content.

About the relation between the first two points, as has already been shown, the editors saw experiments as an important part of mathematical practice, and they suggested the "essential ingredients of a paper published" in *Experimental Mathematics* are twofold: "some experimental aspect, and relevance to mathematics proper" (Epstein et al. 1992). Here the contrast between the new field and "mathematics proper" stands out. To Epstein and Levy, the role of experiments was to lead to insights that could be proved and thus be of interest to colleagues within "mathematics proper." In distinguishing experimental and applied mathematics, the editors reiterated that "focus is on work that will have a theoretical impact and contribute to the development of mathematical ideas" (Epstein et al. 1992, p. 1). Making it more concrete, Epstein and Levy suggested types of themes that would be relevant to the journal (closely adapted from Epstein et al. 1992, p. 2):

- Experiments that give rise to new theorems or new conjectures
- Experiments that lend support to existing conjectures
- Experiments that point to areas worthy of investigation
- New theorems proved with help of experimental results

In general, the letter to authors emphasized the "acceptability" of its results to the broader mathematical community: new theorems should be accompanied by "formal proofs" (which must mean rigorous proofs); computer experiments should be

reported so as to be replicable; and "results of computer experiments should be presented in such a way as to be graspable by humans" (Epstein et al. 1992, p. 2).

Epstein and Levy's conceptualization of the roles of experiments resonate with what leading experimental mathematicians Borwein and Bailey presented in (Borwein and Bailey 2004) and elaborated upon in Bailey et al. (Bailey et al. 2007), giving examples or short explanations. Borwein and Bailey presented eight roles for computation as part of their curiously named "experimental mathology" (Bailey et al. 2007, p. 11):

1. "Gaining insight and intuition – or just knowledge" from fact-gathering and aggregating data
2. "Discovering new facts, patterns, and relationships" from identifying uses of computation-intensive but well-known tools such as generating functions which might show relations even from sparse data
3. "Graphing to expose mathematical facts, structures, or principles" whenever techniques of visualization can be found, thus drawing on other perception systems
4. "Rigorously testing and especially falsifying conjectures," which draws on observations to either gain additional empirical support for a conjecture or refute it (see below)
5. "Exploring a possible result to see if it merits formal proof," which as a tool for strategic planning of proofs can focus human cognitive resources on proof sketches that pass preliminary exploration
6. "Suggesting approaches for formal proof," which complements the previous point of planning proof sketches
7. "Computing replacing lengthy hand derivations," which might have been an original point of using computers, but which can also be seen as outsourcing part of a proof sketch
8. "Confirming analytically derived results," which complements the search for refutations of conjectures, which can be settled by a single counterexample, and empirical support for conjectures, which do not (usually) provide epistemic warrant (see below) (Bailey et al. 2007, quotes from p. 13–15).

Previously, I have analyzed these eight roles as spanning heuristics (pt. 1–2), concept and result formation (pt. 3–4), and justification (pt. 5–8) (Sørensen 2016, p. 146). But in the following I will focus more on the role of experimentation and observation.

4 Experimental Mathematics Characterized by Practice

In 1994–1996, some of the early proponents of *experimental mathematics* had an exchange of views in general mathematical journals (*the Mathematical Intelligencer* and *Notices of the AMS*). Zeilberger, Borwein and Bailey, and Epstein and Levy all had something to say about the nature of *experimental mathematics*, which closely

reflect views discussed above. This debate ran at the same time as the Jaffe-Quinn debate, and they were closely related; yet, this exchange dealt more directly with the definition and nature of *experimental mathematics*.

4.1 Zeilberger and Andrews on Warrant from Unlikely Circumstances

This debate was initiated by Zeilberger's provocative opinion piece in the *Mathematical Intelligencer* entitled "Theorems for a Price" (Zeilberger 1994). In it, Zeilberger briefly presented some complicated-looking identities, which he claimed were "trivial" in *Wilf-Zeilberger theory (WZ)* (Petkovšek et al. 1996) once the use of computers had been properly accepted. Among the claims that Zeilberger made was that in the "day after tomorrow," theorems would come at a price and that "almost certainty" could be achieved for "an (epsilon) of the cost." The central step in his argument was to use Wilf-Zeilberger theory to transform identities into overdetermined systems of numerical ones, and then claim the solvability of such overdetermined systems as a "very good indication that the identity is indeed true":

> As absolute truth becomes more and more expensive, we would sooner or later come to grips with the fact that few nontrivial results could be known with old-fashioned certainty. (Zeilberger 1994, p. 14)

Zeilberger was knowingly provocative, and one of the first responses came from one of his collaborators, George Andrews. Aiming directly at the aspect of knowing *which* of the infinitely many identities are relevant, Andrews attached the autonomy of computer-discovered proofs:

> They [(Karl) Dilcher and the Borweins (Jon and Peter)] provide in the remainder of *BBD* absolute certainty about what is going on, and they provide concomitantly great insight and, dare I say it, beauty. Their paper is an almost perfect example of the computer aiding crucially in the discovery of facts but not in their proof. (Andrews 1994, p. 17)

In Zeilberger's account, Andrews not only missed the relevance of computer-discovered identities but also the more general insights usually gained from new results. To him, an identity might be a fact, but it does not (in itself) constitute mathematical knowledge – and herein lies the gist of the controversy, I think.

Zeilberger's vision of the "day after tomorrow" – with the hype and controversy so often shrouding predictions about computers and AI – was then made concrete by Andrews:

> Until Zeilberger can provide identities which are (1) discovered by his computer, (2) important to some mathematical work external to pure identity tracking, and (3) too complicated to allow an actual proof using his algorithm, then he has produced exactly no evidence that his Brave New World is on its way. (Andrews 1994, p. 17)

4.2 Epstein and Levy on Warrant from Systematic Experimentation

The next stage in the debate came when the editors of *Experimental Mathematics* tried to bring the epistemology of experimental mathematics on a firmer grounding. Calling upon Carl Friedrich Gauss, as they often have done, Epstein and Levy called for "systematic experimentation" as a way to discover new insights (David Epstein and Silvio Levy 1995, p. 670). And they advanced some of the arguments from the inauguration of *Experimental Mathematics*: Such processes should be made public, also to encourage sharing and using expertise in the community, an elaboration of the argument that tied it in with the Jaffe-Quinn debate.

> Note that we do value proofs: experimentally inspired results that can be proved are more desirable than conjectural ones. However, we do publish significant conjectures or explorations in the hope of inspiring other, perhaps better-equipped researchers to carry on the investigation. The objective of Experimental Mathematics is to play a role in the discovery of formal proofs, not to displace them. (David Epstein and Silvio Levy 1995, p. 671)

Addressing the role of computers, Epstein and Levy argued that "the introduction of the computer into mathematics has made more of a quantitative than a qualitative difference" (David Epstein and Silvio Levy 1995, p. 671). Using computers allows mathematicians to visualize and compute properties of, e.g., three-dimensional manifolds in the line of Bill Thurston, which have been published in traditional prominent journals (David Epstein and Silvio Levy 1995, p. 672).

Finally addressing the epistemic opacity of computer results, they again voiced an argument that have since become quite standard in the social epistemology of mathematics: "The danger of taking the results of a computer program on trust are not that different from taking a traditional result on trust" (David Epstein and Silvio Levy 1995, p. 672).

Thus, the editors of *Experimental Mathematics* made three points about *experimental mathematics* and their *Experimental Mathematics*:

1. Experimental mathematics is not (just) about looking for shiny objects but has a methodology of *systematic experimentation*.
2. Experimentally derived results can be formulated into traditional proofs that can be published in ordinary journals.
3. The introduction of computers is not quite the fundamental game-changer as some commentators had come to believe.

4.3 Borwein Et al. on Warrant from Crucial Experiments

The final piece of the debate to be included here is another attempt at providing philosophical grounding for *experimental mathematics*, this time coming out of the Borwein-Bailey group, whose purpose was to "[make] sense of experimental

mathematics" (Borwein et al. 1996). Framed as discussions about "discovery and verification" and "experiment and theory," they explicitly tied in with Zeilberger's piece but also, again, with the Jaffe-Quinn debate.

Some of the central points of this opinion piece were:

1. Experimentation (open ended) allows for serendipity.
2. As criticism of Zeilberger, they followed Andrews in emphasizing that mathematics "is all about insight" – not knowledge, individual facts, or computations.
3. The distinction of four different types of experiments introduced by the biologist and Nobel laureate Peter Medawar, who differentiated between
   (a) Kantian experiments which seek out the consequences of changing assumptions (or axioms)
   (b) Baconian experiments which try things out without any ulterior motive or direction
   (c) Aristotelian experiments which serve as demonstrations of already established facts
   (d) Galilean experiments which discriminate between possibilities, which provide confidence, also called *crucial experiments* (Borwein 2008; Medawar 1979).

Returning to their analyses of the roles of computers in mathematics, these points make it interesting to focus on the roles of data generation from uses of computers and experiments. On this issue, Borwein and Bailey and their coauthors also presented ideas in their books in the early 2000s, some of which are presented in a flowchart (Bailey et al. 2007, p. 7) showing the role of experiments in scientific practices as adapted to mathematical experiments. The flowchart is not thoroughly explained, but the following provides an interpretation useful in the subsequent analysis.

In their circular flowchart, four experiments are shown, each informed by "empirical knowledge": The first experiment has the form of observation and provides unstructured data. This leads to a second experiment in the form of intuition, which provides suppositions and hunches. Further, a third experiment has the form of scrutiny, resulting in questioning of the assumptions. Finally, a fourth experiment leads to results, which are also connected to formalism in the form of axioms and models through proofs and deductions. All experiments are informed by empirical knowledge and previous experiments. The results can either have the form of confirmation, giving "(almost) certainty about the validity of a conjecture" with a nod to Zeilberger, uncertainty in which no decision about a conjecture is reached, or refutation of a conjecture. However, the formation of the conjecture mentioned as the key result is nowhere included in the graph. Thus, one is led to interpret experiments 1 as connected to fact-gathering, experiment 2 as part of conjecture formation, experiment 3 as exploration of the conjecture, and experiment 4 as connecting experimental and deductive information about the conjecture.

As has been illustrated, practicing mathematicians were simultaneously engaging in advancing research in *experimental mathematics* and providing their approach

with a legitimization and a position within the mathematical community. One of the important consolidating factors for *experimental mathematics* was – as has been claimed above – the founding of the journal *Experimental Mathematics* in 1992. In the following, a combination of quantitative analyses and qualitative examples will provide a firmer grasp of what that journal intended and has achieved.

5 Empirical Studies of *Experimental Mathematics*

Over the past decades, philosophy of mathematical practice has come to establish itself on the spectrum of philosophy of science, drawing upon methods from a variety of disciplines including philosophy of mathematics, mathematics education, and sociology to answer philosophical questions about mathematical research practices (see Carter 2019; Rittberg 2019). These empirically informed studies of mathematical practices (or cultures as they are sometimes referred to) (see Löwe 2016; Kant et al. 2021) have drawn upon empirical methods such as interviews, observations, or questionnaires (see Inglis and Aberdein 2014; Misfeldt and Johansen 2015; Aberdein and Inglis 2019) and case studies (for a discussion, see Rittberg and Van Kerkhove 2019). These methods each shed important, new light on otherwise intractable aspects of mathematics as practiced by actual mathematicians. With the growth in available methods, the field is now at a stage, where the different approaches also should be capable of triangulation (see Löwe and Van Kerkhove 2019).

The present chapter uses another empirical method to study mathematical practice: that of corpus analysis. Drawing upon large corpora of mathematical texts combined with methods from statistics and machine learning, we can begin to access the vast *variety* of mathematical works, many of which are not singled out for case studies. In this vein, the digital study of mathematical practices allows us to begin to appreciate what is sometimes referred to as the *great unread*: We can begin to study dissemination of mathematical ideas, social aspects of mathematical practice, or the structure of mathematical corpora anew, this time grounded in large datasets and a new combination of quantitative and qualitative methods (for an exposition of this methodology, see Sørensen and Johansen 2024).

To get an impression and allow for analyses of the content of the journal *Experimental Mathematics*, I will use methods from this empirically grounded philosophy of mathematical practice based on large corpora of text. In the present case, I will work my way from general, summary statistics toward closer analyses. This combination of different tools then allows for some closer, but short, qualitative analyses, picked out by quantitative methods, thus giving them some empirical reasons for focus without claiming any kind of statistical representativity (Sørensen and Johansen 2024).

Thus, the method will be using bibliometric information about publications in the journal and linguistic analyses focusing on the information contained in titles and abstracts of individual papers. These methods allow us to address some of the ambitions of Epstein and Levy in founding the journal and some of the demarcations

of *experimental mathematics* that the journal has offered. Thus, I will be presenting analyses to answer questions such as:

- Who published in the journal? Which mathematical fields were represented in the journal?
- How was the experimental component of the journal framed?

Some of these questions arise from the background analyses in the previous sections, whereas others are more serendipitous in that they arose in the process of gathering and processing empirical data. For instance, it turned out that there might be something interesting to say from the data about the editors' goal of early sharing of conjectures and results.

Obviously, building quantitative analyses on short extracts such as titles and abstracts raises some methodological concerns: It is not immediately clear that abstracts reflect neither author attitudes nor philosophically interesting nuances of papers. To the first point, as argued in Hyland (2009), abstracts serve important roles in communication, in particular in framing and hedging results, so it is to be assumed that authors *and* editors will have invested in placing the right pieces of information in the very concise format of an abstract. This would also cover objections of insufficient language abilities of authors as a very large proportion of authors are native English speaking, and the editors invest effort in publishing a journal of high quality and interest to the community.

To the second point, the distinction between experiments and observations that interest me here may be an uninteresting distinction to mathematicians – if it is even recognized as a distinction to make. Practicing mathematicians may simply take the two to be synonymous, even without reflecting any ontological stances. Indeed, there are instances in the corpus where the two words occur in ways where they seem to be synonyms, but to address those, it is first required that we build a (quantitative) overview of the corpus.

The corpus from which the above questions are to be analyzed and answered consists of all research publications in the *Experimental Mathematics* between its beginning in 1992 (vol. 1, issue 1) and the last issue of 2022 (vol. 31, issue 4), a total of 124 issues containing 1175 articles with all the relevant metadata available. This corpus has an intermediate size for analyses: It is perhaps too big for human-only analysis, yet too small to draw definitive statistically significant conclusions about general representability. Thus, I am aiming for qualitative insights that can be backed by quantitative observations and visualizations.

The corpus was built by fetching the unique digital object identifiers (DOIs) of publications from the journal's webpage, and combining it with metadata such as title, authors, and abstract from the webpage of the individual article and metadata about the subject classification retrieved from the public API of *zbMATH* provided by the European Mathematical Society (EMS).

5.1 The Journal and Its Authors

The very first issue of *Experimental Mathematics* was published in 1992 and it contained eight research articles together with the information from its editors Epstein and Levy presented and discussed above. About a decade, Epstein and Levy could be proud of what might "well be the best issue" of the journal to date. Its volume had grown by almost 50% – from 420 pages to 640 pages – to make room for a large backlog (Epstein and Levy 2000). The evolution of the size of the journal is shown in Fig. 2 from which the increase in the *number* of papers during the first decade and the subsequent stabilization is clearly shown. This plateau after an early growth might indicate that the backlog has gradually been exhausted and, perhaps, that the acceptance criteria might have become higher. The papers in the very first issue stand out for three, interrelated, reasons: Some of their authors had already established great mathematical fame, the papers came from a great variety of mathematical fields, and there was little focus on the *experimental* aspects of the papers.

Among the authors in the very first issue, one finds such prominent names in contemporary mathematics as Barry Mazur and fields medalist John Milnor (Mazur 1992; Milnor 1992), whose main contributions to mathematics do not fall within

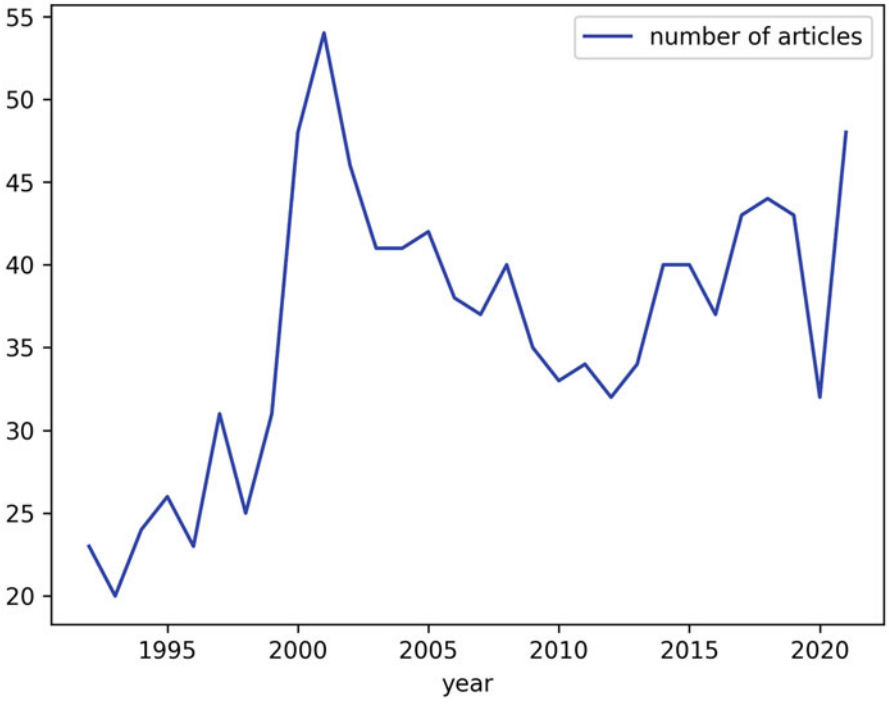


Fig. 2 Number of articles published in *Experimental Mathematics* 1992–2021. Publication metadata are extracted from the *Experimental Mathematics* journal webpage

Table 1 List of authors with five or more publications in *Experimental Mathematics*. Author names are extracted from the metadata and disambiguation has been done by hand

Place author	First	Latest	Number
1 Jonathan M. Borwein	1994	2018	17
2 David H. Bailey	1994	2018	12
3 Michael Pohst	1992	2011	10
4 Richard E. Crandall	1994	2010	9
5 Robert S. Strichartz	1992	2022	9
6 Igor E. Shparlinski	2008	2019	8
7 Richard Evan Schwartz	2001	2022	7
8 Avner ash	1992	2022	7
9 István Gaál	2001	2021	7
10 Benjamin A. Burton	2004	2022	6
11 John Milnor	1992	2007	6
12 David ford	1992	1998	5
13 Jesús Guillera	2003	2012	5
14 Neil Dummigan	2001	2021	5
15 Derek F. Holt	1992	2013	5
16 Kevin G. Hare	2004	2022	5
17 George Havas	1998	2007	5
18 Gert Almkvist	1998	2012	5

experimental mathematics. Whereas Mazur's paper addressed conjectures and computational aspects, the paper by Milnor was more of a summary of previous research. Such contributions were no doubt partly influential in gaining immediate recognition of the journal within the broader discipline.

Other mathematicians, who became prominent in the *experimental mathematics* field, are noticeable for their absence. For instance, Zeilberger did not publish in *Experimental Mathematics* until 2006. The group around Borwein and Bailey also did not publish in *Experimental Mathematics* straight away. But when they started publishing there, they quickly rose to become the most prolific authors in the journal. As Table 1 shows, Borwein, Bailey, and their frequent collaborator Richard Crandall all published extensively, ranking top 5 as the most frequent contributors to the *Experimental Mathematics* among a total of more than 1900 distinct authors.

5.2 A Cross-Field Enterprise

The first issue also points to an interesting evolution in the distribution of fields over the first volumes. In the first issue, the mathematical subjects covered encompassed number theory and algebraic geometry which would be stable foci for the following decades. But they also included fields such as geometry (general or convex and dynamical) as well as dynamical systems and several complex variables, which would later figure less. This might suggest that the first issue included more fields where visualization was key than the later focus on fields where algebraic

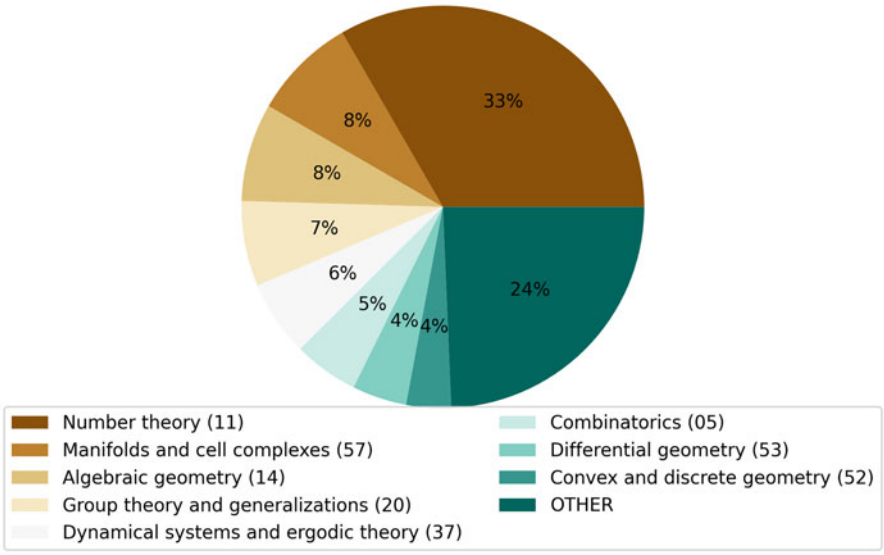


Fig. 3 Distribution of primary MSC classifications in *Experimental Mathematics*. The classification is drawn as the primary classification according to the MSC scheme as listed on zbMATH, using only the first two digits to indicate the major mathematical subject field

manipulation constituted a larger role of what computers contribute to mathematical research.

Figure 3 shows the total distribution of subjects in the *Experimental Mathematics*. Subjects are mapped by the primary classification according to the MSC scheme, and the first two digits of this primary classification are used to identify the mathematical subject field at the most general level; obviously this focus on primary classifications precludes other interesting and more fine-grained analyses of the mathematics included in the analyses. The category "OTHER" refers to classifications that each amount to 2% or less of the total corpus.

When the distribution is plotted over time, it is difficult to see general patterns, except for the dominance of number theory in the journal (see Fig. 4): If number theory – for some reason – plays a lesser role in a volume, the graph looks more erratic.

The most important information gathered from this quantitative analysis is visualization of the very broad spectrum of fields covered in the journal over the first years. Indeed, the first volume stands out as more varied, but over the entire time span, a clear dominance of number theory emerges in the journal with 33% of the articles classified as primarily belonging in that field (MSC 11) (for comparison, less than 7% of the articles on arxiv fall in the number theory classification, see more below). It is important to notice that MSC 11 is one of a number of fields to explicitly include a computational subfield: 11Yxx: "Computational number theory." Yet, a

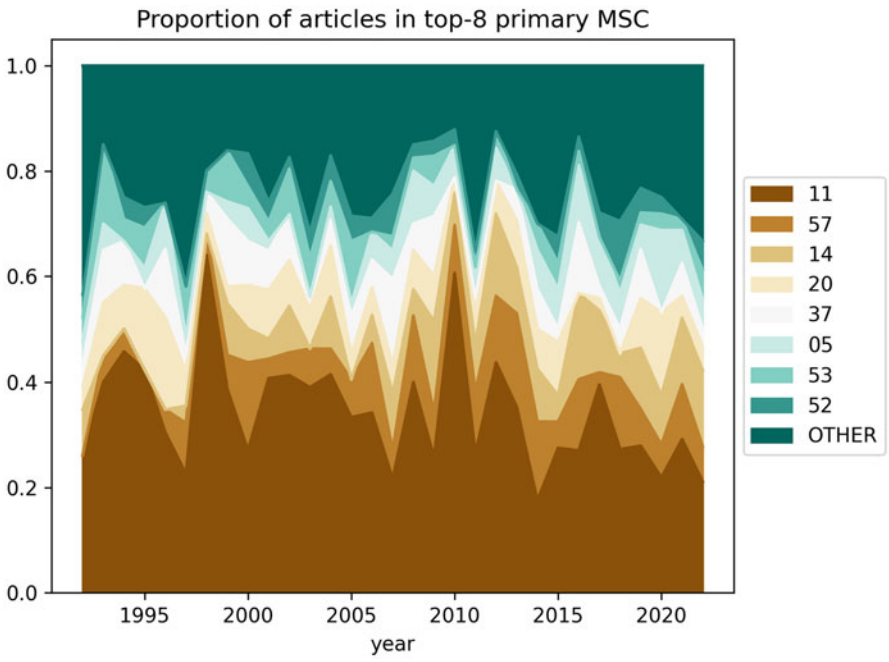


Fig. 4 Development of subject composition in *Experimental Mathematics* over time. Colors refer to primary MSC codes, see Fig. 3

Table 2 Indicators for `experiment` and `observation` written as regular expressions. All matches are case in-sensitive. Examples of matches given here are not exhaustive

Indicator	Regular expression	Example matches at start of word
Experiment	`r'\bexperiment\w*'`	`Experiment, experiments, experimental, . . .`
Observation	`r'\b(empirical\w*|data|observ\w*)`	`Empirical, empirically, data, observation, observe, . . .`

steady proportion of 25% of the articles published did not fall within the top 8 subjects which were each responsible for 4% or more of the corpus.

5.3 Experiments and Inductions: Quantitative Analyses

The primary theoretical focus of this section is the relationship between experiments and observations. To get at quantitative impression from the *Experimental Mathematics*, I chose to focus first on the framing that occurs in titles and abstracts of published papers, and I devised a set of indicators to pick out word-proxies for the topics of interest.

Two indicators were defined via regular expressions to identify experiments and observations (see Table 2). Two different approaches were then open: First to

compute and analyze the mere prevalence of these indicators, and second to use simple, but automated, linguistic analyses to see what objects and actions these indicator proxies point to.

These first quantitative forays show that the *experiment indicator* is found in 17% of the papers in the *Experimental Mathematics* corpus; in other words, almost every one in five papers explicitly refer to experiments or experimental methods or approaches in their titles or abstracts. Similar results show that fewer papers, 11%, refer to observations in their titles or abstracts, with observations being defined as in Table 2 (the frequencies should not be compared as the indicators are defined only to catch proxies).

Visualizations also show that the prevalence of the experiment indicator is slowly declining, whereas the observation indicator is slowly on the increase. The first part of this observation may be (at least partly) explained by a phenomenon like *obliteration by consolidation* when the experimental nature of the journal alone would render it superfluous to flag each paper as experimental. If this interpretation has any merit, the increase in the observation indicator becomes even more interesting.

In order to assess the extent to which these observations are specific to *Experimental Mathematics*, `arXiv` provides an easily accessible corpus of mathematical texts, albeit with the major limitation that `arXv` is *not* a journal, and papers are not refereed. Therefore this quality check is missing, and as refereeing is not only checking the validity of claims but also involves framing and narrative (Andersen 2017; Andersen et al. 2021b), this may be quite an important bias.

With this in mind, running the indicators on 450,000 mathematics papers uploaded to the `arXiv` during the period 1992–2022 can provide room for comparison. Extracting papers from these 31 years and filtering out the proxy categories for physics `(math.MP)` and computer science `(math.CS)` and the general categories math.HO and math.GM as well as the statistics category (math.ST), rather different pictures are found (for the `arXiv` categories, see https://arxiv.org/archive/math). Using the same indicators and metrics as above, one finds a proportion of 1.7% of papers having the experiment indicator and 6.2% having the observation indicator. Furthermore, the prevalence of both indicators are on a rise, steeper than seen in *Experimental Mathematics*. If anything, the first of these observations seem to lend support to the obliteration interpretation for *Experimental Mathematics* and point to this aspect of the `arXiv` as another question worthy of further study. It is *not* the case in general, that experimental (or observational) indications in titles and abstracts are on the decrease – rather the opposite.

5.4 Experiment and Observation: Linguistic Analyses

To get from the level of individual words – captured through the proxy indicators – to a higher semantic level, one might apply linguistic tools derived from natural language processing (NLP) tools. These tools are probabilistic in nature, and allow for the identification of different functions of tokens in a sentence through

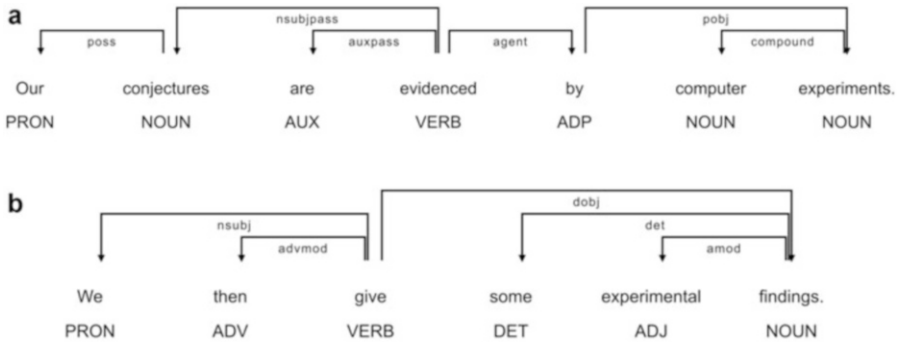


Fig. 5 Dependency graphs offer a way of analyzing linguistic relations on the level of sentences (de Marneffe and Manning 2008). In these two simple examples, a word matching the experiment indicator is shown with its dependencies. Similar graphs for verbs and adverbs can also be produced but these forms are rarer in the corpus. An arrow from *A* to *B* indicates that *A* is the *head* for *B*. (**a**) dependency graph of a noun; example from Mellit (2022) and (**b**) dependency graph of an adjective; example from Katz (2021)

Table 3 When experiments are adjoined in a noun chunk, the numerical or computational form of the experiment can be extracted. A third class contains those indescript chunks that refer to "some" or "our" experiments, etc.

Category	Extensive	Count
Experiment	4	32
Computer experiment	4	28
Numerical experiment	2	24

position-of-speech tagging (PoS) and structural-level analysis of sentences through *dependency graphs* as illustrated in Fig. 5 (spaCy 2021).

Focusing on the *experiment indicator*, it matches all nouns, verbs, adjectives, and adverbs. However, "to experiment" or "experimenting" are very rare in the corpus (total three occurrences), so the indicator is rarely a verb. The adverbial form "experimentally" is also quite rare (total 19 occurrences), referring to such experimental activities as studying, finding, confirming, discovering, exploring, or observing.

Among the 70 instances where the experiment indicator functions as a noun, we find their associated verbs which indicate the functions of experiments to include describe, suggest, study, or even lead. If there is an adjective connected to the noun, resulting in a so-called *noun chunk*, it is mostly to signify the "numerical" or "computational" nature of the experiment (see Table 3); some of these experiments were stressed to be particularly extensive.

Finally, among the 91 situations where "experimental" is an adjective, the most frequent associated noun is either evidence, data, or result, with an associated verb of

Table 4 A grounded coding led to six categories of roles of experiments as extracted from the tokens dependent on the experiment indicator. Here, the categories are shown together with their three most frequent forms

Category	Form	Total
Study	Study	14
Study	Explore	4
Study	Conduct	4
Result	Result	12
Result	Describe	9
Result	Find	8
Proof	Show	5
Proof	Confirm	4
Proof	Certify	1
Evidence	Evidence	19
Evidence	Datum	16
Evidence	Observe	6
Heuristics	Suggest	6
Heuristics	Inspire	3
Heuristics	Approach	2
OTHER	Lead	4
OTHER	Method	2
OTHER	Include	2

present, give, describe, or similar words. This may be indicating sentences describing how specific kinds of experiments (as indicated by adjectives) provide (as per the verbs) the mathematician with various items of interest.

To facilitate brief contextual and qualitative readings, a preliminary, grounded coding of the tokens dependent on the experiment indicator resulted in six categories and a miscellaneous code (see Fig. 4, Table 4).

Shifting attention to the *observation indicator*, it is similarly found that observations and empirical elements occur in the corpus. As adjectives, the token "empirical" is the most frequent match of the observation indicator (78%), and typically it is connected as "empirical evidence presents" or similar. However, the single instance of "observable" is interesting, as it encompasses many perspectives of this analysis:

> We confirm that several of the features that are predicted by the mathematical results are observable by numerical computation. (de la Llave and Petrov 2002)

From this one sentence, we see many of the indicators working together: the authors *confirm*, features can be *predicted* by *results*, and these *features* can be *observed* using computers. Thus, in what they call in their title an "experimental study," they use established results to conjecture extensions which are then shown to be in agreement with numerical observations. It might be that the authors are simply conflating different epistemic steps, and from the paper it is difficult to ascertain what constitutes the *experimental* nature of their study: The word is never explicitly used, but the paper *is* structured around the predictions, and evidence in the form of

tables and graphs is given *en masse*. Thus, the content supports the summary of the argumentative structure given here.

An even stronger claim for experimental and numerical warrants for results can be found in another abstract:

> They [certain "explicit algorithms"] are followed by numerical examples, which experimentally confirm main results of the article. (Enolski and Fedorov 2018)

This abstract is indicative of a specific type of numerical warrant that has to do with algorithms. Often, numerical experiments are undertaken to determine the efficiency or run time of algorithms (see also Sørensen et al. 2024). Here, the numerical experimentation actually confirms the *results* of the algorithms, and again the authors might simply have meant that the experiments are in agreement with the proved results. Even that would be very interesting, as one – at least from certain classical philosophical positions – might wonder what purpose an experimental verification could serve *after* the result had been formally (rigorously) proved. But the authors are not imprecise. In the paper, the experimental confirmation is given a context:

> Then we calculate [(defining) equations] by using the main Theorem 4.2 and, finally, see that the period matrices of the Jacobians and the curves share the same sets of absolute invariants. This provides an experimental confirmation of the main results of the paper. (Enolski and Fedorov 2018)

So the confirmation is not of a particular (even exemplary) case but a result of computing two defining properties and noticing that they correspond.

These two instances have shown a variation in epistemic warrants drawn from what seems to be inductive, at least numerical, examples. However, the corpus is also full of more heuristic uses of observed evidence in forming conjectures and beliefs. Here, two examples will have to suffice. In the following abstract, it is described how experimental evidence is found for certain properties that would merit proof:

> We also identify stationary solutions and periodic solutions in an associated problem, and find experimental evidence of hyperbolic behavior. (Colliander et al. 2013)

Here, the thing-to-be-proved seems rather well defined without the experiments. Another form of the heuristic use of evidence can be seen from this abstract:

> A conjectural density is derived from a single heuristic assumption, and experimental data supporting this assumption are given. (Stevenhagen 1993)

Here, the thing-to-be-proved is the result of an even vaguer form of heuristics – a single assumption – and experimental data serves to support the assumption and thus suggest the feasibility of reaching a proof.

A final example, and perhaps in a qualitative reading the most interesting, arises from the following indicator, as the verb "led" stood out and could not be given meaning from the quantitative, sentence (or abstract)-level data:

> Such experiments led to the results above. (Chas and Phillips 2010)

In the paper, it becomes more clear what roles the experiments played:

> Here is some numerical evidence, computed using the algorithm given in [Cohen and Lustig 87]; see [Chas and Krongold (09)] for a more detailed presentation (Aberdein and Inglis 2019). This evidence was in fact the motivation for the research presented here. (Chas and Phillips 2010, p. 131)

A footnote at the point (Aberdein and Inglis 2019) then points out where the Java program can be found online. The numerical evidence leads to two "Experimental Theorems," which serve as motivational theorems, called so but remaining to be proved (Chas and Phillips 2010, p. 131). The main proof is a long and complicated one, split into several partial results with proofs. At the end of the main theorem comes an entire section devoted to "experimental results and conjectures" (Chas and Phillips 2010, p. 144), containing another four "experimental theorems." The use of the term *experimental theorem* in this paper calls for some analysis, as it is – even in *Experimental Mathematics* – a rather unusual (and perhaps provocative) substitute for or adjugation to "conjecture." They stand out because they are labeled *theorems*, but *no* direct evidence is given for them. Instead, they formulate, in completely general terms, properties about the topological objects being studied. In fact, the statements could probably have been *real* theorems, if any evidence beyond induction from computational experimentation was known.

6 Conclusions and Perspectives

The first part of this chapter was devoted to present a background history of the debates and – sometimes – controversies that surrounded *experimental mathematics* in the infancy of computer-assisted experimental practices in mathematics. Particular attention was paid to outlining three different views on the possible epistemic warrants of experiments.

As analyzed, Zeilberger held that methods of checking the solvability of vastly overdetermined systems of equations provide "almost certainty," and that this would be sufficient except for *very* important mathematical results. Thereby, Zeilberger essentially drew on a very strong version of the position that "there are no true coincidences in mathematics" (see, e.g., Lange 2010), strong enough to serve as an epistemic warrant. As will be obvious, this position is a metaphysical one as much as any other position about the warrants of plausible or experimental reasoning.

From the side of *Experimental Mathematics*, another view was put out, when Epstein and Levy argued that *systematic* experimentation would be a sine non qua of

experimental mathematics and the (only) way toward drawing epistemic warrants from experiments. However, as can be extracted from Lakatos' reconstructions (e.g., Lakatos 1976), the right domain for systematic experimentation is difficult to ascertain beforehand. And the systematic approach may run into other types of bottlenecks when the domain becomes very large. To this latter point, the advent of increasingly powerful desktop computers may have alleviated some of the early challenges (Sørensen 2016), but it remains an unwinnable race between technology and mathematical ingenuity.

The third position, that of the group around Borwein, sought a middle ground by presenting what was subsequently incorporated as *exploratory experimentation*: A practice of experimentation that essentially transgressed the old divide between contexts of discovery (heuristics) and justification (proof). Initially thinking about mathematical experiments as similar to critical experiments that determine what is true and what is not, Borwein's philosophical point was limited to refuting conjectures. With the practice described as exploratory, the experiments could be guided toward forming useful concepts and proof sketches that would eventually lead to formal (i.e., rigorous) proofs. At least, this is a philosophically more coherent reading of the eight roles of computers in mathematics, that Borwein and Bailey repeatedly presented (Sørensen 2010).

Having outlined these positions in their larger historical complexity, focus shifted toward the role played by the journal *Experimental Mathematics* in shaping *experimental mathematics*, and particular importance was given to the relations between experiments and observations. But first, summary statistics allowed for answering some of the more general questions about the journal during its first 30 years of existence.

First, it was shown that on a number of parameters, set out by its initial editors, *Experimental Mathematics* has grown to be a success. Three decades into its life, it continues to publish papers of high quality from leading mathematicians – for instance, the recent fields medalist Peter Scholze has a recent paper published in the journal (Scholze 2022).

What is more, the almost 1200 papers published during these decades have a total of almost 2000 authors, indicating that the journal's body of contributors is quite large and stable. There have been some prolific authors like Borwein and Bailey, but more than anything, these numbers seem to indicate a high degree of collaborative authorship in *Experimental Mathematics*, although the average number of authors on papers in *Experimental Mathematics* is 2.07, which does not differ much from the value for all of mathematics when measured in the `arXiv`.

As a third quantitative finding, it was revealed that the broad interest from many different mathematical subjects that was part of the initial motivation has indeed been reached, and reached quite early on. Number theory has dominated the pages of the journal since the beginning, but a good number of other branches of mathematics also regularly find way into the journals, eight subject fields each accounting for more than 4% of the papers. Despite the high prevalence of number theory, this information still supports an interpretation of *experimental mathematics* not being

tied to any specific mathematical discipline, neither being a subfield of any mathematical field.

There are many more interesting phenomena around *Experimental Mathematics* that are open to quantitative analysis: For instance, the occurrence of an indicator *computer* seems to suggest a steady, but increasing reliance on computers in the journal, more so than in the `arXiv` corpus. This is not a stable finding in itself, but it does challenge the interpretation of *obliteration* invoked above, as computers surely are becoming more and more ubiquitous in mathematical research practice. And perhaps the proper way to gauge any such dependency would be to inquire about software rather than hardware, something that the corpora are also becoming more open to. Similarly, an indicator for *sharing incomplete or early results* seems to suggest that also on this point Epstein and Levy were successful. It would be interesting to know if publication of conjectures or preliminary results set *Experimental Mathematics* apart from other, otherwise comparable journals (if such even exist).

Among the many other interesting questions to pursue from this study, let me just outline one: In analyzing the subject fields covered in the very first issues of *Experimental Mathematics*, I found indications of a potential shift from some fields that mainly used computers to visualize objects toward fields where algebraic manipulation was a more important task for computers. This observation is only a hunch from the present data, but it aligns with overall historical developments at the time. So it may be worth further investigation, which, however, cannot be done from the great distance of titles and abstracts. Instead, we will need to closely analyze key papers, the identification of which is, however, prepared by the quantitative tools presented and used here.

Returning to the second research question posed above about observations and experiments, the quantitative analyses are – as expected – less productive by themselves, so they must be complemented by more qualitative aspects. This interaction between quantitative and qualitative methods is one of the strengths of an integrated approach to using digital and computational tools in studying mathematical practices. It allowed me to single out cases for closer inspection based on indicators suited to identify instances of (strong) epistemic warrants drawn from observation through induction or from (computational) experimental approach (i.e., the intersection *AD* in Fig. 1).

In studying a few cases which are not representative but still identified using general criteria, I found that indeed both experiments and observations can provide warrants in contemporary mathematical practice. Observations were often signaled by terms like “evidence” or “data,” thus highlighting the noninteractive role of the mathematician in accessing this knowledge. Rather, the computer (or other induction-producing practices) yielded data that serves its epistemic role as supporting or verifying conjectures.

Complementarily, experiments also ascribed epistemic role and warrant, both as a heuristic but also as supporting or confirming results. It is noteworthy that experimentation is here often used as a more *open* form of using computational (or other similar) practices. In general, the roles of experiments identified in these examples

closely mimic the roles which Borwein envisioned for computers in mathematical practice. And this should perhaps not come as a surprise, since Borwein and *Experimental Mathematics* were so connected.

In the end, the challenge of what can be learned from these analyses of *moderate size corpora* – larger than a human can survey in a reasonable time, yet not large enough to only be accessible statistically. The process has been to zoom in on the interesting philosophical practice from the outside, first through summary quantitative means, then through indicators leading to interesting instances, and finally to closer analyses on the level of paragraphs (abstract) or papers. One version of the challenge raised whether the phenomenon being studied is not just a mirage – an artifact brought about by accidental factors such as imprecise language or lack of philosophical finesse. As answer to such a challenge it helps to notice that mathematicians are often refined in using the terms, in particular in putting them together to bring about nuance in their argument. Thus, it is appropriate that the opening words of the abstract of Gronau and Wagenmakers (2018) function as an overall conclusion of this chapter:

> Many questions in experimental mathematics are fundamentally inductive in nature. (Gronau and Wagenmakers 2018)

7 Cross-References

- ▶ Experimental Mathematics in Mathematical Practice
- ▶ Experiments in Mathematics: Fact, Fiction, or the Future?
- ▶ Heuristics and Mathematical Practice
- ▶ Mathematical Experiments on Paper and Computer
- ▶ Non-deductive Justification in Mathematics
- ▶ The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts

References

Aberdein A, Inglis M (eds) (2019) Advances in experimental philosophy of logic and mathematics. Adv Exper Philos. Bloomsbury

Andersen LE (2017) On the nature and role of peer review in mathematics. Account Res Policies Quality Assurance 24(3):177–192. https://doi.org/10.1080/08989621.2016.1274885

Andersen LE, Andersen H, Sørensen HK (2021a) The role of testimony in mathematics. Synthese 199:859–870. https://doi.org/10.1007/s11229-020-02734-9

Andersen LE, Johansen MW, Sørensen HK (2021b) Mathematicians writing for mathematicians. Synthese 198:6233–6250. https://doi.org/10.1007/s11229-019-02145-5

Andrews GE (1994) The death of proof? Semi-rigorous mathematics? You've got to be kidding! Math Intell 16(4):16–18

Atiyah M et al (1994) Responses to "theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics", by a. Jaffe and F. Quinn. Bull Am Math Soc 30(2): 178–207

Bailey DH, Borwein JM (2011) Exploratory experimentation and computation. Not Am Math Soc 58(10):1410–1419
Bailey DH et al (2007) Experimental mathematics in action. A K Peters, Wellesley
Baker A (2010) Mathematical induction and explanation. Analysis 70(4):681–689. https://doi.org/10.1093/analys/anq074
Baker A (2007) Is there a problem of induction for mathematics? In: Leng M, Paseau A, Potter M (eds) Mathematical knowledge. Oxford University Press, Oxford, pp 59–73
Baker A (2008) Experimental mathematics. Erkenntnis 68(3):331–344
Baker A (2009) Non-deductive methods in mathematics. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. http://plato.stanford.edu/archives/fall2009/entries/mathematics-nondeductive/
Borwein J et al (1996) Making sense of experimental mathematics. Math Intell 18(4):12–18
Borwein JM, Borwein PB (1992) Strange series and high precision fraud. Am Math Mon 99(7):622–640
Borwein J (2008) Implications of experimental mathematics for the philosophy of mathematics. In: Gold B, Simons RA (eds) Proof & other dilemmas: mathematics and philosophy. Mathematical Association of America, Washington, DC pp 33–59. Chap. 2
Borwein J, Bailey D (2004) Mathematics by experiment: plausible reasoning in the 21st century. A K Peters, Natick
Carter J (2019) Philosophy of mathematical practice. Motivations, themes and prospects. Philos Math (3) 27(1):1–32. https://doi.org/10.1093/philmat/nkz002
Chas M, Phillips A (2010) Self-intersection numbers of curves on the punctured torus. https://doi.org/10.1080/10586458.2010.10129073
Colliander JE et al (2013) Behavior of a model dynamical system with applications to weak turbulence. Exp Math 22(3):250. https://doi.org/10.1080/10586458.2013.793110
Davis PJ, Hersh R (1981) The mathematical experience. The Harvester Press, Boston
de la Llave R, Petrov NP (2002) Regularity of conjugacies between critical circle maps: an 897 experimental study. Exp Math 11(2):219. https://doi.org/10.1080/10586458.2002.10504688
De Mol L (2014) The proof is in the process. A preamble for a philosophy of computer-assisted mathematics. In: Galavotti MC et al (eds) New directions in the philosophy of science. The Philosophy of Science in a European Perspective 5. Springer International Publishing, pp 15–33. https://doi.org/10.1007/978-3-319-04382-1_2
Ekhad SB, Zeilberger D (1996) The number of solutions of $X^2 = 0$ in triangular matrices over $GF(q)$. Electron J Comb 3(R2):1–2
Enolski VZ, Fedorov YN (2018) Algebraic description of Jacobians Isogeneous to certain Prym varieties with polarization (1,2). Exp Math 27(2):147. https://doi.org/10.1080/10586458.2016.1236357
Epstein D, Levy S (1995) Experimentation and proof in mathematics. Not Am Math Soc 42(6):670–674
Epstein D, Levy S (2000) To the reader. Exp Math 9(1):1
Epstein D, Levy S, de la Llave R (1992) About this journal. Exp Math 1(1):1–3
Franklin LR (2005) Exploratory experiments. Philos Sci 72:888–899
Giardino V (2022) Experimenting with triangles. Axiomathes 32(S1):55–77. https://doi.org/10.1007/s10516-022-09639-x
Goldstein C (2007) How to generate mathematical experimentation, and does it provide mathematical knowledge. In: Feest U, Hon G (eds) Generating experimental knowledge, Preprint 340 edn. Max-Planck-Institut für Wissenschaftsgeschichte, Berlin, pp 61–85
Gronau QF, Wagenmakers E-J (2018) Bayesian evidence accumulation in experimental mathematics: a case study of four irrational numbers. Exp Math 27(3):277. https://doi.org/10.1080/10586458.2016.1256006
Hazewinkel M (1985) Experimental mathematics. Math Modell 6:175–211
Hersh R (1991) Mathematics has a front and a Back. Synthese 88:127–133

Hyland K (2009) Academic discourse. English in a global context. Continuum discourse. Continuum International Publishing Group, London/New York
Inglis M, Aberdein A (2014) Beauty is not simplicity. An analysis of mathematicians' proof appraisals. Philos Math (3), 23(1):87–109. https://doi.org/10.1093/philmat/nku014
Jaffe A, Quinn F (1993) "Theoretical mathematics": toward a cultural synthesis of mathematics and theoretical physics. Bull Am Math Soc 29(1):1–13
Johansen MW, Misfeldt M (2016) Computers as a source of a posteriori knowledge in mathematics. Int Stud Philos Sci 30(2):111–127. https://doi.org/10.1080/02698595.2016.1265862
Kant D, Pérez-Escobar JA, Sarikaya D (2021) Three roles of empirical information in philosophy. Intuitions on mathematics do not come for free. KRITERION – J Philos 35(3):247–278. https://doi.org/10.1515/krt-2021-0025
Katz NM (2021) On a question of Zannier. Exp Math 30(3):422. https://doi.org/10.1080/10586458.2018.1551818
Lakatos I (1976) In: Worrall J, Zahar E (eds) Proofs and refutations. The logic of mathematical discovery. Cambridge University Press, Cambridge, UK
Lakatos I (1986) A renaissance of empiricism in the recent philosophy of mathematics? In: Tymoczko T (ed) New directions in the philosophy of mathematics. An anthology. Birkhäuser, Boston, pp 29–48
Lange M (2010) What are mathematical coincidences (and why does it matter). Mind 119(474): 307–340. https://doi.org/10.1093/mind/fzq013
Lolli G (2008) Experimental methods in proofs. In: Lupacchini R, Corsi G (eds) Deduction, computation, experiment. Exploring the effectiveness of proof. Springer, Berlin et al, pp 65–79
Löwe B (2016) Philosophy or not. The study of cultures and practices of mathematics. In: Cultures of mathematics and logic. Springer International Publishing, pp 23–42. https://doi.org/10.1007/978-3-319-31502-7_2
Löwe B, Van Kerkhove B (2019) Methodological triangulation in empirical philosophy (of mathematics). In: Aberdein A and Inglis M (eds) Advances in experimental philosophy of logic and mathematics. Advances in Experimental Philosophy. Bloomsbury, pp 15–37. Chap. 2
Lupacchini R, Corsi G (eds) (2008) Deduction, computation, experiment. Exploring the effectiveness of proof. Springer, Berlin
de Marneffe MC, Manning CD (2008) Stanford typed dependencies manual. https://downloads.cs.stanford.edu/nlp/software/dependencies_manual.pdf
Mazur B (1992) The topology of rational points. Exp Math 1(1):35. https://doi.org/10.1080/10586458.1992.10504244
McEvoy M (2008) The epistemological status of computer-assisted proofs. Philos Math 16(3): 374–387
McEvoy M (2013) Experimental mathematics, computers and the a priori. Synthese 190(3): 397–412. https://doi.org/10.1007/s11229-011-0035-1
Medawar PB (1979) Advice to a young scientist. Harper & Row, San Francisco. ISBN: 9780465000920
Mellit A (2022) Rationality proofs by curve counting. Exp Math 31(3):773. https://doi.org/10.1080/10586458.2019.1691088
Milnor J (1992) Remarks on iterated cubic maps. Exp Math 1(1). https://doi.org/10.1080/10586458.1992.10504242
Misfeldt M, Johansen MW (2015) Research mathematicians' practices in selecting mathematical problems. Educ Stud Math 89(3):357–373
Misfeldt M, Danielsen K, Sørensen HK (2015) Formal proof and exploratory experimentation: a Lakatosian view on the interplay between examples and deductive proof practices in upper-secondary school. In: Barbin E, Jankvist UT, Kjeldsen TH (eds) History and epistemology in mathematics education. Proceedings of the seventh European Summer University. Danish School of Education, Aarhus University, pp 433–446
Petkovšek M, Wilf HS, Zeilberger D (1996) A = B. A K Peters, Wellesley. ISBN: 1-56881-063-6

Polya G (1968) Mathematics and plausible reasoning: patterns of plausible inference. Princeton University Press, Princeton

Quinn F (1995) Roadkill on the electronic highway: the threat to the mathematical literature. Not Am Math Soc 42(1):53–56

Rittberg CJ (2019) On the contemporary practice of philosophy of mathematics. Acta Balt Hist Philos Sci 7(1):5–26. https://doi.org/10.11590/abhps.2019.1.01

Rittberg CJ, Van Kerkhove B (2019) Studying mathematical practices. The dilemma of case studies. ZDM 51(5):857–868. https://doi.org/10.1007/s11858-019-01038-8

Scholze P (2022) Liquid tensor experiment. Exp Math 31(2):349. https://doi.org/10.1080/10586458.2021.1926016

Sørensen HK (2010) Exploratory experimentation in experimental mathematics. A glimpse at the PSLQ algorithm. In: Löwe B, Müller T (eds) PhiMSAMP. Philosophy of mathematics: sociological aspects and mathematical practice. Texts in philosophy 11. College Publications, London, pp 341–360. http://www.lib.unibonn.de/PhiMSAMP/Book/

Sørensen HK (2016) 'The End of Proof'? The integration of different mathematical cultures as experimental mathematics comes of age. In: Larvor B (ed) Mathematical cultures. The London meetings 2012–2014. Trends in the history of science. Birkhäuser, pp 139–160. https://doi.org/10.1007/978-3-319-28582-5_9

Sørensen HK, Johansen MW (2024) Copenhagen program for digital philosophy of mathematical practice. To appear in PLOSone 2024

Sørensen HK, Mathiasen SK, Johansen MW (2024) What is an experiment in mathematics? New evidence from mining the Mathematical Reviews. To appear in Synthese 2024

spaCy (2021) Industrial-strength natural language processing in python. Explosion AI. https://spacy.io

Steinle F (2016) Exploratory experiments. Ampère, faraday, and the origins of electrodynamics (trans: Alex Levine). University of Pittsburgh Press, Pittsburgh

Stevenhagen P (1993) The number of real quadratic fields having units of negative norm. Exp Math 2(2):121. https://doi.org/10.1080/10586458.1993.10504272

Stöltzner M (2005) Theoretical mathematics: on the philosophical significance of the Jaffe-Quinn debate. In: Boniolo G, Budinich P, Trobok M (eds) The role of mathematics in the physical sciences. Interdisciplinary and philosophical aspects. Springer, Dordrecht, pp 197–222

Tymoczko T (1979) The four-color problem and its philosophical significance. J Philos 76(2):57–83. http://www.jstor.org/stable/2025976

Van Bendegem JP (1993) Foundations of mathematics or mathematical practice: is one forced to choose. In: Restivo S (ed) Math Worlds: philosophical and social studies of mathematics and mathematics education. State University of New York Press, Albany, pp 21–38. Chap. 2, ISBN: 0791413292

Van Bendegem JP (1998) What, if anything, is an experiment in mathematics? In: Anapolitanos D, Baltas A, Tsinorema S (eds) Philosophy and the many faces of science. Rowan & Littlefield Publishes, Lanham, pp 172–182. Chap. 14

Zeilberger D (1994) Theorems for a price: tomorrow's semi-rigorous mathematical culture. Math Intell 16(4):11–14, 76

Revisiting "The New 4CT Problem"

Mark McEvoy

Contents

1 The Traditional View of Proof and the (Old) Four Color Problem 2459
2 Tymoczko's Conclusions: The New Four Color Problem 2460
3 Response 1: Denying the Legitimacy of the CAP of the 4CT 2462
4 Response 2: Nothing New Under the Sun 2464
5 Response 3: Proofs Are A Priori if They Involve Only A Priori Steps 2467
6 Response 4. Unsurveyable CAPS (Can) Give A Priori Knowledge 2470
7 Response 5: The Future 2472
8 Response 6: The CAP of the 4CT Is Surveyable, in a Sense 2474
9 Certainty 2477
References 2477

Keywords

Computer proofs · Apriorism · Tymoczko · Unsurveyability · Certainty · Four color problem · Mathematical proof · Knowledge

1 The Traditional View of Proof and the (Old) Four Color Problem

According to what one might label the traditional view of proof in mathematics, proofs have the following characteristics. They are knowable a priori, the knowledge they provide is certain, rather than merely probable, they are surveyable, and, because of these other features, a mathematical proof is convincing to one who understands it. Opponents of this view typically drew their motivation not from the study of mathematics, but rather from a more general antipathy to apriority in epistemology and necessity in metaphysics (Mill, Putnam and Quine all spring to

M. McEvoy (✉)
Hofstra University, Hempstead, NY, USA
e-mail: Mark.V.McEvoy@hofstra.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_38

mind here). Tymoczko (1979) provides a different sort of challenge, taking as its cue a development within mathematics itself, rather than being based on philosophical scruples over apriority or necessity.

Tymoczko's challenge famously uses the (Appel et al. 1977) computer-assisted proof (CAP) of the Four Color Theorem (4CT) as its starting point. The problem to which the 4CT provided the solution was the problem of whether all countries on a map could be colored with only four colors in such a way that no neighboring countries were of the same color. The positive answer to this question, established in 1977, was provided, in part, with the essential aid of a computer. Appel, Haken and Koch built on work by earlier researchers, to the effect that certain configurations of countries (the "reducible" configurations) are not possible in the smallest five-colored maps, because they can be replaced by smaller configurations (resulting, contradictorily, in a yet smaller map). Note next that a configuration is said to be unavoidable if every map contains a member of the set. (Appel 1984, 36–7) notes that the theorem would be proved once an unavoidable set of reducible configurations could be obtained. He continues:

> Our proof consisted of providing such a set of reducible configurations (and checking by computer that all members of the set were reducible) and showing by hand that the set was unavoidable.

Put simply, the CAP of the 4CT shows that "every map contains a reducible configuration and hence there can be no smallest map requiring five colors, thus no map at all can require five colors" (Appel 1984, 36). Crucially, because there were between 1400 and 1500 configurations, and 1200 computer hours (at the time) were needed to check these configurations, the proof is too long to be surveyed by a human mathematician.

2 Tymoczko's Conclusions: The New Four Color Problem

The old Four Color Problem has thus been solved. The New Four Color Problem, arising from consideration of certain aspects of the 4CT is a philosophical question about what a proof is — whether it must be surveyable, formalizable, convincing, and whether our understanding of proof must be revised in the aftermath of the 4CT, and, if so, how (and if not, why not). Broadly speaking, it is a philosophical question about the essential nature of proof and how that nature is affected by the role of computers in mathematics. Tymoczko (1979, 57) tells us that his goal is "to raise the question of whether the 4CT is really a theorem." He is quick to add that he agrees that the CAP of the 4CT has definitively established its result, and that he has no desire to dictate to mathematicians what is or is not a theorem. However, Tymoczko (1979, 58) suggests that:

> [I]f we accept the 4CT as a theorem, we are committed to changing the sense of 'theorem', or, more to the point, to changing the sense of the underlying concept of "proof.

As Tymoczko (1979, 59) sees it, the three main criteria of traditional proofs are that they are convincing, surveyable, and formalizable. One way in which the 4CT creates trouble for the traditional view of proof as Tymoczko has characterized it is that it "does drive a wedge" (Tymoczko 1979, 59) between the latter two criteria. More troubling still for the traditional view of proof (Tymoczko 1979, 63) finds the appeal to an unsurveyable CAP means that the 4CT has been established "on grounds that are in part empirical."

Thus, on the assumption that the 4CT is indeed a genuine proof, a number of what the traditional view of proof held to be essential characteristics of proof seem to fall away. Most obviously, the proof does not seem to be surveyable. And, if surveyability is required for apriority, then the proof does not seem to be knowable a priori. Nor can it be said to be convincing, at least not in any sense beyond being convinced that the software and hardware have functioned correctly to give us the correct answer to the problem. Moreover, due to the possibility that an undiscovered error may lurk somewhere in the unsurveyable proof, it seems that the knowledge provided by the theorem is merely probable rather than certain. The new 4CT problem, then, is that of how we should understand proof in the light of such unsurveyable CAPs being accepted as proofs.

Before proceeding to the main body of the article which examines the various responses made to this new 4CT problem, I want to briefly consider an analogy Tymoczko uses in support of the claim that unsurveyable computer results must be established on empirical grounds, and thus are not knowable a priori (Thanks are due to an anonymous referee for pushing me to include discussion of the Simon analogy here.). Tymoczko (1979, 71) considers the hypothetical history of Martian mathematics, which proceeds much as our own, until the arrival on Mars of a genius named Simon. Under Simon's guidance, Martian mathematics made great strides. Eventually, however, Simon began producing proofs which included phrases such as "Proof is too long to include here, but I have verified it myself." On occasion, Martian mathematicians could not reconstruct the reasoning which had gone on behind these new "lemmas," but such was Simon's prestige that they accepted these results under the general heading "Simon Says." Tymoczko concludes firstly that Martian mathematics under Simon is not a legitimate development of mathematics, and secondly, that we are in a similar position with respect to our appeal to unsurveyable computer proofs.

Given Tymoczko's description of the case, the first of his conclusions seems unobjectionable. The second, however, is dubious. A natural response here is to note (as Tymoczko 1979, 72–74) does) that there are important differences between the position we are in with regard to computers, and the one that the Martians are in with regard to Simon. One such difference is that we (i.e., members of our community) designed the computer. The fact that we designed the computer gives rise to other, more significant differences between the computer case and the Simon case. The computer is designed to do a certain job: it is designed to follow programs. We know what program the computer is supposed to be executing. We can, therefore, as often as we choose, check to ensure that the computer is in fact executing this program. Furthermore, the proofs a given computer discovers can be checked by others.

Random portions of these proofs can be hand-checked by human mathematicians. The program may be tested by having a computer verify already established proofs. If the computer failed such tests, we would no longer believe its results. In the case of Simon, the Martian mathematicians did not design him, and at least some of the tests available in the case of computers are not available in his case. They cannot, for example, open him up and examine his inner workings. Nor can they assume, from the fact that he has given the correct answers in verifiable cases, that his answers in the unverifiable cases are likewise correct. He may be lying. He may have lost his ability, and be attempting to protect his prestige. It is true that, in the case of the computer, it might somehow have lost its ability, but this is exactly the sort of thing that can be checked, if desired, simply by running tests on whether it is executing its program. No such check is available in the case of Simon, since no-one knows what his "program" might be.

Having considered some of these differences between Simon's case and our own, Tymoczko (1979, 72) observes that computers are, unlike Simon, not just authority, but *warranted* authority. And this leads us to the real point of the analogy: if we accept unsurveyable CAPS but not Simon's appeals to his own authority, that can only be because we have evidence for the reliability of the former, but not the latter. We thus "admit evidence for the reliability of computers into a philosophical account of computer-assisted proofs" (Tymoczko 1979, 72). And, for Tymoczko (1979, 74), that evidence, being largely a matter of engineering, physics and program checking, is empirical. These claims, about whether and how the ineliminable use of computers introduces an empirical element in our warrant for unsurveyable CAPS will be taken up in detail in Sects. 4, 5, and 6 below.

3 Response 1: Denying the Legitimacy of the CAP of the 4CT

Faced with the claim that we now have an unsurveyable proof knowable only a posteriori, one forthright defense of the traditional view of proof in mathematics would be to simply deny that the 4CT is a proof. (Brown 1972, xii), for example, claims that the CAP of the 4CT is not a proof, since it has failed to establish a crucial proposition. However, as Swart (1980, 698) notes, Brown doesn't say which proposition it is that the CAP fails to establish, so we cannot evaluate his reasoning here. Shanker (1986, 410) also denies that the CAP of the 4CT is a proof. He takes a Wittgensteinean view according to which the role of proof in mathematical language is to "establish the rules for applying mathematical concepts." To play this role, proofs must be surveyable. Since this CAP is unsurveyable, it cannot play the essential role in mathematics that proof does, and thus is not a proof at all:

> (I)t makes no sense to speak, as Tymoczko supposes, of Appel and Haken 'proving' the 'four-colour theorem,' let alone of their 'proof' forcing us to modify our understanding of the concepts of proof and theorem.

Azzouni (1994, 170) also expresses doubts that CAPs are proofs at all. For Azzouni, unsurveyable verifications such as that of the 4CT "are really not proofs at all." Rather, computers supply:

> empirical existence claims for mathematical proofs, existence claims that, because of the limitations of human capacity, cannot even be checked *by the mathematical community as a whole* by examining the proof the computer has purportedly generated.

Rather than being genuine proofs, unsurveyable CAPs are "empirical demonstrations that certain proofs exist."

A less strident version of this strategy is to hold that even if the 4CT is indeed a proof, it is not a proof in the fullest sense — perhaps because it is insufficiently enlightening, or fails to shows unexpected connections between previously disparate areas of mathematics, or seems unlikely to prompt further fruitful research. For example, Rota (1997, 186) observes that verifications (i.e., arguments that list all possible cases relevant to establishing the truth of a particular assertion, and show that, for each case, the assertion holds true) are not all accepted as proofs. In particular, for our interest here, the 4CT is a verification which "falls short of being accepted as a proof—despite its undeniable correctness." For Rota, what is missing for such a verification is that it fails to give "satisfying reasons why a conjecture should be true" (Rota is sometimes willing to apply the term 'proof' to a verification like that of the 4CT, but he always qualifies such usage: "Verification is proof, but verification may not give the reason." (i.e., a verification may not give satisfying reasons why the relevant conjecture is true (Rota 1997, 220). Sometimes quotation marks are called in to do the qualifying: "This "proof" was the first verification of a major mathematical theorem by computer."). Rota (1997, 190) compares Appel and Haken's CAP with Wiles's proof of Fermat's last theorem:

> The actual value of what Wiles and his collaborators did is far greater then the mere proof of a whimsical conjecture. The point of the proof of Fermat's last theorem is to open up new possibilities for mathematics. . . . They develop a host of new techniques that will lead to new connections between number theory and algebraic geometry. Mathematicians of the future will benefit from Wiles's lead. They will find untold new applications of these newly found methods to the solution of other problems, even to problems of great practical interest. The value of Wile's proof lies not in what it proves, but in what it opens up, in what it makes possible.

In contrast, Appel and Haken's CAP "fails to open up any significant mathematical possibilities."

Along similar lines, Rav (1999, 2) sees the "essence of mathematics" as residing in:

> [I]nventing methods, tools, strategies and concepts for *solving problems* which happen to be on the current internal research agenda or suggested by some external application. [note elided] But conceptual and methodological innovations are inextricably bound to the search for and the discovery of proofs, thereby establishing links between theories, systematising knowledge, and spurring further developments.

Merely arriving at a truth value for a putative theorem by means of an unsurveyable CAP, we fail to "convey the links, inter-theoretical connections, strategies and overall mathematical knowledge embodied in these methods" (Rav 1999, 26) ((Hirsch et al. 1994, 189) is another who sees the 4CT as a diminished form of proof. In his response to (Jaffe and Quinn 1993), he asks rhetorically, "Are we now to consider the 4-color theorem as proved in the same sense as, say, the prime number theorem?" Though (Hirsch et al. 1994, 188) is also skeptical of the human produced, but unsurveyed Finite Simple Groups proof (Mackenzie 2001) and (Mackenzie 1999) are good sources for other views of this kind. (Mackenzie 2001, 321) quotes Yuri I. Manin: "A good proof is a proof that makes us wiser. If the heart of the proof is a voluminious (sic) search or a long string of identities, it is probably a bad proof." Paul Halmos is cited as remarking that "The present proof relies in effect on an Oracle, and I say down with Oracles! They are not mathematics'" (Mackenzie 2001, 102). A final example is taken from Bariwise, who, while discussing automated formal proofs in general rather than the 4CT specifically, observes that: (I)t could be that the best proofs (in the sense of being most enlightening or easiest to understand) of a program's correctness will use methods, like symmetry considerations, that are not adequately modeled in the logician's notion of formal proof, and so which would not be deemed correct by some automated proof checker designed around the formalist's model (Barwise 1989, 849).).

However, whatever view one takes as to whether the CAP of the 4CT is or is not a proof, denying that it is one doesn't so much dissolve the problem as relocate it. For, suppose we agree that the CAP of the 4CT is not a proof. Then Tymoczko's claim that we must rethink proof doesn't follow —there is, ex hypothesi, no empirical, less-than-certain unsurveyable *proof* which forces such a rethinking. But we are left facing similar issues about mathematical methodology in general. For even if not a proof, the CAP has established the truth of the theorem in a way that is arguably empirical, less than certain and unsurveyable. We must rethink our view of mathematical methodology to accommodate the undeniable fact that mathematicians accept this way of doing things (even if it is not their preferred method).

Unless, of course, we think the 4CT hasn't introduced into mathematical practice anything significantly novel.

4 Response 2: Nothing New Under the Sun

A second common line of response to Tymoczko's paper argue that the features Tymoczko sees as raising new challenges to our view of proof are in fact already present in human produced mathematics ((Prawitz 2008, 87–88) and (Secco 2017, 294) observe that one can categorize the majority of early responses to 4CT according to whether the author thinks the CAP introduces anything new into mathematics, and, amongst those who deny that it does, according to their view as to whether mathematics has always been either a posteriori or a priori.). Teller (1980) is an early example of a response to the effect that the 4CT does not introduce any philosophically new element into mathematics. For Teller (1980, 798), the role of the

computer in the 4CT is simply an extension to our surveying capacities — not a change in our concept of proof. Such novelty as there is in the 4CT "is in the means of surveying, not in the means of proof" (Teller 1980, 799). Levin (1981, 84) points out that even unsurveyability is not new. After all, "the whole of traditional mathematics may already be unsurveyable by one man in a lifetime" (Moreover, there are humanly produced proofs which have not been surveyed by any individual human. See for example, the proof of the Classification of the Finite Simple Groups. This proof was announced in 1979 by a group of mathematicians lead by Daniel Gorenstein. The proof consisted of a collection of connected results written down in various places, totaling over 10,000 pages (Barendregt and Wiedijk 2005, 2532).). For Levin (1981, 86), the moral of the introduction into mathematics of CAPs, whether unsurveyable or not, is merely that " the speed of computation has increased, but no qualitative change has taken place in mathematical method." One might here object, with Bassler (2006, 117) that such moves pay insufficient attention to the tension Tymoczko finds between surveyability and formalizability. Indeed to argue in the way Teller and Levin do here is in effect to reject Tymoczko's claim that surveyability is properly conceived of as a criterion of traditional proof. Teller (1980, 802), however, does not see unaided human surveyability as being criterial for what counts as a proof, and offers a slippery slope argument to the effect that it cannot be:

> Consider a race of organisms very like humans who do something which certainly looks just like what we call mathematics, and a second race similar to the first, and so on until we get to objects that look more like computers than people. Is mathematics a sliding conception, or are we really to suppose that we can draw the line somewhere and say the creatures on our side of the line do mathematics and those on the other side do something else?

Teller (1980, 798) is explicit: "surveyability is not part of what it is to be a proof in our accustomed sense."

However, regardless of whether surveyability is properly taken to be criterial for proof, unsurveyability does raise the possibility of there being hidden errors within the proof. Here, defenders of the traditional view (e.g., (Teller 1980, 800) have once again denied that this introduces a novel factor into mathematics; after all that we might raise similar worries about humanly surveyable proofs. Swart (1980, 703) too tells us that the possibility of hidden errors due to bugs or hardware problems are "nothing other than errors of logic, no different in essence from errors that crop up in proofs that have nothing to do with computers." And, while the computer is fallible, and we can raise worries about software and hardware malfunctions, Teller 1980, 801) observes that "Tymoczko gives us no grounds for thinking that the fallibility of computers and mathematicians are relevantly different."

Levin (1981, 82) argues that any argument from the possibility of error due to machine or program error would apply just as easily to traditional mathematics:

> This undesirably strong consequence is due to his emphasis on the fact that the computer is a physical object. For surely human beings and their workings are also physical objects, just as subject to physical laws as computers. If error is possible in computer results because of the

> (highly confirmed but) unproven theories that describe them and their programs, then the same possibility presents itself whenever human mathematicians work out proofs by themselves. . . . . Our wiring is subject to physical laws; indeed, we humans are prey to all sorts of mishaps- inattention, boredom, forgetfulness, slips of the pen or tongue, unconscious urges to fail – that computers escape.

(Levin notes that this line of argument might appeal to some empiricists (as of course it does to Detlefsen and Luker). Though she does not argue that such a position is false, she notes that Tymoczko could not accept it, as he is arguing that the 4CT gives us a *novel* line of argument against apriority.) Detlefsen and Luker (1980, 817) agree that CAPs introduce nothing essentially new, but contra Swart and Levin, they hold that this is because human proof has always been empirical in character. Approvingly citing work by Cerutti and Davis, they note that properties that are "in principle quite like those attributed by Tymoczko . . . to computer-assisted proofs can, with equal warrant, be attributed to human proofs" (Detlefsen and Luker 1980, 806). In particular, while we might worry, regarding CAPs, about faulty or buggy programs, or false initial data, human mathematicians are subject to fatigue, and to limits regarding memory and knowledge ((Resnik 1997, 151) makes a similar claim to argue that all of mathematics is empirical.).

Detlefsen and Luker (1980, 807–808) argue further for the claim that the introduction of computers to mathematics does not change its epistemological status on the grounds that both calculation and the act of surveying a proof are empirical. With respect to calculation, they note that, when a mathematician follows (e.g.,) Gauss's proof that the sum of the first 100 integers is 5050, she undertakes a "substantial" amount of calculations. And, for Detlefsen and Luker (1980, 808), our confidence in these calculations is based on four assumptions:

(i) That the underlying algorithm to be used is mathematically sound
(ii) That the particular program to be used is a correct implementation of this algorithm
(iii) That the computing agent correctly executes the program
(iv) That the reported result was actually obtained

Whether we are talking about a computer or human produced proof, (ii)–(iv) concern facts ascertainable only empirically. If these are part of the warrant in a proof like Gauss's, it would follow that proofs involving calculation — whether undertaken by human or computer — are empirical, and thus that at least much of traditional, human produced mathematics is empirical.

Moving to the argument from the act of surveying a proof, Detlefsen and Luker (1980, 816) argue as follows:

> If, as Tymoczko says, a proof is a unit of reasoning that needs nothing outside itself to be convincing and if, as seems clear, in order for a proof surveyed first hand to be convincing we must believe in the reliability of our survey, then, since that reliability is not something amenable to traditional, surveyable proof, we may conclude that the appeal to first-hand

> survey introduces the same kind of method into mathematics as is introduced by the appeal to computers.

If, then both computation and surveying a proof are themselves empirical forms of warrant, then all of traditional mathematics cannot avoid "the presence of empirical elements in mathematical proof" (Detlefsen and Luker 1980, 817).

Detlefsen and Luker's arguments that both the 4CT and traditional mathematics are empirical cannot be assessed without paying closer attention to the issue of what makes a sequence of reasoning a priori or empirical. It is to this latter issue that we now turn.

5 Response 3: Proofs Are A Priori if They Involve Only A Priori Steps

One obvious way of understanding a piece of reasoning to be a priori is to count it as such provided all of the steps contained within it are a priori licensed steps, which make no warranting use of sensory observation. Swart (1980, 701), for example, has it that mathematical theorems are a priori truths, regardless of how they are discovered, where for him, such truths are *in principle*, if not in practice, discoverable "without recourse to sense experience of the physical world." Various authors have issued a challenge to make precise what is meant by this type of 'in-principle' clause. For example, Kripke (1972, 34–35) asks what is meant when we say that something may possibly be known a priori: "And possible for whom? For God? For the Martians? Or for People with minds like ours?" Resnik (1997, 155), along similar lines, urges that the issue about apriority is an issue about what *we*, rather than some ideal being (or some in-principle version of us), can know a priori. Resnik grants that ideal beings could have a priori knowledge of humanly unsurveyable proofs, but insists that *we* cannot. However, Swart has a response here. For Swart, what makes a sequence of reasoning a posteriori or a priori is the nature of the steps in the reasoning. If all of the steps in a given proof are a priori, then that proof — even if unsurveyable — is a priori (I have defended a move of this type (though without Swart's commitment to this being equivalent to "in principle" a priori knowable). (Johansen and Misfeldt 2016, 119) criticize this kind of view as an unacceptable form of logicism. Mathematics, according to this reading of my view, is a matter of proofs and axioms, and the latter are true by definition. Thus, CAPs are a priori, because they are analytically true. In the two papers they cite (McEvoy 2008; McEvoy 2013), my main goal was to defend the apriority of mathematics, and I thus did not focus on where mathematics fell on the analytic/synthetic divide. But I am not committed to a logicism about arithmetic. For consider, axioms, may be synthetic, rather than analytic. This may hold, moreover, not only for established axioms, but for novel candidate axioms, which owe (a significant part of) their plausibility to fruitfulness considerations. See, e.g., (Gödel 1983), (Katz 1998, 197).). Moreover, for Swart (1980, 701), if we do not add the "in principle"

clause to our understanding of apriority, then mathematics has been a posteriori since early mathematicians first drew drawings in the sand.

How does this view of apriority handle Detlefsen and Luker's arguments that calculation and surveyability (whether undertaken by human or computer) render mathematics empirical? Let us take these in order. To begin with the former, it is difficult to see why the presence of calculation of instances should render empirical a piece of mathematical reasoning. After all, there seems no reason why one calculating instances involved in a relatively brief proof would not have a priori knowledge of the conclusion of that proof. One could assure oneself, for example, that there is only one even prime in the first one hundred natural numbers, and one could do this by mentally calculating instances. The knowledge one would arrive at by this process seems archetypical a priori knowledge ((Prawitz 2008, 92) holds that once the number of calculations required in a proof gets to the point where they cannot simply be incorporated into the body of a humanly surveyable proof, but instead require a computer "then the proof is not entirely deductive, and there is the undeniably epistemological consequence that the theorem is known only a posteriori." At least some of the critics of Tymoczko (e.g., (Swart 1980), (Levin 1981), (McEvoy 2008), and (Krakowski 1980)) do not contest the claim the proof is known only a posteriori (even while arguing that is nonetheless an a priori proof). But why think the proof is not deductive? (Prawitz 2008, 84) operates with a view of proof that is tightly connected to our grounds for making assertions (rather than with a view of proof as being a formal structure). For him, the "most crucial epistemological function of deductive inferences and proofs" is that they "furnish us with conclusive grounds for asserting the conclusions." And, given that nobody could survey the proof in its entirety, no-one has such conclusive grounds. In this connection, Prawitz holds that derivations are "codifications of proofs" rather than "answering the question of what a proof is." However, if one does not share Prawitz's view of the relative centrality to proof of grounding assertions as opposed to formalization, one need not concede that the 4CT is not deductive.). And, with respect to factors (ii)–(iv) from Detlefsen and Luker's list above, we can here, following Burge (1998) distinguish between what is actually part of the warrant in a given proof or sequence of reasoning, and what merely allows one access to that warrant. In a proof, the warrant is based on the inferential connections between lines (and the apriority of the premises); if these are a priori, then so is the proof. The appeal to the reliability of the computer—or to the reliability of the human mathematician's brain—is a background assumption which "need not be a part of the warrant for our mathematical result" (Burge 1998, 6).

What of the claim that surveying a proof itself renders warrant a posteriori? Similar remarks hold here; the fact that, in order to access the proof, a mathematician needs a functioning brain, or computer processor to survey the proof is not itself part of the warrant. Burge (1998, 12–13) offers useful remarks on the difference between warrant and the role of perception in accessing that warrant:

> We do have to hear the words of others, just as we have to read the print-outs of computers. But perceiving the words need not be, and normally is not, part of the fundamental

> entitlement for understanding what intentional content they convey, or for understanding them as presentations-as-true. The words are like diagrams or symbols that help one see the point of a mathematical claim: They call to the recipient's mind the thoughts of the interlocutor that they express. . ..But perception of them does not normally enter into the recipient's entitlement for relying on understanding of them or their assertive force.

This line can be pressed into service here: the act of surveying the proof, while an empirical act, is not itself part of the warrant for a proof. Surveying a proof merely allows one to access the relevant warranting inferential relations in the proof.

This view of apriority must address the claim of Tymoczko (1979, 63) that an appeal to computers is an empirical premise, a report *on an experiment*. Here the defender of the a priori must ask whether the 4CT is accurately described as an experiment, and, if so, whether this would imply that it was a posteriori, proof. Tymoczko (1979, 78) himself notes one strange consequence of seeing the 4CT along the lines of an experiment. Having claimed that the 4CT blurs the distinction between an alleged a priori discipline like mathematics and natural science, Tymoczko runs up against the problem of what to say about experiments in pure mathematics:

> It is easy to see how experiments play a role in the arguments of physical theory. The physical theory can predict phenomena of space-time which equipment can be designed to register. Are we to say that the computer registered a phenomenon of mathematical space? If not, then how else are we to explain the role of experiment in mathematics?

Tymoczko (1979, 78), naturally, does not conclude that the 4CT has indeed registered a phenomenon in mathematical space, instead opting to "simply note these puzzles as among the consequences of the 4CT." But the fact is that these puzzles are among the consequences of the 4CT only if that theorem really is accurately describable as an experiment. If we do not classify it as such, the puzzles disappear. Moreover, there are stark disanalogies between physical experiments and the use of computers in proving the 4CT (See, e.g., Levin (1981, 81), Swart (1980, 706), Teller (1980, 802). A rough characterization of what an experiment is in the physical sciences would be that it is an attempt to control and isolate part of nature, in order to learn, through observation, about that part of nature, or about some other part of nature to which the controlled and isolated part is causally connected. That is, an experiment, in the sciences in which the notion has a reasonably well-understood use, involves all of the following: the control and isolation of nature, observation, prediction, and the correlation of causes and effects. None of these features is present in the appeal to computers in the 4CT.

As to the second challenge, as one might expect, the proponent of this view of the a priori argues just that if the proof of the 4CT consists of a priori true propositions connected logically by valid a priori rules of inference, then the proof is a priori, even if one wishes to count it as being in some sense experimental (e.g., if one wished to count it as belonging to the subfield of experimental mathematics) (For a useful introduction to the field of experimental mathematics, see (Baker 2008), reprinted in (Marcus and McEvoy 2016). (Swart 1980) and (McEvoy 2013) offer detailed

arguments to the effect that CAPs cannot be considered experimental in any sense that occludes their apriority, even if they may be said to be experimental in certain senses consistent with that apriority.).

6 Response 4. Unsurveyable CAPS (Can) Give A Priori Knowledge

Tymoczko finds it obvious that the appeal to a computer renders one's knowledge of the 4CT a posteriori. As we have just seen, some of his critics argue that, while our knowledge of the proof is indeed a posteriori, due to the appeal to the computer report, the proof itself is an a priori proof, since each of the steps in the proof (assuming it is sound) is licensed by an a priori rule of inference. However, some philosophers have argued that we can actually (and not just in-principle) know a priori unsurveyable CAPs. For example, Burge (1998, 12–13) grounds his response to the New Four Color Problem on his Burge (1993) view of testimony. According to Burge (1998, 4), testimony, under the right conditions, transmits a priori warrant:

> If one's source knows a proposition a priori and there is no reason to doubt the source, one can in certain cases know it a priori by accepting the source's word**.**

For Burge, given that the source is sufficiently reliable, it "preserves content" and thus allows the hearer to have a priori warrant of a report, provided the source has a priori entitlement to its report. Barberousse and Vorms (2014, 3601) explain the point in regard to mathematical knowledge transmitted by human testimony or memory:

> The content of the knowledge is not altered by its being acquired through a process relying on memory, or testimony. . . . As a consequence, acquiring mathematical knowledge through memory or testimony does not ipso facto make the pieces of knowledge so obtained empirical. It remains a priori, whenever its justification (before its being transmitted through memory or testimony) was itself a priori.

Thus, if my mathematics teacher has a priori understanding of the proof of the infinity of primes, and informs me of this, then, on the assumption that I have no reason to doubt this testimony, I can come to have a priori warrant for my belief in the conclusion of this proof, despite not having come to know the proof by reading and understanding it.

How is this applied to a case like the 4CT, where only computers have surveyed the entire proof? Well, again, assuming the proof is in fact sound, and thus that the computer has correctly carried out the steps, then the computer's output report can "preserve content" for the reader who can thus come to have a priori warrant for her belief in the soundness of the theorem (See (Barberousse and Vorms 2014) for an attempt to extend Burge's arguments to computer simulations.).

Burge's view is controversial. There are two main sources of criticism. First, one may, with Hume, see testimonial knowledge as resting on empirical grounds, and thus reject the idea that testimony yields a priori knowledge under any circumstances. But even if one is willing to accept that human testimony can provide a hearer with a priori warrant, there are reasons to balk at so doing when the "testimony" is due to a computer**.** Burge (1998) discusses a number of such reasons, including the fact that the computer is not an autonomously rational agent, and the possibility of error in running the program. With respect to the latter, Burge offers a similar line to that we have already seen above: such errors are analogous to errors in humanly produced proofs. With respect to the former, Burge (1998, 9) appeals to the rationality and intentionality the computer derives from its human designers:

> The meaning of the machine's activities, and indeed the nature of its rational powers, are derivative from intentionality and rationality of rational beings. Such activities provide an amplification of the designers' and programmers' rationality. In understanding the machine's offerings, we have an apriori entitlement to rely on the rationality of these offerings, and the justifications backing these offerings.

Symons and Alvarado (2019, 50–51) raise another difference between human and computer "testimony," which they see as undermining claims of apriority. In a discussion of computer simulations in science they argue that because mathematical content is discretized by computers, simulations cannot count as preserving content in a way relevantly similar to human testimony and so, on Burge's account, cannot yield a priori knowledge. Thus:

> [C]omputer simulations cannot [preserve content] given Burge's characterization because justificatory elements distinct from the ones warranting the original content do in fact enhance, decrease or constitute the epistemic warrants of the manipulated content.

And, for Symons and Alvarado, given that content preservation is required for a priori testimonial knowledge, one's warrant from a computer simulation cannot be a priori. Symons and Alvarado (2019, 50 n.24) extend the argument to CAPs. First they observe that the alteration of content involved in coding introduces an empirical element. They then note a remark from McEvoy (2008, 380) that "[i]f the method of inference for any of the steps in the proof is a posteriori, [the proof] is a posteriori." Symons and Alvarado take these points together to concede the point that CAPs are a posteriori.

There are two parts to this objection, one specific to Burge's version of apriorism, with its attendant notion of content preservation and one more general, to the effect that the process of coding required to translate the premises of a proof into computer code (and from computer code back into mathematical propositions) will involve inferential steps that are not a priori mathematical steps, but are instead empirically grounded pieces of computer science.

Faced with the objection that CAPs do not preserve content, one might give up Burge's notion of content preservation. One might maintain instead the view (presented in Sect. 5) that once the steps of the proof are all a priori inferences from a

priori propositions, the proof itself remains a priori (even if our knowledge of it is a posteriori). This view does still face the more general objection, however. But the fundamental question for this approach that is raised by issues of transformation of content between formal proofs and coding is whether what the encoded program produces is a sound proof. If it is, then the proof is warranted by the a priori inferential steps between premises, and is a priori. If not, then we do not have a proof (and so the question of its apriority is moot).

Thus the view outlined in Sect. 5 above survives the objection. It is not, however, consistent with the general thrust of this section, that we can have a priori knowledge of unsurveyable proofs. One wishing to maintain this might instead turn to Arkoudas and Bringsjord (2007, 186) who argue that unsurveyable CAPs do not differ in terms of warrant from human produced surveyable proofs " as long as the algorithms that the computers execute are known to be correct in the traditional sense" ((Arkoudas and Bringsjord 2007, 195) point out "that rigorous, surveyable, and convincing proofs of correctness are possible for some algorithms is not controversial; it has been established by the work of theoretical computer scientists and logicians such as Hoare, Dijkstra, Manna, and Scott."). Once this condition is satisfied, then, granted that surveyable proofs can provide a priori mathematical warrant, unsurveyable CAPS can too (Arkoudas and Bringsjord 2007, 189). Arkoudas and Brinsjord exploit the fact that modern proof checking programs are compact enough to be themselves surveyable, allowing a priori knowledge of their correctness. They argue that if one knows a priori that the checker is correct, and one also knows that it has established some theorem—even by means of a humanly unsurveyable proof, then one can know the theorem a priori. They are quick to acknowledge that Appel and Haken's proof of the 4CT does not satisfy the condition; however, they maintain that the proof due to Gonthier (2005) — using the mechanized COQ proof checker — does (See also (Gonthier 2008).). Knowledge that is obtained in such manner, Arkoudas and Bringsjord (2007, 193) label "computational a priori":

> Knowledge obtained by computations is a priori as long as we know what the relevant algorithms do and can prove mathematically—using traditional, surveyable proofs—that the algorithms produce correct results. If that is the case, then the justificational structure of our knowledge is similar to that which obtains for knowledge produced by surveyable proofs."

7 Response 5: The Future

A different way of arguing that the proof of the 4CT is a priori despite being unsurveyable, and despite currently allowing only a posteriori knowledge of its conclusion is to appeal to future surveyable versions of the same proof. Here, for example, is Swart (1980, 700):

> As of now the 4CT itself (must be carried out by computer). But no one has proved that this is a permanent categorization, ... and the development of more powerful mathematical techniques may, in the course of time, change the existing categorization of theorems out of all recognition. Who can say what might not happen in a hundred or even fifty years' time?

> . . . Are we to assume that the 4CT is, as of now, an a posteriori theorem but that it may in time become an a priori one? This is surely a most odd use of the term a priori.

McEvoy (2008, 377–9) uses the possibility of notation change to illustrate the same issue. He notes that the (surveyable) CAP of the Robbins problem, originally formulated in standard Boolean notation, has been translated by Kauffman (2000) into a variant of C.S. Peirce's graph logic (For detailed and discussion of Peirce's graph logic, see (Kauffman 2000), and (Corfield 2003, pp. 53–55).). Kauffman's reformulation of the proof of the Robbins problem translates some extremely complex and lengthy Boolean algebra into far shorter and more easily comprehensible Peircean graphs. Kauffman himself holds that his translation "makes a mathematical domain accessible to human beings that has heretofore been only accessible to computers'" (Kauffman 2000). Thus, suitable notation change may render surveyable some previously unsurveyable proofs. For any such reformulated proof, a human mathematician might well survey, and thus come to know a priori the warranting inferential relations that obtain between premises and conclusions. It seems there are three things we might say about this situation

(a) The proof is and always has been an a priori proof.
(b) The proof *was* an a posteriori proof, but it has now become an a priori proof.
(c) The proof was and remains a posteriori.

Unless we have additional grounds for arguing that proofs are a posteriori (i.e., grounds independent of surveyability) (c) is clearly false, as the proof now has been reformulated as a surveyable a priori proof.

(b) seems equally implausible. We are here considering two formulations of the very same proof. The difference, apart from the difference in surveyability, is only a difference in notation. But the fact that the same sequence of reasoning expressed in different notation can be surveyed and so known a priori shows that the inferential relations obtaining between premises and conclusion in that sequence of reasoning are a priori. If the original proof were not a priori, merely translating equivalent symbols could not make it so.

Since (b) and (c) are both false, we are left with (a): the proof is and always was a priori, despite the unsurveyability of the original formulation of the proof. The falsity of (b) and (c) thus shows that the surveyability of a proof is not a necessary condition on its apriority. If this line of reasoning is correct, this tends to support the view outlined above in Sect. 5 that what makes a proof a priori is not the contingent fact that it can be surveyed and understood by beings like ourselves, but is instead the fact that the each of the steps in the proof is a priori warranted.

Such a line of thought clearly does not accept surveyability as being criterial for proof (or for apriority). And we have seen already Bassler (2006, 117) argue that this is mistaken (and that one fundamental insight of Tymoczko (1979) is that the 4CT drives a wedge between two of the main criteria of proof, surveyability and formalizability. The next section thus examines whether the 4CT might be in some sense surveyable.

8 Response 6: The CAP of the 4CT Is Surveyable, in a Sense

We have thus far been following along with the seemingly quite obvious assumption that the proof of the 4CT is unsurveyable. However, a number of authors have challenged this, at least when the unsurveyability of the 4CT is to be understood as the introduction of a new element in mathematics. Margarita Levin, for example, notes that the entirety of human-produced mathematics is, taken as a whole, unsurveyable by any one human mathematician. Of course, given that it is human-produced mathematics, it all has been surveyed by different mathematicians. Something similar could conceivably be done with the CAP of the 4CT. It is unsurveyable by any one human mathematician, but not by many of them acting together, each surveying a relatively small portion. Thus, the 4CT is, like the full set of human produced mathematics, unsurveyable by any one human mathematician, but surveyable by many acting together (Various other authors have suggested this as a method of surveying proofs too lengthy to be surveyed by any individual human mathematician. See, e.g., (Swart 1980, 705), (McEvoy 2008, 381).). Were this to be carried out in practice, the proof would be in the same boat as the above mentioned Finite Simple Groups proof, which, while fully human-produced, has not been surveyed by any individual mathematician, and consists of "connected results written down in various places, totaling 10,000 pages" (Barendregt and Wiedijk 2005, 2532).

More realistically, the 4CT may be — and indeed has been — surveyed by other computers. Swart (1980, 703) notes that even by that time, Appel and Haken's proof had been "indirectly surveyed in a manner very little different from the hand surveying of other proofs involving a large amount of case testing, namely, by means of independently written computer programs run on different computers." Swart concludes that the proof has been "fairly effectively surveyed" and finds the proof to be "no worse off in this regard than theorems concerning other equally intractable problems" (Swart 1980, 704); see also (Teller 1980, 798–9).

Krakowski (1980, 93) argues likewise:

> [I]nsofar as I understand the notion of surveyability, computers can survey proofs and that the proof of the crucial lemma has been surveyed. What is essentially required is that it be checked that each step follows from the previous one; and this the computer certainly has done.

Indeed, Krakowski (1980, 92) finds the suggestion that the proof has not been surveyed (because not humanly surveyed) to be chauvinistic.

Some critics have noted that this kind of computer verification of lengthy CAPs sets us off on a regress. Mackenzie (2005) notes that:

> [A]utomated theorem provers are themselves quite complicated computer programs which may contain faults in design or implementation. In interviews for this research, designers of automated theorem provers often reported experience of 'bugs' in their systems that would have allowed 'theorems' that they knew to be false nevertheless to be proven.

((Mackenzie 2001, 288) cites an interview with Alan Robinson, developer of the fundamental theorem-proving procedure of 'resolution'), in which Robinson colorfully inveighs against computer verification: "You've got to prove the theorem-proving program correct. You're in a regression, aren't you? And that's what people don't seem to realize when they get into verification. They have a hairy great thing they're in doubt about, so they produce another hairy great thing which is the proof that this one's ok. Now what about this one which you've just used . . . ? I say that serves them jolly well right. They shouldn't be doing it that way.") That is, if our worry is that an unsurveyable CAP may contain errors due to software or hardware errors, then surely these worries arise also for the computer program that carries out the verification. However, this particular worry may be overblown, as contemporary automated theorem checkers have relatively short programs that can fairly easily be surveyed to ensure that they are error free and working correctly. Arkoudas and Bringsjord (2007, 186) describe "small and simple" proof-checking algorithms which "emit low-level correctness proofs—called "certificates"—that can be verified by simple proof checkers." Once such a certificate is produced, then, we need "no longer care how the search was performed, in how many pages of C code it was encoded, or what resources it used; our only care is whether the certificate is valid" (Arkoudas and Bringsjord 2007, 199). The computer then need only "know" how to check the validity of the certificates, which can be done in first order logic with identity. The required proof-checking algorithm, they note is "of trivial conceptual and computational complexity. It can be implemented in a page or two of an advanced programming language such as ML or Haskell. . ." (Indeed, as the authors note, Georges Gonthier's 2005 proof of the 4CT, can be checked by small and simple proof checkers. The proof is available here: http://www2.tcs.ifi.lmu.de/~abel/lehre/WS07-08/CAFR/4colproof.pdf).

Here is how Hales et al. (2011) puts the point in connection with the HOL Light proof assistant on the Flyspeck homepage:

> The system is built on a small kernel, and the type checking insures that no fatal bugs can occur outside the kernel. That is, if you leave a bug in your code, the worst that can happen is that you will fail to prove a theorem you hoped to prove. You can never put a bug in your code that yields a false statement. The system becomes self-checking after the first thousand lines of code. The source code is short and quite readable (Hales (2008, pp. 1375–1376) notes that the kernel for HOL light involves fewer than 500 lines of computer code. See also Harrison (2008, 1405) and Pollack (1997).).

Once the kernel of the system has been shown to be free of bugs, one can be certain that a formal proof found to be sound by the assistant is in fact sound. Thus, with the advent of proof assistants, there is surveyability of a kind even for very long proofs: one surveys the proof assistant, and the proof assistant surveys the proof. Moreover, proofs produced or checked by one proof assistant can be translated and "exported" for checking by a different system (Hales 2008, 1376; Pollack 1997, 3).

To the extent that one shares Tymoczko's view that surveyability is important not only because it allows detection of error, but also *because a surveyed well understood proof is convincing*, these responses may be felt to miss the point. A recent

attempt to argue that the CAP of the 4CT is surveyable in a sense which may allow for a human mathematician to survey, understand and thus become convinced of the proof is due to Bassler (2006). (Bassler (2006, 100) first distinguishes between global and local surveyability:

> (L)ocal surveyability requires the surveying of each of the individual steps in a proof in some order, while global surveyability requires the surveying of the entire proof as a comprehensible whole.

This may not seem to take us very far. After all, one might well think that in order to survey an entire proof as a comprehensible whole, one must first survey (and understand) each of the individual steps. This sort of understanding of the relation between global and local surveyability, Bassler (2006 112) rejects as reductive, preferring instead a relation between the two types of surveyability that allows for a proof to be globally surveyable without being locally surveyable. Indeed, the CAP of the 4CT is taken to be an example of a globally, even if not locally, surveyable (Bassler 2006, 126). What makes the CAP of the 4CT unsurveyable are the computer calculations for the unavoidable configurations. As Bassler (2006, 123) observes, Appel and Haken have surveyed the proof with the exception of these calculations. Assuming that the calculations are sufficiently routine, then the proof is globally surveyable — that is, distinguishing between mere calculation instances and proof steps, the whole proof is surveyable as a comprehensive whole, even though it is not locally surveyable (where this *would* require surveying each individual calculation). Bassler (2006, 125) takes Haken to have offered an article length description of the proof that is "conceptually surveyable":

> In the course of an article-length description, Haken has arrived at a global description of the strategy by way of which he and Appel set up the proof of the 4CT. . . . The logical structure of the argumentation at the level of Haken's expository article is easily globally surveyable: we see what tasks are required in a global sense and why these tasks will collectively provide a proof of the 4CT.

For one convinced that human surveyability of a kind is criterial for proof, Bassler's view seems quite promising. If Bassler's non-reductive understanding of the distinction between global and local surveyability can be maintained, it allows for the elimination of Tymoczko's wedge between formalizability and (at least global) human surveyability. What, on this view, we are to say about whether the 4CT is a priori or empirical will depend on the details of the relationship between apriority, global surveyability and local surveyability. On the one hand, it may still make most sense for the apriorist to say that the 4CT is an a priori proof but is known only empirically. On the other hand, however, Bassler's view may make room for the idea that one can have a priori knowledge of the globally surveyed proof. And, if the calculations are sufficiently routine, and if one understands the overall character of these calculations (despite being unable to survey them all), globally surveying the proof might allow one, after all, to arrive at a kind of a priori knowledge of the 4CT.

9 Certainty

I will close with some brief remarks on certainty. Tymoczko (1979, 77–8) tells us that that our knowledge must be qualified by the uncertainty of our instruments. And, there are, he tells us, truths from electrical engineering which have a higher degree of certainty than the 4CT. Tymoczko takes our lack of certainty regarding our instruments undermine that aspect of the traditional view of mathematics that its truths are certain, in contrast to those of the empirical sciences. "Certainty" here may mean one of two things. On the first reading, "certainty" means objective certainty. To say, in this sense, that the truths of mathematics are certain is to say that their truth is objectively certain, regardless of what we happen to think about them (For more on the notion of objective certainty, see (Katz 1998, 63–64).). It is objectively certain, for example, that seven times seven is 49, and this truth would be certain even if everyone came to believe otherwise. In this sense of certainty, the truths of mathematics are objectively certain, and thus cannot be less certain than any truths of electrical engineering. On the assumption that the proof of the 4CT is a truth of mathematics, it cannot be less objectively certain than any other.

On the other hand, by "certainty" Tymoczko may have in mind our degree of confidence in the truth of the 4CT. In this sense, he is correct to say that there are truths of engineering with a higher degree of certainty than the 4CT. But it isn't clear that this represents a novel threat to the traditional view. Our confidence may be higher in certain obvious empirical truths than in the correctness of very long proofs, but this is true even for humanly produced proofs, like the Finite Simple Groups proof. Moreover, as others have noted (e.g., Arkoudas and Bringsjord (2007, 200), Barberousse and Vorms (2014, 3602)), the computing powers of modern computers, in comparison to the human capacity for fatigue, distraction, and so on, suggest that computers are more reliable than humans for many tasks. Granted that the production of lengthy proofs is one such task, we should actually have more confidence (i.e., hold in a higher degree of certainty) in CAPs than in lengthy and complex human proofs.

References

Appel K (1984) The use of the computer in the proof of the four color theorem. Proc Am Philos Soc 128(1):35–39

Appel K, Haken W, Koch J (1977) Every planar map is four colorable. Ill J Math 21:429–567

Arkoudas K, Bringsjord S (2007) Computers, justification, and mathematical knowledge. Mind Mach 17(2):185–202

Azzouni J (1994) Metaphysical myths, mathematical practice. Cambridge University Press, New York

Baker A (2008) Experimental mathematics. Erkenntnis 68(3):331–344

Barberousse A, Vorms M (2014) About the warrants of computer-based empirical knowledge. Synthese 191(15):3595–3620

Barendregt H, Wiedijk F (2005) The challenge of computer mathematics. Philos Trans 363(1835): 2351–2375. The Nature of Mathematical Proof

Barwise J (1989) Mathematical proofs of computer system correctness. Not Am Math Soc 36: 844–851
Bassler OB (2006) The surveyability of mathematical proof: a historical perspective. Synthese 148(1):99–133
Brown GS (1972) Laws of form. Br J Philos Sci 23(3):291–292
Burge T (1993) Content preservation. Philos Rev 102:457–488
Burge T (1998) Computer proof, apriori knowledge, and other minds. Philos Perspect 12:1–37
Corfield D (2003) Towards a philosophy of real mathematics. Cambridge University Press, Cambridge
Detlefsen M, Luker M (1980) The four-color theorem and mathematical proof. J Philos 77(12): 803–820
Gödel K (1983) What is cantor's continuum problem? In: Paul B, Hilary P (eds). Cambridge University Press, Cambridge, pp 470–485. (Reprinted in Philosophy of Mathematics, 2nd ed.)
Gonthier G (2005) A computer-checked proof of the four colour theorem. http://www2.tcs.ifi.lmu.de/~abel/lehre/WS07-08/CAFR/4colproof.pdf
Gonthier G (2008) Formal proof – the four-color theorem. Notices AMS 55(11):1382–1393
Hales T (2008) Formal proof. Notices AMS 55(11):1370–1380
Hales T et al (2011) The flyspeck project. http://code.google.com/p/flyspeck/wiki/FlyspeckFactSheet
Harrison J (2008) Formal proof – theory and practice. Notices AMS 55(11):1395–1406
Hirsch M et al (1994) Responses to "theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics". In: Jaffe A, Quinn F (eds) Bulletin of the American Mathematical Society Volume 30, Number 2, pp 178–207
Jaffe A, Quinn F (1993) 'Theoretical mathematics': toward a cultural synthesis of mathematics and theoretical physics. Bulletin (New Series) AMS 29(1):1–13
Johansen MW, Misfeldt M (2016) Computers as a source of a posteriori knowledge in mathematics. Int Stud Philos Sci 30(2):111–127
Katz JJ (1998) Realistic rationalism, Bradford
Kauffman L (2000) The Robbins Problem – computer proofs and human proofs. http://homepages.math.uic.edu/~kauffman/Robbins.htm
Krakowski I (1980) The four color problem reconsidered. Philos Stud 38(1):91–96
Kripke S (1972) Naming and necessity. Harvard University Press, Cambridge, MA
Levin M (1981) On Tymoczko's argument for mathematical empiricism. Philos Stud 39:79–86
MacKenzie D (1999) Slaying the kraken: the sociohistory of a mathematical proof. Soc Stud Sci 29(1):7–60
MacKenzie D (2001) Mechanizing proof: computing, risk and trust. MIT Press, Mass
MacKenzie D (2005) Computing and the cultures of proving. Philos Trans A Math Phys Eng Sci 363(1835):2335–2347; discussion 2347–50. PMID: 16188609. https://doi.org/10.1098/rsta.2005.1649
Marcus R, McEvoy M (eds) (2016) An historical introduction to the philosophy of mathematics. Bloomsbury Academic
McEvoy M (2008) The epistemological status of computer proofs. Philos Math 16:374–387
McEvoy M (2013) Experimental mathematics, computers and the a priori. Synthese 190(3): 397–412
Pollack R (1997) How to believe a machine-checked proof. Basic Res Comput Sci 97(Report Series):1–18
Prawitz D (2008) Proofs verifying programs and programs producing proofs. In: Lupacchini R, Corsi G (eds) Deduction, computation, experiment: exploring the effectiveness of proof. Springer, Berlin/Helderberg/New York, pp 81–94
Rav Y (1999) Why do we prove theorems? Philos Math 7(1):5–41
Resnik M (1997) Mathematics as a science of patterns. Oxford University Press, New York
Rota G (1997) The phenomenology of mathematical proof. Synthese 111(2):183–196

Secco and Pereira GD (2017) Proofs versus experiments: wittgensteinian themes surrounding the four-color theorem. In: Silva M (ed) How Colours matter to philosophy. Springer, Cham, pp 289–307. https://www.semanticscholar.org/paper/Proofs-Versus-Experiments%3A-Wittgensteinian-Themes-Secco-Pereira/15f28edc77c7f5c54f0dea0db67e8fea8816710f

Shanker SG (1986) The Appel-Haken solution of the four-color problem. In: Shanker (ed) Ludwig Wittgenstein: critical assessments, vol 3. Croom Helm, London, pp 395–412

Swart ER (1980) The philosophical implications of the four-color problem. Am Math Mon 87(9): 697–707

Symons J, Alvarado R (2019) Epistemic entitlements and the practice of computer simulation. Mind Mach 29(1):37–60

Teller P (1980) Computer proof. J Philos 77(12):797–803

Tymoczko T (1979) The four-color problem and its philosophical significance. J Philos 76:57–82

Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice

Giovanni Ferraro

Contents

1 Summary and Introduction 2482
2 Proofs by Example 2482
3 Examples as Persuasion Tools 2493
Reference 2501

Abstract

In this chapter, I will discuss the role of examples in Euler's mathematical practice. In modern mathematics, examples have the function of clarifying findings, procedures, and theorems; instead, in Euler's writings, they could also have two other different functions. Firstly, he used examples as a proof for a general statement, namely, he considered a proposition as proved if verified for some examples. Euler felt the need of general and rigorous proofs even pointed out that proofs by example could give rise to errors; however, he never rejected proofs by example entirely and limited himself to advising caution in their use. Secondly, Euler used examples as verification and persuasion tools. In certain cases, he checked the validity of a proof by appropriate examples. In these cases, examples did not substitute the general proof of a statement; rather their aim was to convince the reader of the correctness of certain procedures of the calculus and of the internal coherence of the various branches of the analysis of infinities.

Keywords

Euler · Proofs by example · Arbitrary exemplification · Inductive generalizations · Proofs and persuasion

G. Ferraro (✉)
University of Molise, Campobasso, Italy
e-mail: giovanni.ferraro@unimol.it

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_13

1 Summary and Introduction

In this chapter, I will discuss the role of examples in Euler's mathematical practice. (For a discussion of the main features of Euler's mathematics, and, in particular, of what can be termed as Euler's analytical program, see Capobianco et alia (2017); Ferraro (2020).) In modern mathematics examples have the function of clarifying findings, procedures, theorems; instead, in Euler's writings they could also have other different functions: more precisely, Euler used examples as proofs and as persuasion tools.

In Sect. 2, I will deal with the use of examples as a proof for a general statement, namely, I will investigate the praxis by which a proposition was considered as proved if verified for some examples. I will also show that Euler felt the need of general and rigorous proofs, namely, the need of deductions strictly derived by principles and not by the observations of particular cases. Euler even pointed out that proofs by example could give rise to errors; however, he never rejected proofs by example entirely and limited himself to advising caution in their use.

In Sect. 3, I will investigate the use of examples as verification and persuasion tools. In certain cases, Euler gave a proof of a theorem, but then he checked the validity (the "truth" – he said) of that theorem by appropriate examples. More specifically, given a theorem *T*, Euler first applied this theorem to some specific problems, which could also be solved by means of another theorem *Q* (usually, a more elementary theorem than *T*). Then he verified that that the different theorems gave the same result. In these cases, examples did not substitute the general proof of a statement; rather their aim was to convince the reader of the correctness of certain procedures of the calculus (though the doubts on the foundations of this field of mathematics). The use of examples as persuasion and verification tools also showed the internal coherence of the various branches of the analysis of infinities.

2 Proofs by Example

In his *The Calculus as Algebraic Analysis*, Craig G. Fraser investigated the mathematical work of Euler and Lagrange and observed:

> A theorem is often regarded as demonstrated if verified for several examples, the assumption being that the reasoning in question could be adapted to any other example one chose to consider. (Fraser (1989, p. 328))

(On Lagrange's, see Ferraro and Panza (2012)) Fraser's words highlight a remarkable feature of Euler's mathematical practice: the proof by example. However, it is to be observed that the use of an example as a proof did not necessarily imply an inappropriate generalization. Indeed, a distinction should be drawn between arbitrary exemplifications and inductive generalizations. Let us take the function $y = x^3$. In order to calculate the differential ratio $\frac{dy}{dx}$ of this function, a seventeenth-century mathematician could consider the quantity. (In this chapter, following Euler, I use the term "differential ratio" to denote what today is named

as "derivatives." Euler understood the differential ratio of the function $y(x)$ as the ratio of two differentials, dy and dx.)

$$y(x+\omega) = (x+\omega)^3.$$

Then he could calculate the ratio

$$\frac{(x+\omega)^3 - x^3}{\omega} = 3x^2 + 3x\omega + \omega^2$$

Set $\omega = dx$, he could applied the principle of cancellation of infinitesimals and so he would obtain the differential ratio:

$$\frac{dy}{dx} = 3x^2.$$

In the above procedure no specific property of the natural number 3 is used; for this reason, the number 3 can be replaced by any other natural number n and so the above reasoning can be considered as the proof that the differential ratio of $y = x^n$ is

$$\frac{dy}{dx} = nx^{n-1}$$

for any natural number n. Indeed, it suffice to calculate the nth power of $(x+\omega)^n$ – by using Pascal's triangle – and to repeat the procedure.

This is a case of arbitrary exemplification: one presents a typical example, or a few examples, in place of a general proof. In an arbitrary exemplification the demonstration of a general proposition

$$P(x), \forall x \in \mathrm{A},$$

where A is a given set, is reduced to the proof of some particular proposition

$$P(a_1), P(a_2), P(a_3), \ldots, P(a_n)$$

where a_1, a_2, a_3, ..., a_n are sufficiently generic or paradigmatic examples without any apparent particularity toward the statement.

Now let us consider the function $y = x^\alpha$ where α is a real number. (I use the term "real number" for the sake of simplicity. The notion of real numbers did not exist in seventeenth century. Euler would have spoken of any quantity (rational or irrational), but this is not essential in our discussion.) The above described procedure cannot easily be generalized to the case where the exponent is a real number α. The function $y = x^3$ is not an arbitrary exemplification of $y = x^\alpha$ when α is a real number. Therefore, if one deduces the rule

$$\frac{dy}{dx} = \alpha x^{\alpha-1}$$

for any real number α, from the example $\frac{dy}{dx} = 3x^2$, then one does not perform an arbitrary exemplification. In this case, it is preferable to speak of inductive generalization.

In arbitrary exemplifications, particular examples are used in place of a general proof of a theorem: however, the selected examples are sufficiently generic and have no any apparent particularity toward the statement to be proved. In other terms, the examples have properties (at least as concerns the properties used in the proof) that make them indistinguishable from any other example that one could be selected. Consequently, the reasoning or procedure used in those examples is repeatable in every other example and the validity of the general statement is guaranteed.

Instead, in inductive generalizations, a general statement arises from the investigation of some particular examples that are not sufficiently generic, that is, they are not representative of all possible cases. Inductive generalization is, in fact, a form of induction similar, in some way, to that of the empirical sciences.

Proofs by example were certainly typical of the algebraic tradition. (For example, see (Bombelli 1966).) In Euler's mathematics, the use of examples in place of a proof (both inductive generalizations and arbitrary exemplifications) could take a rather complex form.

Let us consider Euler's discussion of the partial fraction decomposition of rational functions in his *Introductio in analysin infinitorum*. For instance, in proposition 39, Euler considered a proper rational function $\frac{f(x)}{g(x)}$ such that $g(x) = p(x) \cdot q(x)$ and $p(x)$ and $q(x)$ have no common factor. He stated that any rational function of this type may be decomposed into the sum of two proper rational functions, the denominators of which are equal to $p(x)$ and $q(x)$, respectively. He also noted that such a decomposition is unique (Euler 1748, I, p. 23).

Euler did not provide a general demonstration of this proposition since "the truth of this matter may be understood more clearly from an example than by reasoning." ("rei veritas ex exemplo clarius quam per ratiocinum perspicietur" (Euler 1748, I, p. 23).) As an example Euler (1748, I, pp. 23–24) considered the function

$$\frac{1 - 2z + 3z^2 - 4z^3}{1 + 4z^4} \tag{1}$$

He stated that since $1 + 4z^4 = (1 + 2z + 2z^2)(1 - 2z + 2z^2)$, this function could be written as

$$\frac{\alpha + \beta z}{1 + 2z + 2z^2} + \frac{\gamma + \delta z}{1 - 2z + 2z^2},$$

where α, β, γ, δ were appropriate constants. Indeed, by supposing that

$$\frac{1-2z+3z^2-4z^3}{1+4z^4}=\frac{\alpha+\beta z}{1+2z+2z^2}+\frac{\gamma+\delta z}{1-2z+2z^2}.$$

Euler solved this equality for α, β, γ, δ, by equating the coefficients of terms involving the powers of x and found

$$\alpha=\frac{1}{2};\beta=-\frac{5}{4};\gamma=\frac{1}{2};\delta=-\frac{3}{4}.$$

In this example Euler did not rely on the specific properties of the particular rational function $\frac{1-2z+3z^2-4z^3}{1+4z^4}$, which played the role of a generic function satisfying the hypothesis of the theorem. It was implicit that the procedure used in that example could be repeated in any other example. In the subsequent sections of the *Introductio in analysin infinitorum* (1748, I, pp. 24–35), Euler modified the hypothesis on the function $\frac{f(x)}{g(x)}$ and thus provided other theorems on the decomposition of rational functions. He never gave general proofs of these propositions but always illustrated them by appropriate examples.

A case of inductive generalization can be found in *Consideratio progressionis cuiusdam ad circuli quadraturam inveniendam idoneae*. In this paper (1739, pp. 357–359), Euler computed the approximate value of π by means of the asymptotic series. He considered $\arctan t=\int_0^t\frac{dt}{1+t^2}$ as a sum of the type $\sum\frac{dt}{1+t^2}$ and, using relations concerning the sums of powers of integer numbers, derived

$$\pi=\frac{4n}{n^2+1}+\frac{4n}{n^2+4}+\ldots+\frac{4n}{n^2+n^2}+\frac{1}{n}+\sum_{h=o}^{\infty}(-1)^h B_{4h+2}\frac{1}{2^{2h}(2h+1)n^{4h+2}} \quad (2)$$

where B_n are the Bernoulli numbers. (Bernoulli numbers B_r are defined by the relation $\frac{t}{e^t-1}=\sum_{r=0}^{\infty}\frac{B_r}{r!}t^r$.) Euler understood the asymptotic nature of series (2). He observed that, although (2) seems to be the more convergent the greater the number n, it actually "converges" (he meant that the terms decrease in absolute value) only up to a certain term, after which it begins to diverge (namely, the terms increase). (This statement is based upon the observation that $\frac{B_{2h+2}}{B_{2h}}>\frac{(h-1)(2h-3)}{2\pi^2}$ (for $h>0$) and that if $h=\infty$ the ratio holds $\frac{h^2}{\pi^2}$ (i.e., $\frac{B_{2h+2}}{h^2B_{2h}}$ approaches asymptotically $\frac{1}{\pi^2}$). On these relations, see Euler (1755, pp. 428–430).) Euler suggested using (2) by calculating its first terms (according to Euler's terminology: the terms that "converge") and neglecting them as soon as they "begin to diverge" (1739, p. 357). More specifically, the terms of (2) must be considered up to the index $s\approx\frac{\pi n}{\sqrt{2}}$ (approximately $s\approx 2n$).

Euler gave the value 1, 3, 5 to n. He noted that, in such cases, the number of exact figures of π was about $3n$. Basing only on these three examples, Euler stated that, by induction (*per inductionem*), formula (2) provided about $3n$ exact figures of π. Euler

validated this inductive generalization by considering three other examples (precisely $n = 2, 4, 6$).

A more complex case of inductive generalization is found in chapter I of the first part of the *Institutiones calculi differentialis*, where the validity of a theorem P_f for a generic function f is inferred from the validity of the proposition P_f for some functions (Euler, 1755, pp. 3–38). In this chapter of the *Institutiones*, Euler dealt with the first differences of a function $y(x)$

$$\Delta y = y(x + \omega) - y(x)$$

where ω is the finite increment of the independent variable x. He stated that if $y(x)$ is any function, both algebraic and transcendent, then $\Delta y = y(x + \omega) = y(x)$ can be expressed in the form:

$$\Delta y = P\omega + Q\omega^2 + R\omega^3 + S\omega^4 + \text{etc.} \tag{3}$$

where P, R, S, ... are appropriate functions of x.

Formula (3) played a crucial role in the Eulerian foundation of the calculus, since the definition of differential ratio $\frac{dy}{dx}$ was based upon it. I do not dwell on this. Here I limit myself to observing that Euler derived (3) by induction on some specific cases.

Indeed, Euler firstly proved formula (3) for the functions $y = x^n$, where n was an integer, by expanding $(x + \omega)^n$ (see Euler 1755, pp. 13–16). Because of the additivity of the operator Δ, this was sufficient to prove that (3) held for integer functions (Euler 1755, p. 11).

Secondly, Euler (1755, p. 18) stated that formula (3) could be generalized to the functions $y = x^n$ for any number n. In effect, he (1755, 19) provided eight examples:

$$y = x^{-1}, \quad y = x^{-2}, \quad y = x^{-3}, \quad y = x^{-4},$$

$$y = x^{\frac{1}{2}}, \quad y = x^{\frac{1}{3}}, \quad y = x^{-\frac{1}{2}}, \quad y == x^{-\frac{1}{3}}.$$

Thirdly, Euler (1755, pp. 20–22) showed that (3) was valid for the functions

$$\frac{1}{a^2 + x^2} \text{ and } \frac{1}{\sqrt{a^2 + x^2}}.$$

He thought that this fact was sufficient to infer (*colligimus*) that any algebraic functions could be expressed in the form (3) (Euler 1755, p. 22).

Finally, Euler stated that (3) was valid for transcendental functions, too:

> Nor are the differences of transcendental functions excluded from this form, what will appear clearer from the following examples. (Neque etiam ex hac forma differentiae functionum trascendentium excluduntur, id quod ex sequentibus exemplis clarius apparebit (Euler 1755, p. 22).)

The examples concerned the logarithmic, exponential, sine, and cosine functions. Even if Euler's argument was affected by his conception of a function, (In Euler's conception, the only functions in the strict sense of the term were elementary functions and their composition (see Fraser 1989; Panza 1996, Ferraro 2007a).) it is clear that a general proof of (3) was lacking. The validity of (3) for any function was a mere inductive generalization grounded upon several examples.

Euler had often used inductive procedures in his early years, but after 1750 these procedures were felt to be problematic and not entirely reliable. For instance, let us consider the binomial series

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{1\cdot 2}x^2 + \frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}x^3 + \dots \tag{4}$$

Formula (4) was already well known in the case that n was a natural number when Newton generalized to the case where n was any number. Newton had given no proof of (4) for a real number n and had derived it by inductive generalization (see Newton 1712, pp. 67–86 and 1967, 1, pp. 104–111).

In his *Demonstratio theorematis newtoniani* (1774), Euler attempted to prove (4). He preceded his proof with an explanation why an inductive generalization was not a sufficient proof of the binomial theorem and it was necessary what he called a "rigorous" proof. He (1774, pp. 104–105) noted that certain formulas $F(n)$ were true for a natural number n; however, they became false when n was a non-integer number. This is the case of the series:

$$\begin{aligned}&\frac{1-a^n}{1-a} + \frac{(1-a^n)(1-a^{n-1})}{1-a^2} + \frac{(1-a^n)(1-a^{n-1})(1-a^{n-2})}{1-a^3}\\ &+ \frac{(1-a^n)(1-a^{n-1})(1-a^{n-2})(1-a^{n-3})}{1-a^4} + \dots\end{aligned} \tag{5}$$

Euler calculated the sum $f(n)$ of (5) for $n = 0, 1, 2, 3$, and found:

$$f(0) = 0;$$

$$f(1) = \frac{1-a}{1-a} = 1;$$

$$f(2) = \frac{1-a^2}{1-a} + \frac{(1-a^2)(1-a)}{1-a^2} = 1 + a + 1 - a = 2$$

$$\begin{aligned}f(3) &= \frac{1-a^3}{1-a} + \frac{(1-a^3)(1-a^2)}{1-a^2} + \frac{(1-a^3)(1-a^2)(1-a)}{1-a^3}\\ &= 1 - a + a^2 + 1 - a^3 + 1 - a - a^2 + a^3 = 3.\end{aligned}$$

In the same manner – he stated – one could show that

$$f(n) = n \text{ for any natural number } n.$$

An inductive generalization would lead us to the conclusion

$$f(n) = n \text{ for any } n.$$

But Euler observed $f\left(\frac{1}{2}\right)$ was not equal to $\frac{1}{2}$. It was clear that inductive generalization did not hold in this case. For this reason, a rigorous proof of (4) was necessary when n was not a natural number.

To modern eyes, example (5) appears unnecessarily complicated. In order to have a function that equals to n for an integer n and differs from n for a non-integer n, it suffices to consider

$$g(n) = n + \sin 2\pi n.$$

This function was known by Euler already when he wrote *De serierum determinazione seu nova methodus inveniendi terminos generales serierum* (1750). In this paper Euler had noticed that not only could the general term $y(n)$ of the series

$$1 + 2 + 3 + 4 + 5 + \ldots$$

be expressed by means of the function

$$y(n) = n \text{ for a positive integer } n,$$

but also it could be given by a formula of the type

$$y = n + \sum_{i=0}^{\infty} \mathrm{F}_i(n) \sin i\pi n,$$

where $\mathrm{F}_i(n)$ is any function of n. In the last series, when n is not integer, $y(n)$ assumes different values depending on how we choose the function $\mathrm{F}_i(n)$.

Another interesting case is the following. In his *Institutiones calculi differentialis*, Euler derived the formula today known as Euler-Maclaurin summation formula. (On Euler-Maclaurin summation formula, see (Ferraro 1998).) In section 111, he obtained the series

$$S(z) = \int z dx + \alpha z + \beta \frac{dz}{dx} + \gamma \frac{d^2 z}{dx^2} + \delta \frac{d^3 z}{dx^3} + \varepsilon \frac{d^4 X}{dx^4} + \ldots \tag{6}$$

where $z(x)$ was a generic function (of course, the term "function" has to be understood in Eulerian sense), $S(z) = \sum_{n=1}^{x} z(n)$ and α, β, γ, δ, ε ... were appropriate coefficients (Euler 1755, pp. 408–409). These coefficients were iteratively defined by an infinite sequence of equations. Euler (1755, pp. 409–410) explicitly provided the first six equations:

$$\alpha - \frac{1}{2} = 0,$$

$$\beta - \frac{\alpha}{2} + \frac{1}{6} = 0,$$

$$\gamma - \frac{\beta}{2} + \frac{\alpha}{6} - \frac{1}{24} = 0,$$

$$\delta - \frac{\gamma}{2} + \frac{\beta}{6} - \frac{\alpha}{24} + \frac{1}{120} = 0,$$

$$\varepsilon - \frac{\delta}{2} + \frac{\gamma}{6} - \frac{\beta}{24} + \frac{\alpha}{120} - \frac{1}{720} = 0,$$

$$\zeta - \frac{\varepsilon}{2} + \frac{\delta}{6} - \frac{\gamma}{24} + \frac{\beta}{120} - \frac{\alpha}{720} + \frac{1}{5040} = 0.$$

Thus Euler was able to compute the first five coefficients α, β, γ, δ, ε:

$$\alpha = \frac{1}{2}, \beta = \frac{1}{12}, \gamma = 0, \delta = -\frac{1}{720}, \varepsilon = 0.$$

The observation of the first terms of the sequence of the coefficients led Euler to state that the subsequent odd terms were equal to zero (1755, p. 410). However, according to Euler, "this sequence of values seems to violate the law of continuity" (haec valorum series contra legem continuitatis impingere videtur (Euler 1755, p. 410).).

At a first glance, induction seemed to be in contrast with continuity (understood in Euler's sense, since the law of formation of coefficients could not be reduced to one single formula). (Euler understood the law of continuity as the uniqueness of the law that expresses a certain relationship. See (Ferraro 2000a; Panza 1996).) For this reason, Euler thought that in this case induction was not sufficient and it was necessary to prove rigorously that all odd terms, except for the first, necessarily vanished (Euler 1755, p. 410).

In order to obtain a "rigorous" demonstration, Euler set

$$V = 1 + \alpha u + \beta u^2 + \gamma u^3 + \delta u^4 + \varepsilon u^5 + \ldots,$$

where α, β, γ, δ, ε, ... were the coefficients of (6), and showed that

$$V - \frac{1}{2}u = \frac{u\left(e^{\frac{u}{2}} + e^{-\frac{u}{2}}\right)}{2\left(e^{\frac{u}{2}} - e^{-\frac{u}{2}}\right)}$$

Since $\frac{u\left(e^{\frac{u}{2}}+e^{-\frac{u}{2}}\right)}{2\left(e^{\frac{u}{2}}-e^{-\frac{u}{2}}\right)}$ was equal to

$$\frac{1+\frac{u}{2\cdot 4}+\frac{u^4}{2\cdot 4\cdot 6\cdot 8}+\frac{u^6}{2\cdot 4\cdot\ldots\cdot 12}+\frac{u^8}{2\cdot 4\cdot\ldots\cdot 16}+\ldots}{1+\frac{u^2}{4\cdot 6}+\frac{u^4}{4\cdot 6\cdot 8\cdot 10}+\frac{u^6}{4\cdot\ldots\cdot 14}+\frac{u^8}{2\cdot 4\cdot\ldots\cdot 18}+\ldots}$$

and there was no odd power in this fraction, the expansion of $V-\frac{1}{2}u$ had only even powers of u. In this way, according to Euler, it was proved that all odd terms of (6), except for the first, were equal to zero (Euler 1755, pp. 410–413).

Euler was aware that induction can lead to errors. He himself pointed out a "remarkable example of false induction" (*exemplum memorabile inductionis fallacis*) in his *Observationes analyticae* (1765). In this paper, Euler investigated the coefficients of x^j in the expansion of $(1+x+x^2)^n$. Today these coefficients are termed as trinomial coefficients and are denoted by $\binom{n}{k}_2$, with $n \geq 0$ and $-n \leq k \leq n$. In particular, Euler investigated the central trinomial coefficient $\binom{n}{0}_2$ and proved that

$$\binom{n}{0}_2 = 1 + nx + \frac{n(n-1)}{1\cdot 1} + \frac{n(n-1)(n-2)(n-3)}{1\cdot 2\cdot 1\cdot 2} + \frac{n(n-1)(n-3)(n-4)(n-5)}{1\cdot 2\cdot 3\cdot 1\cdot 2\cdot 3} + \ldots$$

Euler considered the first 10 terms of the sequence $\binom{n}{0}_2$ and wrote the rows:

1	1	3	7	19	51	141	393	1107	3139	...(T)
	3	3	9	21	57	153	423	1179	3321	… (B)
	2	0	2	2	6	12	30	72	182	… (C)

The row (T) is the sequence $\binom{n}{0}_2$ for $n = 0, 1, 2, \ldots, 9$. The nth entry of the row (B) is obtained multiplying (T) by 3 and shifting one to the right. The third row is the difference between the first and the second row starting with the two-entry column. According to Euler, "not without reason it seems to turn out" (*non sine ratione evenire videtur*) that all the numbers in the row (C) are of the type $m^2 + m$, where m is equal

$$1 \quad 0 \quad 1 \quad 1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \ldots (\mathrm{F})$$

Euler observed that (F) is a recurrent sequence, and any term is the sum of two preceding terms (in effect, it is the Fibonacci sequence with the first two terms $F_{-1} = 1$ and $F_0 = 0$). (Euler did not use the expression "Fibonacci sequence.") Therefore, the investigation of the first 10 terms of the given sequence leads one to a rule, which in modern terms can be written as

$$3\binom{n+1}{0}_2 - \binom{n+2}{0}_2 = F_n(F_n+1), \quad n=-1,0,1,2,\ldots \tag{7}$$

Euler rhetorically asked:

> "since the rule is obtained by examination of the first ten terms of the given sequence, who can ever doubt that the same rule holds for the whole of series? Less certain inductions were often successful" (Qui ordo cum in decem primoribus terminis deprehendatur, quis dubitaverit eundem universae seriei tribuere? Saepe profecto inductiones minus certae successu non fuerunt destitutae (Euler 1765, p. 54).)

By using the property of the Fibonacci sequence, Euler derived the following direct formula for $\binom{n}{0}_2$:

$$\binom{n}{0}_2 = \frac{3^n+(-1)^n}{10} + \frac{1}{5}\left(\frac{3+\sqrt{5}}{2}\right)^n + \frac{1}{5}\left(\frac{3-\sqrt{5}}{2}\right)^n + \frac{1}{5}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{1}{5}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

Then, he provided a recursive formula

$$\binom{n}{0}_2 = 6\binom{n-1}{0}_2 - 8\binom{n-2}{0}_2 - 8\binom{n-3}{0}_2 + 14\binom{n-4}{0}_2 + 4\binom{n-5}{0}_2 - 5\binom{n-6}{0}_2$$

At this point, Euler observed that these formulas, obtained by induction, were false. Indeed, the 11th term of the sequence the sequence $\binom{n}{0}_2$ is 8953; however, the formula (7) is false. Indeed

$$8953 - 3\cdot 3139 = 464 \neq 21^2 + 21 = 462 \text{ (Euler, 1765, p.55)}.$$

(On Euler's example of misleading induction, see Andrews (1990).) Euler returned to this question in Varia artificia in serierum indolem inquirendi (1783). In this paper, he observed that the above example of false induction "was worthy of being noted because no similar cases had been found in which a so beautiful induction failed." Since induction is misleading, Euler investigated the sequence of the central trinomial coefficients by its same nature (*ex ipsa eius natura*) (Euler 1783, p. 386). In effect Euler was able to prove easily that:

$$\binom{n}{0}_2 = \binom{n-1}{0}_2 + \frac{n-1}{n}\left(\binom{n-1}{0}_2 + 3\binom{n-2}{0}_2\right).$$

However, Euler did not reject this kind of induction completely, rather he preferred to limit himself to advising caution in using it (Euler 1783, p. 51).

Now, let us go to examine how Euler dealt with mathematical induction. Let us consider the following theorem. Given the continued fraction

$$a_1 + \cfrac{b_1}{a_2 + \cfrac{b}{a_3 + \cfrac{b_3}{a_4 + \ldots}}}$$

and let

$$C_k = \frac{p_k}{q_k} = a_1 + \cfrac{b_1}{a_2 + \cfrac{b}{\ldots + \cfrac{b_{k-1}}{a_k}}},$$

then the numerator p_k and the denominator q_k of the convergent C_k are defined by the iterative relationship

$$\begin{cases} p_k = a_k p_{k-1} + b_{k-1} p_{k-2} \\ q_k = a_k q_{k-1} + b_{k-1} q_{k-2} \end{cases}, \tag{8}$$

with $p_1 = a_1$, $p_0 = 1$, $q_1 = 1$, $q_0 = 0$.

In *De fractionibus continuis dissertatio*, Euler (1737, p. 103) proved the theorem by computing the first four convergents C_1, C_2, C_3, C_4. Euler wrote the convergents according the following Scheme **S** (of course, Euler did not use the indicial notation but the 10 letters a, b, c, d, e, α, β, γ, δ, ε):

$$\begin{array}{ccccc} a_1 & a_2 & a_3 & a_4 & a_5 \\ \frac{1}{0} & \frac{a_1}{1} & \frac{a_1a_2+b_1}{a_2} & \frac{a_1a_2a_3+a_3b_1+a_1b_2}{a_2a_3+b_2} & \frac{a_1a_2a_3a_4+a_3a_4b_1+a_1a_4b_2+a_1a_2b_3+b_1b_3}{a_2a_3a_4+a_4b_2+a_2b_3} \\ b_1 & b_2 & b_3 & b_4 & b_5 \end{array}$$

where the first fraction $C_0 = \frac{1}{0}$ gives the initial conditions $p_0 = 1$, $q_0 = 0$.

Euler (1737, p. 104) formulated the recursive law (8) verbally: the product of a numerator with the corresponding upper indicator, added to the product of the preceding numerator with the lower indicator, yields the succeeding numerator. A similar rule applies to the denominators. Then, Euler stated:

> This law is easily observed from the inspection of these fractions if they are continued further, but it is possible to deduce their nature from the continued fractions themselves; however, I believe there is no point in including the proof here. (Lex quidem haec ex ipsa inspectione harum fractionum, si ulterius continentur, facile observatur; sed eadem ex ipsa fractionum continuarum natura deduci potest; quam demonstrationem autem hic apponere superfluum iudico (Euler 1737, p. 104).)

This is the usual Eulerian approach toward demonstrations by mathematical induction: he limited himself to verifying the validity of the proposition $P(n)$ in some cases (in the *De fractionibus continuis dissertatio*, Euler verified the validity of theorem for $k = 1, 2, 3,$ and 4). He thought that this was sufficient to infer $(P(n), \forall\, n \in N)$. The mention of a demonstration that there was no point in including in the paper seems to indicate that Euler grasped the difference between mathematical induction and arbitrary exemplifications. In other terms, Euler understood the need of proving that the procedure is effectively repeatable (namely, $P(n) \rightarrow P(n + 1)$). However, he did not provide such a proof (see Netz 1999, p. 269). On the contrary, when Euler dealt with this theorem in his *Introductio in analysin infinitorum* (see Euler 1748, I, pp. 296–299), he no longer mentioned the possibility of including a demonstration and seemed to consider the Scheme **S** as clearly satisfactory.

3 Examples as Persuasion Tools

In this section, I dwell upon the use of example as persuasion tool. In Euler's mathematics, and in particular in his great treatises, *Introductio in analisyn infinitorum*, *Institutionum calculi integralis*, and *Institutiones calculi differentialis*, the mere logical reasoning was often accompanied by a myriad of examples and various applications: These examples played an important role: they showed the coincidence of results obtained with different techniques. In this way, the examples contributed to convince the reader of correctness and internal coherence of the calculus, which at the time was still a young and not well-founded science.

In order to accomplish his goal, Euler sometimes did not hesitate to use circular reasoning. In chapter IV of the second part of the *Institutiones calculi differentialis* (1755, pp. 359–402), Euler derived the series expansions of various functions by using the Taylor series

$$y(x+\omega) = y(x) + \frac{\omega dy}{dx} + \frac{\omega^2 d^2 y}{2dx^2} + \frac{\omega^3 d^3 y}{6dx^3} + \frac{\omega^4 d^4 y}{24dx^4} - \ldots, \tag{9}$$

which he proved in the previous chapter III (Euler 1755, pp. 335–336). In particular, Euler obtained the developments of several elementary functions, such as $(x + \omega)^n$, *log x*, a^x, *cos x*. However, in the context of Euler's treatises, the derivation of these expansions was vitiated by *petitio principii*.

In fact, to obtain the expansion of an elementary function by applying (9), one must know the differential ratio $\frac{dy}{dx}$ of that function. In the *Institutiones calculi differentialis*, part I, chapter IV (1755, pp. 98–107), Euler derived the differential ratio $\frac{dy}{dx}$ of an elementary function starting from (3). Euler set $\omega = dx$ in (3), namely, he considered the finite increment ω as an infinitesimal *dx*, and applied the principle of omission. (The principle of omission of differentials is the rule according to which

$dx^n + dx^m = dx^n$ if n < m.) Therefore, in Euler's construction, (3) was a necessary premise to the derivation of the differential ratio $\frac{dy}{dx}$ of a given function $y(x)$.

As seen above, in the chap. I of the *Institutiones calculi differentialis* (1755, pp. 3–38) Euler obtained (3) by using the power series expansions of elementary functions that he had proved in the *Introductio in analysin infinitorum*. In other terms, Euler derived the expansions of elementary functions by applying (9); however, these expansions were necessary to prove (3) and (9).

For instance, in the second part of the *Institutiones calculi differentialis*, chapter IV (1755, pp. 367–369), Euler derived the expansion of the function $y = \log_{10} x$, by calculating the differential ratios

$$\frac{d \log_{10} x}{dx}, \frac{d^2 \log_{10} x}{dx^2}, \frac{d^3 \log_{10} x}{dx^3}, \ldots.$$

and so obtained

$$\log_{10}(x + \omega) = \log_{10} x + \frac{m\omega}{x} - \frac{m\omega^2}{2x^2} + \frac{m\omega^3}{3x^3} + \ldots, \tag{10}$$

where $m = \log_{10} e = 0.4342944819032$. However, in the first part of the same book, chapter 1 (1755, pp. 22–23), Euler had derived the first difference of the logarithmic function in this way. He had noted that

$$\Delta y = \log(x + \omega) - \log x = \log\left(1 + \frac{\omega}{x}\right) \tag{11}$$

In the *Introductio in analisyn infinitorum* (1748, I, pp. 85–90), Euler had proved that

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \ldots \tag{12}$$

Euler applied (12) to (11) and so got

$$\Delta y = \frac{\omega}{x} - \frac{\omega^2}{2x^2} + \frac{\omega^3}{3x^3} - \frac{\omega^4}{4x^4} + \ldots \tag{13}$$

In chapter VI of the first part of the *Institutiones calculi differentialis* (1755, p. 152), Euler did not apply formula (13) in order to derive the differential ratio of the logarithmic function. He preferred to replace ω by dx in (11) so to obtain

$$dy = \log(x + dx) - \log x = \log\left(1 + \frac{dx}{x}\right) \tag{14}$$

Then he applied (12), (13), and (14) and so got

$$dy = \frac{dx}{x} - \frac{dx^2}{2x^2} + \frac{dx^3}{3x^3} - \frac{dx^4}{4x^4} + \ldots$$

At this juncture, he used the principle of omission, namely, he neglected high order infinitesimals, and got

$$d(\log x) = dy = \frac{dx}{x}.$$

However, (10) is only a trivial transformation of (12), since

$$\log_{10}(x+\omega) = \log_{10}\left[x\left(1+\frac{\omega}{x}\right)\right] = \log_{10}x + \log_{1\,0}\left(1+\frac{\omega}{x}\right)$$
$$= \log_{10}x + m\log\left(1+\frac{\omega}{x}\right)$$

In conclusion, the knowledge of the expansion of the logarithmic function presupposes the knowledge of the differential of the logarithmic function; however, the differential of the logarithmic function presupposes the knowledge of the expansion of the logarithmic function.

Euler's procedure is circular. It, however, is to be seen in the attempt to show that the various parts of analysis formed a coherent structure. Given the ease with which (10) could directly be deduced from (12), it is probable that the goal of Euler's derivation was to convince the reader of the substantial correctness of the computations, showing the uniqueness and non-contradiction of the findings of the calculus, although its foundations could appear uncertain.

This approach is often found in Euler's writings. It is closely linked to Euler's use of the examples as tools for verification of results and persuasion of the reader. In fact, even there existed a proof of certain theorems, Euler sometimes made use of examples to verify and check the validity of these theorems. These examples, well-placed in some critical points of his works, also had a rhetorical effect that contributed to promote the new analysis of infinities. They connected the various parts of analysis and showed its overall coherence also through the a posteriori agreement of the results.

Let us consider the Eulerian treatment of the so-called de l'Hôpital rule. Already in *Analyse des Infiniment petits* (1696, p. 146), the Marquis de l'Hôpital had clarified the rule that today bears his name by applying it to the function $y = \frac{aa-xx}{a-\sqrt{ax}}$ for $x = a$. He had then showed that the same result could also be obtained by algebraic ways and that, therefore, the results of the new Leibnizian calculus coincided with the results obtained by now more consolidated methods.

Euler approached the matter in a similar way. In his paper (1755, pp. 746–750) Euler used the de l'Hôpital rule to calculate the value of rational functions such as

$$\frac{x-x^n}{1-x}, \frac{x-x^{2n+1}}{1-x^2}, \frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2}, \ldots, \text{for} x = 1.$$

He found that the sought values were n, n, $\frac{n(n+1)}{2}$, ... In order to better persuade the reader of the correctness of these results, Euler showed that they could also be obtained by employing the following elementary identities:

$$\frac{x - x^n}{1 - x} = x + x^2 + \ldots + x^{n+1},$$
$$\frac{x - x^{2n+1}}{1 - x^2} = x + x^3 + \ldots + x^{2n-1},$$
$$\frac{x - (n+1)x^{n+1} + nx^{n+2}}{(1-x)^2} = x + 2x^2 + \ldots + n\,x^n,$$
$$\ldots$$

For $x = 1$, the right-hand sides of these identities, respectively, becomes

$$1 + 1 + 1 + \ldots + 1 = n,$$
$$1 + 1 + 1 + \ldots + 1 = n,$$
$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2},$$
$$\ldots$$

Euler explicitly observed:

> Since ... the same sums are observed by using other methods, the truth of this method will be further evident from the agreement [of the results]. (Quoniam ... eaedem summae aliunde constant, ex consensu veritas huius methodi magis elucebit Euler (1755, p. 747).)

After validating the 'truth' of the de l'Hospital rule by the a posteriori agreement of this rule and elementary methods, Euler applied the rule to transcendent functions. He thus showed the effectiveness of the rule by means of another series of examples, which now only performed the function of clarifying the method (and no longer the function of showing its correctness):

> The same method will be required to be used, if in the given fraction P/Q, the numerator or the denominator or both were a transcendent quantity. In order to explain these operations more clearly, it has been considered to add the following examples. (Eadem methodo erit utendum, si in fractione proposita P/Q vel numerator vel denominator vel uterque fuerit quantitas trascendens. Quae operationes, quo clarius explicentur, sequentia exempla adiicere visum est (Euler 1755, pp. 750–751).)

Another case of the use of examples as a tool of persuasion and verification is found in Euler (1755, p. 336). Immediately after having proved formula (9), Euler stated:

the truth of this expression can checked in examples in which the higher differentials of y vanish at last; for, in these cases the number of terms of the above expression will become finite. (veritas huius expressionis comprobari poterit eiusmodi exemplis, quibus differentialia altiora ipsius y tandem evanescunt: his enim casibus numerus terminorum superioris espressionis fiet finitus (Euler 1755, p. 336).)

Euler (1755, pp. 336–338) provided three examples concerning the functions

$$x^2 - x, \quad x^3 + x^2 + x, \quad x^2 + 3x + 1$$

He computed the values of these functions when one set $x + 1$, $x + 2$, x-3 in place of x in $x^2 - x$, $x^3 + x^2 + x$, $x^2 + 3\,x + 1$, respectively. The computation was performed twice. Firstly, Euler used formula (9). Secondly, he employed elementary algebra. Both in the case, he got the same results and this was considered as a validation of expansion (9).

For instance, let us consider the function $y = x^2 - x$. The differential ratio of this function are

$$\frac{dy}{dx} = 2x - 1, \quad \frac{d^2y}{dx^2} = 2, \quad \frac{d^3y}{dx^3} = 0, \ldots$$

By applying (9), one gets

$$y(x + 1) = x^2 - x + 1(2x - 1) + \frac{1}{2} \cdot 2 = x^2 + x$$

The same result is obtained setting $x + 1$ in place of x. Indeed x^2 changes into $(x + 1)^2$ and x into $x + 1$; then $(x + 1)^2 - (x + 1) = x^2 + x$.

In (1755), Euler used other examples to show the 'truth' of some infinite series which he had derived from the Taylor series. More precisely, Euler, first, formally manipulated the Taylor series and obtained an infinite series. Then, in order to persuade the reader of the correctness of his investigations, he applied his finding to some particular cases in which the infinite series was reduced to a finite series or to the sum of the series was already known.

For instance, in his (1755, p. 353) Euler noted that if one put $\omega = -x$, formula (9) changed into

$$y(0) = y(x) - \frac{xdy}{dx} + \frac{x^2d^2y}{2dx^2} - \frac{x^3d^3y}{6dx^3} + \frac{x^4d^4y}{24dx^4} - \ldots \tag{15}$$

He asserted that "the following examples will illustrate the truth of which (15)" ("cuius veritatem sequentia exempla illustrabunt" (Euler 1755, p. 353).). The examples concern the functions

$$x^2 + ax + ab, \; x^3 - 2x + 3, \; \frac{x}{1 - x}, \quad e^x, \sin x$$

(Euler 1755, pp. 353–356). I do not dwell on Euler's interpretation of (9) and (15), which was very different from the modern interpretation. I limit myself to observing that, according to Euler, his interpretation of (15) was true because it was verified by five functions.

For instance, in case of the function $y \sin x$, Euler showed that the right-hand side of (15) was precisely equal to the value of the function at x 0. In the same way, Euler (1755, p. 357) provided the formula

$$v(t) = y(x) + \frac{(t-x)dy}{dx} + \frac{(t-x)^2 d^2 y}{2dx^2} + \frac{(t-x)^3 d^3 y}{6dx^3} + \frac{(t-x)^4 d^4 y}{24dx^4} + \ldots,$$

where $v(t)$ is the function arising from $y(x)$ on setting t in place of x, and stated that "the truth of which can be proved by any example" ("cuius veritas quibuscunque exemplis comprobari potest" (Euler 1755, p. 357)).

In his *Institutiones calculi differentialis*, Euler also applied (9) in order to derive what he termed "the well-known Newtonian expression," namely, the Taylor series of $y = x^n$:

$$(x+\omega)^n = x^n + \frac{n}{1}x^{n-1}\omega + \frac{n(n-1)}{1\cdot 2}x^{n-2}\omega^2 + \frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}x^{n-3}\omega^3 + \ldots, \tag{16}$$

When Euler proved this formula at pp. 360–361 of his treatise on differential calculus (1755), he had already used it several times in the previous pages. This derivation is to be considered a further verification of (9) and, more generally, of the validity of the calculus. Indeed, this derivation was superfluous in the context of the second part of the *Institutiones calculi differentialis*. In section 73, chapter IV (1755, p. 361), Euler applied (16) in order to determine the following expansion of $(x+u)^n$

$$(x+u)^n = x^n + \frac{nx^n u}{1(x+u)} + \frac{n(n+1)x^n u^2}{1\cdot 2(x+u)^2} + \frac{n(n+1)(n+2)u^3 x^n}{1\cdot 2\cdot 3\cdot (x+u)^3} + \ldots \tag{17}$$

(In fact, formula (17) was derived from (16) by the substitution $\omega = -\frac{ux}{x+u}$). Expansion (17), useful for numerical applications, was the final objective of chapter IV of the *Institutiones calculi differentialis*. However, (16) was not essential to derive (17): in the subsequent section 74 (1755, pp. 362–363), Euler himself gave another proof of (17) where (16) was not applied.

Expansion (16) was, therefore, superfluous whether Euler's goal was the derivation of (17). More significantly, the proof of (16) was also vitiated by a circular

reasoning, as Euler himself noted some years later. Indeed, in his *Demonstratio theorematis newtoniani de evolutione potestatum binomii* (1774, pp. 105–106) he stated that this proof of (16) derived from analysis of infinite quantities; however, analysis of infinite was grounded upon this theorem. (The binomial theorem was used to prove (9) and (3).) For this reason, he asserted:

> I now recognize that this *petitio principi* is to be completely rejected. (eam tanquam petitionem principia penitus reiiciendam nunc agnosco (Euler 1774, p. 106).)

In (1774) Euler provided a new demonstration of (16) free of this error and observed: (Another demonstration of the binomial theorem was later given by Euler in his (1787). On the history of the binomial theorem, see Pensivy (1987).)

> I am confident that geometers will like my demonstration, because nothing is based on induction in it (meam demonstrationem Geometris eo gratiorem fore confido, quod in ea nihil plane induction tribuitur (Euler 1774, p. 106)).

Therefore, in the *Institutiones calculi differentialis*, the proof of (17) can be viewed as a rhetorical tool to convince the reader of the internal coherence of the analysis of the infinities and of the interlocking of its various parts. Thus, when in the *Demonstratio theorematis newtoniani* Euler admitted the existence of a circular reasoning, it seems that he felt the need to overcome certain methods of persuasion.

Euler used the above-described approach in other cases. Sometimes, the selected examples are not appropriate to substantiate the general proposition. This occured in *De usu functionum discontinarum in Analysi* (1765b), where Euler gave one only example in order to eliminate any possible doubt about the differential ratio of discontinuous functions (the example is the function $y = ax^2 + bx + c$):

> Since this notion of differentials and the investigation of the ratio between evanescent quantities seem to be extremely suspect to most, all doubts disappear with an example. (Cum plerisque haec differentialum notio, et rationis, quae inter quantitates evanescentes intercedit, investigatio, maxime suspecta videri soleat, unico exemplo omnio dubia evanescet (Euler 1765b, p. 12).)

In his *De summis serierum reciprocarum* (1734), Euler determined the sums of the series

$$\sum_{i=1}^{\infty} \frac{1}{n^{2k}} (k = 1, 2, \ldots)$$

(on Euler's theory of series, see Ferraro (2007b)). His procedure was the object of various criticisms. In his *Démonstration de la somme de cette suite* $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \frac{1}{36} +$ etc. (1743a), Euler replied to these criticisms and stated:

> The method I gave in the Comments of the Petersburg Academy to find the sum of this sequence, when the exponent n is an even number

$$1 + \frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{4^n} + \frac{1}{5^n} + \frac{1}{6^n} + etc.$$

has something extraordinary, because it is drawn from a principle that has not been used much in the researches of this nature. However, it is as sure and grounded as any other method which is ordinarily used in the summation of infinite sequences: I have also shown this by the perfect agreement of some cases already known elsewhere and by the approximations, which provide us with an easy way to examine the truth in practice.

> La méthode que j'ai donnée dans les Commentaires de l'Académie de Pétersbourg pour trouver la somme de cette suite, lorque l'exponant n est un nombre pair
>
> $$1 + \frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{4^n} + \frac{1}{5^n} + \frac{1}{6^n} + etc.$$
>
> a quelque chose d'extraordinaire, parcequ'elle est tirée d'un principe dont on n'a pas encore fait beaucop d'usage dans les recherches de cette nature. Elle est cependant aussi sûre et aussi fondée que toute autre méthode dont on se serve ordinairement dans la sommation des suites infinies: ce que j'ai fait voir aussi par le parfait accord de quelques cas déjà connus d'ailleurs et par les approximations, qui nous fournissent une manière aisée d'examiner la vérité dans la pratique (Euler (1743a, p. 177).

In (1743a) – but in other papers too (for example Euler (1743b)) – Euler validated his method by performing the same sums by a different method; the coincidence of the results should have persuaded of the validity of the method. No one, in reality, doubted the accuracy of the sums

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6};$$
$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90};$$
$$\dots$$

Criticisms concerned the proof, not the result (Cf. Euler (1743b, pp. 138–141).). To modern eyes, it is clear that exhibiting new different demonstrations, in itself, is not a valid defense of the original demonstration. Euler again made an a posteriori justification which was a persuasion tool for the reader and checked the validity of the obtained results. (In a similar way in his *De seriebus divergentibus*, Euler used four different metods in order to sum 1–1 + 2–6 + 24 + ... The different methods achieved the same result: thus each of these methods supported the others (on this, see Ferraro 2000b).)

Reference

Andrews GE (1990) Euler's "Exemplum memorabile inductionis fallacis" and Q-trinomial coefficients. J Am Math Soc 3(3):653–669. Accessed January 12, 2020. https://doi.org/10.2307/1990932

Bombelli R (1966) L'Algebra, Milano, Feltrinelli, 1966

Capobianco G, Enea MR, Ferraro G (2017) Geometry and analysis in Euler's integral calculus. Arch Hist Exact Sci 71(2017):1–38

Euler L (1734–35) De summis serierum reciprocarum, *Commentarii academiae scientiarum Petropolitanae* 7 (1734–35): 123–134; in Euler (OO, ser 1, vol 14, pp 73–86)

Euler L (1737) De fractionibus continuis dissertatio, "Commentarii academiae scientiarum Petropolitanae" 9, 1737, pubbl.1744, pp 98–137; in Euler (OO, ser 1, vol 14, pp 197–215)

Euler L (1739) Consideratio progressionis cuiusdam ad circuli quadraturam inveniendam idoneae. Commentarii academiae scientiarum Petropolitanae 11 in Euler (OO, ser 1, vol 4, pp 350–365)

Euler L (1743a) Démonstration de la somme de cette suite $1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\frac{1}{36}+$ etc. Journal littéraire d'Allemagne, de Suisse et du Nord in Euler (OO, ser 1, vol 14, pp 176–186)

Euler L (1743b) De summis serierum reciprocarum ex potestatibus numerorum naturalium ortarum dissertatio altera in qua eaedem summationes ex fonte maxime diverso derivantur. Miscellanea Berolinensia, in Euler (OO, ser 1, vol 14, pp 138–159)

Euler L (1748) Introductio in analysin infinitorum. Lausannae: M.M. Bousquet et Soc. 1748

Euler L (1750) De serierum determinatione seu nova methodus inveniendi terminos generales serierum. Novi Commentarii academiae scientiarum imperialis Petropolitanae, 3 (1750–51), pp 8–10 e 36–85 in Euler (OO, 14: 463–515)

Euler L (1754) De seriebus divergentibus. Novi Commentarii academiae scientiarum imperialis Petropolitanae 5 (1754–55): 19–23 e 205–237, in Euler (OO, ser 1, vol 14: 583–617)

Euler L (1755) Institutiones calculi differentialis cum eius usu in analysi finitorum ac doctrina serierum. Impensis Academiae Imperialis Scientiarum, Petropoli, 1755

Euler L (1765) Observationes analyticae. Novi Commentarii academiae scientiarum Petropolitanae, 8 (1765), pubbl.1765, pp 124–143; in Euler (OO, ser 1, vol 15, pp 50–69)

Euler L (1765b) De usu functionum discontinarum in Analysi. Novi Commentarii academiae scientiarum Petropolitanae 11 (1765), pubbl.1767, pp 3–27

Euler L (1774) Demonstratio theorematis newtoniani de evolutione potestatum binomii pro casibus quibus exponentes non sunt numeri integri. Novi Commentarii academiae scientiarum Petropolitanae 19:103–111, in Euler (OO, ser 1, vol 15, pp 207–216)

Euler L (1783) Varia artificia in serierum indolem inquirendi. Opuscula analytica 48–63, in Euler (OO ser 1, vol 15, pp 383–398)

Euler L (1787) Nova demonstratio quod evolutio potestatum binomii newtoniana etiam pro exponentibus fractis valeat. Nova acta academiae scientiarum Petropolitanae 5:52–59

Euler L (OO) Leonhardi Euleri Opera omnia. Series I: Opera mathematica, Leipzig, Teubner, Zurich, Füssli, Bâle. Birkhäuser, 1911–...

Ferraro G (1998) Some aspects of Euler's theory of series. Inexplicable functions and the Euler-Maclaurin summation formula, Historia mathematica, 25 (1998), pp. 290–317

Ferraro G (2000a) Functions, functional relations and the laws of continuity in Euler. Historia mathematica 27:107–132

Ferraro G (2000b) The value of an infinite sum. Some Observations on the Eulerian Theory of Series. Sciences et Techniques en Perspective 4:73–113

Ferraro G (2007a) The foundational aspects of Gauss's work on the hypergeometric, factorial and digamma functions. Arch Hist Exact Sci 61(2007):457–518

Ferraro G (2007b) Convergence and formal manipulation in the theory of series from 1730 to 1815. Hist Math 34:62–88

Ferraro G (2020) Euler and the structure of mathematics. Hist Math 50 (2020):2–24

Ferraro G, Panza M (2012) Lagrange's theory of analytical functions and his ideal of purity of method. Arch Hist Exact Sci 66(2012):95–197

Fraser CG (1989) The Calculus as algebraic analysis: some observations on mathematical analysis in the 18th century. Archive for History of Exact Science 39:317–335
Netz R (1999) The shaping of deduction in Greek mathematics. Cambridge University Press, Cambridge
Newton I (1712) Commercium epistolicum d. Johannis Collins et aliorium de analysi promota. Iussu Soc. Regoae, typis Paeersoniani, London
Newton I (1967) The Mathematical papers of Isaac Newton, DT Whiteside (ed), 8 vols., Cambridge University Press, Cambridge
Panza M (1996) Concept of function, between quantity and form, in the 18th century. In: Jahnke HN, Knoche N, Otte M (eds) History of mathematics and educations: ideas and experiences. Vandenhoeck & Ruprecht, Göttingen, pp 241–274
Pensivy M (1987) Jalons historique pour une épistémologie de la série du binôme. Sciences et techniques en perspective 14

Mathematical Experiments on Paper and Computer

Dirk Schlimm and Juan Fernández González

Contents

1 Introduction 2504
2 Mathematical Experiments and Experimental Research Projects 2504
2.1 Mathematical Experiments 2504
2.2 Mathematical Experimental Research Projects 2507
2.3 Experiments with Diagrams 2508
3 The Genesis of the Midpoint Path Construction 2509
3.1 Genesis and Explanation of the Construction 2510
3.2 Geometric Experiments on Paper 2513
3.3 Geometric Experiments with Computers 2516
3.4 The Broader Experimental Research Project 2518
4 Conclusion 2519
References 2520

Abstract

We propose a characterization of mathematical experiments in terms of a setup, a process with an outcome, and an interpretation. Using a broad notion of process, this allows us to consider arithmetic calculations and geometric constructions as components of mathematical experiments. Moreover, we argue that mathematical experiments should be considered within a broader context of an experimental research project. Finally, we present a particular case study of the genesis of a geometric construction to illustrate the experimental use of hand drawings and computers in mathematical practice.

D. Schlimm (✉)
Department of Philosophy, McGill University, Montreal, QC, Canada
e-mail: dirk.schlimm@mcgill.ca

J. Fernández González
Graduate School of Design, Harvard University, Cambridge, MA, USA
e-mail: jfernandezgonzalez@gsd.harvard.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_112

Keywords

Mathematical experiments · Geometry · Midpoint paths · Philosophy of mathematics · Diagrams · Computers in mathematics

1 Introduction

Mathematics can be regarded as an a priori science that proceeds deductively from certain indubitable truths. However, while this might be an idealized generalization of how a particular mathematical subject is presented, it is quite far from being an adequate characterization of how mathematics is practiced. Indeed, not much acquaintance with mathematics is needed to realize that establishing new mathematical results frequently starts with a generalization of some systematically obtained instances (Polya 1954), that the formulation of a system of axioms is frequently the eventual outcome of trying out various possible sets of candidates (Schlimm 2013), and that definitions are frequently obtained by a gradual honing in on the those formulations that are deemed most fruitful (Tappenden 2008). In addition, the increased use of computers in mathematical research in recent years has emphasized again certain analogies between mathematics and the natural sciences, and motivated mathematicians and philosophers to reconsider their view of mathematics (Secco 2017; McEvoy 2023). This challenge is highlighted by the claim that experiments, which are deemed a cornerstone of the scientific method, also play a crucial role in mathematics (Borwein and Bailey 2004). In reaction to this claim, philosophers of mathematics have suggested and discussed various characterizations of the notion of a *mathematical experiment* (Van Bendegem 1998; Baker 2008; Sørensen 2010). In this chapter, we first propose a refinement of these characterizations into different components. In addition, we extend the notion of computation used in one of these components to include geometric constructions, and we embed mathematical experiments into a broader context of experimentation (Sect. 2). Then, in the second part, we illustrate our framework by presenting the genesis of a geometric construction (the Midpoint Path Construction) with a focus on how mathematical experiments done on paper and computer contributed to its development (Sect. 3).

2 Mathematical Experiments and Experimental Research Projects

2.1 Mathematical Experiments

To set the stage, let us consider an individual *mathematical experiment* to consist of the following three components: (i) a setup, (ii) a process with an outcome, and (iii) a conclusion, i.e., an interpretation of the outcome. For the sake of simplicity, this characterization omits further layers of complexity that have been addressed by philosophers of science, such as the theory-ladenness of observations, the

interactions with artifacts and measuring instruments, and the role of models.[1] Nevertheless, we maintain that our three components provide a focused and useful framework for the further discussion of experiments in mathematics. Let us now take a closer look at these components.

The **(i) setup** of a mathematical experiment consists in the specification of an initial mathematical situation or configuration, such as a formula or algorithm with some initial parameters. In addition, the setup can also include certain expectations about the outcome of the experiment. Such expectations are typically expressed in terms of a *hypothesis* that is to be confirmed or rejected on the basis of the outcome and which is embedded in a set of background assumptions, such as a theory, which is formulated more or less explicitly. Because experiments that are intended to test a hypothesis by providing evidence for its acceptance or rejection are typically employed in the natural sciences, we shall refer to them as *scientific* experiments. In contrast, experiments that are conducted without any particular expectations regarding their outcome are *exploratory* experiments (Sørensen 2010).[2] The latter are frequently performed at the beginning of a new research project and are used to gather additional data for determining regularities and for subsequent conjecturing.

The **(ii) process** of an experiment is the part about which the experimenter should have no influence beyond setting the initial conditions and specifying the procedure.[3] In the case of a mathematical experiment, the process typically consists of a computation or construction that yields a certain **outcome**, i.e., a mathematical situation or configuration that results from applying a determinate procedure to the setup.[4] However, the nature of the experimental process can vary considerably: it can be carried out by the mathematician on paper, on a computer, or in the mind (e.g., as a thought experiment[5]), or it can be carried out completely automatically on a computer. In Baker's account of mathematical experiments, these processes are the "means" of calculation (Baker 2008, 340). In the natural sciences, the process traditionally involves some physical interaction with the world, but it has also

[1] As examples for the rich taxonomies of scientific experimentation developed by philosophers of science, see Hacking (1988, 508) and Galison (1988, 525), in which many different relations between experimentation, theory, and instrumentation are discussed. For a more recent overview, see Feest and Steinle (2016). In light of the turn in philosophy of science towards research processes and knowledge generation, one can see the present chapter as a contribution towards bringing together analyses of scientific and mathematical experimental practices.

[2] In philosophy of science, we find this distinction under the labels of "theory-driven" and "exploratory" experiments (Steinle 1997, S70); for a discussion of the epistemological significance of exploratory experiments, see Steinle (1997, S72).

[3] If the experiment is carried out by hand or as a thought experiment, the experimenter is directly involved in the process, but this should not affect the outcome.

[4] Notice that a computation, understood as a process, is only a *component* of a mathematical experiment; this might answer in part Van Bendegem's worry that "it seems that it is very difficult to make any sense out of the idea of considering a computation (whether aiming for a numerical result or a visual image) as a form of mathematical experiment" (Van Bendegem 1998, 178).

[5] See Starikova and Giaquinto (2018) and Starikova (2023) for a discussion of mathematical thought experiments.

become common practice to regard explorations using mathematical models and computer simulations as experiments.[6]

Finally, for an activity to count as an experiment, its outcome has to lead to a certain **(iii) conclusion**, i.e., an appropriate interpretation of the outcome. For example, we can interpret the outcome as new data to be studied in the case of exploratory experiments or as evidence for the confirmation or rejection of a hypothesis in the case of scientific experiments. What distinguishes an experiment from, say, an exercise is the nature of the conclusion that is drawn from the outcome. In an exercise, the outcome is simply accepted as the result, reflecting the correct or incorrect application of the process; in an experiment, the outcome is interpreted in relation to a hypothesis or taken as a data point for further investigations. In the case of scientific experiments, and in contrast to exploratory ones, the experimenter needs to have a clear conception of the expected outcome (hypothesis) and to retain a principled openness with regard to the possibility that the actual outcome might violate this expectation.

To illustrate the various components of mathematical experiments discussed above, let us briefly consider a simple activity and how it can be considered to feature in a mathematical experiment. As *setup*, we have two collections of pebbles that we have individually counted to consist of 3 and 5 pebbles. The experimental *process* consists in putting these two individual collections together and counting the pebbles in the new collection, resulting in the outcome of a count of 8.[7] Now, if we had the expectation that the all pebbles together will yield a count of 8 as part of the setup, then the outcome corroborates this hypothesis and we arrive at the *conclusion* that, when counting pebbles, $3 + 5 = 8$. Notice, however, that the same process and outcome could be interpreted in different ways. On the one hand, we might already be convinced that $3 + 5 = 8$, but we are not sure whether this arithmetical fact is applicable to counting pebbles. In this case, the conclusion of the experiment is support that pebbles are indeed things that obey the laws of arithmetic. Further support would be gained by repeating the experiment. On the other hand, the process might be part of an exploratory experiment, where the setup does not contain any expectations about the outcome. This could be the case when we are only beginning to learn about addition of numbers and we genuinely do not know what the result is of adding three things to five things. In this case, the experiment gives us an additional data point that we can use in support of the general fact that $3 + 5 = 8$, which may lead to further generalizations or investigations. These three different interpretations of the same process and outcome show that the conclusion drawn from a particular outcome depends to some extent on what kind of assumptions and expectations are part of the initial setup. In particular, this illustrates that our

[6] See Morgan's discussion of different "materialities" of experiments in Morgan (2003): material, semi-material, nonmaterial, and mathematical; for a discussion of computer simulations, see Barberousse et al. (2009).

[7] Van Bendegem considers such "real-world-experiments" but concludes that "they are quite uninteresting" for mathematics, "rather trivial," and that "no serious implications should be drawn from them" (Van Bendegem 1998, 179).

characterization of experiments implies that the same process and outcome can be considered to constitute an experiment in some cases but not an experiment in others (e.g., when no particular conclusion is drawn from the outcome). In other words, whether something counts as an experiment depends on the context determined by the setup and the conclusion.[8]

2.2 Mathematical Experimental Research Projects

While it can be useful to consider individual mathematical experiments in isolation in order to identify their characteristics (as we did above), we also believe that this idealization actually hinders our understanding of their role in mathematical practice. Indeed, experiments are rarely conducted in isolation but are typically part of an experimental research project embedded in much more elaborate practices, including background assumptions, techniques, and goals.[9] For example, in practice, the setup of an experiment is rarely based on an arbitrary hypothesis, but on one that is part of a *guided investigation*, which typically consists of numerous iterations of experimenting and theorizing.[10] Such an investigation could proceed as follows: The initial spark might be a fortuitous observation of some pattern that is generalized to form a conjecture; this leads to an experiment with a slight variation of the original setup, which is expected to result in a pattern that also falls under the generalization; the result of this experiment, besides supporting a confirmation or refutation of the hypothesis, yields new data that can be used to suggest conceptual innovations or refinements as well as new generalizations; and so on. This iterative account of what we shall call a *mathematical experimental research project* provides us with a rough sketch for discussing the context of mathematical experiments.[11] Such a context is usually acknowledged also by researchers who focus on the characterization of individual experiments: for example, Sørensen speaks of "an *integrated process* of experimentation, concept formation, and conjecturing" (Sørensen 2010, 354; emphasis added).

In the literature, mathematical experimental research has been discussed frequently in relation to mathematical discovery, in particular in the seminal work of George Polya. For him, the iterative character of such research is analogous to the experimental practice in the natural sciences, which he describes as follows:

[8] This view about the context-dependency of experiments is not shared by everyone; Van Bendegem, for example, considers it "slightly, if not simply, odd" (Van Bendegem 1998, 175).

[9] See Ferreirós (2015) and Wagner (2017) for general accounts of mathematical practices.

[10] In the case of investigations directed at the solution of specific questions or problems, Detlefsen and Arana speak of "directed investigations" (Detlefsen and Arana 2011, 9); see also the notion of a "trial" in a mathematical thought experiment (Starikova and Giaquinto 2018, 263).

[11] The iterative nature of experimental research in science has been emphasized by Chang (2004); see also Elliott (2012) for a discussion.

> If a naturalist observes a striking regularity, which cannot be reasonably attributed to mere chance, he conjectures that the regularity extends beyond the limits of his actual observations. [...] The following step may be prediction. On the basis of his former observations and their concordance with conjectural law, the naturalist predicts the result of his next observation. (Polya 1954, 49)

In his discussion of Euler's formula for polyhedra, which is satisfied by nine different polyhedra, Polya explicitly draws an analogy between research in mathematics and the scientific practice of a naturalist, who will doubt even a plausible hypothesis that has been occasionally verified: the scientist "will question it and collect new observations or design new experiments to test it." Accordingly, Polya suggests to his mathematical reader: "We may do the very same thing" (Polya 1954, 38). Even after further corroboration of the hypothesis, "a conscientious naturalist would feel satisfaction over the success of his experiments, but would go on devising further experiments," in particular ones that are more likely to refute it (Polya 1954, 38–41).[12] Similar accounts of mathematical research can easily be found in mathematicians' own reflections, as well as in studies of their research practices. An example of the latter is Haffner's recent investigation of unpublished manuscripts by Dedekind that document the path towards his introduction of lattices (or, as he called them, "Dualgruppen"), which she describes as "Dedekind's *systematic experimentations*" (Haffner 2021, 15; emphasis added).[13] Haffner explicitly identifies several phases of explorations in Dedekind's work using "computations and diagrams as ways to experiment to better understand the laws of module theory by manipulating simple cases, testing on progressively more general cases, trying to identify the laws susceptible to hold in the general case, and the ones that do not." These were followed by conceptual developments, which were "then followed by new phases of experimentation and clarification," to finally end in his theory of Dualgruppen (Haffner 2021, 10–11).

2.3 Experiments with Diagrams

Recall that a crucial component in our characterization of mathematical experiments was a *process with an outcome*. Now, in the field of "experimental mathematics," as understood by Borwein and Bailey (2004), this process is carried out by a computer. However, as Baker has plausibly argued, the use of computers should not be considered a necessary condition for mathematical experiments. We agree with Baker that the process of a mathematical experiment may consist of computations by computer or by hand, but we push his conclusion even further by suggesting also the use of other means of mathematical reasoning. In general, we can understand mathematical reasoning as the drawing of consequences from mathematical

[12] These passages are also referred to in Lakatos (1976, 7).

[13] Similarly, Dedekind's contemporary Cantor also speaks of his "experiments [Versuche] which have lasted many years" (Cantor 1883, quoted from the translation in Ewald 1996, 890).

situations within certain prescribed boundaries. Because mathematical situations can quickly become too complex for us to keep before our mind's eye, so to speak, we represent them externally by symbols, formulas, or diagrams. In doing so, we extend our cognitive capacities, but such representations also allow us to keep a durable and shareable record of the reasoning process and they can be used to determine the boundaries of accepted inferences, such as the inferential steps within a regimented language or the restriction to a set of tools, e.g., compass and straightedge. Moreover, rather than just representing a mental reasoning process, external representations can themselves be the vehicles of epistemic inferences.[14]

An impressive example of mathematical experimental research using geometric representations has been reported about the late Fields medalist Maryam Mirzakhani: "As she thought about mathematics, Mirzakhani constantly doodled, drawing surfaces and other images related to her research" (Klarreich 2018, 151). Since hyperbolic surfaces with slight spatial variations could have different geodesics, Mirzakhani produced numerous drawings of slightly different variations to study them. According to her husband, Jan Vondrák, "[s]he has these huge pieces of paper on the floor and spends hours and hours drawing what look to me like the same picture over and over" (Klarreich 2018, 151). The fact that each drawing is part of an overarching experimental project comes out in the "over and over" passage of the previous quote. As we shall see in our case study below, using computers can greatly facilitate and enhance this aspect of an experimental practice.

3 The Genesis of the Midpoint Path Construction

Geometric constructions with compass and straightedge date back to at least the times of the ancient Greeks. These are based on combinations of the following constructions, which correspond to the first three postulates of Euclid's *Elements*:[15] drawing a line segment between two points, extending a line segment by any length in a straight line, and drawing a circle with a given point as center that also passes through another given point. A more complex construction from antiquity, which in addition also relies on drawing parallel lines and which is justified by the Intercept Theorem (also known as Thales' Theorem), allows one to draw a line segment whose length is a rational multiple of the length of a given segment. Recently, an alternative construction that yields the same result was put forward (Fernández

[14] See, e.g., Giardino (2018), Carter (2019), and Vold and Schlimm (2020). Because of the predominance of presenting mathematical reasoning in terms of external representations, some philosophers, such as Peirce and Wittgenstein, have claimed that *all* mathematical reasoning is ultimately based on observations.

[15] Heath warns against interpreting the object of the postulates purely as limiting possible constructions (Heath 1968, 124).

González 2016).[16] The Midpoint Path Construction makes it possible to multiply the length of a line segment by a rational number between 0 and 1 by constructing only midpoints and a straight line. It is also based on the use of compass and straightedge but does not require the drawing of parallel lines.

In the present section, we describe the genesis of the Midpoint Path Construction as a case study to illustrate and discuss the use of experiments in mathematics. Section 3.1 presents the construction and the main stages of its discovery process. Then, in Sects. 3.2 and 3.3, we take a closer look at the use of mathematical experiments by hand and by computer in the emergence of the construction.

3.1 Genesis and Explanation of the Construction

We present here a simplified narrative of the process of discovery of the Midpoint Path Construction. According to Juan Fernández González, the second author of the present chapter who developed the construction in question, the process began when he doodled in his school agenda without expecting any particular result and found what seemed to be a way to divide a line segment into three equal parts.[17] The diagram on the right of Fig. 1 shows this initial construction. Given a line segment $[UV]$ with endpoints U and V, and an arbitrary point M_0 not on $[UV]$, a series of

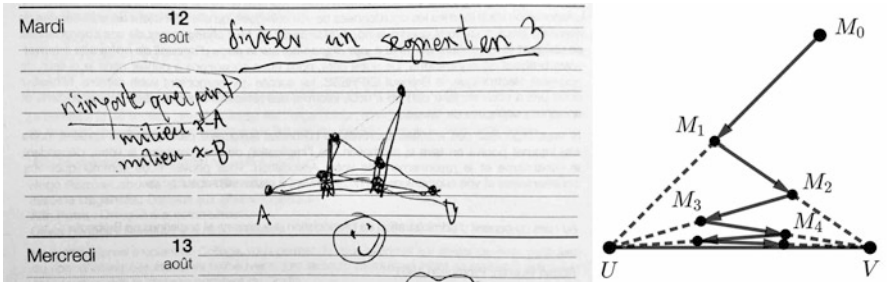


Fig. 1 Left: One of Juan's first doodles showing a construction to divide a line segment into three equal parts, drawn on his school agenda. Right: Computer drawing of the doodle in which A and B were respectively renamed U and V to be consistent with Fernández González (2016). (This and other figures were used in Fernández González and Schlimm (2023) and are used here with permission of the authors)

[16] The background of the development of this result is presented in the form of an interview in Fernández González and Schlimm (2023). Some of the reproductions from Juan's notebooks shown in that article are also used below with permission from the authors.

[17] As reported by Martin Gardner (1983, 82–83), Stanislaw Ulam's discovery of particular patterns in the distribution of prime numbers (Stein et al. 1964) began in a similar, absent-minded kind of way: During a "long and very boring" conference talk, Ulam "doodled a grid of horizontal and vertical lines on a sheet of paper," then numbered them in a specific way, circled the prime numbers, and "to his surprise the primes seemed to have an uncanny tendency to crowd into straight lines." We are thankful to David Waszek for pointing us to this example.

midpoints is constructed: M_1 is the midpoint of $[M_0U]$, M_2 is the midpoint of $[M_1V]$, M_3 is the midpoint of $[M_2U]$, M_4 is the midpoint of $[M_3V]$, and so on, always alternating between U and V in each step. The series of points M_0, M_1, M_2, ..., which was later called a *midpoint path*, will get closer and closer to the line segment $[UV]$. If this construction would be continued forever, Juan conjectured that the points would end on the line segment $[UV]$, dividing it into three equal parts.

By varying the construction process (i.e., by not only simply alternating between U and V in the construction of the midpoints), Juan was able to obtain different outcomes, such as a division of the line segment by the midpoint path into five and seven equal parts (see Fig. 2). For example, to divide $[UV]$ into five equal parts, the midpoint path must go twice toward V, then twice toward U, and so on. In other words, the midpoint path goes toward V, V, U, U, ... repeated indefinitely, as opposed to V, U, ... repeated indefinitely, as was the case in the construction that divides the segment into three equal parts. The representation of a midpoint path in terms of the endpoints of the segments of each constructed midpoint, such as "V, V, U, U, ..." turned out to be fruitful later, because it is independent of the location of the starting point M_0 and is called a *midpoint series*. In particular, since a midpoint series consists of only two symbols, it suggested a link to the representation of numbers in binary notation, which was pointed out by Gilbert Labelle.[18] After many such variations, Juan found other midpoint series that made it possible to divide a line segment into different whole numbers other than 3, 5, and 7.

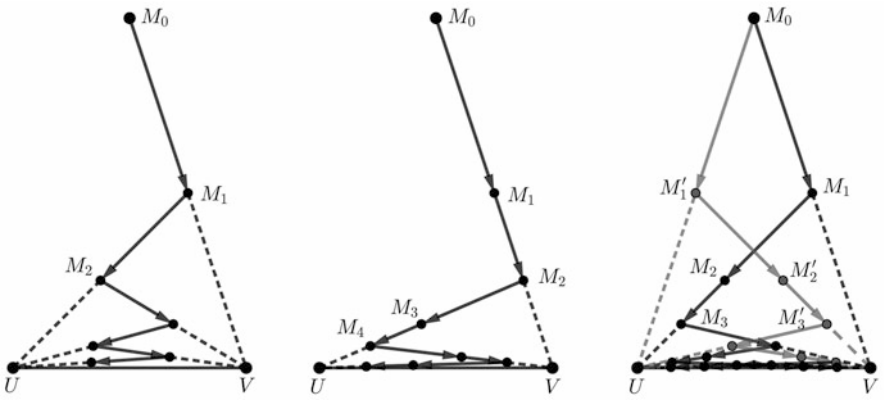


Fig. 2 Constructions using midpoint paths whose lengths tend to infinity to divide a line segment into three, five, and seven equal parts. Each midpoint path starts at point M_0 and is composed of successive constructions of midpoints M_i, which are either the midpoints of $[M_{i-1}U]$ or $[M_{i-1}V]$, for $i = 1, 2, \ldots$. The rightmost construction shows two midpoint paths: one composed of points M_0, M_1, M_2, ..., and one composed of points $M_0, M_1', M_2', \ldots$. Both are needed to divide $[UV]$ into seven equal parts

[18] Gilbert Labelle, emeritus professor at the Department of Mathematics at the *Université du Québec à Montréal*, helped Juan with the writing of the published article.

A drawback of these constructions was that they ultimately required an infinite number of midpoints. Thus, while the end results in the limit could be *seen* in the diagrams, they could actually never be reached after only a finite number of construction steps. Through inspection of various instances of such constructions, Juan realized that the same outcome could be obtained after only a finite number of steps. In fact, he noticed that in the diagrams, some of the midpoints seemed to be aligned, such as M_0, M_2, M_4, M_6, M_8, ... in Fig. 2 (left). Thus, it seemed that one needed to construct only the first two points of this midpoint path, namely M_1 and M_2, and then to draw a straight line passing through M_0 and M_2. This *convergence line*, Juan conjectured, would then intersect the segment $[UV]$ at one-third of its length (see Fig. 3; also for the constructions of 1/5 and 1/7). With this insight, it became possible to use only a finite number of steps to obtain the same result that originally had required an infinite process. Moreover, the construction was also simplified in another respect: to construct the point at 1/3 of $[UV]$, one did not have to simultaneously also construct another point at 2/3.

After much experimentation (i.e., varying the setup, carrying out the same construction or process, and evaluating the results), the idea of a special type of midpoint paths, which start and end at the same point, emerged. More precisely, their existence was observed when Juan drew midpoint paths that started at the newly constructed point P. These paths eventually named *midpoint loops* and were essential to further develop the construction. By drawing midpoint paths and midpoint loops in the same diagram, new spatial relationships became visible, which helped Juan formulate new conjectures and advance his research. Figure 4 shows the same examples from Fig. 3, but with added midpoint loops on the segment $[UV]$. In the diagram on the left of Fig. 4, we see a midpoint path consisting of points M_0, M_1, and M_2, and the corresponding midpoint loop with points P and P_1 (such that the loop begins and ends in P). Both are defined by the midpoint series V, U, such that P_1 is the midpoint of $[PV]$ and P is the midpoint of $[P_1U]$. From observing the behavior of midpoint loops, Juan realized that for any midpoint series, there are infinitely many

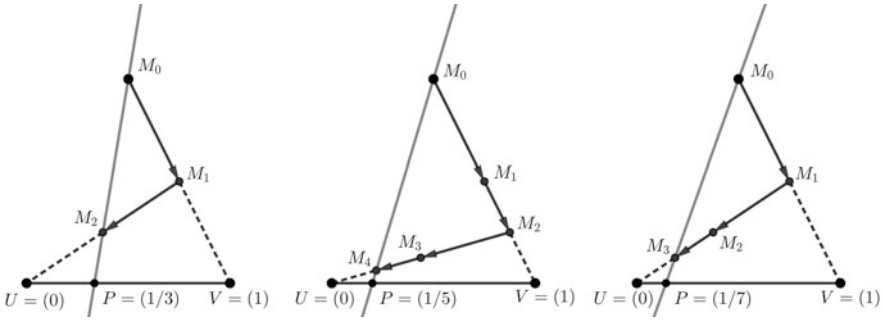


Fig. 3 Constructions using midpoint paths with a finite number of points M_i to construct the points $P = (1/3)$, $P = (1/5)$, and $P = (1/7)$, respectively, from left to right, with $U = (0)$ and $V = (1)$. This is done with the help of a convergence line

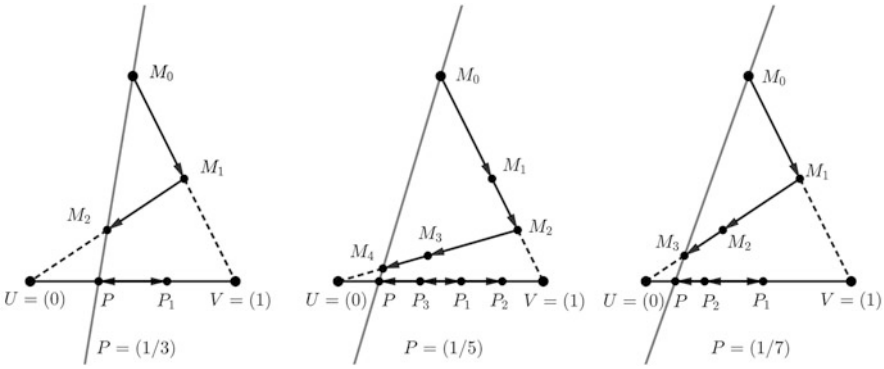


Fig. 4 Three midpoint paths (composed of points M_i) and their corresponding midpoint loops (composed of points P_i, each starting and ending at point P). These midpoint paths are the same as those in Fig. 3

midpoint paths (because the starting point M_0 could have an arbitrary position not on the straight line (UV)) but only one midpoint loop.

The developments sketched in the previous paragraphs ultimately led to a further generalization, called *homothetic paths*. For these, Juan was able to prove the Homothetic Path Theorem, which establishes a relationship between an infinite number of midpoint paths and a single midpoint loop. The Midpoint Path Theorem establishes the link between a midpoint series and the point P on a line segment $[UV]$, such that the length of $[UP]$ is a fixed rational multiple of the length of $[UV]$.[19]

3.2 Geometric Experiments on Paper

Let us now look in more detail at the role played by mathematical experiments in the development of the Midpoint Path Construction. While it is not difficult to think of a single line segment and its midpoint, imagining a series of midpoint constructions is more challenging, and so it should not come as a surprise that Juan's constructions were initially drawn on paper. This allowed him to visually represent complex mathematical situations, to trace the construction process, and to study the result of the outcome. In fact, Juan drew on various media for his experiments, such as his school agenda, napkins, and a bus's frozen window, but the great majority of the creative process behind the discovery was documented in small notebooks that were easy to carry around (Figs. 5 and 6 show spreads these notebooks). It is a characteristic feature of Juan's notebooks that they are very detailed, perhaps more so than the notes of a more experienced researcher, whose guided investigations would be honed by previous experiences.

[19] These results are presented in Fernández González (2016).

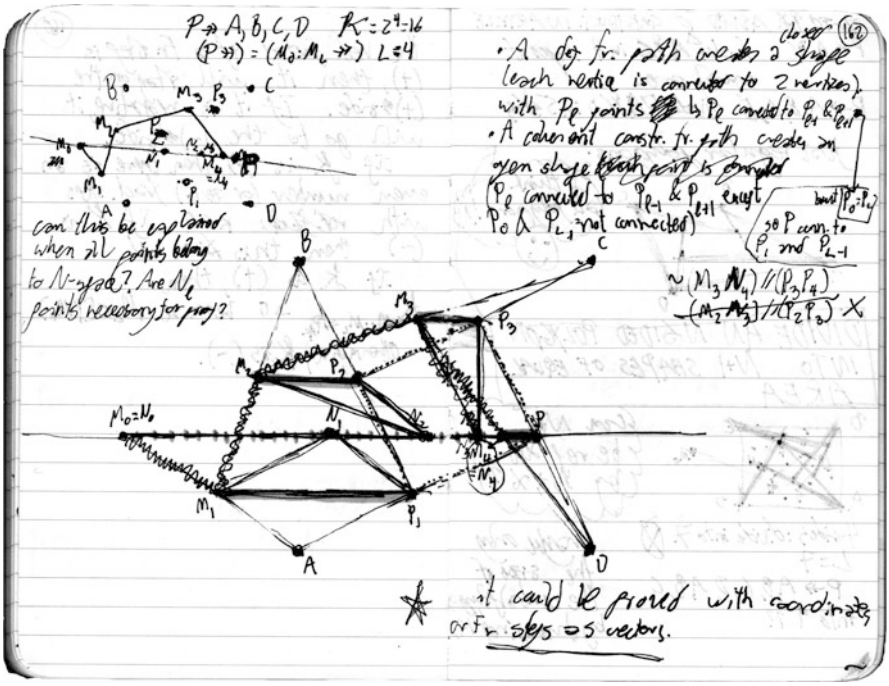


Fig. 5 In the central drawing, one can see a midpoint path M_0, M_1, M_2, M_3, M_4 and a midpoint loop P, P_1, P_2, P_3, P. Instead of having two points U, V defining the construction of each new midpoint like in the previous examples, there are four arbitrary noncollinear points A, B, C, and D that play the same role. As a result, the points of the midpoint loop P, P_1, ... are not collinear either. This allowed Juan to visualize similar triangles (such as PM_0A and P_1M_1A) that were essential to his first proof of the Homothetic Path Theorem

The very first drawing that sparked Juan's investigations was a simple doodle, drawn almost absentmindedly without any expectations about a particular outcome or insight (see Fig. 1, on the left). Recognizing a possible pattern formed by the points and lines was a serendipitous discovery (for a similar case, see Sørensen 2010, 354). From then on, however, each drawing was motivated by particular intentions: either to explore the outcome of variations of the construction (e.g., repeating the pattern U, U, V, V, instead of U, V) or to test a certain conjecture about the outcome. Thus, Juan's drawings fit the characterization of a mathematical experiment (first exploratory, then scientific) given in Sect. 2.1 in terms of a setup (i.e., the initial configuration of points and lines, and the instructions for the construction), process and outcome (i.e., the carrying out of the geometric construction), and a conclusion as the interpretation of the result. Evidence for the latter is frequently found in the text that accompanies a drawing in the notebooks. Juan's interpretations of the results are formulated in sentences that express general propositions. Reversing the well-known dictum that a picture is worth a thousand words (see Larkin and Simon 1987), Juan suggested that "a theorem is worth a thousand

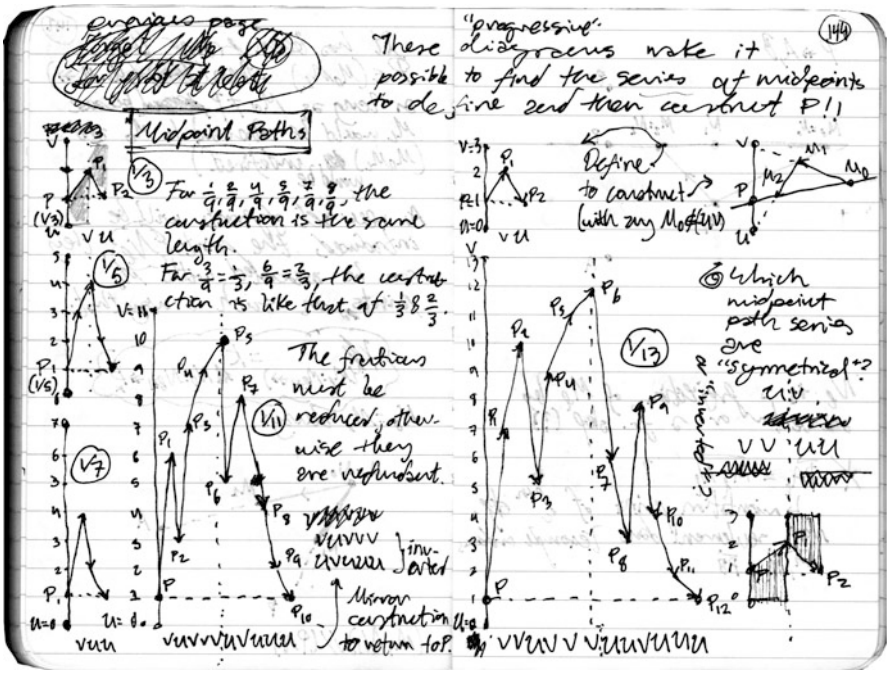


Fig. 6 This spread shows some hand-drawn experiments in which Juan aimed to determine the midpoint series associated to different rational numbers

doodles" (Fernández González and Schlimm 2023). Here, a single theorem can describe many (indeed, infinitely many) instances or individual geometric constructions, thus encapsulating the numerous experiments that motivated its formulation.

We now turn our attention to some particular aspects of the mathematical experiments that led to the formulation of the Midpoint Path Theorem, namely the permanent nature of drawings that allows for the careful study of the process and outcome, and how the experiments contributed to the observation of new patterns and the formation of concepts.

In contrast to experiments that yield an outcome but leave no trace of the process that led to it, geometric experiments that involve the drawing of a diagram produce an artifact that can be inspected after its construction process (similar considerations also apply to written computations, etc.). On the one hand, this allows us to retrace the entire process with special attention on how particular features of the setup affect the result. On the other hand, the diagram itself might reveal certain patterns that can be observed and studied. In Juan's case, the permanent nature of the drawings is further emphasized by his preference of drawing with a pen in order to avoid anything to be erased: "It is not trial and error, because nothing is wrong. Since it's all a creative process, I believe that 'trial and better' is a more fair description" (Fernández González and Schlimm 2023).

The following three examples illustrate how new patterns emerged from Juan's experiments that lead to new concept formations. One of the first patterns is the one that led to the notion of a convergence line. While it might seem obvious in retrospect, when Juan conceived of his construction as a potentially infinite process, he was not yet aware of the alignment of certain points of the midpoint paths. Only after observing these alignments in various diagrams did he conjecture that some of the points of a midpoint path lay on a line. He then went back to the earlier diagrams to corroborate this insight and he added the convergence lines (e.g., the darker lines on Fig. 1 were added to the original drawing during this process).

A second example involves the formation of the concept of midpoint loops. These are particular midpoint paths that begin and end at the same point (see Sect. 3.1). By simultaneously drawing a midpoint path and a midpoint loop in the same figure, new patterns appeared. To bring out these patterns, Juan connected the midpoint paths and midpoint loops by a series of similar triangles (as shown in Fig. 5). The idea of constructing these triangles came to Juan only during the experimental process itself and turned out to be crucial for the later proof of one of the theorems. The notebook spread shown in Fig. 5 documents the evolution of this approach: The smaller drawing on the top-left of the spread was drawn first in order to quickly capture the general idea and is a less elaborate version of the main drawing. Moreover, Juan captured his observations immediately after drawing the diagrams in the accompanying text: on the top-right of the spread he discusses the difference between midpoint paths and midpoint loops. These observations again proved to be essential for the formulation and proof of some of the theorems in Fernández González (2016).

Our third example of the emergence of new patterns illustrates how a change of representation (here, into a different type of diagram) can bring out new insights. As can be seen in Fig. 4, all the points of a midpoint loop are on the line segment $[UV]$. However, as the number of points increases, more and more arrows need to be drawn on the segment, which makes it increasingly difficult to keep them apart visually and to understand the diagram. To improve this situation Juan developed a two-dimensional representation of midpoint loops, as shown in Fig. 6. Despite lacking geometrical accuracy, these diagrams provide a more perspicuous visualization of midpoint loops and thereby helped to advance the investigations. On the top-right corner of the shown spread, Juan wrote: "These 'progressive' diagrams make it possible to find the series of midpoints to define and then construct P!!", where P corresponds to the start and end point of a midpoint loop.

3.3 Geometric Experiments with Computers

The drawings shown in the previous figures are mainly qualitative in nature: the lines are not straight and the points have irregular shapes. Indeed, the drawings were not intended to approximate a geometric ideal, but to capture a particular instance of a geometric construction and to exhibit the relevant relations between the setup and the outcome of the experiment. Nevertheless, to improve the accuracy, Juan began at

some point to use the computer program GeoGebra for his experiments.[20] In addition to the higher accuracy that can be obtained by the use of computers in the experimental process, they also allow for quick manipulations of the setup and process, thereby considerably increasing the speed in which different constructions could be carried out. This, in turn, also increased the number of experiments that could be performed. Note that in our discussion, the computer is used as a tool for conducting experiments interactively, not for independently carrying out calculations or simulations.

For Juan's constructions, only four tools were required from GeoGebra's numerous functions in the *Geometry* two-dimensional workspace: (1) *point*, which places a point at any position, or at the intersection of lines and/or curves; (2) *line segment*, which draws a line segment between two given points; (3) *straight line*, which draws a straight line passing through two given points; and (4) *midpoint*, which places a midpoint between two given points. With these tools, any drawing of a Midpoint Path Construction could be reproduced on the screen. Thus, like with the hand drawings described above, GeoGebra allowed Juan to visually represent the mathematical setup, to perform the construction, and then to observe the outcomes.

In addition, because GeoGebra allows the user to zoom in and out in a diagram, several operations became possible that had been inconceivable with drawings on paper.[21] For example, whereas drawing several midpoints on a piece of paper could quickly lead to overlaps in the drawing, this never happened with the computer, because it was always possible to zoom in until the points were far enough apart. The zoom function also made it possible to get as close as desired to a point in order to determine its coordinates. Juan reports of having been fascinated by zooming into a point with coordinate 0.333333333333, because the precision increased with the zooming, which strongly supported that it was indeed at 1/3 of the length of the desired segment. Through numerical calculations performed by the software, which gave the coordinates of points, it became possible to simply read off the positions of points. Thus, in addition to providing an enhanced means for the visual representation, GeoGebra provided numerical accuracy at a level unattainable by constructions on paper.

The use of the computer also facilitated greatly the drawing of multiple constructions that started with different spatial configurations. For example, Fig. 7 shows three different constructions of a midpoint path that divide a line segment $[UV]$ at point P, located at 1/3 of its length. First, these diagrams demonstrate that the construction works indeed for those three different positions of the starting point M_0. Second, moving the starting point around the workspace by clicking on it and progressively dragging it to different locations generated an array of further instances, which could easily be manipulated and studied. This added a dynamic understanding of the relations in the diagram that could otherwise only be obtained

[20] For more information about this freeware software package, see https://www.geogebra.org

[21] See Starikova and Giaquinto (2018, 273–274) for a discussion of the importance of zooming in and out in mathematical research.

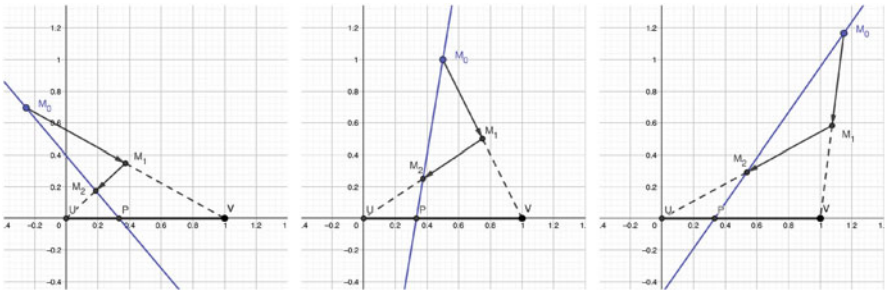


Fig. 7 Midpoint Path Constructions that multiply the length of a line segment [*UV*] by 1/3 at point *P* that differ in the position of M_0, without affecting the position of *P*

by mentally combining several static instances drawn on paper. Because the construction did not have to be redrawn each time, as was the case for drawings on paper, this significantly increased the speed with which conjectures could be corroborated. In fact, Juan's overall confidence in the correctness of the conjectures was greatly increased by the geometric experiments by computer.[22]

3.4 The Broader Experimental Research Project

Juan's drawings, both by hand and computer, can be considered to be parts of mathematical experiments according to our discussion in Sect. 2.1. However, to understand the process that led to the formulation of the Midpoint Path Construction, the experiments cannot be considered in isolation. In fact, the individual diagrams presented above were part of a larger context, namely Juan's mathematical experimental research project, which ultimately led to the publication of the article *Chemins homothétiques* (2016).

Throughout Juan's notebooks one easily notices that each drawing is often found among a group of similar drawings, illustrating how each experiment was part of a guided investigation. For example, Fig. 6 shows a spread in which different experiments are drawn side by side. These particular drawings were used to determine the series of midpoints that were used for the Midpoint Path Construction. They illustrate the series of midpoints needed to multiply the length of a line segment by 1/3, 1/5, 1/7, 1/11, and 1/13. This systematic increase of the values of the denominator illustrates the iterative nature of the experimental research, as described by Polya.

The hand drawings in Fig. 5, and most notably the text that accompanies the drawings, also illustrate the broader context of Juan's mathematical experimental research project. On this spread, he aimed to study the relationship between a

[22] This last remark echoes a sentiment about mathematical proofs expressed in Hume's *Treatise* (Norton and Norton 2007, 121).

midpoint path and a midpoint loop, both defined by the same midpoint series. His notes on the spread suggest future experiments, such as "Are N_I points necessary for proof?" In fact, the N_I points, which correspond to projections of the points of a midpoint path on its convergence line, disappeared in the following research stages, because they did not contribute to any new insight. Discarding unsuccessful avenues of exploration was important during the process of discovery. Another note on the same spread shows the emergence of a new idea: it states that "vectors" could be used to study these configurations. Indeed, linear algebra and vectors were a crucial ingredient of the presentation of the published proofs, whereas similar triangles, which are used in the diagram, play only a minor role in the published paper. Thus, the observations and suggestions captured in Juan's notebooks show the links between the individual experiments, the context of experimentation, the conceptualization of the process, and later stages of research.

4 Conclusion

Building on previous work on experiments in mathematics by Baker (2008), Sørensen (2010), and others, we began this chapter with a characterization of mathematical experiments in terms of three different components: a setup, a process with an outcome, and a conclusion. To remind, the (i) setup consists in the specification of an initial mathematical situation or configuration, as well as possible expectations regarding the outcome (hypotheses, conjectures); the (ii) process is, in a sense, the part of the experiment that is set in motion once the setup is specified and that is independent of the researcher (e.g., an algorithm, computation, or construction); and the (iii) conclusion is the interpretation of the outcome, which can be the corroboration or refutation of a conjecture, or simply new data to be considered. Using this simple framework (which can easily be refined to include additional aspects of mathematical experiments), we discussed the distinction between *scientific* experiments, which are closely tied to a hypothesis, and *exploratory* experiments, which are conducted without particular expectations regarding their outcome. *Thought experiments* also fit nicely into this framework, as they can be characterized by the absence of any empirical restrictions on the experimental process. For example, this absence allows to consider the result of completing an infinite process. However, the conclusion is more speculative in such cases, due to fewer constraints on the outcome. Moreover, due to the essentially human nature of thoughts, thought experiments are not completely independent of the researcher. We argued that mathematical experiments are rarely conducted in isolation, but that they are typically embedded in a wider context of mathematical experimental research projects. To illustrate various aspects of geometric experiments in mathematical practice, we discussed some steps in the genesis of the Midpoint Path Construction (Fernández González 2016), drawing on handwritten notes and computer-drawn diagrams that were produced by its author during the discovery and elaboration of the mathematical result.

In general, we have seen in our case study how mathematical experiments contributed to increasing confidence in conjectures and to rejecting hypotheses about the outcome of certain processes. Complex mathematical situations can be made accessible through the drawn figures, which also constitute traces of the geometric constructions and which can be studied after the construction process itself is finished. In particular, the visual traces of the experiments in our case study (i.e., the diagrams) gave rise to novel patterns that could be observed and that led to further conceptual innovations (e.g., the convergence line). Experiments with different initial conditions allow to explore how such variations affect the outcome of the construction processes. The use of computers in mathematical experiments enhances the accuracy of the drawings and, by increasing the speed with which a construction can be performed, allows for a considerably greater number of experiments. In addition to these practical advantages, the interactive use of computers can also introduce new qualitative dimensions, for example, by allowing for the direct manipulation of figures. Through such manipulations of figures on the screen, variations of the experimental process can be displayed dynamically. Similarly, computer renderings of diagrams can easily been zoomed in and out, which allows again for dynamic interactions with figures that would otherwise be possible only in thought. While our case study focused on geometric experiments, the general framework proposed in Sect. 2 easily accommodates many other kinds of mathematical experiments, e.g., in which the process consists mainly of calculations or graphical visualizations. Indeed, since experiments in the natural sciences also fall under the notion of experiment proposed in this chapter, some similarities between research in mathematics and in science, which were also emphasized by Polya, are made apparent. By considering mathematical experimental research within a broader context of experimental research projects, the analogy to experiments in the natural sciences becomes even more pronounced.[23]

Acknowledgments The authors would like to thank Jessica Carter, Valeria Giardino, Fenner Tanswell, and David Waszek for comments on a previous version of this chapter. This work was supported in part by the Social Sciences and Humanities Research Council of Canada.

References

Baker A (2008) Experimental mathematics. Erkenntnis 68:331–344

Barberousse A, Franceschelli BWS, Imbert C (2009) Computer simulations as experiments. Synthese 169:557–573

Borwein J, Bailey D (2004) Mathematics by experiment. Plausible reasoning in the 21st century. A. K. Peters, Natick

Cantor G (1883) Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen. Teubner, Leipzig. Translation by William Ewald in Ewald (1996), pp 878–920

[23] This is in line with the application of Lakatos' account of scientific research programs to mathematics, as discussed in Hallett (1979).

Carter J (2019) Exploring the fruitfulness of diagrams in mathematics. Synthese 196:4011–4032
Chang H (2004) Inventing temperature: measurement and scientific progress. Oxford University Press, Oxford
Detlefsen M, Arana A (2011) Purity of methods. Philos Impr 11(2):1–20
Elliott KC (2012) Epistemic and methodological iteration in scientific research. Stud Hist Phil Sci 43:376–382
Ewald W (1996) From Kant to Hilbert: a source book in mathematics. Clarendon Press, Oxford. Two volumes
Feest U, Steinle F (2016) Experiment. In: Humphreys P (ed) The Oxford handbook of philosophy of science. Oxford University Press, pp 274–295
Fernández González J (2016) Chemins homothétiques. Bulletin de l'Association Mathématique du Québec 56(2):36–50
Fernández González J, Schlimm D (2023) From a doodle to a theorem: a case study in mathematical discovery. J Humanist Math 13(1):4–35
Ferreirós J (2015) Mathematical knowledge and the interplay of practices. Princeton University Press, Princeton
Galison P (1988) Philosophy in the laboratory. J Philos 85(10):525–527
Gardner M (1983) Martin Gardner's sixth book of mathematical diversions from scientific American. University of Chicago Press, Chicago/London
Giardino V (2018) Manipulative imagination: how to move things around in mathematics. Theoria 33(2):345–360
Hacking I (1988) On the stability of the laboratory sciences. J Philos 85(10):507–514
Haffner E (2021) The shaping of Dedekind's rigorous mathematics: what do Dedekind's drafts tell us about his ideal of rigor? Notre Dame J Form Log 62(1):5–31
Hallett M (1979) Towards a theory of mathematical research programmes I and II. Br J Philos Sci 30(1–25):135–159
Heath TL (ed) (1968) The thirteen books of Euclid's elements, volume 1: introduction and books I, II. Dover Publications. Translated from the text of Heiberg with introduction and commentary by Sir Thomas L. Heath
Klarreich E (2018) A tenacious explorer of abstract surfaces. In: Lin T (ed) The prime number consipracy: the biggest ideas in math from quanta. MIT Press, pp 147–154
Lakatos I (1976) In: Worrall J, Zahar E (eds) Proofs and refutations. Cambridge University Press, Cambridge
Larkin JH, Simon HA (1987) Why a diagram is (sometimes) worth ten thousand words. Cogn Sci 11:65–99
McEvoy M (2023) The new four color problem revisited. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer
Morgan M (2003) Experiments without material intervention: model experiments, virtual experiments, and virtually experiments. In: Radder H (ed) The philosophy of scientific experimentation. University of Pittsburgh Press, pp 216–235
Norton DF, Norton MJ (eds) (2007) David Hume: a treatise of human nature. A critical edition, vol 1. Clarendon Press, Oxford
Polya G (1954) Mathematics and plausible reasoning, volume 1: induction and analogy in mathematics. Princeton University Press, Princeton
Schlimm D (2013) Axioms in mathematical practice. Philos Math 21(1):37–92
Secco G (2017) Proofs versus experiments: Wittgensteinian themes surrounding the four-color theorem. In: Silva M (ed) How colours matter to philosophy. Springer, Cham, pp 289–307
Sørensen HK (2010) Exploratory experimentation in experimental mathematics: a glimpse at the PSLQ algorithm. In: Löwe B, Müller T (eds) PhiMSAMP. Philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London, pp 341–360
Starikova I (2023) Thought experiments in mathematics: from fiction to facts. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer

Starikova I, Giaquinto M (2018) Thought experiments in mathematics. In: Stuart MT, Fehige YJH, Brown JR (eds) The Routledge companion to thought experiments. Routledge, pp 257–278

Stein ML, Ulam SM, Wells MB (1964) A visual display of some properties of the distribution of primes. Am Math Mon 71(5):516–520

Steinle F (1997) Entering new fields: exploratory uses of experimentation. Philos Sci 64(S1): S65–S74

Tappenden J (2008) Chapter 9: mathematical concepts and definitions. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, pp 256–275

Van Bendegem JP (1998) What, if anything, is an experiment in mathematics? In: Anapolitanos D, Baltas A, Tsinorema S (eds) Philosophy and the many faces of science. Rowan & Littlefield Publishers, Ltd, Lanham, pp 172–182

Vold K, Schlimm D (2020) Extended mathematical cognition: external representations with non-derived content. Synthese 179(9):3757–3777

Wagner R (2017) Making and breaking mathematical sense. Histories and philosophies of mathematical practice. Princeton University Press, Princeton/Oxford

Thought Experiments in Mathematics: From Fiction to Facts

Irina Starikova

Contents

1 Introduction 2524
2 Mathematics 2525
2.1 Use of Sensory Experience 2525
3 Thought 2527
3.1 Mathematical Thought Experiments and Imagined Proofs and Calculations 2527
3.2 Mathematical Thought Experiments, Illustrated Arguments and Picture Proofs 2534
3.3 Fiction in Mathematical Thought Experiments 2539
3.4 Impossible Elements in Mathematical Thought Experiments 2541
3.5 *Thick* Concepts 2542
3.6 A Veridical Fictional Imagination 2542
4 Experiment 2543
4.1 Mathematical Thought Experiments and Mathematical Experiments 2544
4.2 Variety of Experience 2545
5 Conclusions 2548
References 2549

Abstract

As in science and philosophy, thought experiments in mathematics link a problem to new epistemic resources that are unavailable in a given practice, e.g., Euclidean geometry. Thought experiments invite us to perform an imaginary scenario involving counterfactual, deductive and sensory elements. This chapter aims to pinpoint the beneficial peculiarities of thought experiments in mathematics in comparison with inferences, diagrams and calculative procedures. Reflection about thought experiments assists us to realize both the limits and opportunities in mathematical thinking. Henceforth, the analysis of examples suggests a broader understanding of *mathematical*, *thought* and *experiment*.

I. Starikova (✉)
UFMG, Belo Horizonte, Brazil

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_45

Keywords

Thought experiments · Mathematical experimenting · Picture proofs · Sensory imagination · Mathematical practice · Informal proofs · Mathematical intuition

1 Introduction

Mathematical thought experiments are a method that allows epistemic progress through experimental thinking, which extends the rules prescribed by a given practice. They often involve sensory imagination as a means of eliciting the benefits of past perceptual experience. Usually, thought experiments are undistinguished from other types of reasoning that use pictures or models. However, they deserve to be valued as an independent method. The importance and the special status of mathematical thought experiments has been stressed by both philosophers and mathematicians. According to Árpád Szabó (1958), as cited in Lakatos (1976): "Thought-experiment [...] was the most ancient pattern of mathematical proof. It prevailed in pre-Euclidean Greek mathematics."[1] More recently, Yuri Manin (2002, p. 6) pointed out that "Successful thought experiments produce mathematical truths which, after being accepted, solidify into axioms, and the latter start working on the treadmill of logical deductions."[2] The intriguing question, which will be addressed in this chapter, is to investigate what is so special about "the most ancient pattern."

"A mathematical thought experiment" appears to be a philosophically challenging concept. It implies that mathematics could be experimental, and moreover, experimental in thought. The two following *observations* raise doubts.

(1) An *experiment* is an a posteriori method, whereas mathematics is usually conceived as an a priori discipline.
(2) Pure mathematics is generally understood to be "processed in thought" anyway; and if *thought* is refined down to "imagination" we are back to square one, because the latter is not an a priori source for gaining strict mathematical knowledge.

No surprise therefore that philosophers question the very existence of mathematical thought experiments, and tend to fit apparent candidates into traditional mathematical species such as proofs. For example, Norton (2004) regards all thought experiments as illustrated arguments, and Brown (1991, 2007) takes picture proofs as analogues for thought experiments in science. Observation (1) about *experiments in mathematics* is addressed by Van Bendegem in this volume. This chapter however is concerned with Observation (2), and will demonstrate that thought experiments do exist in mathematics as an original form of reasoning.

[1] This encouraging claim of Szabó is quoted in Lakatos (1976) footnote 10.
[2] Manin (2002, p.6).

The question to start with is therefore, do the independent species of mathematical thought experiments, as praised by Manin and Szabó, exist at all? In short, the answer to this question has been inclining from No to Yes along with the development of each constituent part of the concept: *mathematical*, *thought* and *experiment* (M-T-E). In this chapter, the understanding of these terms will help us to realize that mathematical thought experiments contain features that inferences, visual proofs and calculations do not have. In particular, they are:

- *Mathematical* for solving mathematical problems, but not restricted to a priori evidence.
- *Thought* for the variety of narratives, imaginary tools and experience which can be involved, but not reduced to the prescribed algorithms of a given practice.
- *Experiments* with varied epistemic aims, but not only for accepting/refusing a hypothesis.

To unfold these properties, one can start with phenomenology. What is noticeably distinctive of thought experiments is that they are typically expressed in the form of a *narrative* starting with "Imagine. . .." The narratives in mathematical thought experiments, which guide the reader via the imaginary scenario, are often based on visualizations. This has been stressed by Brown and Fehige (2019) and Buzzoni (2011). Starikova and Giaquinto (2018) pinpoint thought experiments in mathematics that involve physical properties similar to those in science. Others, such as Heinzmann (2022) and Glas (2009) accept also formula-based mathematical thought experiments. The examples will demonstrate a variety of narratives in mathematical thought experiments based on diagrams (understood as visualizations of an established practice), pictures (free-style visualizations), metaphors, fictions and formulas. This chapter will be structured around the M-T-E and the types of narratives. The properties specific for mathematical thought experiments will be discussed under this structure.

2 Mathematics

2.1 Use of Sensory Experience

It is usually assumed that mathematical beliefs cannot be justified by empirical experience, which would imply either deductive or heuristic but not evidential roles of mathematical thought experiments. Meanwhile, mathematical facts can be known without empirical justification or with it, for example, in replying to the question: How many vertices does a cube have? Inspection of a material (or imaginary) cube and counting its vertices provides evidence for the conclusion that a cube has eight vertices. But is this really a *mathematical* fact? Why not? It is a very simple fact, but we can extend the problem: Do all platonic solids have the same Euler characteristic? Surely the answer to this question is a mathematical fact. And it can be verified in the same way. Physical models of each of the five platonic solids can be visually inspected to find out whether $V - E + F$ is the same for all of them.

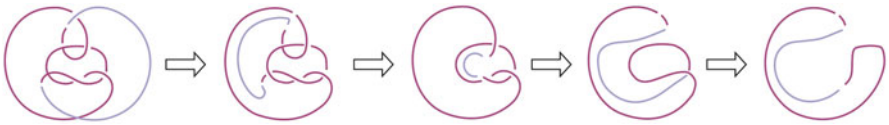


Fig. 1 This diagram depicts a possible thought experiment: imagine if you were step by step unfolding the initial diagram; after a few steps you end up with a diagram without crossings, which is known as an unknot

The visual inspection provides empirical evidence for a positive answer. Similarly, single-digit addition facts can be acquired empirically, from experiences of counting.[3] Such empirical evidence can be also provided by thought experiments. In the following, more examples are presented.

2.1.1 Knot or Unknot?

In knot theory, by trying a series of moves, one can find out whether the original diagram is a diagram of the unknot or not (Fig. 1).

Here we accompany the thought experiment by the diagram to demonstrate all the steps, but one should really experiment on one's own: performing a series of actions in order to test the hypothesis that the original diagram is a diagram of the unknot. In this example we are imagining a *physical* possibility, at least partly based on experience with string, yarn, thread etc. Imagining makes it reasonable to believe the mathematical claim and leads to the discovery. This is possible because visual thought experiments take place in the context of background knowledge about links between idealized physical objects and transformations that can be visualized. These links belong to the foundational aspects of knot theory, and often the exposition of the foundations reveals that the mathematical definitions are tailored to represent the former. In imagination we apply our past sensory experience of dealing with objects similar to the diagram's objects, to make the unfolding moves. In the following mathematical thought experiment, the diagram is also treated as a physical object.

2.1.2 Visual Symmetry in Graphs

Below is a graph and we want to know whether it is homogeneous, i.e., all the vertices have the same structure in terms of outgoing and incoming edges. This can be quickly done by a simple thought experiment. Imagine rotating the graph in respect to its center in the plane, where it is depicted. The vertices will move around but the graph will stay the same (Fig. 2).

Treating a diagram as a physical rigid object is widely used in mathematical practice and is very similar to scientific experiments. For example, referring to experimental studies of mental transformations, Shepard (2008, p. 5) points out that "[T]he cognitive processes in the cases of physics and mathematics are often

[3] Starikova and Giaquinto (2018).

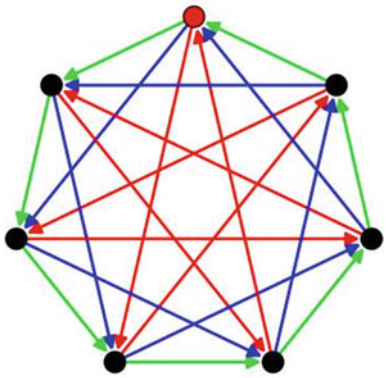


Fig. 2 An imaginary rotation of this graph helps us to notice that all the vertices are the same, in respect to the number, the colors and directions of the edges

more similar than the traditional distinction between mathematical and the empirical sciences would suggest."

To sum up, *mathematical* thought experiments do not have to be limited to a priori evidence, but instead can be understood in a broader sense: i.e., searching for evidence to resolve a *mathematical* problem. More examples will follow; as for a *non*-example, *Achilles and the Tortoise* by Zeno is a good one. It provokes new insights into our understanding of motion and space-time relation, but it contributes more to metaphysics and physics than it does to mathematics (Buzzoni 2011).

3 Thought

3.1 Mathematical Thought Experiments and Imagined Proofs and Calculations

Thought experiments are often understood as being limited to performances of physical experiments in the mind. Such a reading leaves no room for thought experiments in pure mathematics, because they lack real physical counterparts. This view has been defended by Buzzoni (2008) and revised in Buzzoni (2011).[4] Although almost

[4] As Buzzoni (2008, p. 71) put it: "Mathematical counterexamples, however, do not arise from experimental practices, because they involve no interaction between our body (or the computer as its extension) and mathematical objects." And also "The actual performance of a mathematical thought experiment intended to demonstrate or to offer an 'argument' for a theorem coincides with the proof of, or argument for that theorem, thus leaving no room for a separate real performance of the experiment." Buzzoni (2008, 77). Later Buzzoni accepted mathematical thought experiments under certain conditions: "1) The distinction between logical-formal thinking and experimental-operational thinking must not be underestimated. 2) The separation between the context of discovery and that of justification – a separation advocated by logical empiricists but difficult to support after the "pragmatic turn" – must be rejected. 3) The distinction between mathematical thought experiments and formal proofs must be regarded as one of degree, and not as a distinction in principle, however important this distinction may be de facto for particular or local purposes." Buzzoni (2011, p. 6).

any experiment is preconceived, "...[t]hought experiments should be distinguished from thinking about experiments, from merely imagining any experiments to be conducted outside the imagination. . ." Brown and Fehige (2019). Not only are some thought experiments physically unrealizable even in principle, as in *Schrödinger's Cat* or the *Ship of Theseus*, they are designed as unrealizable (Sorensen 1992, p. 74).

Understanding a *thought experiment* as being just an "experimenting in mind" misses the point, as the mind's lab for thought experiments can be as well-equipped as any lab, but with its own instruments. (Mathematical) thought experiments involve sensory experience, imagination and various types of reasoning, such as counterfactual, inferential and analogical. In each particular case, these come in degrees. For example, Balcerak Jackson (2016) develops the view that different thought experiments have different functions and depending on the function they can be based more on imagining, supposing or conceiving. Murphy (2020) also argues for a variety of thought experiment types and cognitive modalities involved, depending on the problem. A fundamental function of thought experiments is to evoke relevant past experience or "tacit knowledge," as defined in Polanyi (1966). He argued that this type of knowledge may come from any part of one's personal background, but is often related to sensory (i.e., visual) activity and once articulated, may lead to a crucial insight. Imagination itself can be of various types, each with its own role (e.g., Arcangeli 2010).

To see that "mathematical thought experiments" are more than just a preview of a proof, it may be useful to start with a variety of narratives that they can have. The two examples above have a narrative based on diagrams. The following cases demonstrate a richer diversity of the physical elements involved.

3.1.1 A Graph Diameter

Graphs, that are considered in the mathematical field of graph theory, are combinatorial objects containing collections of vertices and edges. It is further possible to introduce the notion of *distance* between two vertices: the number of edges in the shortest path between them. Take a cyclic graph as depicted in Fig. 3. One can also introduce the notion of a *diameter* of a graph, being the maximum of all distances. In this small graph one can count the diameter ostensibly as "one, two, three," but a

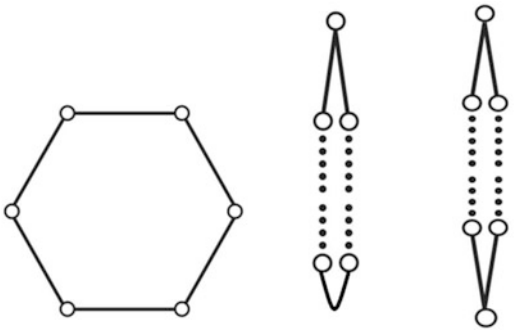


Fig. 3 The first figure is the canonical visualization of a cyclic graph of six vertices. The other two figures demonstrate the two cases: odd and even number of vertices

thought experiment reveals a general formula. Just imagine that the graph is heavy, like a necklace – hold it by any pearl, and let the rest of it go. What will happen? The rest will fall as far as the thread will let it, and we observe that the maximum distance between the top pearl (the held pearl) and any other will be *n/2* or *(n-1)/2*, where n is the number of vertices in the graph. Since the top vertex is chosen randomly and we assume that all the edges have the same length, the result does not depend on the choice of pearl held.

Why are there two solutions? They appear from a finer-grained examination of the graph. Finer-grained imagery results from taking into account the coupling of pairs, i.e., that the graph's vertices will hang down in pairs as in the figure above. The shapes of the lower end of the necklace in both cases are revealed to us by visualization. This series of cognitive actions performed helps us to find the shapes that the necklace would take, and we use these shapes to find the general formula.

Furthermore, we do not know which vertex is the *antipode* (the most distant from the one chosen). When we dangle it down, it is the *gravity* that helps to articulate the paired arrangement of the vertices and the antipode. The sensory experience reveals something new: that there are two options for the number of vertices, one odd and one even.

One may ask if thought experiments are possible with objects that do not have a visual representation within mathematical practice. One option is first to visualize such objects, e.g., as graphs, and then some new perspectives may occur. Here is an advanced case where graphs are used as *mediators* between algebraic groups and geometric spaces.[5]

3.1.2 The Large Scale Geometry of Groups

The large-scale geometry of groups is a revolutionary idea that originated in the work of Mikhail Gromov and gave rise to a new powerful branch of mathematics, geometric group theory.[6] Let G be any infinite group that has different finite sets of generators S, T and maybe others, where generators are some elements of the group. Each set generates the group with a particular structure, which can be depicted by a Cayley graph (after Arthur Cayley, to whom we owe this important concept connecting groups and graphs). In a Cayley graph each vertex corresponds to a group element, and each arrow corresponds to the multiplication of an element by a generator, and each generator is represented by a unique color.

From a visual representation of a Cayley graph of generated groups (G, S) or (G, T) – it does not matter which – we can try zooming out in our imagination, in so far as the fine details of the Cayley graph get lost and the features of the object's large-scale geometry now come into view. The picture below represents a typical section of three Cayley graphs of the group of integers over addition generated by

[5] The mediating function of graphs appears in various cases. For example, it has been exploited by Arthur Cayley and John Sylvester in applying group theory to studying chemical structures. More recently, Carter (2018) provided an example of directed graphs serving as mediating objects between graph algebras and their K-groups.

[6] For a more detailed philosophical discussion of this example see Starikova and Giaquinto (2018).

sets {1}; {1,3} and {1,2} correspondingly, i.e., Cayley graphs of **Z**(+,{1}); **Z**(+,{1,3}) and **Z**(+,{1,2}).

The thought experiment invites us to imagine that from faraway all three look roughly linear, i.e., roughly similar to a section of the real line. Mathematically speaking there is a quasi-equivalence between integers and reals, which may sound nonsense at first, but this reasoning can go even further. While we are zooming out from the picture, the large-scale geometry looks the same for whichever Cayley graph we start with, that is, regardless of a generating set. At least to a geometer's imagination it is possible. This suggests that properties that are invariant with respect to a generating set, can be discovered by attending to the coarse geometry of the object we reach by zooming out. In Gromov's (1993) own words:

> This space [the space of the Cayley graph of $\Gamma = (G, S)$ with the shortest path metric] may appear boring and uneventful to a geometer's eye since it is discrete and the traditional local (e.g., topological and infinitesimal) machinery does not run in Γ. To regain the geometric perspective one has to change his/her position and move the observation point far away from Γ. Then the metric in Γ seen from a distance d becomes the original distance divided by d and for $d \to \infty$ the points in Γ coalesce into a connected continuous solid unity which occupies the visual horizon without any gaps or holes and fills the geometer's heart with joy.

To get Gromov's idea one must bear in mind that neither groups, nor the corresponding Cayley graphs, are geometric objects. *Edges* (in set-theoretic terms) are not line segments, but just pairs of vertices and so contain no points between endpoints. In order to make a graph a metric space we can introduce a *metric* in a graph as the length of the shortest path between any two vertices (given that each edge has a unit length). For example, the distance between elements 1 and 4 in the Cayley graph of Z(+,{1}) in Fig. 4 is equal to 3. Still this metric is discrete and does

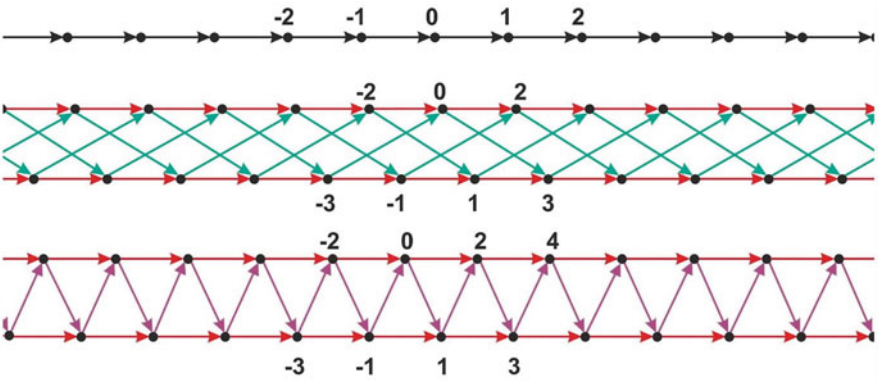


Fig. 4 Here are three Cayley graphs of the group of integers over addition **Z**(+) generated by sets {1}; {1,3} and {1,2} correspondingly: **Z**(+,{1}); **Z**(+,{1,3}) and **Z**(+,{1,2}). Although the different generating sets determine the different structure of the graphs (in respect to the number and the colors of outgoing and incoming edges), from a distance, all three have the same linear appearance, i.e.,, the same large-scale geometry

not help us to discover any interesting geometric relation, as the distance between any two vertices is just an integer (an example will be given in the next subsection). However, by zooming out from the group Γ in our imagination, the discrete object can be transformed in appearance into a space with continuous metric, i.e., having (non-negative) real values. The Cayley graph of integers transforms into a continuous line (Fig. 5), and in the two-dimensional case, which corresponds to a grid of integers, it transforms into a plane (Fig. 6).

This thought experiment is not limited to Euclidean metric spaces. Zooming out can result also in a non-Euclidean metric space. Look at the Cayley graph of the group F_2, which is the free group with two generators, *a* and *b* (Fig. 7 on the left). It is a graph that depicts the set-theoretic information about the group and curvature is not among this information. We may be reminded however of the Poincaré disc model of the hyperbolic plane H_2. The Poincaré open disc model of hyperbolic geometry represents geodesics (i.e., the shortest paths) as diameters or arcs of circles orthogonal to the boundary. The unit distance is represented by ever shorter *Euclidean*

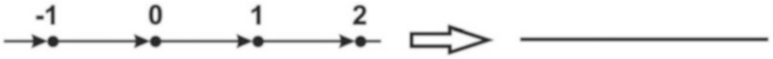


Fig. 5 This figure suggests the similarity between a Cayley graph of (**Z**, +, {1}) and the line **R**

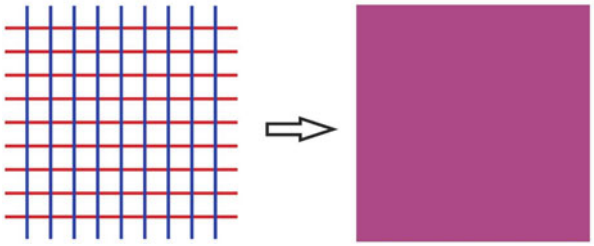


Fig. 6 This figure suggests the similarity between a Cayley graph of ($\mathbf{Z}^2$, +, {1}) and the Euclidean plane $\mathbf{R}^2$

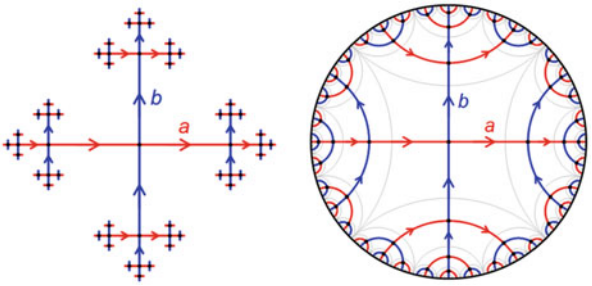


Fig. 7 From https://en.wikipedia.org/wiki/User:Jim.belk/Draft:Free_group. This figure depicts mapping of the Cayley graph of the group F_2 to the hyperbolic Poincaré disc model of the hyperbolic plane H_2. It suggests that the graph, and therefore the correspondent group F_2, have a hyperbolic structure

distances as one moves from the center out towards the boundary (where the boundary is not part of the model). In the graph drawing too we have to decrease the length of edges to avoid overlaps. If we zoom out from the graph we can imagine it embedded in a Poincaré disc (Fig. 7 on the right). This leads to an insight that F_2 is similar to a hyperbolic plane H_2 regarding their large-scale geometry.

3.1.3 Zooming into a Hyperbolic Group

As soon as we talk about a hyperbolic space we have to obey *the triangle condition*, i.e., that the sum of the inner angles of a hypo-triangle is less than π, in contrast to a Euclidean triangle. Except there are no *triangles* in graphs in the sense of geometry. Nevertheless the following thought experiment helps to circumvent the discrete nature of groups (and their Cayley graphs) and see that a graph does not need to be a geometric object to be *hyperbolic*.

As well as zooming out, we can zoom back into the details of a graph, but this time with the geometric perspective in mind. Imagine that you pinch an arbitrary Euclidean triangle by its side's middle points towards the center, as in the left-hand side Fig. 8. It will become *hyperbolic*. Then pinch it until the middle points collapse into one. The already thin hyperbolic triangle will become a tripod. Now zoom into the graph of F_2 and try to tune your eye into a geometric perspective: you may notice or imagine that the graph consists of tripods, i.e., 0-thin triangles, as in the right-hand side Fig. 8. As all triangles are thin in any hyperbolic space (with at least one side pinched in), then all triangles in this graph are also 0-thin. Therefore, the graph is hyperbolic.

To put it in mathematical terms, the Cayley graph of a finitely generated group can be treated as a metric space. The metric space S is *hyperbolic* if and only if for some non-negative real δ, all triangles of S are 'δ-thin' and a triangle is δ-thin if and only if for each side of the triangle every point on that side is no greater than δ from some point on one of the other two sides. This definition makes room for spaces with all triangles 0-thin, which includes the Cayley graph of F_2.

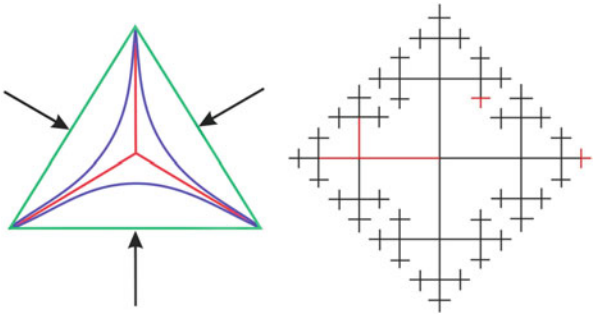


Fig. 8 In our imagination, we pinch the middle points of the green triangle's sides towards its center. Then we observe that this will lead to a thin hyperbolic triangle and, if pinched further, it will lead to a tripod. Such tripods are similar to segments of a Calyley graph of the free group F_2, as depicted on the right

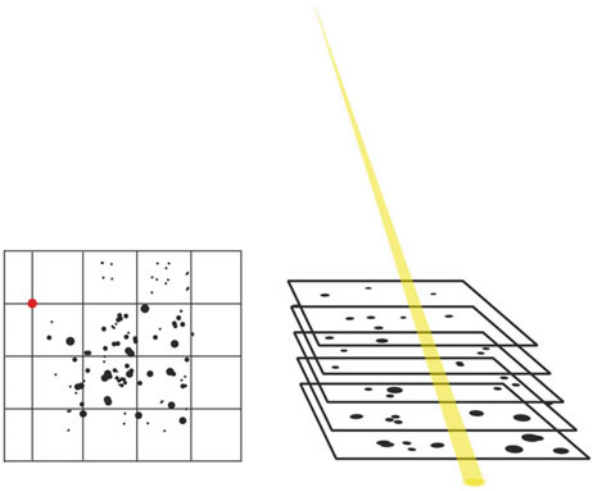


Fig. 9 The first figure depicts the ink blot on the grid with step n. In the second figure, the plane is cut into squares along the grid and they are stacked using only parallel translation. Imagine looking through them to the light. There must be at least a tiny ink-free spot, through which the light will go. Otherwise, the total area of the blots would be no less than n^2, and would contradict the condition of the problem

The graph diagrams in this example are only a basis for mathematical thought experiments, visual representations of the abstract structure. It is not at all obvious how one can drop all the important properties of the graphs that closely look so different. It is zooming out that helps us to see something as a whole and it is the zooming in that makes the triangles noticeable. Finally, it is pinching the triangle that reminds us of the alternative (*thinness*) test for hyperbolicity, and helps to make sense of the *triangles* in a graph.[7]

3.1.4 The Ink Blot Problem

In the case above, the set of objects to experiment with has been formed by visualized graphs. Are mathematical thought experiments still possible in non-diagramatic domains such as algebra or number theory? Here is an example to consider.

On the transparent plane, on which a rectangular grid with step n is applied, ink is poured in the form of a multitude of blots of different sizes and shapes (Fig. 9). The total area of the blots is less than the area of one square. Is it always possible to move the whole array of blots so that ink does not cover any grid point? The thought experiment is the following: cut the plane along the grid into separate squares and put all the squares with ink in a stack (using only parallel translation to do so). Look through them to the light. If we do not see at least a tiny ink-free spot, then the total area of the blots is no less than n^2, which contradicts the condition of the problem. So now, thinking of all the grid squares reassembled as at the start, each grid square has an ink-free area in the same relative position as every other grid square, positioned from its lower left corner (for example), x units to the right and y units above. Now visualize moving the whole array of ink spots rigidly as x units left and y units down.

[7] For more details on this case see Giaquinto (2020).

As every grid node (intersection point of grid lines) is the lower left corner of some grid square, the result of the move will put every grid node into an ink-free spot.

This puzzle is a variation of the famous *Minkowski's theorem,* which gave rise to the geometry of numbers[8]; and the thought experiment is a reconstruction of the argument of the proof. Independently of whether Minkowski used this or something similar to discover the argument or not, the thought experiment explains the proof and brings conceptual clarity. It helps solving the puzzle too. The procedure of cutting and stocking squares corresponds to mapping the plane to the factor-space with points which differ by a given shift n. This procedure is often included in textbooks to explain the notion of a factor group. Through the transparent paper, we can see all points by looking only at a single square (the fundamental domain), which represents the factor group. The experiment links the concept of set with the concept of area: if the fundamental domain is fully covered, the area is at least n^2.

This mathematical thought experiment is based on the visual metaphor of an ink blot, and the previous examples are based on visual representations. These are what Brown compares to picture proofs: "...[I]n mathematics picture proofs are the analogs to thought experiments" (Brown 2007, p. 3). However, the following examples will demonstrate that mathematical thought experiments do not reduce to diagrammatic proofs or picture proofs.

3.2 Mathematical Thought Experiments, Illustrated Arguments and Picture Proofs

As in all the cases above, a picture or a diagram is just the basis for mathematical thought experiments. They come with a flexibility of imagination which is unavailable in static pictures. At the same time, some mathematical thought experiments extend inference to the experimental search for an invariant. For example, the rotation of a symmetric graph provides the evidence that the graph is homogeneous (i.e., invariant under variation of vertices). The following example plays an important role both in mathematics and in philosophical discussions about formal and informal proofs (Starikova and Van Bendegem 2021). It has been presented as a

[8] Bensimon (2009, p. 2) draws attention to the important shift of thinking due to this theorem: "In 1891, mathematician Hermann Minkowski (1864–1909) gave a lecture in Halle, Germany, where he introduced the lattice as the collection of points with integer coordinates in the coordinate system with perpendicular axes. This talk indicated his shift in focus from solving problems in Number Theory with geometric intuition towards "Geometrie der Zahlen," or, "the Geometry of Numbers." He defined this term to mean geometrical investigations of the lattice and its associated bodies. In an 1893 paper, Minkowski presents a theorem that defines parameters for the volume of a certain body, in order for that body to contain a lattice point in the rectangular Euclidean coordinate system. He begins his paper, "In Number Theory, as in all other fields of analysis, the inspiration often comes from geometrical considerations even though at the end maybe only the analytical verification is shown." Indeed, from Minkowski's "geometrical considerations" a new generation of mathematical knowledge was derived. In essence, Minkowski laid the foundation for the modern theory of convexity."

picture proof (Brown 1991, p. 50). There is also a view that it is an illustrated argument, and the picture is an auxiliary (Norton 2004). However, it can be seen as a genuine mathematical thought experiment.

3.2.1 The Mutilated Chessboard Problem

A chessboard has 8 × 8 or 64 squares. If you remove two opposite corner squares, 62 are left. Is it possible to tile the board with 31 non-overlapping dominoes, where a domino covers two adjacent squares? The answer is "No," because both of the absent corners are white, and since each domino covers black and white adjacent squares, we are left with two black squares uncovered. Could this reasoning constitute a thought experiment, and thereby exceed an argument and a picture proof? (Fig. 10)

According to Brown's realist view, we grasp the solution immediately thanks to the special capacity of intuition: we just look at the picture and "see" the answer through the mind's eye. This is how a picture proof works (Brown 1997). However, this view provokes some buts and doubts.

John Norton, contrary to Brown, claims that it cannot be immediate, and many agree that there is some reasoning leading us to our conclusion. What could that be? Norton's view is extreme: such picture proofs and thought experiments are in fact "decorated" arguments, and decorations play no role in the argument, so a specialist could simply write out the argument omitting the picture. With respect to a similar example, Norton says: "This may take more or less time according to our mathematical abilities but will be rapid in the case of someone with a mathematical background" (Norton 1996, p. 352).

Interestingly, different cognitive approaches to solving the mutilated chessboard problem have been investigated by psychologists Kaplan and Simon (1990). The researchers report that although some deductive and counterfactual reasoning has been registered in read-out loud protocols, all participants "noticed parity first, and later made an inference that dominos must cover one square of each type" (Ibid, p. 390). The parity invariant (that each domino always covers or *matches* one white and one black square) is the key to the informal argument. This key does not follow

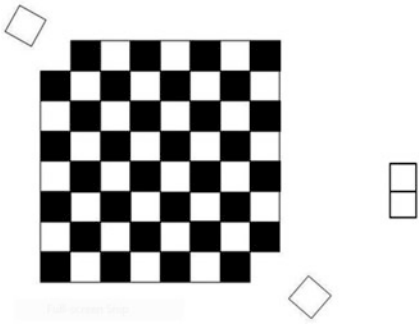


Fig. 10 Is it possible to tile the 8 × 8 board by 1 × 2 rectangles? Imagine that this is a 8 × 8 chessboard with two diagonal corners deleted. The chessboard representation gives a crucial hint for solving this problem, namely the parity of the black and white cells. Each domino must cover two adjacent cells in both colors, but our mutilated board is missing two cells of the same color

from the "hidden in the decorations" inference, but from the sensory experience.[9] Also, contrary to Norton's prediction, attempts to approach this problem mathematically have significantly slowed down performance. Only 12 out of 25 science students found the solution without some sort of a hint. The skill of recognizing invariants turned out to be crucial, whether they come from mathematical expertise or not: "Every single fast subject generated one or more perceptual invariants within the first 10 min, while none of the slow subjects did so!" (Ibid, p. 409–410) Moreover Kaplan and Simon report that *perceptual* invariants (related to color and the position of removed corners) turned out to be more effective than non-perceptual invariants (i.e., related to strategies such as decomposition, mathematical properties etc): "[F]ast subjects possess a heuristic that the slow subjects do not – of paying attention to the perceptual features of the problem." Another experiment by Bilalić et al. (2019) showed that chess players were faster than non-chess players in solving the problem. This is explained by the chess players' easier retrieval of the required color information (figuratively speaking, chess players know every corner of the chessboard). This suggests that sensory experience can play a crucial role in noticing invariants. As soon as the key invariant is found, the deductive part does not take much time.

On the other hand, immediate intuition of the parity invariant, due to just seeing the picture as Brown suggests, also seems to be a rare case. Indeed, it would mean seeing the mismatch of the board and the dominos. The same picture alone however can equally lead us to a clever solution, or to an exhaustive search (i.e., a simulation of a mathematical experiment).[10] To discover the clever argument, we must change from tiling to matching mode of thinking. Presumably we must start trying to tile, then stop it and notice that each white square matches a black one. What could motivate this conceptual shift? Some experimentation in imagination should take place, and here are reasons for that.[11]

According to Kaplan and Simon "fast subjects are noticing not only more things, but a wider variety of things" (p. 412) and that the principal difficulty for slow participants was to *distract* them from the tiling mind set: "Subjects are blocked from exploring parity by their preoccupation with other cues and representations" (Ibid, 392). This suggests that a counterfactual situation where the tiling is already done could free our attention for other (namely set-theoretic) features. This is what Kuhn (1981 [1963]) called a *conceptual change*, and he recognized it as one of the principal goals of thought experiments.

[9] Starikova and Van Bendegem (2021) argue that although formalization is possible and there are even many of them, in order to improve such a formalization and connect the problem to a broader context, e.g., solving related problems, one constantly returns to the image.

[10] Kaplan and Simon (1990) investigated performance on this problem with 25 graduate science students. The think-aloud protocols produced by the participants indicated that initially *all* began by placing the dominos on the board, and eventually *all* tried to find a solution based on the recognition that each domino always covered black and white neighboring squares.

[11] For example, in contrast to immediate intuition, Kant thought that it is thanks to the construction of the concept in our imagination that we come to a correct judgment.

The conceptual change can be done by a thought experiment, which can go as follows. Looking at the picture, we imagine that the truncated board has been mostly tiled from the middle, and we are on our way to the absent corners. *Before* imagining an almost fully-covered board, one may not even pay attention to their colors, which is normal, because coloring plays no role in exhaustive tiling. As soon as we are left with only a few squares, our attention shifts from the global task, i.e., tiling the board, to the local one of fitting individual tiles. At this moment the parity stands out and we see it as matching, i.e., in a set-theoretic mind set.

What is the condition for discovering invariants? They become visible in the background of some *variations*, often present in *experiments*. Mach and Herbert (1976, p 139) characterized experimentation as "the method of variation." In particular, he points out: "By varying the conditions (continuously if possible), the scope of ideas (expectations) tied to them is extended: by modifying and specializing the conditions where we modify and specialize the ideas, making them more determinate, and the two processes alternate." Pictures do present the information, but sometimes one has to experiment with them in order to recognize invariants. For example Cayley graphs alone do not speak about their coarse geometry, but by changing the zoom we notice that some properties do not change with distance.

In the mutilated chessboard problem, one can experiment with laying down the dominos, the size of the board and the labeling of the board's squares. Max Black (1946), who first posed the problem, did not present a diagram (he left it to the reader's imagination); and Claude Berge (1958) gave it a most difficult appearance as a monochromatic grid, which would be a canonical diagram in the tiling theory. This theory studies the coverings of two-dimensional regions with various shapes, so that there are no gaps or overlaps (like floor tilings). The mutilated chessboard problem is one of many such problems. It asks whether it is possible to cover the $n \times n$ region with 1×2 rectangles. The brilliant analogy of the chessboard and dominos problem is the thought experiment. It introduces perceptual parameters (i.e., the colors) unavailable to the standard formulation of the mathematical problem. Finally, Kaplan and Simon, searching for a better prompt for their subjects, were experimenting with the boards using the words "pink" and "black," and boards with "bread" and "butter" in the place of black and white cells. Also variations of tiling attempts and especially on a smaller board, e.g., 4×4, eventually lead to isolated black squares and prompt the color (parity) invariant.

The flexibility of our imagination is an important advantage of mathematical thought experiments. Often a physical image cannot represent features *simultaneously*, whereas imagination can. In the mutilated chessboard example, imagining an almost covered chessboard, we can "see" the colors "through" the dominoes and collate them with the colors of the missing corners. Similarly, a receding physical image of a Cayley graph would simply collapse into a point. Zooming out from it in the thought experiment, we can purposely let the plane grow and the horizontal lines extend, while the graph details are becoming smaller. This can hardly be done by changes in the physical picture, but it can be by harnessing the power of the

imagination.[12] Crucially, in most cases we can modify our imaginations, even if our experience is always of concrete particulars. According to Berkeley's *Principles*, we can imagine a triangle only of one determinate shape, namely an equilateral triangle, but in our imaginations we can exchange it for any other type of triangle in accordance with our aims.

This leads us to an interesting observation: in the mutilated chessboard problem and especially the knot diagrammatic sequences, picture proofs can be seen as resulting *freeze frames* or a visual synopsis of mathematical thought experiments. Perhaps for this reason, although picture proofs are more universally recognized, it is less well understood that they are shortcuts to the latter. Picture proofs and thought experiments are similar in the sense that they are not just inferences, yet they can provide evidence for a mathematical hypothesis. They differ in that in a picture proof, visual transformations are limited by the layout, whereas in a thought experiment, imaginary transformations would go beyond the media in which the diagram is realized, e.g., (a) the paper becomes transparent or rigid, or perhaps it grows as long as we are zooming out, or (b) we can embed a chessboard on a torus, or (c) we can add or remove dimensions and physical properties. A good example is doubling a square from Plato's *Meno*, in which we can distinguish three options (although in other examples the difference may be small):

1. A diagrammatic proof – e.g., solving problems with the use of diagrams in accordance with the given practice;
2. A picture proof – (immediate) grasping the idea of the proof from a picture,
3. A thought experiment, based on a diagram or picture and involving changes which lead to some sort of invariant.

The left figure below can represent a diagrammatic proof in Euclidean geometry (of the proposition that the areas are doubled), namely by looking at the correspondent triangles using propositions, definitions and axioms. The second figure, with geometrically atypical red arrows, is a picture proof; but both proofs can be preceded by a thought experiment, i.e., imagining the square as a napkin and bending its corners to meet the center (Fig. 11).

Besides, diagrams and pictures are only visuals, whereas thought experiments allow for multiple types of stimuli apart from the visual, e.g., motor-sensory as in the necklace example. In fact, thought experiments can motivate access to the whole

[12] There can be cases when mathematical thought experiments do not have enough demonstrative power, because of imaginary limits. For example, in order to grasp symmetries in group theory, it may be more effective to use physical models. Thus Klein advised us: "Turning now to the configuration of the dihedron, I beg the reader, here and in the similar developments in the paragraphs which follow, to make the corresponding diagrams, or to think out for himself directly by aid a model–which is easily constructed–the properties under consideration. For we are treating of concrete matters, which may easily be conceived with the assistance of the suggested aids, but which may occasionally offer difficulties if these are neglected." Klein (1888, Chap. 1, Sect. 4 to 12).

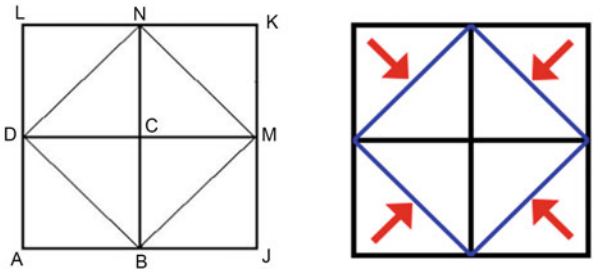


Fig. 11 The first figure is a part of a diagrammatic proof. Four adjacent copies of a given square ABCD give four times bigger square AJKL. Drawing diagonals splits it into eight equal triangles and gives us the inner square BMND. Since the ABCD contains two such triangles and BMND contains four, the latter is double the former. The second figure is a picture proof, which helps one grasp that the blue square is twice as big as any of the small black squares. In fact, both proofs can be based on the following thought experiment. Imagine a square napkin, folded in four (corresponding to the size of ABCD). Unfold the napkin (now corresponding to AJKL), and then fold the corners to coincide in the center. The result is the required doubled square (BMND)

personal experience, and this will be discussed later in this chapter. Diagrams and pictures are the most common basis for mathematical thought experiments, but the latter can easily incorporate artwork elements such as those in the *ink blot problem* or in the *dividing of 17 camels*, coming up next. To sum up, the mathematical thought experiments above extend both our deductive and visual intuition. In particular, they involve experimental variations of the settings that help us to find invariance.

3.3 Fiction in Mathematical Thought Experiments

What makes a big difference between a thought experiment and mere thinking about a real experiment is such a peculiar attribute as fiction. The former can involve a narrative and impossible elements. Indeed, when planning a proper scientific experiment, one usually does not create a story about killing a cat or riding a light beam, but rather uses theoretic terms describing settings such as "a body moving at the speed of light." Thought experiments have a selective, local and momental character, namely they have "evidential significance only historically and locally [. . .]" (McAllister 1996, p. 248). Each thought experiment is unique as a particular story, and many have a name, e.g., the *Balls and Vase Problem,* or *Hilbert's Paradox of the Grand Hotel*. This makes thought experiments closer to literature, theater, cinema and arts. The design of a thought experiment depends upon the creator's personal experience, professional background, and style of thinking. The art of thought experiments is then to tailor a scenario and choose whatever details make them effective, explanatory and revelational.

The fiction-style thought experiments exist in mathematics, and make it more engaging and accessible for non-mathematicians. But crucially, they evoke a broad

range of our tacit knowledge and experience, and direct our thinking to non-standard solutions or questions. Below are two examples.

3.3.1 The Art of Dividing 17 Camels

A father left 17 camels to his three sons, and according to the will the eldest son should be given half of all the camels, the middle son one third, and the youngest son one ninth. This is hard to do, but a wise man helped the sons: he *added* his own camel, so the eldest son took 18/2 or 9 camels, the second son took 18/3 or 6 camels, the third son 18/9 or 2 camels, while the wise man took back his own camel and went away.[13] To approach this problem, we first try to divide 17 by 2, which fails, as does dividing by 3 and by 9. So instead we try a mathematical thought experiment, which leads to a solution. The formal way to put it is $1/2 + 1/3 + 1/9 = 9/18 + 6/18 + 2/18 = 17/18$.

Again, someone like Norton would say that the story is an unnecessary illustration, since the problem implies that it should be solved in integers (representing camels). However, solving the problem without this thought experiment would require at least some mathematical skill and knowledge of the smallest common divisor, whereas the thought experiment requires none of these. The advantage of the story-narrative is that it points clearly to the key condition that the problem has to be solved in integers, as the division of living camels is unacceptable (for this problem). Therefore, the narrative guides us towards experimenting with recombining sets, rather than performing a standard algorithm.[14]

3.3.2 Flatland

Another well-known example of a story-narrative is *Flatland*, a satirical novella by Edwin Abbott (2009), where we are invited to imagine a 2D world inhabited by planimetric figures.[15] What mathematics would they have? How could we explain our 3D world to 2D inhabitants? How can they possibly understand that a space is curved or that the universe is finite but has no boundary? *Flatland* challenges the reader to imagine particular 2D life situations in detail. It also helps us to revise our default intuitions about geometry, and understand how to relate to different dimensions. The details of fiction-based mathematical thought experiments may not play a direct role in the argument, but they are not mere decorations either. Rather they are purposefully tuned to guide us towards what needs to be looked at.

[13] The oldest written version of this problem found by Ageron (2013) was written in Persian by the Iranian mathematician Muhammad Mahdi al-Naraqi (1716–1795).

[14] This popular mathematical parable became an iconic example of the method of adding and removing structure. For example, a mathematician refers to this strategy in a proof of an important lemma from graph theory and in the apportionment of procedures called divisor methods. https://mathoverflow.net/questions/271608/17-camels-trick

[15] Carl Sagan (2009) in his YouTube video discusses the possibilities of higher dimensions to the physical universe based on the *Flatland* concept, and proposes his thought experiment solutions.

3.4 Impossible Elements in Mathematical Thought Experiments

For thought experiments in general, McGinnis (2018) distinguishes *idealized* and *fictional* thought experiments on the basis "that the former at least give the appearance that they are physically possible." For him, examples of fictional thought experiments are zombie worlds or persons splitting and recombining, but Einstein riding the light beam is an idealization. But what about Abbot's 2D persons in 3D worlds? And what about heavy graphs, 0-thin triangles and hyperbolic groups? In what sense are they fictional or impossible?

In mathematics, "physically possible" is not a criterion to distinguish between *idealized* and *fictional* thought experiments, and moreover it is not a purpose. For example, Glas (2009, p. 60), discussing the ingenious thought experiment of Archimedes, says: "Archimedes 'thought of' line segments and areas as possessing weight, but this is not essentially different from normal geometrical practice, in which lines are 'thought of' as possessing length but not breadth, etc." Glas used this analogy to rightly argue that the thought experiment is as rigorous as any other geometric construction: "Archimedes also based his statics proper (the theory of the lever) on postulates – not empirical laws – in exactly the same way as Euclid had axiomatized plane geometry." However, "de-idealization," as Sorensen (1992, p 100) calls it, is not a one-to-one but a one-to-many relation: e.g., an "ideal" mathematical line may return its breadth, but it can also have weight, color or conductivity. So this phenomenon is rather a "re-idealization." While in Euclid's geometry, a geometric line is a simplification by ignoring properties, namely breadth, Archimedes is on the opposite end: *adding* properties. Besides, not all lines in the same diagram consistently hold weight (the parabola in Archimedes' example itself does not). This kind of selective and creative idealization is non-standard, applying neither to geometry nor to statics. Also, a breadthless line is (and already was before Euclid) a common geometric concept, in contrast to a heavy line.

So, there is a difference, but it is irrelevant to its physical actuality. To remind the reader, usually the goal of thought experiments is to challenge standard thinking and to introduce a conceptual change. Fiction is a good way to do this by pushing idealization beyond reality, and here is how it works. For example, in Thompson's thought experiment, "Imagine you woke up bodily, attached to a famous unconscious violinist. . ." can be modified from "Imagine you got pregnant, and it has been statistically calculated that your future child will be a famous violinist." The latter is not as effective as the former, because it does not provide a necessary conceptual change (from "the right to life" to "the right to use another person's body"), which makes an epistemic difference. Similarly, the conceptual change in Archimedes' experiment is thanks to the addition of physical properties. As mathematical concepts are already idealized in a particular way, a conceptual change requires a different type of idealization. In mathematics, "reality" is a given realm or practice with its own type of idealization. Fiction is "impossible" in respect to the current mathematical practice of using (visual) representations.

3.5 *Thick* Concepts

The fictional or "re-idealized" concepts in mathematical thought experiments can also be characterized as "thick," as suggested by Brown (2019) in the sense that they are concepts simultaneously combining both physical and mathematical aspects. As he explains further on, "[...] thick mathematical concepts are drawn (typically) from the physical realm and the pure mathematical realm" (Ibid, p. 12). To give us a flavor of what he means, Brown quotes Michael Spivak: "[Y]ou can feel the second derivative when you sit in an accelerating car" (Ibid). The concept of the second derivative itself is a *thick* concept, thanks to intuitions coming from physics and mathematics. Brown argues that thick concepts may encourage us to make progress with difficult problems, e.g. the continuum hypothesis. What is relevant is that thick concepts are drawn from two realms, and thought experiments are a way to combine them fruitfully. A *derivative* is a thick concept too in the sense that it can be defined analytically (as a limit of *df/dt*) and also geometrically (as the slope of a line tangent to a function's graph). A *hyperbolic group* considered above is another example of a thick concept because it combines geometry and algebra). A combinatorial *group*, if enriched by geometric intuitions such as 0-thin triangles, gives rise to a *hyperbolic group*.

Can we find thick concepts elsewhere in mathematics, say within number theory? We can definitely make them richer by discovering new facts about them. For example, 4 is the area of a square with the side 2 and also $4 = 2 + 2 = 2 \times 2 = 1 + 3$ (*sum of primes*) $= 2! + 2!$ Analogously, $2023 = 7 \times 17 \times 17 = 7^7$ *mod* $7!$ expresses two different facts about the number 2023 in terms of different arithmetical practices, where the latter is mostly used in computing. These are different operational practices, yet remain within pure mathematics: addition, multiplication, factorization, etc. Combining them enriches the concept of the number. In this respect, fiction in mathematical thought experiments can be understood as a different type of idealization borrowed from a different realm, whether it be physics, geometry or another related practice. Euler's proof of the Basel problem will demonstrate this next, just after some considerations about the reliability of imagination.

3.6 A Veridical Fictional Imagination

Use of imagination obviously raises doubts about the reliability of (mathematical) thought experiments. For example, Norton (2018, p. 454) points out that "We give the thought experimenter extensive latitude in introducing idealizations" and considers it as one of the sources of error. Can we really make discoveries that rely heavily on visual imagination which is fictional? Are we not simply replacing proper experiments with "fantasies of the imagination"?

In scientific experiments, the instruments and experimental settings are selected by a methodology which must take into account possible biases and risks. In thought experiments, the selection of relevant properties is an art, and at the same time a task-oriented veridical thinking. Structured and controlled imaginary transformations are

widely recognized as an epistemically valuable component of thought experiments (Kujundzic 1998). This places mathematical thought experiments among other mathematical procedures, all liable to error. Fictional imagination for veridical mathematical reasoning is a very special feature of mathematical thought experiments. Importantly, in the mathematical context, the fictional element is one of the ways that allows us to stretch prescribed procedures, and find insights in another practice or past experience. This feature of thought experiments has been discussed in Balcerak Jackson (2016), Brown (2017), Buzzoni (2011) and Starikova and Giaquinto (2018) among others. The thing is that there is an important distinction between veridical imagining and fantastical imagining. Veridical imagining is aimed at finding out the true answer to certain questions, and is ruled by the accumulated effects of past perceptual experience. This does not make veridical imagining infallible; the adjective "veridical" is intended to describe the aim, not the result, of imagining. Fantastical imagining is not constrained in that way however, as it is not aimed at answering a question, as it serves psychological ends such as wish-fulfillment, horror thrill, sensory fascination and so on. Veridical imagining is widespread and useful. Wanting a desk for a particular room for example, you visit a furniture showroom; there you see an attractive desk, somewhat bigger than what you had in mind. Would it fit reasonably well into the stated room with its other furniture? In such a situation, it is reasonable to visualize the room to reach a judgment – this is veridical imagining. We do in fact rely on sensory imagination to answer such questions, and we get correct answers frequently enough for this practice to persist.

4 Experiment

So far we have discussed how *mathematics* can be experimental, and how the *thought* can be rich in respect to the narratives and cognitive instruments used in mathematical thought experiments. Finally, *experiment* as a scientific method is essentially understood as an *empirical validation or a refutation of a hypothesis*. As previously mentioned, according to the received view, empirical validation is not considered to be a method in pure mathematics. On this basis, mathematical thought experiments, as valued by Manin and Szabó, were reduced to illustrated proofs or "decorated arguments" in Norton's words.[16] However, there is another meaning of *experiment*, which comes from the Latin *experi*, which is "trying" a new method as opposed to those accepted, established and indisputable. This sort of experimenting

[16] As a reminder, Norton (1996, 2004) holds a 'strong argument view' about thought experiments in general. Buzzoni (2008) states that since no empirical evidence can disprove a mathematical fact, mathematical thought experiments, at least as far as validation is concerned, do not exist. Similarly Van Bendegem (2003) claims that mathematical thought experiments are "anything but proofs," meaning that since mathematics aims at proofs, any reasoning towards an insight of a possible proof can be seen as a mathematical thought experiment.

does not have to result in validating a hypothesis. It may however result in discovery, a better understanding of the problem and making the hypothesis plausible.[17]

Trying out something *new* (in respect to the given domain) is a key feature of mathematical thought experiments. For example, Glas (2009) characterizes thought experiments as "relying on information which is... *not assimilated by the traditional* mode of dealing with the world" (p. 57). Starikova and Giaquinto (2018) identify the type of mathematical thought experiment that is closest to the classical thought experiments, and characterized it by two features:

(a) the *experimental* thinking goes beyond application of mathematically prescribed rules, and (b) by using sensory imagination as a way of eliciting the benefits of past perceptual experience.

While the first condition can be taken as necessary, the second is to pinpoint reasoning with imaginary physical objects and properties. Therefore, this can be taken as a good marker for thought experiments. Heinzmann (2022), talking about mathematical thought experiments as "deviation tools," relates them to a sort of sensory intuition: "These deviations involve, generally, an appeal to vision, which is supposed to provide a form of intuition." This implies that the second condition is not purely subsidiary, and sensory experience does play an epistemic role in mathematical thought experiments. Therefore the concept of *experiment* in mathematics can refer to the observation of sensory data, as in counting a cube's edges. It can also refer to the observation of variations in a diagram when we search for invariants, as in the mutilated chessboard problem.

4.1 Mathematical Thought Experiments and Mathematical Experiments

Both mathematical thought experiments and mathematical experiments can help in searching numerical patterns and structures. The former however give us freedom to extend the prescribed numeric rules and use resources from another practice, as the following example will demonstrate.

4.1.1 Basel Problem and Infinite Polynomials

The famous Basel problem asks us to find the value of the following infinite sum:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots \frac{1}{n^2} + \ldots$$

[17] Brown (2017) gives a list of such roles for thought experiments; for the various roles assigned to *mathematical* thought experiments, see also Starikova and Giaquinto (2018), Van Bendegem (this volume), and Buzzoni (2011).

Drawing from history, see Ayoub R (1974), in 1731 Euler calculated the sum of the first terms, and recognized that the sum should be $\pi^2/6$. Earlier approximations already supported his guesswork. Motivated by this evidence, in 1734 Euler published a proof conceived as the approximation of the given sum by *sin (x)/x* and treating it as a polynomial. Applying a factorization by roots, he then compared the coefficients of x^2 terms with the Taylor series of expansion. At that time, such factorization could not be applied to infinite polynomials, so Euler's bold experiment went beyond the legitimate practice. Is it a thought experiment?

Lacking a typical imaginary scenario, this can still be considered as a mathematical thought experiment. The inductive approximations of the sum suggest a hypothetical number, but are insufficient for a general proof. In order to test the hypothesis in its full generality, Euler supposes infinite polynomials and then observes how the problem "behaves'' under the new conditions. For Heinzmann, "If, in its inferences, the demonstration uses deviant tools, it is called a mathematical thought experiment, which later may become an experiment with respect to the standard formalism of the time" (Ibid). The calculative approximations therefore are mathematical experiments, and Euler's approach is a mathematical *thought* experiment due to the non-standard move of taking infinite polynomials for granted.

A similar mathematical thought experiment is described in Glas (2009). Bombelli in 1572 found a way to circumvent the appearance of $\sqrt{-1}$ in solving cubic equations. He applied algebraic procedures, which cancel "impossible numbers" on the way to solutions. According to Glas, the thought experiment is used to hypothesize this outcome and by ignoring the obstacle, to pursue the calculation. Glas describes it "as *post hoc* reasoning from the required result to the argumentative steps that were necessary to attain it" (Ibid, p. 62). In both Bombelli's and Euler's approaches, as in the examples with visualizations, the hypothesis is made plausible first on the level of observation (i.e., of a formula or an image), and next on the level of formalization.

4.2 Variety of Experience

It is important to emphasize that the whole wealth of our experience and knowledge becomes available in mathematical thought experiments. A recent peculiar proof of the Basel problem, published by the Swedish mathematician Johan Wästlund (2010) will help to demonstrate this point.[18] In contrast to Euler's proof, this one is based on a law of physics and involves visualizations. It starts with an important observation: "Inverse squares occur in physics... The apparent brightness of a light is proportional to the inverse square of its distance" (Ibid, p. 2).

[18] The proof is also nicely explained under the link https://mattbaker.blog/2018/03/14/circles-the-basel-problem-and-the-apparent-brightness-of-stars/

4.2.1 Basel Problem and the Lights Proof

The proof involves the following thought experiment: imagine standing at the origin of the number line where each integer is a light. It is like standing on a road between two villages. There are streetlight to the left and to the right of you. At a unit time, your eye is receiving a quantity of light from all the streetlights. According to the law of light distribution, the sum of brightness from all the streetlights amounts to 1 from the first star, 1/4 from the second, 1/9 from the third, 1/16 from the fourth and so on. How could this help us to test the hypothesis that the total brightness is $\pi^2/6$? It may seem that we have not moved on from the problem statement, but in fact, the physical representation opens up a new path. Roughly speaking, the law implies that the closer streetlights give out the most light – but the distant ones give out less light and can emit it from another position at approximately the same distance. How could we rearrange them to preserve the same total quantity of light? Now imagine that the same lights make a circle around a huge lake. The light from the straight line is approximately the same as that from the big circle. The lights from the other side of the lake emit roughly the same amount of light as the lights next to the villages.[19] Putting it differently, a linear distribution of lights can be seen as an approximation of a very large arc, as in Fig. 8 below. Importantly, the circumference "constellation" is where the number π naturally emerges (Fig. 12).

Also note that if we split a light into two and place them in the corners of a right triangle, then the initial light brightness can be expressed as $1/c^2 = 1/b^2 + 1/b^2$, which is the reverse of the Pythagoras theorem. A sequence of Euclidean constructions with circles and triangles can eventually lead to the desired result (Fig. 13).

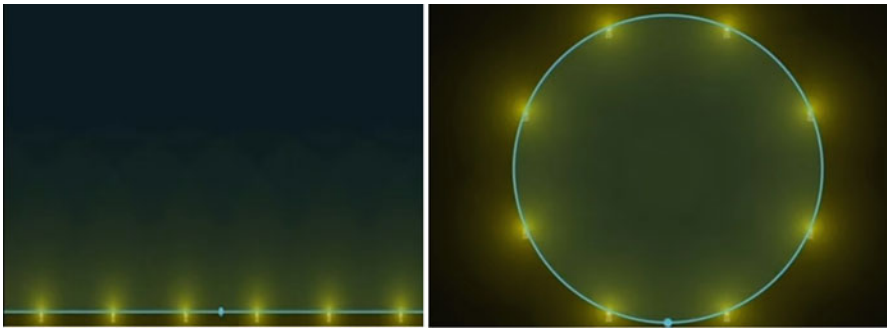


Fig. 12 The law of light distribution suggests the idea that the total amount of light received by an observer at point zero from lights arranged in a straight line and the total amount of light received from lights arranged in a circle are approximately the same

[19] Wästlund mentioned in his blog that he heard a little girl asking what would happen if the universe was finite, and the stars were all in a circle. (http://wastlund.blogspot.com/2018/03/infinite-lake-surrounded-by-lighthouses.html)

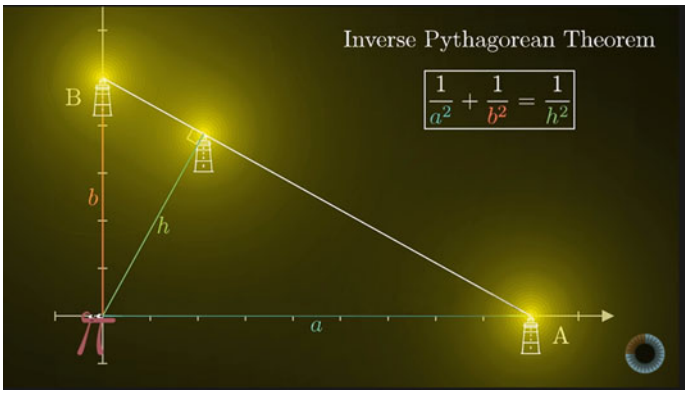


Fig. 13 This screenshot from Sanderson (2018) YouTube video *Why is pi here? And why is it squared? A geometric answer to the Basel problem* demonstrates the use of the inverse Pythagorean theorem

One can object to the conclusion that this constitutes a thought experiment. In the necklace example (in 3.1.1), we do not use knowledge of a physical fact as a premise, or lemma in a mathematical deduction about a physical situation. Instead, we simply suppose that the graph is like a necklace, with beads of equal weight spread along a string at equidistant intervals. Then we imagine what would happen to the necklace if we held on to one bead and let the rest of it drop. Finally, we attend to the resulting imagery for a clue to help us find the diameter. Therefore, we conduct an experiment in visual imagination. We get the facts merely from imagery, that we eventually use to get a mathematical answer.

In contrast it seems that in the light proof we rely on physics, but do we actually? No. In this proof, the number series of inverted squares is treated in a similar way as it is in Euler's proof. As in the necklace example, the physical representation – i.e., a heavy necklace – helps to rearrange the studied structure of the cyclic graph. In the light example too, the physical representation helps us to notice the useful rearrangement of the number line into a circle, but the proof is technically geometric and consists of geometric calculations. As Wästlund describes his proof, "What is new in our presentation is the physical interpretation which allows us to replace all trigonometric identities by classical geometry. . ." (2010, p. 10).

All the physics of light emission and reflection are reduced to Euclidean geometry.

The thought experiment consists of picking up the invariant (the inverted square), and observing what happens if it is placed in another context. A reversed square $1/n^2$ is used as an icon or a *sensed pattern*: it is a visual pattern of symbols, which is identified in the context of number theory and can be recognized outside it – i.e., in the context of polynomials as in Euler's proof, or in physics as in Wästlund's proof. In the latter, physics connects the series to a circle, and this explains the presence of π in the famous formula.

Formula-based mathematical thought experiments (Euler's proof of the Basel problem and Bombelli's approach to complex numbers) are mathematical thought experiments, because they go beyond the given calculation algorithms. They do not extend the domain of mathematics, but extend the given practice (e.g., number theory to polynomials, or real numbers calculus to complex numbers). Mathematically, the light argument is outstanding not so much due to its link with physics, but more importantly, with geometry and explaining the presence of π. In both types of mathematical thought experiment, formula-based and image-based, we use either *sensed* patterns or *thick* concepts (e.g., hyperbolic groups, or heavy graphs). The distinctive feature of mathematical thought experiments in comparison to *calculative* experiments is the scope and range of knowledge, experience and creativity that becomes available to the mathematician. The discussed examples demonstrate that mathematical thought experiments play a crucial role in discovery. They are used primarily when other methods do not help. Often, the most valuable mathematical thought experiments are exactly those that produce an *insight* such as zooming out to get the idea of large scale geometry. Zooming out per se is not so much of an extraordinary idea, but as part of a thought experiment it can lead to an extraordinary discovery, i.e., the large-scale geometry of groups. This makes thought experiments strikingly different from experimental calculations.

5 Conclusions

It is common practice for mathematicians to reason with diagrams, but it is rarely mentioned that they also reason with physical objects that for them have a somewhat abstract character. This is possible thanks to mathematical thought experiments that often go unnoticed.

Mathematical thought experiments can be based on diagrams, models, metaphors, fictions and formulas, and can combine counterfactual, deductive and visual/sensory reasonings, but they also open up even more possibilities:

1. Enriching mathematical concepts by associating them with physical properties, as well as with the properties of objects from different practices (namely *thick* concepts).
2. Extending the current algorithmic practice by fictional, but at the same time veridical, imagination.
3. Providing epistemic resources from the different mathematical practices, and the entire personal experience of the experimenter.

They are *mathematical* because they provide a deeper insight into mathematical problems. They are *thought* experiments, not for planning arguments and algorithms in mind, but for extensively exploiting the richness of our cognition and experience. They are *experiments* not necessarily for testing a hypothesis, but also for making a hypothesis plausible, revealing the new invariants and developing concepts.

Acknowledgments This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Financing Code 88887.802894/2023–00

References

Abbott E (2009) Flatland. Broadview Press

Ageron P (2013) Le partage des dix-sept chameaux et autres exploits arithmétiques attribués à l'imam 'Alî: mouvance et circulation de récits de la tradition musulmane chiite. Revue d'histoire des mathématiques 19:1–41

Arcangeli M (2010) Imagination in thought experimentation: Sketching a cognitive approach to thought experiments. In Model-Based Reasoning in Science and Technology: Abduction, Logic, and Computational Discovery. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 571–587

Ayoub R (1974) Euler and the zeta function. Am Math Mon 81:1067–1086

Balcerak-Jackson M (2016) On the epistemic value of imagining, supposing, and conceiving. In: Kind A, Kung P (eds) Knowledge through imagination. Oxford University Press, Oxford, pp 41–60

Berge C (1958) Theorie des Graphes et ses Applications. Dunod, Paris 175–176

Bilalić M, Graf M, Vaci N, Danek AH (2019) When the solution is on the doorstep: better solving performance, but diminished Aha! experience for chess experts on the mutilated checkerboard problem. Cogn Sci 43(8)

Black M (1946) Critical thinking: An introduction to logic and scientific method. Prentice Hall 157:433

Brown JR (1991) The laboratory of the mind. Thought experiments in the natural sciences. Routledge, London

Brown JR (1997) Proofs and pictures. Br J Philos Sci 48(2):161–180. Available via http://www.jstor.org/stable/687743

Brown JR (2007) Thought experiments in science, philosophy, and mathematics. Croat J Philos Sci 7(1):3–27. Available via https://doi.org/10.5840/croatjphil20077128

Brown JR (2017) Proofs and guarantees. Math Intell 39(4):47–50

Brown JR (2019) Ethics and the continuum hypothesis. In: Fillion N et al. (eds) Springer science and business media, LLC, part of Springer Nature, Algorithms and complexity in mathematics, epistemology, and science, fields institute communications 82. Available via https://doi.org/10.1007/978-1-4939-9051-1

Brown JR, Fehige Y (2019) Thought experiments. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. Available via https://plato.stanford.edu/archives/win2019/entries/thought-experiment/

Buzzoni M (2008) Thought experiment in the natural sciences. An operational and reflective-transcendental conception. Königshausen+Neumann, Würzburg, Germany

Buzzoni M (2011) On mathematical thought experiments. Epistemologia XXXIV:61–88

Carter J (2018) Graph-algebras – faithful representations and mediating objects in mathematics, Endeavour, 42, 2–3, 180–188., ISSN 0160–9327. Available via https://doi.org/10.1016/j.endeavour.2018.07.006

Giaquinto M (2020) The epistemology of visual thinking in mathematics. In: Zalta E (ed) The Stanford encyclopedia of philosophy (Spring edition). Available via https://plato.stanford.edu/archives/spr2020/entries/epistemology-visual-thinking/

Glas E (2009) On the role of thought experiments in mathematical discovery. In: Meheus J, Nickles T (eds) Models of discovery and creativity. Springer, Dordrecht

Gromov M (1993) Asymptotic invariants of infinite groups. In: Niblo and Roller (eds) Geometric group theory, Vol 2. London mathematical society lecture notes 182, Cambridge University Press

Heinzmann G (2022) Mathematical understanding by thought experiments. Axiomathes 1–16

Kaplan CA, Simon HA (1990) In search of insight. Cogn Psychol 22(3):374–419
Klein F (1888) Lectures on the icosahedron (transl. Morrice GG)
Bensimon I (2009) Minkowski's convex body theorem. Project for MA341: Appreciation of Number Theory, Boston University
Kuhn T (1981 [1963]) A function for thought experiments. In: The structure of scientific revolutions. Oxford: Oxford University Press
Kujundzic N (1998) Role of variation in thought experiments. Int Stud Philos Sci 12(3):239–243
Lakatos I, Worrall J, Zahar E (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge
Mach E, Hiebert EN (1976) Knowledge and error. In: Knowledge and error. Vienna circle collection, Vol 3. Springer, Dordrecht. Available via https://doi.org/10.1007/978-94-010-1428-1_7
Manin Y (2002) Georg Cantor and his heritage. arXiv:math.AG/0209244 v1.
McAllister JW (1996) The evidential significance of thought experiment in science. Stud Hist Philos Sci A 27(2):233–250
McGinnis J (2018) Experimental thoughts on thought experiments in medieval Islam. In: Stuart MT, Fehige Y, Brown JR (eds) The Routledge companion to thought experiments. Routledge, pp 77–91
Murphy A (2020) Toward a pluralist account of the imagination in science. Philos Sci, 87(5), 957–967. Available via https://doi.org/10.1086/710620.
Norton JD (1996) Are thought experiments just what you thought? Can J Philos 26(3):333–366
Norton J (2004) On thought experiments: is there more to the argument? Philos Sci *71*(5): 1139–1151
Norton J (2018) In: Stuart MT, Brown JR, Fehige Y (eds) The worst thought experiment, the Routledge companion to thought experiments. Routledge, London, pp 454–468
Polanyi M (1966) The tacit dimension. London: Routledge & Kegan Paul
Sagan C (2009) Cosmos – fourth Dimension. Available via https://www.youtube.com/watch?v=UnURElCzGc0&t=18s
Sanderson G (2018) YouTube video Why is pi here? And why is it squared? A geometric answer to the Basel problem. Available via https://www.youtube.com/watch?v=d-o3eB9sfls
Shepard RN (2008) The step to rationality: the efficacy of thought experiments in science, ethics, and free will. Cogn Sci 32(1):3–35
Sorensen RA (1992) Thought experiments and the epistemology of laws. Can J Philos *22*(1):15–44
Starikova I, Giaquinto M (2018) Thoughts experiments in mathematics. In: Stuart MT, Fehige Y, Brown JR (eds) The routledge companion to thought experiments. Routledge, pp 257–278
Starikova I, Van Bendegem JP (2021) Revisiting the mutilated chessboard or the many roles of a picture. Log Anal 255:289–312
Van Bendegem JP (2003) Thought experiments in mathematics: anything but proof. Philosophica 72:9–33
Wästlund J (2010) Summing inverse squares by Euclidean geometry. Available online via http://www.math.chalmers.se/*wastlund/Cosmic.pdf

The Intimate Interplay Between Experimentation and Deduction: Some Classroom Implications

Michael de Villiers and Hans Niels Jahnke

Contents

1 Introduction 2552
2 Philosophical Issues 2555
3 Historical Examples 2557
4 Checking, Guessing, Finding a Proof 2560
5 Refutation in the Classroom 2564
6 Using Scientific Principles in Mathematics 2570
7 Modelling the Real World 2576
8 Conclusion 2582
References 2584

Abstract

We understand experimental mathematics as the systematic experimental investigation of concrete examples of a mathematical structure in the search for conjectures about its properties. Experiments might be done by pencil-and-paper work, building physical models, and, of course, by using available computer programs for doing time-consuming calculations, geometric constructions, and other types of visualizations. The guiding intention of the chapter is the study of the interplay of experimentation and deduction (proof) in the context of the teaching of mathematics at lower and upper secondary schools as well as the education of teachers of mathematics.

The didactical intentions of the chapter make it necessary to take a special philosophical perspective, which might be characterized as a combination of Lakatosian and non-Lakatosian viewpoints (Sect. 2).

M. de Villiers (✉)
Stellenbosch University, Stellenbosch, South Africa
e-mail: profmd1@mweb.co.za

H. N. Jahnke
Faculty of Mathematics, University of Duisberg-Essen, Essen, FR, Germany
e-mail: njahnke@uni-due.de

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_40

In Sect. 3 we shortly discuss three historical examples (Euler, Fourier series, Chinese Remainder Theorem) to give a more realistic image of the interplay between experimentation and deduction in contrast to the somewhat artificial character which didactical examples necessarily involve.

In the following Sects. 4, 5, 6, 7, and 8, we discuss a few classroom examples of mathematical experimentation. Section 4 is on checking, guessing, and finding a proof, and in Sect. 5 we put a particular emphasis on finding and handling counterexamples. These two sections can be characterized as the classical domain of Lakatosian ideas transferred to the classroom. Section 6 explores the potential of viewing elementary Euclidean geometry from the point of view of statics. This approach has had a long tradition in mathematics since ancient times and represents a transition to discussing experimentation with hypotheses in Sect. 7 on modelling. The latter is definitely a non-Lakatosian viewpoint.

Lastly, in Sect. 8 it is concluded that the interplay between experimentation and deduction needs to be incorporated in primary education and extended throughout for the long-term development of sound mathematical thinking and understanding of students.

Keywords

Experimentation · Deduction · Proof · Checking · Guessing · Refutation · Modelling · Lakatosian and non-Lakatosian points of view

1 Introduction

In a recent textbook for early graduate and advanced undergraduate students, Eilers and Johansen (2017, p. 1) define experimental mathematics to be: "*the systematic investigation of concrete examples of a mathematical structure in the search for conjectures about its properties*." For the aims of the present chapter, this seems to be a sufficiently simple definition which on a concrete level covers the following important functions of experimentation giving an idea of the large scale of possible activities (see De Villiers (2010, p. 206) and Jonathan Borwein (2012, pp. 73–74):

- Discovering new relationships and conjecturing (looking for inductive patterns, generalizations, etc.)
- Visualizing math principles (viewing graphical or other visual representations of analytically derived results)
- Verification (obtaining certainty about the truth or validity of a statement or conjecture)
- Computing replacing lengthy hand derivations (use of computer algebra, data handling, and/or graphical representations)
- Global and heuristic refutation (disproving a false statement by generating a counterexample, as well as reformulating, refining, or polishing a true statement by means of local counterexamples)

- Understanding (gaining insight/intuition, and grasping the meaning of a proposition, concept, or definition or assisting in the discovery of a proof)

The book from which the above definition is taken associates "experimentation" exclusively with the use of computers for doing time-consuming numerical calculations. In the programmatic editorial of the Journal *Experimental Mathematics*, however, the editors write (Epstein et al. 1992, p. 1; see also De Villiers 2021):

> *The word 'experimental' is conceived broadly: many mathematical experiments these days are carried out on computers, but others are still the result of pencil-and-paper work, and there are other experimental techniques, like building physical models.*

The present chapter has the teaching of mathematics at lower and upper secondary schools as well as the education of teachers of mathematics in mind and, consequently, follows this broader outlook. It includes quite a number of examples from geometry comprising the use of dynamic geometry software and above all the modelling of real-world phenomena.

The inevitable companions of experimentation in mathematics are proof and deduction. Like with experimentation, we suppose a sufficiently broad concept and do not precisely distinguish between "proof" and "deduction." For the moment, we refer to the discussion document of the ICMI Study 19 which says that proof is based on the principle

> *to specify clearly the assumptions made and to provide an appropriate argument supported by valid reasoning so as to draw necessary conclusions*. (Hanna et al. 2012, p. 443)

For a more detailed discussion, the reader should consult the next section.

Proving and experimentation are different sides of the same coin, inseparably connected and, yet, to be clearly distinguished. A conjecture which might be suggested by a mathematical experiment has to be proved, and, conversely, there must be some sort of a theoretical idea behind any experiment as already the very first words "systematic investigation" of the above definition show. Indeed, there is no experiment without a question, neither in science nor in mathematics. The guiding idea of the present chapter can be expressed by saying that *there is no deduction without experimentation and no experimentation without deduction*.

Experimental mathematics in the educational context can be considered as a component of "Inquiry-based learning" (IBL) (for a survey, see Artigue and Blomhøj 2013). As Bruder and Prescott (2013, p. 820) sum up in an overview of empirical studies, IBL is not per se superior to other more traditional approaches to teaching. Rather, IBL proves to be superior only when the teacher finds a right way to guide the work of the students, and positive effects are also dependent on the type of knowledge taught.

Seen from a pedagogical point of view, the relation between experimentation and proof in mathematics is among others a problem of finding a balance. When a mathematics course at university proceeds in a purely axiomatic-deductive style

without sufficiently discussing examples and counterexamples and without giving opportunities for mathematical experimentation, the image of mathematics which might arise from such "anti-didactical inversion" (Freudenthal 1973) are grossly misleading. With a high probability, the teacher trained in this way will convey the respective bad practices also to his/her students. This means:

(1) Learners are not given the opportunity to explore and to experiment but are given the conjecture as fait accompli in the form "Prove that . . ." Instead of that, there should be enough room to explore, experiment, and make conjectures on patterns or invariances. Such situations should be open-ended and divergent.
(2) Learners are exclusively provided with well-formulated conditional statements "Prove that p implies q." Instead of that, learners should themselves have the opportunity to transform conjectures into conditional statements.
(3) The instruction "Prove that . . ." implies beforehand that the result is true and learners develop the habit to accept the truth of statements on the authority of the teacher and the textbook. Instead of that, they should be given the experience to explore the truth of a statement by themselves searching for examples/counterexamples or any other kind of argument to support it.

To arrive at a good balance between proof and experimentation is a difficult and never-ending task of education. It requires looking at one and the same topic of teaching from different perspectives, and this multiperspectivity much too often meets the limitations of overloaded curricula and available time, external examinations, class size, limited facilities such as basic experimental artifacts (such as compass & ruler, calculators, computers, etc.), the philosophical perspectives of teachers and textbook authors, appropriate past pedagogical experiences of teachers, etc.

At present, in the curricula of many countries mathematical topics are implemented which strengthen the role of experimentation without lowering the role of proof and deduction and thus point in the direction which is elaborated in the present chapter. We think here of:

(a) Pattern recognition and arguing with patterns (see the chapter on proof in the NCTM standards 2000 and, for example, Wittmann and Müller 1990)
(b) Problem solving with a vast amount of didactical literature (see f. e. Schoenfeld 1985)
(c) Modelling (For the state of the art in this field, see the ICMI study 14 (Blum et al. 2007))

In what follows, we concentrate on exploring the *interface of proof and experimentation* and on *studying by means of historical and educational examples how they interrelate*. Concerning the educational examples, most of them are based without any claim of originality on experiences of the authors as educators of mathematics teachers at lower and upper secondary levels and on teaching experiments at schools. Some of them might be found directly helpful. But since any

teaching example is bound to special situational conditions and in the light of the abovementioned empirical studies on IBL, such a direct transfer will be possible only in rare cases. We aim more at encouraging the reader to develop their own experiments. It is in support of this more indirect transfer that in the next section we attempt at clarifying our general views on the topic by some philosophical considerations.

2 Philosophical Issues

Somehow as a counterpoint to the formalistic trends promoted by the New Math movement, Imre Lakatos (1922–1974) published in 1963/1964 an essay on the philosophy of mathematics an extended version of which appeared in 1976 as a book under the title "Proofs and refutations." In it, he substantially and famously criticized the traditional neglect of the experimental, inductive, and heuristic processes that have historically accompanied mathematical developments:

> *Mathematics is presented as an ever-increasing set of eternal, immutable truths. Counter examples, refutations, criticism cannot possibly enter. An authoritarian air is secured for the subject by beginning with disguised monster-barring and proof-generated definitions and with the fully-fledged theorem, and by suppressing the primitive conjecture, the refutations, and the criticism of the proof. The deductivist style hides the struggle, hides the adventure. The whole story vanishes, the successive tentative formulations of the theorem in the course of the proof-procedure are doomed to oblivion while the end result is exalted into sacred infallibility.* Lakatos (1976, p. 142)

On the other hand, he rightly rejected the position that mathematical discovery "is a completely nonrational affair" (l.c., 143). In fact, it was exactly his analysis of the intimate interplay of experimentation and deduction which made his case-study on the history of Euler's polyhedron formula so philosophically innovative and attractive to many readers. To this contributed also the terminology Lakatos invented for describing this interplay: proof-generated definition, proof-generated theorem, global and local counterexamples, etc.

In this general sense, Lakatos' philosophy of mathematics serves as a background to the following analysis of examples of teaching. We use also part of Lakatos' terminology but do not relate his views on mathematics to other approaches to the philosophy of science in the 1960s (see Lakatos and Musgrave 1970).

However, we find it necessary to extend Lakatos' approach in another regard, and this is connected with our broad view on proof and our only vague distinction between proof and deduction at the beginning of this chapter. It was Árpád Szabó, himself a teacher and close friend of Lakatos, who in his seminal study on the genesis of the axiomatic method in ancient Greece (Szabó 1960) pointed at another sense of "experimentation" Lakatos had in mind. One can call this "experimenting with hypotheses." Szabó showed in his study how the axiomatic approach originated from pre-Euclidean and pre-Platonic dialectics that is the art of exchanging arguments and counterarguments in a dialogue debating a controversial subject. In any

debate, be it philosophical, political, or scientific, it is a frequent procedure to agree for a moment to a certain proposition in order to critically examine its consequences. Participants take the proposition as given, derive consequences, and accept or reject the proposition depending on whether the consequences are plausible (desired) or implausible (not desired). On a general level, this is the procedure of theory-building in science, and today we call this "modelling," and one can find many instances of this already in ancient Greek science (Jahnke 2010). In Sect. 7, we discuss some concrete examples providing a more detailed and sophisticated view on modelling.

Even though some philosophers might feel uneasy with such a postulational perspective of mathematics, the inclusion of modelling in our understanding of mathematical experimentation seems inescapable for pedagogical reasons. There is a variety of (teaching) situations which require the introduction of assumptions which later on might be revised (Bruder and Müller 1983; Jahnke and Ufer 2015, p. 6). Such assumptions could be:

(1) Evident statements in order to convince other persons of something. The evidence might be drawn from visual impressions or from a number of special cases or from shared knowledge in the classroom.
(2) A preliminary assumption needed for the completion of an argument. Later on, one might find a backing for this assumption.
(3) Stating an assumption in order to refute it. When one arrives at a logical contradiction, one has devised an indirect proof.
(4) Stating an assumption of whose truth one is convinced but which one cannot prove yet. This case is similar to (3), but this time the intention is to see whether the consequences are plausible. Classic case: Riemann's hypothesis.
(5) Thought experiments, especially in physics. Famous example: Galileo's thought experiment for refuting Aristotle's theory of free fall. Also similar to case (3).
(6) A statement from physics in the context of a mathematical theory. Famous examples: Archimedes' application of the law of lever (see below), Dirichlet's principle, etc.

This list could of course be extended and refined. In all cases, the assumptions made have the status of a hypothesis which means it is a statement whose evaluation as true or false is insecure. Nevertheless, they function as a condition of a deductive argument which can be qualified as proof. Thus, experimenting with hypotheses is an essential tool for furthering the awareness and consciousness of a learner for the conditions of mathematical theorems. Boero et al. (2010) rightly speak of the "conditionality" of mathematical statements. It would be a pedagogical sin not to use the potential inherent in this variety of deductive argumentations.

To sum up, we think that from the first beginnings of primary education, experimentation should be a central concern of mathematics teaching guided by the question "why." Inherent in this, there is also the question "what if when things were different," i.e., what if we start with a different hypothesis? There is a natural

curiosity even of small children to imagine a world different from the one we live in. We should further encourage this curiosity also in mathematics teaching.

3 Historical Examples

As a prelude to our teaching examples, we present in this section three historical cases where the interaction of deduction and experimentation shows up in a remarkable and instructive way. There is an important and fundamental difference, however, to the situation in teaching: time. In general, we cannot expect that the interaction of deduction and experimentation takes place in short periods of time. There are such cases, but they are not the rule. The Lakatos example of the polyhedron formula, for instance, which he compressed in a fictitious dialogue, covered in real history a time span of about 150 years.

Our examples refer to.

(1) The mathematical work of Leonhard Euler
(2) The history of Fourier series
(3) The Chinese remainder theorem

(1) One of the most prolific mathematicians of all times, Leonhard Euler (1707–1783) has been characterized by Calinger (2016, p. 93) as an "experimental" mathematician who carried out numerous calculations in the process of his mathematical research (see also Ferraro's contribution in this handbook). In many of his research papers, he was frequently not above writing down numerical examples for the general formulae he had derived. As a specially gifted person, he used innovative transformations and clever substitutions to develop many new methods to solve problems from both mathematics and physics, which in some cases took up more than a century for other scientists and mathematicians to properly grasp or to fully justify. Sometimes Euler would set aside a problem if he was making no progress, allowing for what has been called the *incubation* period by Hadamard (1945), until eventually *illumination* would allow him to return and finally solve it. Several of Euler's discoveries were left to future generations to establish rigorous foundations. Perhaps one of the most famous of these is Euler's zeta function, which was extended further by Bernard Riemann (1826–1866), and is the basis of one of the greatest, still outstanding unresolved problems of mathematics, namely, the Riemann's Hypothesis (Derbyshire 2003).
(2) The second example is that of the historical development of Fourier series. Jean-Baptiste Joseph Fourier (1768–1830) (on his biography, see Grattan-Guinness and Ravetz 1970–1990) derived a Partial Differential Equation (PDE) for describing the diffusion of heat in a metal plate and broadly discussed its solution. He arrived at most of his results in a paper which he submitted to the French Academy of Science in 1807 and which was rejected on account of a negative evaluation by

J. L. Lagrange. Only in modern times has the paper been edited by Grattan-Guinness and Ravetz (1972). Nevertheless, Fourier won a prize of the French Academy in 1811 with a paper on heat diffusion, though also this time the jury made some objections in regard to "rigor and generality" probably again under the influence of Lagrange. Finally, Fourier published a comprehensive account of his results in a book in 1822, nine years after Lagrange's death.

As with most PDEs, the solution of the heat equation is dependent on an "arbitrary" function and Fourier established the remarkable result that any "arbitrary" function could be represented by a trigonometric series of sine waves. However, as pointed out by Davis and Hersh (1983, pp. 262–263), Fourier used mathematically problematic reasoning which also at his times would not have been accepted by many mathematicians. Fourier's argument implied two inconsistent assumptions and employed division by a divergent infinite product. Nevertheless, under the guidance of his physical insight, experimental background, and intuition he arrived at a correct result.

Langer (1947, p. 33) has suggested that it was probably Fourier's disregard of rigor and his calculations of many examples (l.c., p. 45) that enabled him to make the huge conceptual leap necessary to arrive at his remarkable conclusion.

But of course it was not simply by intuition and calculations of examples that Fourier found his results. There was a long and impressive prehistory in the eighteenth century comprising substantial theorizing as well as one of the most famous controversies in the history of mathematics. This concerned the PDE for the vibrating string derived by d'Alembert in 1747 (for details and access to the vast literature, see Luetzen 1983; Jahnke 2003). Again this PDE involved "arbitrary" functions, and d'Alembert argued that these functions should be analytic (that is developable in a power series), even, and periodic.

A year later, Euler protested against the restriction to analytic functions saying that for physical reasons one has to include any function which can be drawn by "free hand" and hence interpreted d'Alembert's solution as a rule to construct the solution geometrically starting from a curve drawn by free hand. Euler considered also as a special case functions defined by series of sine waves. For good mathematical reasons, d'Alembert did not accept Euler's idea of curves drawn by free hand since such curves may not have a first or second derivative, a necessary condition for applying the PDE. Then, in 1753 Daniel Bernoulli, son of Johann Bernoulli, entered the scene with the argument that any sound physically consists of a superposition of sine (or cosine) waves, and therefore Euler's special case can be generalized to comprising any function whatsoever. The controversy was never solved in the eighteenth century and received a new and decisive push by Fourier's work.

A first rigorous proof of Fourier's result was given by J. P. G. Dirichlet (1805–1859) who during his stay in Paris 1822–1826 was in close contact with Fourier. In 1829, he proved that any piecewise continuous and piecewise monotonous function can be represented by a Fourier series (Luetzen 2003, 158). Of course, this class of functions was far from arbitrary functions but for the first time included functions whose graph showed finitely many jumps.

It took about another century of effort for mathematicians and two generalizations of the notion of integral by Riemann (1826–1866) and Lebesgue (1875–1941) (see Hawkins 1980; Hochkirchen 2003) to substantially generalize the class of admissible functions and to establish firm foundations for Fourier series analysis. However, few physicists and engineers ever felt the need for a justification of Fourier analysis as its functionality in modelling experimental evidence was what counted for them.

(3) A third example shows that formal generalization and incorporation of intuitive ideas into more and more advanced rigorous structures may eventually obscure or completely hide the primitive, computational origins from which these ideas arose. Take, for example, the Chinese Remainder theorem, which states that if one knows the remainders of the Euclidean division of an integer n by several integers, then one can determine uniquely the remainder of the division of n by the product of these integers, under the condition that the divisors are pairwise coprime. The Chinese Remainder Theorem is best explained by the following party trick. A "magician" asks a participant to choose an arbitrary number less than say 60. The participant is then asked to divide the number, respectively, by 3, 4, and 5, and mention the remainder in each case. With only the knowledge of these three remainders, call them, respectively a, b, and c, the magician is then able to determine the number chosen by the participant. How is the magician able to do that?

The solution is simply that the magician has to divide $40a + 45b + 36c$[1] by 60, and the remainder of that division will be the original number selected by the participant. Suppose, for example, the number 43 was selected. The respective remainders would then be 1, 3, and 3. Therefore, $40 \times 1 + 45 \times 3 + 36 \times 3 = 283$ has to be divided by 60. Since $60 \times 4 = 240$, it follows that the remainder is equal to 43, which is the original number selected.

The Chinese Remainder theorem today has important applications, among others, in Fast Fourier Transforms and RSA secure data transfers, and in the proof of mathematical results like Gödel's incompleteness theorems.

The earliest known formulation of the theorem is that of Sun Tzu in the third century, but no proof or general algorithm was provided. Later in 1202, Leonardo of Pisa (Fibonacci) revisited the result in a slightly more general way, but also describing it in a computational algorithmic way with a specific set of integers. Like Sun Tzu, Fibonacci ignored the possibility of nonuniqueness and provides no proof of why it works. In 1734, Euler was probably the first to formulate a complete algebraic generalization and prove the generalization. However, by comparison, Shockley's 1967 and Prather's 1976 version of the Chinese Remainder theorem, based on

[1] To fully understand the choice of the arbitrary seeming numbers 40, 45, and 36 requires some modular arithmetic and of congruence relations between integers. More information can be found in abstract algebra and/or number theoretic texts.

Gauss' work on congruences and the abstract algebraic concepts of groups and fields, are hardly recognizable (to the novice) in relation to its original form.

As pointed out by Davis and Hersh (1981, p. 194), the modern version of the Chinese Remainder theorem focuses on structure at the expense of the algorithm and ignores the historically motivating question of how, given the remainders, the original number can be determined.

From an educational point of view, this illustrates how it might sometimes be necessary for the benefit of the student in order to develop conceptual understanding to retrace, at least in part, the gradual evolution of the mathematical ideas behind the Chinese Remainder theorem from its primitive, often experimental origins to its eventual abstract formulation.

In the following Sects. 4, 5, 6, 7 and 8, we discuss a few classroom examples of mathematical experimentation. Section 4 is on checking, guessing, finding a proof, and in Sect. 5 we put a particular emphasis on finding and handling counterexamples. These two sections can be characterized as the classical domain of Lakatosian ideas transferred to the classroom. Section 6 explores the potential of viewing elementary Euclidean geometry from the point of view of statics. This approach has had a long tradition in mathematics since ancient times and represents a transition to discussing experimentation with hypotheses in Sect. 7 on modelling. This is definitely a non-Lakatosian viewpoint.

4 Checking, Guessing, Finding a Proof

As should be done with students on all levels, we begin with the simple. Guessing and checking is probably one of the most basic, elementary methods of experimentation in mathematics.

For many years, one of the authors would ask the following simple question (among others) at the start of a course on Differential Equations for 3rd and 4th year university students: Solve for x if $x^2 = 2^x$. Without fail, the students would eagerly start tackling the problem by taking logarithms both sides, but then getting stuck and not knowing what to do next. Attacking the problem by algebraic manipulations is probably a natural approach given their past learning experience. It is, however, restrictive when they do not see or consider alternative approaches.

Only after letting them stew for a little while, it would be suggested that they use "trial-and-error," the most basic numerical method, upon which they then quickly find the two positive roots of $x = 2$ or $x = 4$. Since $0^2 = 0$ and $2^0 = 1$ whereas $-1^2 = 1$ and $2^{-1} = ½$, the curves of x^2 and 2^x cut each other between 0 and -1 telling us that there is a third root between -1 and 0. This third root then can be determined by, for example, computer graphing or other numerical methods. Students at this level should know that for sufficiently large x exponential functions surpass any polynomial function, and thus, they can be fairly sure that, in fact, the above equation has only these three roots. These results are reached by mental calculations within a few minutes. Students should again and again be encouraged to do simple substitutions for getting a first orientation about or intuitive understanding of a mathematical

situation. In this case, the introductory activity is intended as an "advance organizer" (Ausubel et al. 1978) for the later introduction of students to numerical methods for solving differential equations that cannot be solved algebraically.

When we imagine the same situation in an upper secondary classroom, students might do the same substitutions and mental calculations but not know how exponential functions behave for large x. Then the simple idea to organize the results of their calculations, which they might have scribbled on a sheet of paper, into a table will lead to a rigorous proof that in fact they have found all roots of the equation (Table 1).

Between any change of sign from + to – and vice versa, the two curves x^2 and 2^x cut each other and the equation has a root there. For negative numbers -3, -4, -5, etc., x^2 gives the positive squares 9, 16, 25, etc. whereas 2^x produces 1/9, 1/16, 1/25, etc. Another change of signs cannot occur and, thus, no further root. For positive integers, the table teaches us that the differences between successive squares form the sequence of odd integers whereas the value of 2^x is doubled from each integer to the next one. Therefore, x^2 will be smaller than 2^x for every integer >4 and no further root can occur.

Iteration in mathematics, doing the same procedure over and over, with each successive step using the results of the previous step, is a very fundamental concept and procedure in mathematics dating back to ancient times.

The famous method of Hero (who lived in Alexandria, North Africa, round about 100 BC) for finding the square root of a number r is a classic example. One starts the process by guessing a square root x. Therefore $r = x \cdot \frac{r}{x}$ is then a factorization of r into two factors x and $\frac{r}{x}$, one of them being larger, the other smaller than $\sqrt{r}$. Thus, taking the average of x and $\frac{r}{x}$ will give a new and better guess. Repeating this process, one will find the square root of r to whatever level of accuracy is required. In modern notation, Hero's method can be very simply expressed in the following recursive formula where x_{n+1} represents the improved guess from each iteration: $x_{n+1} = \frac{1}{2}\left(x_n + \frac{r}{x_n}\right)$.

Although carrying out repetitive calculations or constructions like these is tedious if done by hand, this has completely changed with the advent of the computational age of electronic calculators and computers. Apart from specialized mathematical software like *Mathematica* and DGS, one can now use ordinary scientific calculators, programmable graphing calculators, and/or spreadsheets to investigate many different kinds of iterations at school level.

Iterative numerical methods like these are indispensable for mathematicians, scientists, and engineers as many problems within mathematics, as well as in the application of mathematics to the real world, cannot be solved algebraically and can

Table 1 Function values for different x

x	-2	-1	0	1	2	3	4	5
x^2	4	1	0	1	4	9	16	25
2^x	1/4	1/2	1	2	4	8	16	32
$x^2 - 2^x$	+	+	–	–	0	+	0	–

only be solved numerically. In general, for example, polynomial equations of order 5 and higher cannot be solved algebraically, and the proof of this by Abel in the nineteenth century is one of the great triumphs of abstract algebra (Derbyshire 2006). And even in the cases of third- and fourth-degree equations, it is often easier to use an iterative algorithm.

As a core example of the fact that "trial and error" is a royal road in mathematics, we discuss an example of a traditional calculus problem that even Grade 8 or 9 learners can and should solve experimentally by only using a calculator and completing a table of values (or using a spreadsheet on computer to speed up the calculations):

A railroad has to be built between two towns A and B as in Fig. 1. The ground to the east of line PQ is marshy, and as a result it costs R20 000 per km to build there, as opposed to the R12 000 per km to the west of this line. If PQ is 80 km, where should X be chosen so that the construction of the railroad is as cheap as possible?

Posing problems like this with students at both school and university, the immediate response is usually to think of minimizing the most expensive cost (building in the marshy terrain), and therefore initially choosing the route AQB. However, after letting them calculate the cost of the route AQB, one could suggest to them to check the minimality of their proposed solution and to repeat the same calculation with other points X between P and Q. Much to their surprise, it will turn out that, for example, for $PX = 5$ km the Total Cost in thousands of rand is $= 20\sqrt{5^2 + 60^2} + 12\sqrt{(80-5)^2 + 60^2} = 2356$, and this is lower than the cost of the route AQB, namely, R2400. Continuing calculating in this manner, as shown in Table 2, leads

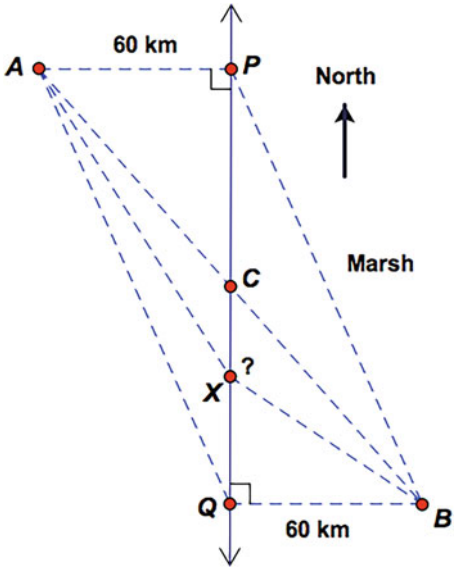


Fig. 1 Railroad building

Table 2 Total cost for different x

PX	5	10	15	20	25	30	35
Total	2356.72	2322.90	2298.44	2283.14	2276.73	2278.87	2289.24

them to eventually conclude that the minimum lies between $QX = 25$ and 30 km, and that they can continue searching within that interval for better and better solutions.

While such an approach may seem very tedious, it pedagogically emphasizes that $PX = x$ is a *numerical variable* that can assume different values and shows how the output varies as the input changes. The substitution and calculation of many values such as this could therefore help to combat the still pervasive "fruit-salad algebra" misconception among high school (and even university) students that letter symbols in algebra represent concrete objects such as apples and oranges (Küchemann 1981; MacGregor 1986).

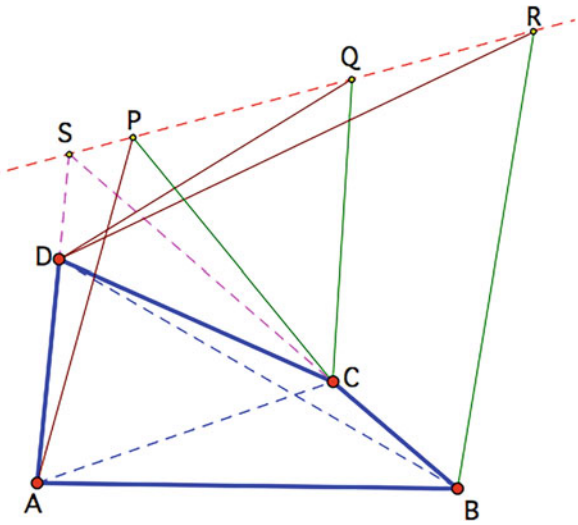
After completion of such a table, it is then more natural and meaningful to introduce an algebraic representation of all those calculations, namely, Cost $(x) = 20\sqrt{x^2 + 60^2} + 12\sqrt{(80 - x)^2 + 60^2}$. Not only is this algebraic representation more economical, but also it now enables one to use the graphing capability of a DGS, or other graphing application, to more quickly and efficiently, graphically, find a minimum. In later grades, one could even differentiate this Total Cost function within a DGS or other symbolic algebra software, draw a graph of the differential function, and then find the minimum more accurately where it cuts the x-axis.

On a higher level, numerical methods can be seen as experimental strategies or algorithms designed to solve mathematical, scientific, or engineering problems in a systematic way, and include error analyses, convergence, and the comparison of different approaches. Though many applications are readily available, numerical analyses may also involve some computer programming. Given the practical importance of numerical methods in modern-day society, it is unfortunate that few elementary methods like decimal search, half-interval or bisection method, iterative formulae, Gaussian elimination, Newton-Raphson, Euler method, etc. have been incorporated into school curricula or undergraduate programs around the world (Shirali 2012).

Experimentation in mathematics not only is useful as a method of discovering new results, but also can sometimes contribute to finding a proof. Here is an example from De Villiers (1997) relating to the following result that had been discovered through experimental investigation with dynamic geometry: If similar triangles APC, DQC, and DRB are constructed on AC, DC, and DB of any quadrilateral $ABCD$ so that $\angle APC = \angle ASB$, where S is the intersection of AD and BC extended, then P, Q, R, and S are collinear (see Fig. 2).

After some unsuccessful attempts at proving this result, the author visually noticed, while manipulating and dragging various points in the configuration, that $\angle ACB$ appeared to be equal to $\angle APQ$. Measuring these two angles confirmed that they were indeed equal, and shortly afterward, a proof of that observed fact followed,

Fig. 2 Collinearity result



which quickly cascaded into a complete proof of the collinearity result. So the empirical observation of the two equal angles was the critical tipping point in this case.

Lew (2006) also gives examples of how the interaction between analysis and dynamic experimental constructions can enable students to arrive at solutions to problems.

5 Refutation in the Classroom

Perhaps one of the most neglected areas in traditional classroom teaching is the involvement of students in the refutation of conjectures and (false) proofs. Most national or provincial mathematics curricula tend to only provide a list of content that needs to be covered in different grades, and invariably such lists are of true results and correct algorithms. When these are translated into textbooks, little opportunity is therefore often given for teachers to deal with false, but plausible conjectures, or with conjectures with restricted domains.

With the exception maybe of a few examples in the junior secondary school of patterns like $6n - 1$ giving a prime number for all n, breaking after a while (e.g., for $n = 6$ it's false), little consistent attention is given to this important aspect of mathematical thinking. In geometry also, one of the few examples that are usually dealt with is the ambiguous case of two triangles with two sides and a "non-included" angle correspondingly equal, which therefore does not guarantee congruency between them.

It is encouraging, however, that the research community in mathematics education appears to be starting to investigate this issue as exemplified by the recent work

of Komatsu (2017), Komatsu et al. (2017), and Komatsu and Jones (2019). In Komatsu and Jones (2019), for example, some task design principles for developing students' heuristic refutation (revising conjectures/proofs through addressing counterexamples) within DGEs are theoretically developed and empirically tested. Using two sets of tasks that were purposefully designed so that with a DGE possible counterexamples could easily emerge, it was found through task-based interviews that these tasks aided students' recognition of contradictions, and assisted them in reformulating their original conjectures.

It is also important to engage students in activities where they are required to produce global counterexamples to false conjectures. Ideally, this might happen when students are encouraged to make their own conjectures, which would give them a greater sense of personal ownership to either disprove or prove their findings. However, the educator may also choose appropriate pedagogical examples, and mixing them in randomly with true conjectures can be quite effective to keep students on their toes.

A particularly instructive example from De Villiers (2010, p. 218) that demonstrates not only the importance of global refutation but also the danger of relying on only experimental evidence is the following investigation. It is adapted from Rotman (1998, p. 3) and can be fruitfully used with mathematics and mathematics education students at school and university:

Investigate whether $S(n) = 991n^2 + 1$ is a perfect square or not. What do you notice? Can you prove your observations?

Since systematic or random calculator or computer investigation for several n strongly suggests that $991n^2 + 1$ is never a perfect square, students are usually easily convinced about the truth of this conjecture, particularly after random testing of the conjecture with a wide range of suggested numbers, including some very large ones on the WolframAlpha site using their prime factorization app at: https://www.wolframalpha.com/input/?i=factor+70560.

After giving them some time to attempt to prove the conjecture, it therefore comes as a great surprise (and therefore an excellent learning experience!) when it is later pointed out to them that the statement is true only for all n until[2]:

$$n = 12\,055\,735\,790\,331\,359\,447\,442\,538\,767$$

For example, as shown in Fig. 3 using the WolframAlpha app above with this input of n, we obtain a number that is a perfect square (as can be seen by the fact that

[2] This is a specific case of a Pell equation, for which solutions were discovered as an offshoot of theoretical work rather than quasi-empirical testing. For example, one can see with a modest amount of experimentation that $x^2 - dy^2 = 1$ is solvable in positive integers when d is a small, positive nonsquare integer and infer (as Indian mathematicians did in the twelfth century) that it is probably solvable for more general d. This led to ad hoc algorithms that worked pretty well (Bhaskara managed the case $d = 61$), and finally, to a theory that produced the present continued-fraction treatment, which is guaranteed to churn out a solution (and will do so with $d = 991$ in fairly short order).

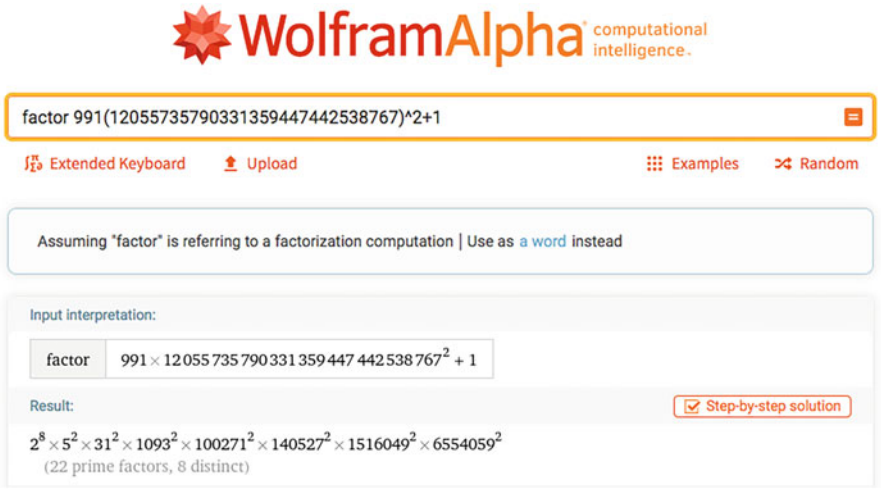


Fig. 3 Refutation with WolframAlpha

all the powers of the prime factors are even), and therefore refutes the original conjecture.

Experimental testing is also useful for identifying incorrect assumptions in otherwise completely valid reasoning. Dynamic geometry software is particularly useful in this regard, as a configuration can be easily and quickly dragged into many different variations in order to check the general validity of one's assumptions. A well-known geometric paradox, namely, "all triangles are isosceles," arises by virtue of construction errors or mistaken assumptions in diagrams (Movshovitz-Hadar and Webb 1998).

A less well-known, yet highly instructive example is the following plausible statement and its "proof" from De Villiers (2009, pp. 25, 112–113, 128–129): "A quadrilateral with one pair of opposite sides equal and one pair of opposite angles equal is a parallelogram."

Proof: Consider Fig. 4. Draw $BX \perp AD$, $DY \perp BC$ and join B and D. Then it is easy to first prove triangles AXB and CYD congruent ($\angle$, $\angle$, s), and then triangles BXD and DYB congruent (90°, s, s). Therefore, $AX = CY$ and $XD = YB$ from which we have $AD = AX + XD = CY + YB = CB$. In other words, $ABCD$ is a parallelogram, since both pairs of opposite sides are equal. QED.

To unsuspecting students, this statement and proof appear quite convincing, and it is usually with great reluctance that they embark on experimentally testing it through a dynamic construction[3] when encouraged to do so. Much to their surprise, as it then turns out, the statement and its proof are false since they are based on the faulty assumption that X and Y will always fall on (the interiors of) AD and CB. This is

[3] Although ideally students ought to do such a dynamic construction for themselves, the instructor could provide them with a ready-made dynamic sketch to save time.

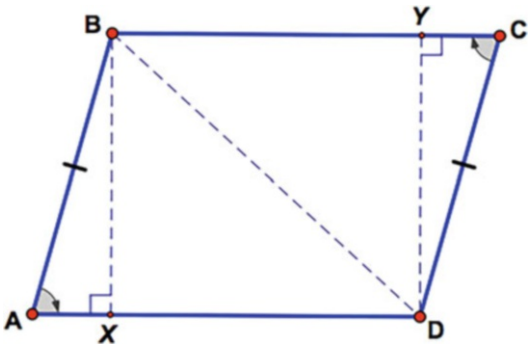


Fig. 4 Plausible conjecture

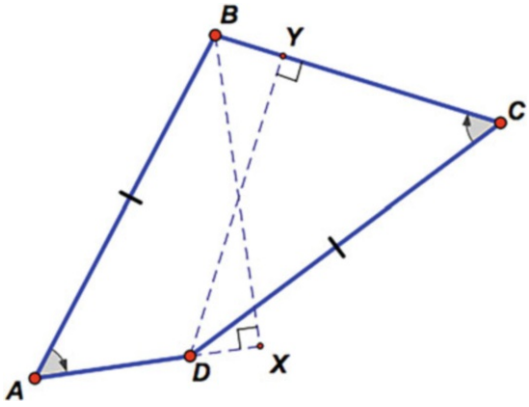


Fig. 5 Counterexample by experimental testing

unfortunately not the case as shown by the counterexample in Fig. 5 where *X* falls outside, on the extension of *AD*, and the above argument is no longer valid. This example therefore illustrates the value of experimentally testing not only one's conjecture, but also one's deductive arguments. It also demonstrates the refutation function of experimentation mentioned at the beginning of the chapter.

One could also propose to students that in order to correct the theorem we would need to add the condition that there must exist at least two opposite vertices with orthogonals to an opposite side falling inside the quadrilateral. Thus, out of a false theorem emerges a correct theorem by the addition of a further condition. Even though this improved theorem is "a bit clumsy," it would give the opportunity for discussing in the classroom that not any true statement is worth becoming a theorem. For example, whether it is worth is a matter of judgment based on criteria such as simplicity, practicability, range of possible applications, etc.

Another example that could be used in class to illustrate the importance of testing one's deductive arguments for faulty assumptions is the following generalization of Viviani's theorem: "the sum of the distances from a point to the sides of an equilateral pentagon is constant." While the result is valid for all convex cases, it

is generally not true for concave equilateral pentagons and is left to the reader to verify at this dynamic sketch online: http://dynamicmathematicslearning.com/viviani-general.html.

An authentic illustration of a heuristic counterexample in the spirit of Lakatos is to have students in class measure and sum a pair of opposite angles of a convex cyclic quadrilateral using dynamic geometry. Then by asking them to drag one of the vertices along the circumference of the circle until a crossed quadrilateral is obtained as shown on the right in Fig. 6, one can create productive cognitive conflict for a fruitful classroom discussion. What is at stake here, of course, is what we choose to understand by the concepts of "quadrilateral" and "interior angle." Such an example therefore naturally stimulates the need to more precisely, and in a more consistent manner, define these concepts (compare Lakatos 1976, p. 16).

Though the easiest way out of the dilemma would be to "monster-bar" the second example as not being a quadrilateral, it is possible to guide students to a mathematically more sophisticated alternative where interior angles are defined more precisely as "directed" angles, and to the following revised formulation of the traditional theorem: "The two sums of the opposite angles of a cyclic quadrilateral are equal (to 180° if it is convex or to 360° when it is crossed)" (compare De Villiers 2015).

Eilers and Johansen (2017, p. 6) have similarly pointed out as follows that "even if a hypotheses is disproved by a counter-example, it may not be necessary to abandon it completely. The common features of a collection of counter-examples will often supply the input to refine the hypothesis into a new form . . . the experimental process can be seen as a stepwise refinement of the initial hypothesis . . ."

Providing students with such simple, accessible Lakatosian learning experiences of modifying, adapting, and reformulating definitions and theorems can potentially

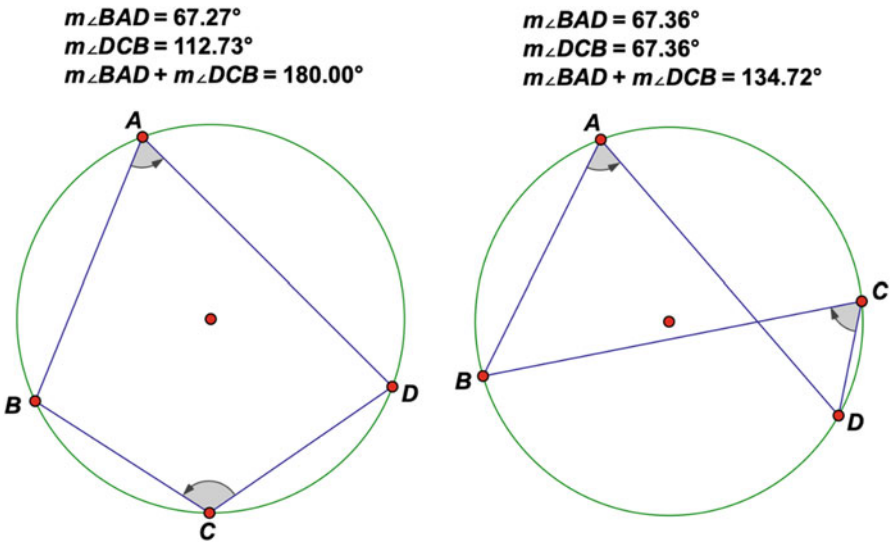


Fig. 6 Counter-example or monster?

help to combat the pervasive myth that definitions and theorems are always static and fixed (Ernest 1991).

Sometimes finding a counterexample with dynamic geometry requires an enlargement (dilation) of the figure by quite a large-scale factor. In 2014, for example, the following conjecture was investigated by De Villiers and Heideman (2014): Start with any triangle ABC and its incircle and incentre *I*. Label the points where the circle touches the sides *BC*, *CA,* and *AB* respectively as A_1, B_1, and C_1 as shown in Fig. 7. Repeat the process with the new triangle $A_1B_1C_1$ and determine its incentre I_1. Then repeating the process twice, it appeared that the four incentres *I*, I_1, I_2, and I_3 were collinear.

Despite the conjecture passing the so-called "drag-test" of dynamic geometry, and initially convincing the authors that it was true, it was only later after some failed attempts at proving it that the authors considered enlarging the diagram by a large factor of 100:1. This conclusively showed that the points were actually not collinear. The reader is invited to dynamically view the enlargement with the interactive sketch using the available slider at: http://www.sineofthetimes.org/refutation-in-a-dynamic-geometry-context/.

As argued and extensively illustrated by Klymchuk (2005, 2007) with many useful classroom examples for introductory calculus, the production of

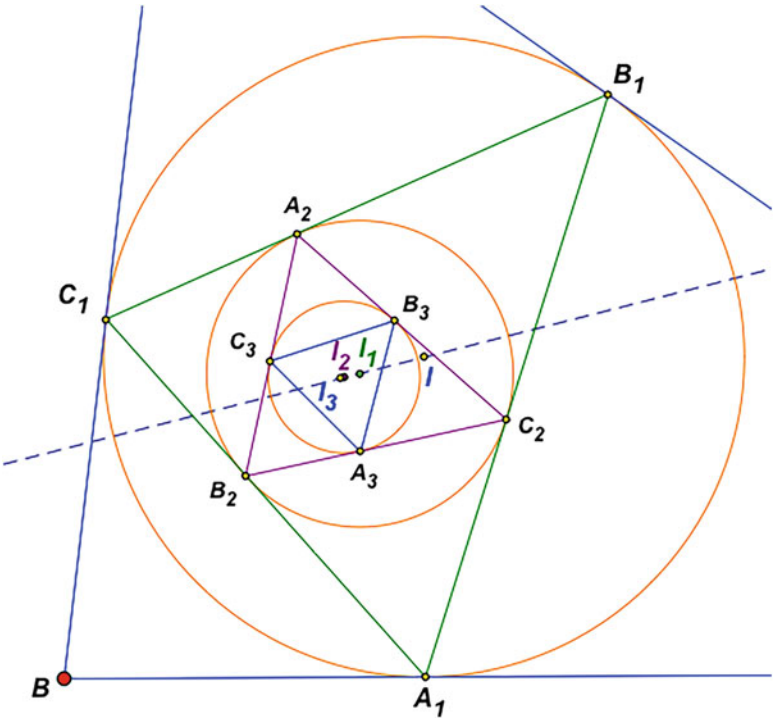


Fig. 7 Iteration of tangent points of incircle

counterexamples by the instructor or students can also deepen understanding, help combat resistant misconceptions, and potentially enhance generic critical thinking.

In the next section, we will briefly examine the use of scientific principles and results within mathematics. In particular, we will focus on the value and use of balancing points (centroids) in the discovery and explaining of some geometric theorems.

6 Using Scientific Principles in Mathematics

The implementation of the principles of physics has since ancient times been a productive approach in many areas of mathematics for not only discovering new results, but also logically explaining (proving) them. Introducing principles of physics into a mathematical theory amounts to adding new axioms/hypotheses to the set of axioms so far accepted. Of course, when a new result is found and proved by such additional principles it still remains an important question to investigate whether it can also be proved in the established axiomatic domain of mathematics proper.

In the following, we restrict ourselves to the application of principles of statics to geometry since these seem to provide the only examples accessible to school teaching.

One of the most famous mathematicians and scientists of antiquity was undoubtedly Archimedes (c.287 – c.212 BC) of Syracuse.

Archimedes used the centroids (centres of gravity) of figures, and the law of the lever, to deduce the volumes of spheres, cones, and pyramids (Heath 1897; Polya 1954; Hawking 2006). Another important figure was the Italian engineer Giovanni Ceva (1647–1734 AD) who in 1678 published a book "De lineis rectis se invicem secantibus statica constructio" exposing a comprehensive approach to elementary geometry by means of static considerations.

The law of the lever can be stated as follows: "Magnitudes are in equilibrium at distances reciprocally proportional to their weights" and is illustrated in Fig. 8. Algebraically, when in balance (equilibrium), it can be formulated as $W_1 \times L_1 = W_2 \times L_2$, where W_i and L_i are, respectively, the weights and distances to the fulcrum. In practice, it is a familiar occurrence to children playing on a seesaw where, for example, a heavy adult would need to be balanced by two or more smaller kids on the other side.

Over the years, the authors have used the experimental and theoretical exploration of balancing points (centroids/centres of gravity) of triangles, quadrilaterals, and other polygons with undergraduate and postgraduate students as well as in

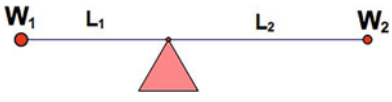


Fig. 8 Law of the balanced lever

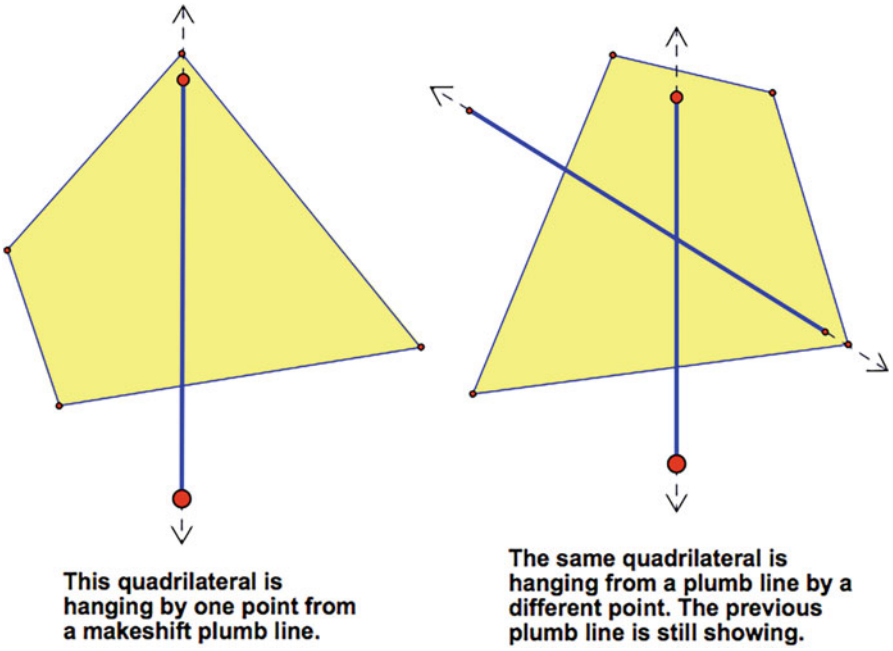


Fig. 9 Finding a centre of gravity with a plumb line

workshops with in-service teachers. The balancing point of a two- or three-dimensional object is called its centre of gravity. In architecture and engineering, accurately locating balancing points is extremely important for designing stable structures that do not collapse (Gordon 1978). The centre of gravity of a two-dimensional object can easily be found through experimentation. For example, in order to locate the approximate centre of gravity of the polygon, students can be asked to balance a cardboard polygon on the tip of a pencil or an eraser.

Another way of experimentally locating the centre of gravity of a cardboard polygon is shown in Fig. 9. The polygon is hanging on a string that is attached near its edge. The string is acting as a carpenter's plumb line, which provides the carpenter with a line perpendicular to the ground. The balancing point or centre of gravity is then located where these two lines cross. When doing experiments like these to locate centres of gravity, the results are obviously subject to some experimental error.

After dealing with the following theoretical method of finding the centroid of a cardboard (lamina[4]) quadrilateral *ABCD* (see Fig. 10), students can be encouraged to carry out the geometric construction on their physical models, and to compare the experimental finding with the theoretical location:

[4] In physics, a lamina is a two-dimensional object with uniform density and negligible thickness.

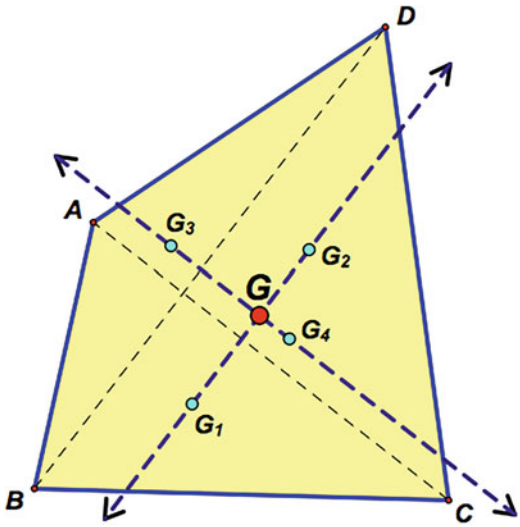


Fig. 10 Finding the centroid of a cardboard quadrilateral

1) Find G_1 and G_2, the respective centroids of triangles *ABC* and *ACD*, then the balancing point of the whole quadrilateral must lie somewhere on the line connecting these two centroids.
2) Find G_3 and G_4, the respective centroids of triangles *ABD* and *BCD*, then the balancing point of the whole quadrilateral must lie somewhere on the line connecting these two centroids.
3) Hence, the balancing point of *ABCD* must lie at the intersection of the two lines in 1) and 2).

Generally, the results of comparing the experimentally found balancing point and the theoretical location agree fairly well, unless of course major experimental or construction errors were made. Of some pedagogical interest too is to have students drag a dynamic quadrilateral with its constructed lamina centroid so that it becomes concave. Much to their surprise, and some cases even with some mild bewilderment, students will find that the balancing point could actually move outside of the quadrilateral as shown in Fig. 11. In order to balance a concave cardboard quadrilateral like that, one would have to attach a thin (comparatively weightless) wire to the centroid in order to balance it. While the balancing point moving outside of a figure is not something uncommon – for example, the rim of a wheel also has its centroid "outside" the rim, which is why it needs to be connected with spokes to its balancing point, i.e., the centre of the circle forming the rim – students nonetheless find this quite a surprising revelation.

From a sporting perspective this observation also provides a scientific explanation why the so-called Fosbury flop is so much more efficient for high jumping than the old-fashioned straddle technique. By bending their bodies in an arc as shown in Fig. 12, high jumpers are able to move their centres of gravity slightly outside their

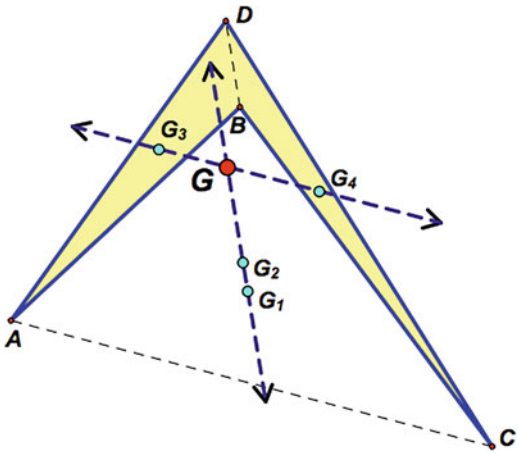


Fig. 11 Balancing point outside concave quadrilateral

Fig. 12 Fosbury flop in high jumping

body as they go over the cross bar, hence saving a few centimeters, which could be the difference between winning or not.

One can also demonstrate the theoretical solution for finding the balancing point of a cardboard quadrilateral quite nicely in another way with dynamic geometry. Since the relative weights of the cardboard triangles ABD and BCD are determined by their areas, we can determine the relative weights, respectively, concentrated at the centroids G_3 and G_4 by measuring the areas of triangles ABD and BCD as shown in Fig. 13. Now using the lever law of Archimedes, we can easily determine the balancing point (centroid) of the weights at G_3 and G_4. In particular, dilating G_3 from G_4 as centre and a scale factor of $\frac{areaABD}{(areaABD+areaBCD)}$ gives us the required balancing point G_3'. And doing this construction on the same sketch as the earlier one in Fig. 10 will show students that it is the same balancing point G as before.

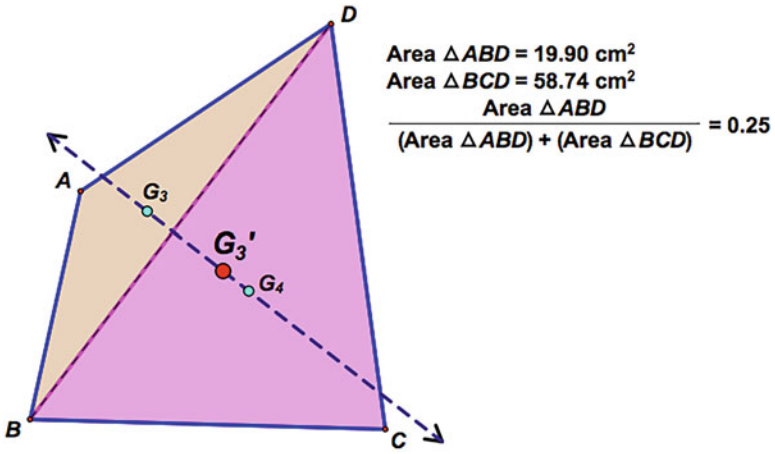


Fig. 13 Using areas as measure of cardboard weight

Saving the theoretical construction of the balancing point of a cardboard quadrilateral centroid as a tool in dynamic geometry also enables us to easily find the balancing point of a cardboard pentagon by dividing it in two different ways into a quadrilateral and a triangle, drawing the two lines between the respective centroids and finding their intersection. Alternatively, one can again use the law of the lever on a division of a pentagon into a quadrilateral and a triangle and determining their respective centroids and areas. This approach is easily extended to hexagons, etc.

A different physical way of analyzing balancing points (centroids) of polygons is to imagine equal point masses located at the vertices, and connected with thin wire of negligible weight to make it a stable structure. Where would such a structure balance? In the case of the triangle, the point mass balancing point is the same as that of the cardboard triangle, lying at the point of concurrency of the medians, but except for a parallelogram, it is not generally the case for a quadrilateral that the point mass centroid coincides with its lamina (cardboard) centroid.

As demonstrated by the French mathematician Pierre Varignon (1654–1722), the point mass centroid of a quadrilateral with equal weights at the vertices is located at the centre of the Varignon parallelogram formed by the midpoints of the sides (e.g., Hanna and Jahnke 2002). As shown in the first figure in Fig. 14, the point mass centroid does not coincide with the lamina centroid. The second figure in Fig. 14 also shows another interesting geometric property that is easy to explain (prove) using arguments from physics. For example, given equal point masses at the vertices of $ABCD$ it follows that the two weights at A and C would balance at the midpoint K of the diagonal AC. Since the same applies to the two weights at B and D, balancing at the midpoint L of the diagonal BD, it follows that altogether as a system, the four weights would balance at the midpoint of KL, which implies that it coincides with the point mass centroid (the Varignon centre of the Varigon parallelogram).

Similarly, by considering different weights at the vertices of a triangle, just like Giovanni Ceva did (see above), we can also deduce, from the principles of mechanics, the celebrated concurrency theorem of Ceva (compare Hanna and Jahnke 2002).

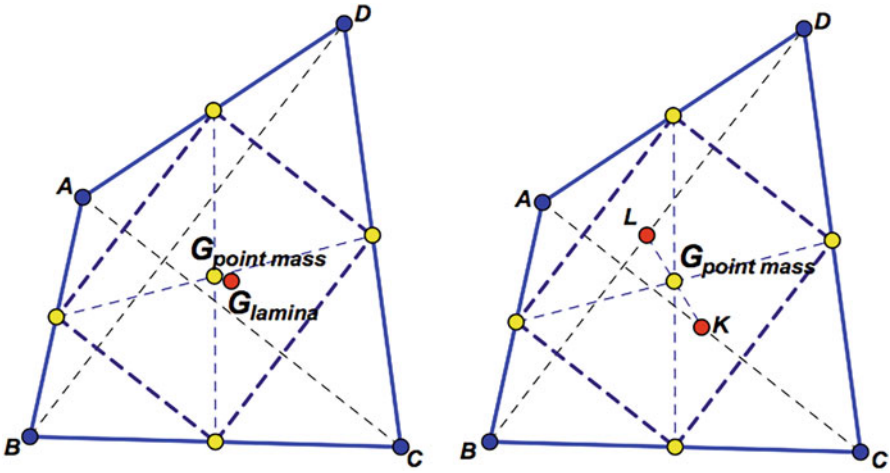


Fig. 14 Point mass centroid and lamina centroid

Fig. 15 Mechanical demonstration of Fermat point

Lastly, it is not hard to design a dynamic geometry exploration for students, and a guided worksheet to prove from coordinate geometry, that the coordinates of the centroid of a triangle can be found from the average of the coordinates of its vertices. This can be further extended by using a weighted average to find the coordinates for the Ceva point for different point mass weights located at the vertices, and easily extended further to higher polygons.

For a short outlook on other ways of using means from physics to understand mathematical relations, we hint at applying physical models and phenomena to illustrate mathematical problems and theorems. This can potentially contribute to making mathematics more meaningful to a lot of students that are not destined to further pursue "pure" mathematics. The Fermat point of a scalene triangle, i.e., the point that minimizes the sum of the distances to the vertices can easily be modeled by weights hanging in balance as shown in Fig. 15. Soap bubble geometry can also be

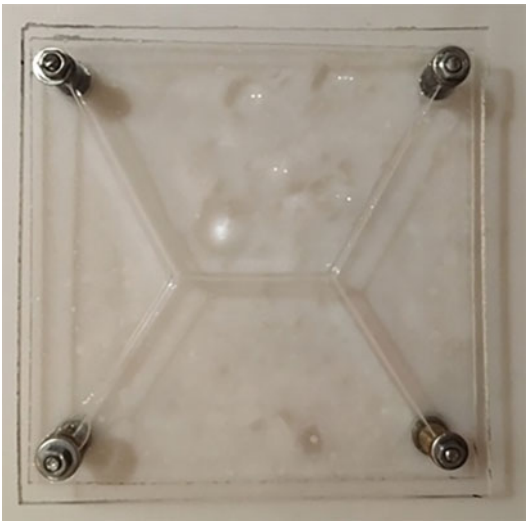


Fig. 16 Soap bubble demonstration of Steiner tree network for a square

used in class and workshops with teachers to quite dramatically illustrate the Fermat point of a scalene triangle, and as shown in Fig. 16, the similar Steiner shortest path network for a square.

7 Modelling the Real World

The inclusion of mathematical modelling as part of the mathematics curricula of many countries provides an excellent opportunity to include more experimentation into the classroom. As explained in Sect. 2, this is an important step beyond the Lakatosian point of view since, with modelling, experimenting with different hypotheses also comes into play.

There exist quite a lot of diagrammatic representations of the circular process of modelling (f. e. Blomhøj and Jensen 2003, p. 125; Kaiser and Stender 2013, p. 279). We use a very simplified diagram which consists of three important steps or stages as illustrated in Fig. 17, namely,

1) Construction of the mathematical model
2) Processing the model
3) Interpretation and evaluation of the results

Out of a practical problem, a mathematical model is constructed comprising a set of data and a set of hypotheses. The latter can be a geometric representation, a number of algebraic formulae or functions or differential equations, respectively, or a mixture of them. In the next step, the mathematical model is processed. This might mean different activities: to insert data into the equations, to manipulate formulae,

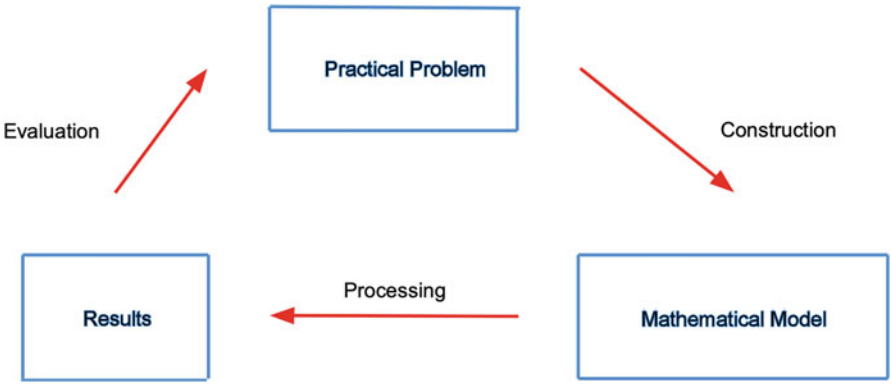


Fig. 17 Modelling stages

and/or to derive logical consequences from the hypotheses of the model which amounts to proving. Then the results of the processing are evaluated in regard to the practical problem, and if necessary, the hypotheses are modified thus starting a run through a new cycle (Fig. 17).

Today, for processing a model many tools such as Dynamic Geometry Systems (DGS) are available for introducing concepts like the derivative in a dynamic way which was not possible before (Ndlovu et al. 2011). Traditionally, much of schooling in the past has consisted of teaching learners Stage 2, namely, how to process given mathematical models by solving equations comprising linear, quadratic, trigonometric, exponential, hyperbolic, etc. functions, but hardly involving them at all in the Stage 1 construction and development of these models from real-world contexts, and much less on Stage 3, i.e., evaluating the solutions and how well they fit "reality." In short, in some countries, it consists of teaching mathematics largely disconnected from the real world, though fortunately in many countries mathematical modelling is now demanded by the official curricula (see Sect. 1 above).

Moreover, the first stage of "model developing" as well as the last "evaluation" stage can provide ample opportunity for including some experimental work. At the first stage, for example, students may simulate Galileo Galilei's experiments of rolling balls down inclined planes and gather data to determine the relationship between distance covered and time. Of course, after the derivation of a model, students could experimentally test how well it fits reality.

Here is a simple example of how "testing" a theoretical solution to a geometry problem against a real-world background led to a new modified solution under certain conditions. In De Villiers (2003, pp. 27–33) as in some other books, the concepts of "circumcentre" and "perpendicular bisector" (as "a locus of equidistant points from the endpoints of a line-segment") are introduced in order to find suitable equidistant points, respectively, from three rural villages A, B, and C (Fig. 18) to build a water reservoir. It was stated at the outset that the villages were all of equal size, while students were expected to identify other implicit assumptions such as that the villages were on the same plane, that the surface was flat, that there were no

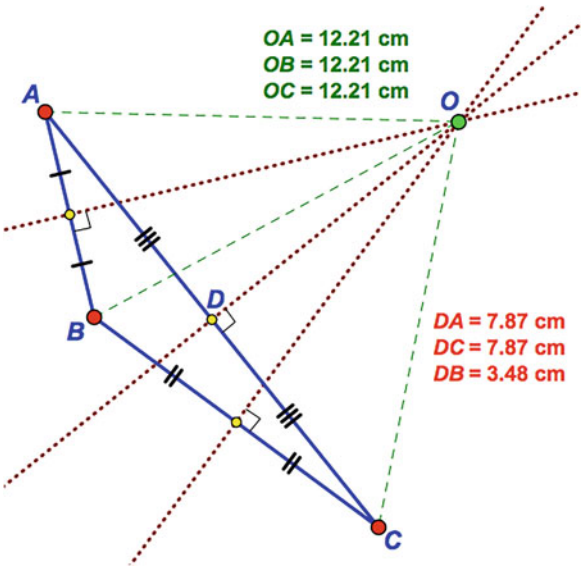


Fig. 18 Modified, improved solution

natural obstacles such as dense forest or rivers, and that the distances were sufficiently large in comparison to the sizes of the villages and the water reservoir that they could geometrically be represented by points.

During 2011, in a class with prospective mathematics teachers, author 1 was discussing with the whole class the results of their investigations that the students had completed during the previous session in the computer lab. Projected on to the screen with the data projector was a triangle with its circumcentre constructed with dynamic geometry as the required equidistant point for the reservoir. To discuss a question on the worksheet, the triangle was dragged into a right triangle, and dragging continued until it became obtuse, so that the circumcentre moved far outside the triangle formed by the three villages as shown in Fig. 18.

Immediately, one of the students objected strongly, saying that position would not be a good solution in the real world as the reservoir would then be too far away from the rural villages. Eventually only after a whole class discussion (compare De Villiers 2013), it was agreed that in an obtuse triangle case such as shown in Fig. 18, a better solution than the circumcentre would be to build the reservoir at the midpoint of the longest side. This makes sense, since in the case of Fig. 18, the midpoint D of the longest side AC is much closer to all three villages (yet still equidistant from villages A and C).

Another example of a theoretical solution in geometry having a restricted domain when interpreted within the real world is the following problem from De Villiers (2003, p. 6): Given an island (approximately) in the shape of a parallelogram, find a point on the island to build a holiday cabin that minimizes the sum of the distances to each of the beaches. It is quite easy to discover and prove by students that

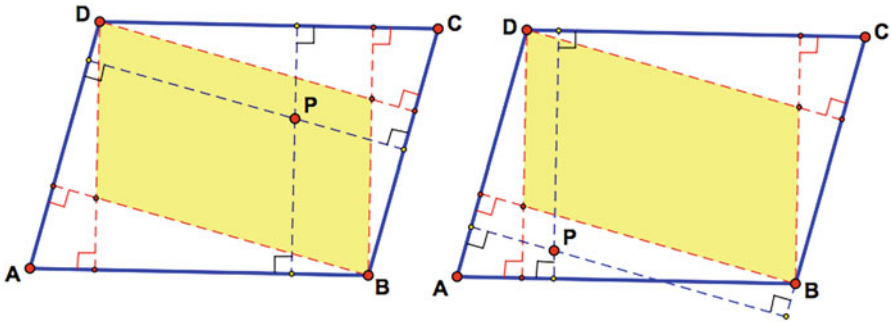


Fig. 19 Restricted solution domain

theoretically the cabin can be built anywhere inside since the sum of the distances from a point *P* to the sides of the parallelogram is constant.

However, in an actual real-world context the solution domain as shown in Fig. 19 would be restricted by the enclosed area formed by the perpendiculars from *B* to the sides *AD* and *CD*, and from *D* to the sides *AB* and *BC*. As shown in the second figure in Fig. 19, if *P* is moved outside the solution domain, one of the distances would fall on an extension of one of the sides outside and would therefore be in the sea and the "distance" would no longer be a distance to a beach. Also note that if the adjacent sides of the parallelogram "island" become extremely tilted toward each other, no feasible domain is possible, in which case one may need to choose a different criterion for building a cabin.

Ignoring the practical, real-world aspects of the problem, students may also be encouraged to drag the point outside the parallelogram to note that according to the DGE software the sum of the distances would no longer be constant. This can then also be used as an opportunity of a local or heuristic counterexample by introducing students to the concept of "directed" distances in order to make the theorem valid outside the parallelogram as well.

Though the two geometry problems above are rather artificial approximations to real-world problems, such problems can still serve an important educational purpose, if teachers point out to students similar constraints may arise when interpreting and testing general theoretical solutions against the real world. This contributes important insights to develop a more adequate understanding of modelling. In performing experiments in school, students should realize that their results are probably crude, and that, due to the limitations of measuring techniques and measuring instruments, "real" scientists in research laboratories also do not obtain "perfect" results (i.e., experimental data that fit a mathematical relationship exactly, for example, lying exactly on a straight line).

Unfortunately, textbooks often convey a distorted perspective in this regard by supplying "typical data" that fit the desired model precisely. They are giving the impression that the data were obtained in an actual experiment, while probably they were generated by substitution in the relationship they are supposedly seeking. Furthermore, by reporting only "perfect" data, the very important scientific activity

of processing and drawing appropriate conclusions from imperfect data seldom, if ever, arises.

By carefully running through modelling cycles from the earliest beginning, students should better and better understand that, since it is not possible to always include all variables in a real-world situation or to accurately quantify them, mathematics can never provide an exact, precise model of the real world. What should have become clear from the simple examples above is also true for science in general. In most cases, simplifying assumptions have to be made, and usually the simplest available mathematical model is preferred as long as it provides a useful, accurate enough approximation for description, prediction, engineering design, etc.

Quite often more than one possible mathematical model can be used to describe the same physical phenomenon from the real world. We can just think here of the complex Ptolemaic system of planetary motion (approx. 150 AD) involving epicycles, yet it was so accurate that it was in use for well over a 1000 years until the arrival of the Copernican system (1543) that was not only much simpler to use, but also growing experimental evidence substantiated it. A neat, illustrative video clip comparing the planetary systems of Ptolemy, Copernicus, and Tycho Brahe is available for students at https://youtu.be/aL_k7NwZ-2Y, but a Google search will produce many more videos that an instructor could use in this regard.

Even in ancient astronomy, the situation arose of two competing models explaining equally well the same set of data. One of the models is mathematically available to interested students of the lower secondary level. In 2010, a respective teaching experiment was performed in two classrooms (see Jahnke 2010; Jahnke and Wambach 2013). It concerns the so-called "anomaly of the sun" which was discovered for the first time by the great astronomer and mathematician Hipparchos (second century BC). Roughly speaking, the term refers to the observation that the half year of summer (spring + summer) is about 1 week longer than the half year of winter. This contradicted the basic conviction of Greek astronomers that all heavenly bodies move with constant velocity in circles around the centre of the earth (see Fig. 20). Therefore, Hipparchos invented a new hypothesis to "save the phenomena" which can be found in Ptolemy's Almagest, book III, section 2 (Toomer 2020, p. 141). In the pre-Hipparchos period, the Greeks assumed that during a year the sun moves with constant velocity on a circle, the so-called ecliptic passing vernal equinox (VE), summer solstice (SS), autumnal equinox (AE), and winter solstice (WS). Since in this model the arcs between these four points which define the four seasons spring, summer, autumn, and winter are equally long, it follows that the seasons last equally long. Hipparchos was the first to explicitly discuss the deviation of astronomical data from this simple model and to propose a new hypothesis. He kept to the assumption that the sun moves on a circular path with constant velocity. His new hypothesis was the assumption that the centre of this circle is not situated in the centre of the earth, but a bit outside (see Fig. 21). Consequently, Ptolemy called this the eccentric hypothesis.

Hipparchos' new hypothesis not only allowed a much more precise calculation of the daily position of the sun, but also implied two remarkable theorems. One of them

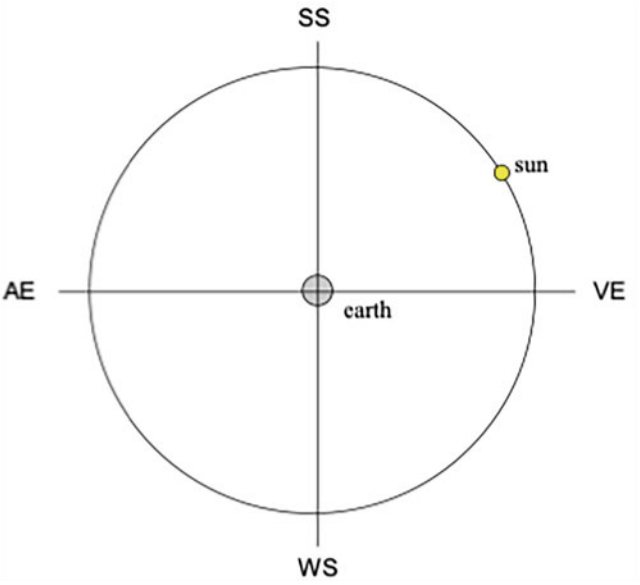


Fig. 20 Path of the sun

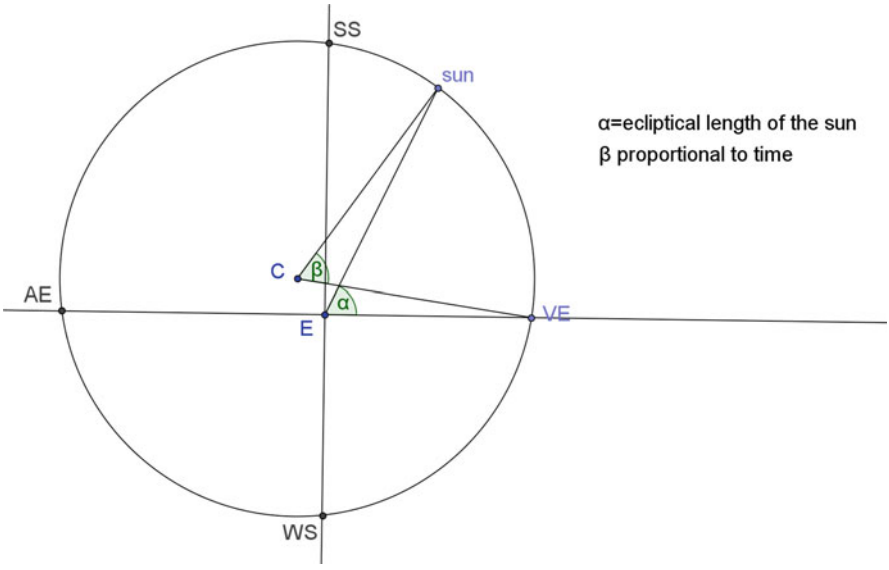


Fig. 21 Hipparchos' eccentric hypothesis. E is the centre of the earth, C the centre of the sun's circular path

says that the distance of the sun from the earth varies over the year and that there is one point where it is nearest to earth (perigee) and one where it is most distant (apogee). The second theorem says that the velocity of the sun also varies with a maximum in perigee and a minimum in apogee.

Hipparchos' model predicts not only the lengths of the seasons, but also the positions of the sun at any point in time. This position can be measured by its so-called "ecliptical length,"which is the angular distance between the vernal equinox and the position of the sun at a certain moment in time observed from the earth (angle α in Fig. 21). Since the sun travels with constant velocity on its circular path, angle β (Fig. 21) measures the time elapsed since the vernal equinox. The model predicts the position of the sun with astonishing precision.

The students constructed, using modern data of the lengths of the seasons, Hipparchos' eccentric model by means of Geogebra. They found that the maximal deviation from modern data is less than 3 angular minutes, by far below observability for the Greeks.

The historical setting of this teaching experiment enabled the students in a special way to reflect on the process of experimenting with hypotheses. The students were aware of the fact that Hipparchos' hypothesis is quite different from the modern theory. Nevertheless, it provided very precise predictions of the yearly path of the sun. They were told that Ptolemy applied also another hypothesis, the epicycle hypothesis, stemming from Appolonios of Perge (third century BC). Since both hypotheses provided the same data, Ptolemy felt free to choose the simpler hypothesis for his practical calculations. Thus, this episode in the history of astronomy uncovers in a particular manner the relativistic and experimental nature of all modelling, in general alien to Greek and European thinking until the twentieth century.

8 Conclusion

In conclusion, here are quotes from two highly respected mathematicians that emphasize the importance of experimentation in mathematics and mathematical research:

> *The mathematician at work makes vague guesses, visualizes broad generalizations, and jumps to unwarranted conclusions. He (*sic*) arranges and rearranges his (*sic*) ideas and becomes convinced of their truth long before he (*sic*) can write down a logical proof. The conviction is not likely to come early – it usually comes after many attempts, many failures, many discouragements, many false starts … experimental work is needed … thought-experiments*. – Paul Halmos (1984:2)

> *… we … practice unashamedly and in broad daylight, certain customs in which we indulge only in the privacy of our offices, and which we never admitted to students: experimentation. To a degree which never appears in the courses we teach, mathematics is an experimental science* … – Henry Pollak (1984:12).

What these professional mathematicians say about the role of experimentation in doing mathematics is true but cannot immediately be applied to situations of

teaching. The mathematicians devote their professional time exclusively to doing mathematics striving for new results. This is of course not the case for students at schools and only in a rather limited manner even for students at the tertiary level. Thus, for developing and keeping alive a balance between experimentation and deduction is an educational task sui generis.

Seen from the point of view of a long-term development of mathematical thinking, experimentation in the primary and lower secondary grades (c. 1 to 8) would be characterized by informal arguments and mostly amount to pattern recognition and arguing with patterns. One could call this a phase of informal thought experiments (Jahnke 2007, 84 pp). From grade 8 onward until the end of upper secondary, students work in small networks of deductively connected mathematical relations. Freudenthal (1973, p.458) speaks of local organization, and if modelling and small physical theories are also considered one could call this the phase of hypothetico-deductive thinking. The concept of proof should explicitly be made a theme, and real measuring should be included which leads to building small networks of theorems based on empirical evidence. At the end of attending school, students should have an idea of what a theory is. Only in a third phase, students (of tertiary level) would work in autonomous mathematical theories, and only then proof would mean what an educated mathematician would understand by proof.

In the present chapter, we have concentrated on examples of the second phase. These included topics of pure mathematics (i.e., number theory, geometry, and calculus). Geometry has a double face. It can be considered as a purely mathematical theory, when, however, seen as a theory of the real space around us, it gets the character of a theory of physics.

These phases are also more or less described by other authors. Wittmann (1981) and Skemp (1976, pp. 54–67) describe a transitional phase from intuitive to reflective thinking and view the latter as a more critical extension of the former. In the Van Hiele model of learning, this corresponds to a transition from the two lower levels of recognition and learning of properties to the third level of logical implications, and beyond (Van Hiele 1973).

Arzarello et al. (2012, pp. 97, 132) have similarly referred to the dynamic tension between the empirical nature of explorations with tools and the deductive nature of the discipline. They also illustrate with several examples the role of experimentation in proving activities in class and discuss some theoretical frameworks for better understanding the cognitive processes of students when involved in such activities. Boero et al. (2010) also distinguish several levels of development in the continuity of student levels of thinking until "cognitive unity" is ultimately achieved.

After exploring the historical interplay between the inductive and deductive, experimental and theoretical, Armella and Sriraman (2005) have proposed the notion of *situated proofs*, which can be defined as the result of a systematic exploration within an (computational) environment. The idea is that situated proofs can span a bridge between situated and formalized knowledge.

Nonetheless, however one conceives of the long-term development of mathematical thinking, experimentation and deductive reasoning should intimately interact. Frequently, it is said that experimentation leads to guessing a theorem and the matter

is finished when we are able to prove it. But this is only half the truth. We do not experiment without an aim or a question, and this requires deductions and considerations before the experiment has started. And the historical examples in Sect. 3 as well as the examples in Sect. 5 (on refutation) and in Sect. 7 (on modelling) show that the story is not finished with a proof. It is often necessary and fruitful to look for examples and possibly counterexamples for the theorem proved.

Also, the intellectual journey from handling concrete matters evolving into abstract theories requires that teachers stimulate their students again and again to think about what would happen when things or conditions were different. Not only the question "why is it as it is," but also the question "what would be if this and that would be different" should be guiding principles of any math teaching. To say it with a metaphor from grammar as another principle of teaching: "more subjunctive in the mathematics classroom!"

In conclusion, experimentation and deductive reasoning each have important, indispensable roles to play in mathematics and its teaching and should complement each other rather than seen to be in opposition to one another.

References

Armella LM, Sriraman B (2005) Structural stability and dynamic geometry: some ideas on situated proofs. ZDM 37(3):130–139

Artigue M, Blomhøj M (2013) Conceptualizing inquiry-based education in mathematics. ZDM 45(6):797–810. https://doi.org/10.1007/s11858-11013-10506-11856

Arzarello F, Bussi MGB, Leung AYL, Mariotti MA, Ian Stevenson I (2012) Experimental approaches to theoretical thinking: artefacts and proofs. In: Hanna G, de Villiers M (eds) Proof and proving in mathematics education, 69 new ICMI study series 15. Springer, New York, pp 97–137. https://doi.org/10.1007/978-94-007-2129-6_5

Ausubel DP, Novak JD, Hanesian H (1978) Educational psychology: a cognitive view. Holt, Rinehart and Winston, New York

Blomhøj M, Jensen TH (2003) Developing mathematical modelling competence: conceptual clarification and educational planning. Teach Math its Appl 22(3):123–139

Blum W, Galbraith PL, Henn H-W, Niss M (eds) (2007) Modelling and applications in mathematics education: the 14th ICMI study. Springer, New York

Boero P, Douek N, Morselli F, Pedemonte B (2010) Argumentation and proof: a contribution to theoretical perspectives and their classroom implementation. In: Pinto MMF, Kawasaki TF (eds) Proceedings of the 34th conference of the International Group for the Psychology of Mathematics Education, vol 1. International Group for the Psychology of Mathematics Education, Belo Horizonte, pp 179–204

Borwein JM (2012) Exploratory experimentation: digitally-assisted discovery and proof. In: Hanna G, de Villiers M (eds) Proof and proving in mathematics education, 69 new ICMI study series 15. Springer, New York, pp 69–96. https://doi.org/10.1007/978-94-007-2129-6_4

Bruder R, Müller H (1983) Zur Entwicklung des Könnens im Lösen von Begründungs- und Beweisaufgaben im Mathematikunterricht. Math Sch 21(12):886–894

Bruder R, Prescott A (2013) Research evidence on the benefits of IBL. ZDM 45(6):811–822. https://doi.org/10.1007/s11858-11013-10542-11852

Calinger R (2016) Leonhard Euler: mathematical genius in the enlightenment. Princeton University Press, Princeton

Davis PJ, Hersh R (1981) The mathematical experience. Birkhäuser

Davis PJ, Hersh R (1983) The mathematical experience. Pelican Books, Harmondsworth

De Villiers M (1997) The role of proof in investigative, computer-based geometry: some personal reflections. In: Schattschneider D, King J (eds) Geometry turned on! MAA, Washington, DC, pp 15–24
De Villiers M (2003) Rethinking proof with sketchpad. Key Curriculum Press, Emeryville
De Villiers M (2009) Some adventures in Euclidean geometry. Lulu Publishers, Morrisville
De Villiers M (2010) Experimentation and proof in mathematics. In: Hanna G, Jahnke HN, Pulte H (eds) Explanation and proof in mathematics: philosophical and educational perspectives. Springer, New York, pp 205–221. https://doi.org/10.1007/978-1-4419-0576-5_14
De Villiers M (2013) Equality is not always 'best'! Learn Teach Math 14:17–21
De Villiers M (2015) Crossed quadrilaterals: a missed Lakatosian opportunity? Philos Math Educ J 29. Accessed on 7 August 2021 at: http://socialsciences.exeter.ac.uk/education/research/centres/stem/publications/pmej/pome29/index.html
De Villiers M (2021) Review of "introduction to experimental mathematics" by Søren Eilers and Rune Johansen, (2017). Math Gaz 105(563):379–380. https://doi.org/10.1017/mag.2021.97
De Villiers M, Heideman N (2014) Conjecturing, refuting and proving within the context of dynamic geometry. Learn Teach Math 17:20–26
Derbyshire J (2003) Prime obsession: Bernhard Riemann and the greatest unsolved problem in mathematics. Joseph Henry Press, Washington, DC
Derbyshire J (2006) Unknown quantity: a real and imaginary history of algebra. Joseph Henry Press, Washington, DC
Eilers S, Johansen R (2017) Introduction to experimental mathematics. Cambridge University Press, Cambridge
Epstein D, Levy S, de la Llave R (1992) About this journal. Exp Math 1(1):1–3
Ernest P (1991) The philosophy of mathematics education. Falmer Press, Hampshire
Freudenthal H (1973) Mathematics as an educational task. Reidel, Dordrecht
Gordon JE (1978) Structures: or why things don't fall down. Penguin Books, London
Grattan-Guinness I, Ravetz JR (1970–1990) Fourier, Jean Baptiste Joseph. In: Gillispie (ed) Complete dictionary of scientific biography. downloaded from encyclopedia.com, New York
Grattan-Guinness I, Ravetz JR (cooperator) (1972) Josef Fourier 1768–1830. A survey of his life and work, based on a critical edition of his monograph on the propagation of heat, presented to the Institut de France in 1807. MIT Press, Cambridge/London
Hadamard J (1945) Essay on the psychology of invention in the mathematical field. Princeton University Press, Princeton
Halmos P (1984) Mathematics as a creative art. In: Campbell D, Higgens J (eds) Mathematics: people, problems, results, vol II. Wadsworth, Belmont, pp 19–29
Hanna G, Jahnke HN (2002) Arguments from physics in mathematical proofs: an educational perspective. Learn Math 22(3):38–45
Hanna G, De Villiers M, Arzarello F, Dreyfus T, Durand-Guerrier V, Jahnke HN, Yevdokimov O (2012) ICMI study 19: proof and proving in mathematics education: discussion document. In: Hanna G, de Villiers M (eds) Proof and proving in mathematics education. The 19th ICMI study. Springer, Dordrecht et al, pp 443–452
Hawking S (2006) The method of Archimedes treating of mechanical problems – to Eratosthenes. In: Hawking S (ed) God created the integers. Penguin Books, London, pp 209–239
Hawkins T (1980) Lebesgue's theory of integration. Its origins and development. University of Wisconsin Press, Madison
Heath TL (1897) The works of Archimedes. Cambridge University Press, Cambridge
Hochkirchen T (2003) Theory of measure and integration from Riemann to Lebesgue. In: Jahnke HN (ed) A history of analysis. American Mathematical Society, London, pp 261–290
Jahnke HN (2003) Algebraic analysis in the 18th century. In: Jahnke HN (ed) A history of analysis. American Mathematical Society, London, pp 105–136
Jahnke HN (2007) Proofs and hypotheses. ZDM 39(1/2):79–86
Jahnke HN (2010) The conjoint origin of proof and theoretical physics. In: Hanna G, Jahnke HN, Pulte H (eds) Explanation and proof in mathematics: philosophical and educational perspectives. Springer, New York, pp 17–32. https://doi.org/10.1007/978-1-4419-0576-5_2

Jahnke HN, Ufer S (2015) Argumentieren und Beweisen. In: Bruder R, Hefendehl-Hebeker L, Schmidt-Thieme B, Weigand H-G (eds) Handbuch der Mathematikdidaktik. Springer, Berlin/Heidelberg, pp 331–355
Jahnke HN, Wambach R (2013) Understanding what a proof is: a classroom-based approach. ZDM 45(3):469–482
Kaiser G, Stender P (2013) Complex modelling problems in co-operative, self-directed learning environments. In: Stillman G, Kaiser G, Blum W, Brown J (eds) Teaching mathematical modelling: connecting to research and practice. Springer, Dordrecht, pp 277–293
Klymchuk S (2005) Counter-examples in teaching/learning of calculus: students' performance. The New Zealand Mathematics Magazine, Vol 42, ISSN: 0549-0510
Klymchuk S (2007) Counter-examples in calculus. Maths Press, Auckland
Komatsu K (2017) Fostering empirical examination after proof construction in secondary school geometry. Educ Stud Math 96(2):129–144
Komatsu K, Jones K (2019) Task design principles for heuristic refutation in dynamic geometry environments. Int J Sci Math Educ 17:801–824. https://doi.org/10.1007/s10763-018-9892-0
Komatsu K, Jones K, Ikeda T, Narazaki A (2017) Proof validation and modification in secondary school geometry. J Math Behav 47:1–15
Küchemann DE (1981) Algebra. In: Hart KM (ed) Children's understanding of mathematics: 11–16. John Murray, London, pp 102–119
Lakatos I (1976) Proofs and refutations. Cambridge University Press, London
Lakatos I, Musgrave A (eds) (1970) Criticism and the growth of knowledge. In: Proceedings of the international colloquium in the philosophy of science, London 1965, vol 4. Cambridge at the University Press, Cambridge
Langer RE (1947) Fourier series. Slaught memorial paper. Am Math Mon Suppl 54:1–86
Lew H (2006) Pappus in modern dynamic geometry: an honest way for deductive proof. Paper presented at the ICMI study conference 17, Digital technologies and mathematics teaching & learning, Hanoi Institute of Technology, 3–8 December 2006
Luetzen J (1983) Euler's vision of a general partial differential calculus for a generalized kind of function. Math Mag 56(5):299–306
Luetzen J (2003) The foundations of analysis in the 19th century. In: Jahnke HN (ed) A history of analysis. American Mathematical Society, London, pp 155–195
MacGregor M (1986) A fresh look at fruit salad algebra. Aust Math Teach 42(3):9–11
Movshovitz-Hadar N, Webb J (1998) One equals zero and other mathematical surprises. Key Curriculum Press, Emeryville
Ndlovu M, Wessels D, De Villiers M (2011) An instrumental approach to modelling the derivative in Sketchpad. Pythagoras 32(2):Art. #52, 15 pages. https://doi.org/10.4102/pythagoras.v32i2.52
Pollak HO (1984) The effects of technology on the mathematics curriculum. In: Olivier AI (ed) The Australian experience: impressions of ICME 5. Centrahil, AMESA
Polya G (1954) Induction & analogy in mathematics, vol 1. Princeton, Princeton, pp 155–158
Rotman J (1998) Journey into mathematics: an introduction to proofs. Prentice Hall, New York
Schoenfeld AH (1985) Mathematical problem solving. Academic Press, Orlando [u.a.]
Shirali SA (2012) Numerical analyses as a topic in school mathematics. Regular lecture presented at the 12th international congress on mathematical education, Seoul
Skemp RR (1976) Relational understanding and instrumental understanding. Math Teach 77:20–26
Szabó Á (1960) Anfange des Euklidischen Axiomensyslems. Arch Hist Exact Sci 1(1):38–106
Toomer GJ (2020) Ptolemy's Almagest, translated and annotated by G. J. Toomer with a foreword by Owen Gingerich. Princeton University Press, Princeton
Van Hiele PM (1973) Begrip en Inzicht. Muusses, Purmerend
Wittmann E (1981) The complementary roles of intuitive and reflective thinking in mathematics teaching. Educ Stud Math 12:389–397
Wittmann EC, Müller GN (1990) Handbuch produktiver Rechenübungen, vol 1 und 2. Klett, Stuttgart

Section XII

Views from Other Domains

Andrew Aberdein

Introduction to Views from Other Domains

Andrew Aberdein

Abstract

The study of mathematical practice has always been an interdisciplinary enterprise and is not confined to history and philosophy. Important contributions have been made by scholars from many domains, including sociology, education, argumentation theory, rhetoric, formal epistemology, and theology. The chapters in this section provide overviews and introductions into the insights that scholars of mathematical practice may glean from these disciplines.

Keywords

Argumentation theory · Education · Formal epistemology · Rhetoric · Sociology · Theology

One of the closest allied disciplines to the history and philosophy of mathematical practice is the sociology of mathematics. Indeed, many of the conferences and collections that helped to define the study of mathematical practice as a field explicitly included sociology within their scope (for example, Van Kerkhove and Van Bendegem 2007; Löwe and Müller 2010). Classic works by sociologists of mathematics have also long attracted significant attention from scholars of mathematical practice (notable examples include Restivo 1992; Heintz 2000; MacKenzie 2001; and Rosental 2008). In their chapter, "Inverting Hierarchies: The Sociology of Mathematical Practice," Michael J. Barany and Milena Kremakova provide a wide-ranging historical survey of the different research traditions in sociology that have been applied to mathematical practice, including positivist, constructivist, Marxist, ethnomethodological, and institutional approaches. They also look forward to some of the ways the sociology of mathematics may develop in the future by way of

A. Aberdein (✉)
School of Arts & Communication, Florida Institute of Technology, Melbourne, FL, USA
e-mail: aberdein@fit.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_122

intersections with the study of other social practices, such as work, gender, education, and technology.

Historiography is another domain of inquiry that is closely allied to the core disciplines of the history and philosophy of mathematical practice. The historiography of mathematics presents distinctive questions, perhaps most conspicuously concerning the proper status of histories of mathematics (and other accounts of mathematical practice) written by professional mathematicians. Many such works have been deservedly influential in the study of mathematical practice, despite occasional departures from methodological orthodoxy. In his chapter, "Practice Makes Perfectoid: Pasts, Futures, and the Historiography of Contemporary Mathematics," Michael J. Barany teases out many of the ways the professional mathematician and the professional historian can diverge. By way of illustration, he reflects on some of his interactions with the number theorist Michael Harris, a doyen of the internal approach (notably in his book, Harris 2015). In particular, Barany contrasts his own historical study of Laurent Schwartz's theory of distributions with Harris' insider account of the much more recent emergence of Peter Scholze's perfectoid spaces, thereby highlighting the strengths and weaknesses of these two perspectives. Barany's chapter itself contrasts with Viktor Blåsjö's chapter in a different section of this *Handbook*, which makes the case for a revival of internalist historiography of mathematics (Blåsjö 2024). For yet further perspectives, a useful recent collection is (Remmert et al. 2016).

Perhaps the most prolific of other domains bearing on the study of mathematical practice is education. There are many more mathematics educationists than historians or philosophers of mathematics; while not all of their research is necessarily of interest to scholars of mathematical practice, there is much that is or should be. Crucially, many mathematics teachers and scholars of mathematics education agree that what goes on in mathematics classrooms should reflect the actual practices of working mathematicians. If this goal is to be met, it is essential for mathematics education to work out what those practices are. In their chapter, "Mathematics Education Research on Mathematical Practice," Keith Weber and Matthew Inglis provide an overview of some of the research mathematics educationists have conducted in this spirit. They focus on two particular research questions: What does the understanding of mathematical concepts comprise and what do mathematicians do when they read proofs? These questions (and others) have inspired significant empirical research from which scholars of mathematical practice may profit. There are many collections with contributions that address the relationship between mathematical practice and mathematics education research (for example, Ernest 2018; Stylianides and Harel 2018; and Kaiser and Presmeg 2019). In particular, Weber is also the coeditor of a recent collection with a focus converse to that of this chapter: introducing scholarship on mathematical practice to mathematics education researchers (the introduction is Weber et al. 2020).

One of the factors that makes mathematics distinctive as a practice is that it embeds a distinctive approach to reasoning. There are fields of mathematics and of traditional philosophy of mathematics that take mathematical reasoning as their subject matter, notably mathematical logic. However, mathematical reasoning also

has significant points of similarity to other forms of reasoning. Argumentation theory is an umbrella term for many approaches that focus on reasoning in nonmathematical contexts. But argumentation theory may also be brought to bear on mathematical practice. In their chapter, "Argumentation in Mathematical Practice," Andrew Aberdein and Zoe Ashton explore some key techniques of argumentation theory and investigate their application to mathematical reasoning. This chapter pays particular attention to two devices which have found widespread application to mathematics: Toulmin layouts and argumentation schemes. Toulmin layouts were first proposed by the philosopher Stephen Toulmin as an analysis of defeasible reasoning (Toulmin 1958). Many mathematics educationists have found this device particularly useful, so this chapter is in some respects complementary to that of Weber and Inglis. An argumentation scheme is a stereotypical pattern of reasoning, such as analogy or appeal to expert opinion. Argumentation schemes have ancient roots but have recently been studied extensively by Douglas N. Walton, a philosopher, and various collaborators, many from computer science (Walton et al. 2008). This chapter also examines the wider role of argumentation in mathematics beyond proof and discusses the social aspects of mathematical argumentation. For further consideration of the topics of this chapter, see (Aberdein and Dove 2013).

As the study of persuasive communication, rhetoric is closely allied to argumentation theory; indeed, it is sometimes considered a proper part of that field. Rhetoric is also related to mathematics, in at least two ways: mathematics can be put to rhetorical use as a way to persuade people of nonmathematical claims; but mathematics also has a rhetoric of its own. Much mathematical reasoning embodies rhetorical features which many treatments of mathematical practice have tended to ignore. Aberdein and Ashton briefly address these relationships between rhetoric and mathematics, but a much more extensive treatment is provided by G. Mitchell Reyes in his chapter, "Rhetorical Approaches to the Study of Mathematical Practice." Reyes pays particular attention to a work familiar to most scholars of mathematical practice, Imre Lakatos' *Proofs and Refutations* (Lakatos 1976). However, as a rhetorician, Reyes brings a novel perspective to this work. Through a close reading, he spotlights some of Lakatos' own rhetorical choices and shows how they in turn work to illuminate the rhetoric implicit in the mathematics behind the development of a proof for the Descartes–Euler conjecture. Reyes is also a coeditor of a recent collection of papers on the relationship between mathematics and rhetoric (Wynn and Reyes 2021).

Bayesian approaches to probability comprise another line of enquiry that is much more familiar in application to nonmathematical reasoning than to mathematical reasoning. This is understandable, since Bayesian epistemology is designed to accommodate uncertainty whereas mathematical proof promises the elimination of uncertainty. Nonetheless, there are important aspects of mathematical practice from which uncertainty cannot be (or, at least, has not yet been) eliminated. James Franklin, in his chapter "Bayesian Perspectives on Mathematical Practice," provides a brief introduction to Bayesian reasoning and discusses some of these aspects, including the means by which mathematical consensus can form about the likely outcome of as yet unproven conjectures, such as the Riemann hypothesis. For a more

extensive, but still accessible introduction to Bayesian epistemology, see (Schupbach 2022).

Formal epistemology is the application of formal methods derived from logic or mathematics to questions from epistemology. As such, it can be seen as incorporating Bayesian epistemology as a component of a broader enterprise. (For a collection of survey articles on formal epistemology, see Pettigrew and Weisberg 2019.) One way of distinguishing the philosophy of mathematical practice from more traditional projects in the philosophy of mathematics is that it tends to pay greater attention to the epistemology of mathematics as opposed to its ontology. Hence, it is natural that formal epistemology should itself be applied to mathematical practice. In their chapter, "Polymath as an Epistemic Community: Analyzing the Digital Traces of the Polymath Project," Patrick Allo, Bart Van Kerkhove, and Jean Paul Van Bendegem apply one such formal method, network theory, to the study of the Polymath project, a much discussed recent example of mathematical practice. The Polymath project originated in a proposal for "massively collaborative mathematics" coauthored by the prominent mathematician Tim Gowers (Gowers and Nielsen 2009). Gowers subsequently helped to orchestrate several crowd-sourced mathematical projects, which have led to genuine progress on some open problems, including papers published under the collective pseudonym "D. H. J. Polymath" (the initials standing for "density Hales-Jewett" after the theorem for which the first Polymath project provided a new elementary proof) (Polymath 2012, 2014). What makes the Polymath project particularly exciting for scholars of mathematical practice is that its collaborations are primarily mediated by publicly accessible web sites, thereby generating a searchable record of research-level mathematics in progress and granting access to a characteristically private process. Allo, Van Kerkhove, and Van Bendegem provide a broad survey of the uses to which this resource has been put by past scholars and offer a novel analysis of their own, grounded in the tools of network theory. This leads to some striking insights into the structures that have emerged within the communities of mathematicians who have participated in the 16 Polymath projects to date.

Theology may seem to be an unpromising source of insights into mathematical practice. It is certainly much less closely allied to mathematics than many of the domains discussed so far. However, Stanisław Krajewski answers the question he poses in the title of his chapter "Is Mathematics Connected to Religion?" firmly in the affirmative. Some mathematicians, from Pythagoras onward, have ascribed a sacred significance to their mathematical work; more widely, theological metaphors have been influential for many mathematicians and mathematical progress has come from theological disputes. Celebrated historical examples include George Berkeley's critique of the calculus which originated as a tu quoque response to the theological views of its early promoters (Moriarty 2018, 2023), and Georg Cantor's surprisingly extensive engagement with Catholic theologians while researching the foundations of transfinite set theory (Dauben 1977). Some philosophers have postulated a deeper connection between mathematics and religion. Silvia Jonas, for example, has argued that abstract beliefs in both contexts present parallel difficulties for rational acceptability (Jonas 2016). Several recent collections offer diverse perspectives on the

interrelation of theology and mathematics: Krajewski is himself the coeditor of a special issue for which he provided a useful introduction (Krajewski 2016); another prominent collection is (Lawrence and McCartney 2015).

There are yet more domains with valuable perspectives on mathematical practice which are not the topic of a chapter in this section. Formal epistemology is not the only recent trend in the study of knowledge that should be of interest to scholars of mathematical practice. Naturalized epistemology, the application of empirical results to epistemological questions, also offers many valuable insights. Experimental philosophy, or "x-phi," can be understood as a closely related trend. Two recent collections each contain several chapters bringing such work to bear on mathematical practice (Aberdein and Inglis 2019; Bangu 2018). Cognitive science is perhaps the empirical field from which the greatest share of the relevant empirical results are drawn. Studies of numerical cognition include classic works such as (Dehaene 1997; Lakoff and Núñez 2000), and there has been a substantial volume of subsequent research. For a recent anthology, see (Cohen Kadosh and Dowker 2015). There is also a considerable overlap between the cognitive science of mathematics and some of the domains addressed by chapters in this section, notably mathematics education.

Researchers in artificial intelligence have long taken an interest in mathematics, particularly in the context of designing theorem proving software. One perspective within such work is that in order to produce an artificial mathematician we must first improve our understanding of what it is that human mathematicians do. Conversely, an artificial mathematician would reveal what was necessary to mathematical practice and what was mere human idiosyncrasy. (That recent breakthroughs in artificial intelligence have come from large language models whose inner workings are notoriously inscrutable does not undermine the value of the work conducted from this opposed tradition, although the application of such techniques to mathematics does present significant philosophical issues of its own.) Alan Bundy, a pioneer in the application of artificial intelligence to mathematics, has long championed an approach that pays close attention to mathematical practice (Bundy 1983). The mathematician turned computer scientist Ursula Martin has also investigated mathematical practice extensively to this end (and was one of the earliest scholars to reify "mathematical practice" as an object of enquiry) (Martin 1999; Martin and Pease 2013). A recent anthology of research on the significance of theorem proving technology for the understanding of mathematics is (Hanna et al. 2019).

The relationship between mathematics and the creative arts is a large enough topic to have inspired another handbook from the same editor as the present collection (Sriraman 2021). Literature, in particular, has a multiplicity of surprisingly deep connections to mathematical practice. Close attention to formal structure is of central importance in many literary media, just as it is to much mathematics. In addition, many of the domains discussed above, such as sociology, education, argumentation, and rhetoric, are also intimately related to literature and its scholarship. Some authors, such as Lewis Carroll and Jorge Luis Borges, are particularly well known for the mathematical sophistication of their work (see Wilson and Moktefi 2019; Bloch 2008). But there are many other writers whose use of mathematical concepts reflects the currency that these ideas had for the educated public of

their era, itself a valuable insight into mathematical practice in the wide sense. In recent years, there has been increasing scholarly attention to many of these phenomena. Individual studies address the relationship between literature and mathematics from the early modern era (e.g., Jarrett 2019) to the twentieth century and beyond (e.g., Engelhardt 2018; Bamford 2020). Other scholars have surveyed a broader chronological range (e.g., Abbott 2023; Hart 2023). A valuable anthology of recent scholarship is (Tubbs et al. 2021).

In sum, interdisciplinarity is at the heart of the study of mathematical practice. Progress relies on insights not only from history and philosophy, but also from many other domains including other humanistic disciplines, as well as social sciences and other fields. While this presents its own challenges, not least in securing suitable venues for publication, it is also a great source of vitality and originality.

1 Cross-References

- ▶ Historiography of Mathematics from the Mathematician's Point of View
- ▶ Mathematical Practices and Written Evidence: General Reflections Based on a Historian's Experience
- ▶ The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts

References

Abbott S (2023) The proof stage: how theater reveals the human truth of mathematics. Princeton University Press, Princeton

Aberdein A, Dove IJ (eds) (2013) The argument of mathematics. Springer, Dordrecht

Aberdein A, Inglis M (eds) (2019) Advances in experimental philosophy of logic and mathematics. Bloomsbury, London

Bamford A (2020) Mathematics and modern literature. New Left Rev 124:107–123

Bangu S (ed) (2018) Naturalizing logico-mathematical knowledge: approaches from philosophy, psychology and cognitive science. Routledge, New York

Blåsjö V (2024) Historiography of mathematics from the mathematician's point of view. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer, Cham

Bloch WG (2008) The unimaginable mathematics of Borges' library of Babel. Oxford University Press, Oxford

Bundy A (1983) The computer modelling of mathematical reasoning. Academic Press, London

Cohen Kadosh R, Dowker A (2015) The Oxford handbook of numerical cognition. Oxford University Press, Oxford

Dauben JW (1977) Georg Cantor and Pope Leo XIII: mathematics, theology, and the infinite. J Hist Ideas 38(1):85–108

Dehaene S (1997) The number sense: how the mind creates mathematics. Oxford University Press, New York

Engelhardt N (2018) Modernism, fiction and mathematics. Edinburgh University Press, Edinburgh

Ernest P (ed) (2018) The philosophy of mathematics education today. Springer, Cham

Gowers T, Nielsen M (2009) Massively collaborative mathematics. Nature 461:879–881

Hanna G, Reid DA, de Villiers M (eds) (2019) Proof technology in mathematics research and teaching. Springer, Cham
Harris M (2015) Mathematics without apologies: portrait of a problematic vocation. Princeton University Press, Princeton
Hart S (2023) Once upon a prime: the wondrous connections between mathematics and literature. Harper Collins, London
Heintz B (2000) Die Innenwelt der Mathematik: Zur Kultur und Praxis einer beweisenden Disziplin. Springer, Vienna
Jarrett J (2019) Mathematics and late Elizabethan drama. Palgrave Macmillan, Cham
Jonas S (2016) On mathematical and religious belief, and on epistemic snobbery. Philosophy 91(1): 69–92
Kaiser G, Presmeg N (eds) (2019) Compendium for early career researchers in mathematics education. Springer, Cham
Krajewski S (2016) Theological metaphors in mathematics. Stud Log Gramm Rhetor 44:13–30
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge
Lakoff G, Núñez RE (2000) Where mathematics comes from: how the embodied mind brings mathematics into being. Basic Books, New York
Lawrence S, McCartney M (eds) (2015) Mathematicians and their gods: interactions between mathematics and religious beliefs. Oxford University Press, Oxford
Löwe B, Müller T (eds) (2010) PhiMSAMP. Philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London
MacKenzie D (2001) Mechanizing proof: computing, risk, and trust. MIT Press, Cambridge, MA
Martin U (1999) Computers, reasoning and mathematical practice. In: Schwichtenberg H (ed) Computational logic. Proceedings of the NATO advanced study institute on computational logic, Marktoberdorf, Germany, 29 July–10 August 1997. Springer, Berlin, pp 301–346
Martin U, Pease A (2013) Mathematical practice, crowdsourcing, and social machines. In: Carette J, Aspinall D, Lange C, Sojka P, Windsteiger W (eds) Intelligent computer mathematics, LNAI, vol 7961. Springer, Berlin, pp 98–119
Moriarty CM (2018) The ad hominem argument of Berkeley's *Analyst*. Br J Hist Philos 26(3): 429–451
Moriarty CM (2023) Duelling catechisms: Berkeley trolls Walton on fluxions and faith. Int Hist Rev 33(2):205–226
Pettigrew R, Weisberg J (eds) (2019) The open handbook of formal epistemology. The PhilPapers Foundation, London
Polymath DHJ (2012) A new proof of the density Hales-Jewett theorem. Ann Math 175:1283–1327
Polymath DHJ (2014) The "bounded gaps between primes" Polymath project: a retrospective. Eur Math Soc Newsl 94:13–23
Remmert VR, Schneider MR, Sørensen HK (eds) (2016) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Cham
Restivo S (1992) Mathematics in society and history: sociological inquiries. Kluwer, Dordrecht
Rosental C (2008) Weaving self-evidence: a sociology of logic. Princeton University Press, Princeton
Schupbach JN (2022) Bayesianism and scientific reasoning. Cambridge University Press, Cambridge
Sriraman B (ed) (2021) Handbook of the mathematics of the arts and sciences. Springer, Cham
Stylianides AJ, Harel G (eds) (2018) Advances in mathematics education research on proof and proving: an international perspective. Springer, Cham
Toulmin S (1958) The uses of argument. Cambridge University Press, Cambridge
Tubbs R, Jenkins A, Engelhardt N (eds) (2021) The Palgrave handbook of literature and mathematics. Palgrave Macmillan, Cham

Van Kerkhove B, Van Bendegem JP (eds) (2007) Perspectives on mathematical practices: bringing together philosophy of mathematics, sociology of mathematics, and mathematics education. Springer, Dordrecht

Walton DN, Reed C, Macagno F (2008) Argumentation schemes. Cambridge University Press, Cambridge

Weber K, Dawkins P, Mejía-Ramos JP (2020) The relationship between mathematical practice and mathematics pedagogy in mathematics education research. ZDM 52(6):1063–1074

Wilson R, Moktefi A (eds) (2019) The mathematical world of Charles L. Dodgson (Lewis Carroll). Oxford University Press, Oxford

Wynn J, Reyes GM (eds) (2021) Arguing with numbers: the intersections of rhetoric and mathematics. Penn State Press, Philadelphia

Inverting Hierarchies: The Sociology of Mathematical Practice

Michael J. Barany and Milena I. Kremakova

Contents

1 Introduction: Mathematical and Social Order 2598
2 Sociologies of Knowledges 2599
3 Making a Hard Case 2601
4 Hard Cases and Soft Targets 2605
5 Roots and Radicals 2607
6 Watching and Listening to Mathematicians 2610
7 Conclusion: Making and Breaking Order 2613
8 Cross-References 2614
References 2614

Abstract

Sociology originated in the mid-nineteenth century from a new confidence in the power of science to explain the world on a mathematical foundation. Both mathematics and sociology transformed over the ensuing century, inverting the hierarchical relationship from sociology as a mathematics-based science of complex human configurations to mathematics as a complex science based on social institutions. That is, where sociology began as the hard case for mathematics, it became possible to see mathematics as the hard case for sociology. In this light, we examine a number of motivations and provocations for the sociology of mathematics, from language and social cognition to Marxist and materialist skepticism of ideology, and sample the evidence and arguments brought to bear on the sociology of mathematics at a range of scales.

M. J. Barany
Science, Technology & Innovation Studies, The University of Edinburgh, Edinburgh, UK

M. I. Kremakova (✉)
The Sociological Review Magazine, The Sociological Review Foundation, London, UK
e-mail: milena.kremakova@gmail.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_87

Keywords

Constructivism · Ethnography · Experience · Institutions · Labor · Language · Marxism · Materialism · Mathematics · Social studies of mathematics · Sociology · Sociology of mathematical practice · Sociology of mathematics · Sociology of scientific knowledge (SSK) · Wittgenstein

1 Introduction: Mathematical and Social Order

Mathematics and society are based upon two radically different but essentially identical forms of order. Where mathematics is made up of *abstract* entities and *formal* structures governed by *universal* laws of reasoning, society counts *human* entities and *social* structures governed by *particular* laws of relation. Both are made from methods and modes of intersubjective agreement and of human groups organizing themselves and their senses of reality around something that is ultimately in their heads, a shared imagination. The evidence of the working of mathematics and of society is in collective episodes of argument and validation and of proposition and correction, by which people make sense of the possibility of having something invisible, immaterial, and greater than themselves in common.

Mathematics is a very old science, arguably dating back thousands of years. Sociology is a very new one, aged just a couple of hundred. The former was integral to the origin of the latter. To give one necessarily fabular tale of beginning: sociology started as the opposite polestar to mathematics in the positive philosophy of Auguste Comte, whose ungainly coinage of *sociologie* urged the possibility of a systematic science of society built on a hierarchy of more fundamental physical and mathematical sciences with abstract mathematics at the very foundation (Comte 1842).

The ambition for a science of society, for a social physics (as Comte and many contemporaries also termed it), grew from early nineteenth-century transformations in the relationship between mathematics and society. Propounded in new scales and uses of comprehensive training in calculus-based engineering (see Belhoste 2003), this transformation instilled a new confidence among certain European elites in the engineerability, and accordingly the mathematical knowability, of the increasingly urbanized social world after the Industrial Revolution. Comte's sociology became one of many sometimes-conflicting mid-nineteenth-century sciences of society – mathematical statistics derived from perhaps an especially notable contemporary variation of "social physics" (Hacking 1990) – each with its own social and scientific legacies and relationships to mathematics.

In this light, the sociology of mathematics has a kind of *ouroboros* character: it circles back; it eats its own tail. If the apparent universality of mathematical order underwrote the promise of a scientific account of the social world, such accounts can have the paradoxical effect of localizing and particularizing socially produced mathematics. Sociologists and mathematicians, alike and in their own ways, revel in paradoxes. They show the limits and the contradictions of what may be agreed, of the terms on which people may come together.

The next section reviews some of the major areas of sociology that have had something to say about mathematics. We then sketch how the sociology of mathematics, as such, has been constructed as one such limit, or contradiction, and develop some of the consequences of that view. Our analysis dwells at some length on the sociology of David Bloor, not necessarily because his work is representative or exemplary of the sociology of mathematics but because it has been a widely shared point of reference in the field, an excellent provocation, and a fruitful illustration of many of the themes discussed here. We identify a thread of interest in collective and societal responsibility in mathematics, which we follow from early Marxist analyses of the discipline through the proposals of Sal Restivo and others motivated by the socially critical potential of sociologizing mathematics. Finally, we consider the varieties of evidence at different scales that have informed such sociologies and some of the conclusions that can be drawn from these, including understandings of small-scale practices, labor, communities, and large-scale networked knowledge and authority.

2 Sociologies of Knowledges

Ideas about the relationship between mathematical and social order are at least as old as the oldest societies that were large and complex enough to put mathematics to social purposes. As an apparatus of recording and accounting, mathematics began as an aid to memory and reconciliation, as far as the oldest surviving mathematical records appear to show. Mathematics stabilized valuable entities across time and among people. Its stabilizing uses made it useful in claims about justice, initially with measurements and resources and later in other domains of contestation and coordination, including (a very long time after the earliest mathematical records) those of civic and religious authority (Robson 2008; Desrosières 1998). These uses and expectations of mathematics constituted claims about its social conditions and implications.

Such claims can be identified in numerous traditions, many of very long standing, including traditions of philosophy as well as the practice of mathematics. They include invocations such as the famous heading to Plato's Academy, forbidding those ignorant of mathematics from entering. The idea of mathematics as a prerequisite to organized knowledge has never been unanimous or uncontested but runs from before Plato's time through to our own times. After the trivium of grammar, logic, and rhetoric, the quadrivial mathematical subjects of arithmetic, geometry, music, and astronomy made up the foundation of the medieval liberal university curriculum. Vernacular arithmetician Robert Recorde claimed in 1543 that numbers were "the ground of all men's affairs" (see Barany 2023a). The pedagogical status of mathematics supported broader claims about the relationship between the subject and the intrinsic characteristics of human thought and perception, notably significant to Kantian metaphysics among many philosophical traditions.

Sociologies of mathematics draw from these long traditions of reasoning about relationships of mathematics and society. These sociologies have come, in turn, from a variety of distinct and often mutually disconnected scholarly traditions concerned

more explicitly and systematically with such society-knowledge (including society-mathematics) relationships. There has been no single or even dominant disciplinary or methodological home for social studies of mathematics, which have ranged across and often combined social scientific, anthropological, humanistic, philosophical, pedagogical, and other modes of inquiry.

For the purposes of this chapter, such an ungainly and unruly cluster of knowledge about knowledge is drawn together by a shared attention to the conditions and labor of producing and reproducing theoretical mathematics in the context of social institutions. That is, the sociology of mathematical knowledge is distinguished by seeking the interpersonal and socially organized dimensions of the most personal and logically governed forms of knowing. This view of the field favors the elements that are closer to the mainstreams of the philosophy of mathematical practice, at the expense of social accounts of the uses of mathematics and quantification, including their social and sociological uses. These latter topics have at times crossed fruitfully with what might be called the sociology of "mathematics itself" (notably, for instance, in the work of Donald MacKenzie 2001) but have tended to distinguish between the socially manifested meanings and implications of mathematical reckoning and the situated production of those reckonings as such.

Institutional contexts of mathematics include formal organizations – educational institutions, research institutes, states, corporations, and so on – as well as the constellations of norms, expectations, and more or less formally organized ways of understanding and being in the world that mediate participation in social groups of various kinds connected to mathematical work (see Barany 2021, 2022; a summary of the main sociological uses of the term "institution" can be found in Miller 2019). Institutions can be understood at a range of scales, from how an individual experiences and navigates institutional forms to how such forms structure roles and interactions at the scale of an entire society or of interacting social groups. The sociology of mathematics has tended to focus on the smaller end of these scales, closer to the activities of intellection and communication of shared interest to the philosophy of mathematical practice. However, especially in sociological studies of methods and systems of education or on patterns of mathematical careers or funding, scales of analysis can grow and interpenetrate, linking the personal to much larger patterns of thought and action.

Sociologies of knowledge and of mathematics interpret institutions from a variety of methodologies and perspectives, reflecting these fields' decidedly mixed parentage from many different branches of sociology, anthropology, psychology, and other humanistic and scientific disciplines. Some generalize and aggregate; some particularize and individualize. Sociologies of mathematics, like philosophies of mathematics, may be unusual in the degree to which individual experiences of cognition or sense-making are taken as general or exemplary on the basis of the presumed generality of laws and principles of mathematical reasoning: many accounts of how mathematics works in general (sociologically) are based on very few social individuals or contexts.

The precise relationship of individual experiences of mathematics to the kinds of phenomena discernible on larger scales can make a great difference, analytically and methodologically. How one understands the nature of mathematical consensus and

of collective reasoning changes what it means for individuals and communities to know and share mathematics. Producers and users of mathematical knowledge manage to agree on a lot, and they often see their agreement as being special in form, character, and even logical necessity. However, the social and historical fact of mathematics is that disagreement, misunderstanding, error, and variant comprehension are regular and routinized parts of mathematical research, learning, and communication. Sociological (particularly microsociological, ethnographic, and microhistorical) approaches cast light on the process of achieving mathematical agreement before disagreements are folded into consensus and erased.

Constructivist sociologies see such variation as evidence that mathematical consensus is only ever provisional and is made in social groups from a variety of ideas, experiences, practices, and values that combine to support particular senses of what one should accept and how one should reason. Among a variety of approaches, one may study such social constructions phenomenologically, examining how mathematics manifests in the lived experience of its users and producers, or one may study these ethnographically at the level of cultural forms and norms and group dynamics of sense-making. Constructivism is often opposed to realism or Platonism, two perspectives that prioritize claims that mathematics derives ultimately and inevitably from stable underlying truths that can be temporarily missed or misunderstood but that are not meaningfully revisable. This distinguishes sociology of mathematics from its object: while being a Platonist or anti-Platonist is typically understood to make no practical difference for a mathematician's daily work, these opposing epistemological perspectives result in entirely different approaches to sociological analyses of mathematics, the study of that daily work. Whether disagreement and negotiation are part of mathematical knowledge or extrinsic matters that circumscribe mathematical knowledge alters what needs to be explained about processes of knowledge making and sharing in mathematics and how these can be explained.

The status of dissensus matters both for settled or "textbook" knowledge and for open, unresolved, dynamic, or contested mathematics in the making. Sociologies of mathematics often focus on the latter as a circumstance where assumptions and processes of reasoning are expected to be more explicit, as people try to develop their understandings and to convince each other. Settled mathematics can be approached with related methods by focusing on contexts of learning, mobilizing the fact that settled knowledge is only ever settled for those who know it, and there is a constant supply of learners who navigate the often unsteady passage from ignorance to acceptance and even comprehension of established knowledge.

3 Making a Hard Case

In Comte's hierarchical philosophy of knowledge that was one of sociology's nineteenth-century points of origin, sociology was the hard case for mathematics, the pinnacle of a system of ever-complexifying structure based on mathematics and verging toward comprehending everything. Over the twentieth century, this relationship inverted. A significant scholarly tradition came to construe mathematics as a

hard case for sociology, perhaps the quintessential hard case for the sociology of knowledge.

Comte called his philosophy Positivism, referring to the combination of empirical facts and reasoning on which he claimed rested all positive knowledge, as opposed to lesser theological or metaphysical knowledge. In the early twentieth century, a critical confluence of philosophical interest in the linguistic foundations of both empirical description and logical reasoning gave rise to a different program in the philosophy of knowledge, also sometimes called positivism. Both positivisms were projects of unification and ordering, explaining how the modern sciences related to each other and to the nature of human thought.

Of their many differences, the most significant for understanding the inversion of mathematics and sociology as hard cases may be the respective positivisms' account of where mathematics comes from. For Comte, mathematics emerged as the first positive science from the problem of indirectly measuring magnitudes, based on the empirical and rational study of magnitudes' relations. For the later positivists, the key relations were between symbols, not magnitudes, and mathematics emerged as a semiotic or representational science built from the logical order and interpretation of written signifiers. This semiotic reorientation placed an emphasis on understanding and on mathematics as an achievement of rational representation (if not necessarily communication, inasmuch as one comprehending mind sufficed).

When mathematics was a science of relations among worldly magnitudes, cognition and comprehension were importantly peripheral to the underlying stuff of the subject. As mathematics came to be located in symbolic representations, language and meaning became more central to characterizing the human production and understanding of mathematics, if not what it was exactly that was being symbolically understood. Language and meaning are social things. The scholarly possibility of sociology of mathematics (in its form as a sociology of knowledge) came from a refocusing of the epistemology of mathematics around individual and collective linguistic representation. The enduring notion of the underlying truth or reality of mathematics as existing outside of language, in worldly relations or in the basic logical forms of knowledge itself, meant that such sociologies of mathematics were founded in a paradoxical project of characterizing the social conditions of knowing something intrinsically independent of the social world.

Ludwig Wittgenstein has come to represent the centrifugal dynamic of mathematics as language in the first half of the twentieth century. It is common to divide his career into an "early" and "late" period. The former, defined by his immersion in the symbolic architectures and paradoxes of turn-of-the-century logicism and represented by his *Tractatus Logico-Philosophicus* (1922), epitomized logical mathematical knowledge as the ultimate product of an orderly propositional system. The latter, shaped by the tumult of interwar philosophy and represented in his *Philosophical Investigations* (1953) and *Remarks on the Foundations of Mathematics* (1956), emphasized the conventionality of linguistic understanding, including the linguistic understandings that underwrite knowledge and agreement in mathematics.

Wittgenstein was no sociologist and a very unconventional philosopher, but his provocative writings, mostly published posthumously, have influenced sociologists

and nucleated major debates about the linguistic and cognitive bases of foundational knowledge in social contexts. The later Wittgenstein's emphasis on contextual and informal understanding and the provisional nature of mathematical agreement was inspirational to a postwar generation of scholars approaching mathematics from a multidisciplinary critical perspective that blended psychology, anthropology, social theory, and other fields into critical reconsiderations of the foundations of knowledge. In Wittgenstein's posthumous aphorisms, read through these approaches, the most banal mathematics – counting in a sequence, multiplying numbers – pulsed with indecision and shed their veneer of inevitability.

One of the most enthusiastic and controversial such re-readers of Wittgenstein was David Bloor, one of the formulators of the strong program in the sociology of knowledge, part of a cluster of constructivist sociologies that came to be known as the Edinburgh School of the sociology of science and technology. The strong program aimed to take a naturalistic view of knowledge, studying it empirically through the terms and manifestations observable wherever knowledge claims are made and collectively challenged or respected (Bloor 1976, pp. 2–3, see also Bloor et al. 1996). Regarding mathematics, the strong program contrasted with a "weak program" of identifying a diversity of mathematical cultures and traditions, seeking sociological explanations for errors or failures in mathematics and its applications but not successes, or otherwise granting the primacy of mathematical reasoning in accounting for mathematical knowledge (Restivo 1993, p. 252). Adopting mathematics as the hard case with which to demonstrate the reach and potential of his program's claims and ambitions, Bloor latched onto Wittgenstein's illustrations of the linguistic embedding of confidence and consensus in mathematics. For Bloor, Wittgenstein demonstrated that mathematics was derived through naturalistically observable social processes of navigating and managing irreducible or undecidable conventions.

These conventions – rules and institutions, in the terms of a later (1997) Bloor book title – were not just social but *social like everything else* (see also Bloor 1983). In Bloor's heterodox interpretation, there was ultimately nothing intrinsically mathematical about the language-mediated determination of right and wrong in mathematics. That "ultimately" is important and a key source of misinterpretation of Bloor's program: this analysis focused on extremal moments of indecision in mathematics where no mathematical principle *in itself* uniquely determined the next step or inference. Asserting the continuity and primacy of social institutions in the fundamental decisions of mathematics, science, and *everything else* was Bloor's answer to postwar sociologists of science who focused on identifying the distinctive (and distinctively virtuous, rigorous, and truth-producing) features of science as a realm apart (Kaiser 1998). As Kaiser (1998) shows, this shaped Bloor's interpretations of earlier sociologies of knowledge (e.g. Bloor 1973), such as the interwar work of Karl Mannheim (e.g., 1979 [1936]), to suggest misleadingly that they assumed an intrinsic difference between science and other forms of knowledge and placed science out of scope for their sociologies.

This posture of framing sociological accounts of mathematics against a backdrop of ignorance, indifference, or presumptive impossibility is itself a long-running

tradition in sociologies of mathematics. Where most sociologies of modern societies could be said to have encountered situations of demystification, mathematics appeared to some to be becoming increasingly mystified. For François and Van Bendegem, "the more we have become dependent upon mathematics, mostly in an indirect or invisible way, the less we are actually understanding its principles" (2010, p. 122). The sociology of mathematics has been called "neglected" (Startup 1983) and "severely underrepresented" (Löwe et al. 2010), characterized by "a queer mixture of devotion and lack of interest" (Heintz 2000, p. 9; noted in Löwe et al. 2010). Mathematics as a subject is, for François (2007), "untouchable and frightening" for the general public, as well as for social scientists, as a result of an alienating technical norm in global curricula. This view of mathematics as marginal to sociology (of knowledge) comes in tandem with the construction of mathematics as a limiting form of knowledge, an extreme that tests what sociologies of knowledge can say. Restivo (1985 [1983], p. 161) called it "the ultimate arbiter of the limits of the sociology of knowledge" and noted the omission of mathematics from a popular 1970s collection on the sociology of knowledge (p. 165). For Bloor it was "the most stubborn obstacle . . . the holy of holies" (Bloor 1976, p. 73; see Barany 2017). For Heintz (2000, p. 10), the "epistemological specificities" of mathematics mean that it is not certain whether the content of mathematics can be studied through constructivist sociology.

Proposing to question the most basic and taken-for-granted elements of mathematics has put these sociologists squarely in the company of foundation-minded mathematicians themselves (see Barany 2023b, pp. 53–56). Going back to basics, revisiting axioms, and rejustifying the seemingly obvious are time-honored narrative and mathematical techniques in both pedagogy and research. At the turn of the twentieth century, Russell and Whitehead famously and at great and inscrutable length rederived the "occasionally useful" proposition that $1 + 1 = 2$. In the middle of the century, Bourbaki presented the number 1 as the product of a symbolic calculus far too tedious and complex to write out in full. The code that makes numbers work in recent computer proof systems contains multitudes, often far beneath the surface of programming interfaces that promise numerical deductions will work as expected.

This tradition is highly misleading, in similar ways, for mathematicians and sociologists alike. Most mathematics looks very little like $1 + 1 = 2$, even for elementary learners. The claim that "*even this* can be understood sociologically" or "*even this* can and must be developed from more fundamental logical principles" obscures how unrepresentative *this* is for most mathematical knowledge and practice. (Barany has argued elsewhere, 2014, that in other contexts such a focus on *the very most elementary* reckoning has derived from and contributed to scientific racism.) As a result, sociologies of mathematics have tended to share many of the biases and fixations of foundational philosophies of mathematics, albeit typically from a very different angle.

4 Hard Cases and Soft Targets

Bloor's strong program unsettled and provoked interlocutors and critics from a whole landscape of fields from anthropology and sociology to philosophy and mathematics. Most provocative was the strong program's principle of symmetry, that the same social causes should explain true and false ideas and theories in science (see esp. Bloor 1976). If truth and falsehood were the products of social processes of investigation and consensus-making, not predetermined manifestations of natural reality, there should be no sociological difference between the success of theories now believed true or now believed false: that they were respectively believed in their own times in their own social contexts made them sociologically equivalent.

The motivation for this approach was deeply empiricist, following a scientistic commitment to the process of deriving knowledge from rigorous observation undiluted, to the extent possible, by metaphysical presuppositions. Realist truth claims were, for Bloor, metaphysical distractions from the empirical question of how people came to agree about what they agreed about the natural or mathematical world. Such questions could and should only be answered by the observation and analysis of social facts. These, Bloor joined many sociologists of knowledge in supposing, should be most visible in episodes of controversy where advocates of rival knowledges had to put their reasoning and argument in the open (see, e.g., Ernest 2023).

Bloor's early articulations of the strong program included an extensive revisionist reading of Imre Lakatos's influential heuristic history of Euler's theorem about polyhedra, presented in the posthumous book *Proofs and Refutations* (1976, see Barany 2017; cf. Lakatos 1980). Lakatos's account was a prime example of the indeterminacy and revisability of less elementary mathematics in Bloor's (1976) *Knowledge and Social Imagery* and the centerpiece of a speculative and provocative 1978 article mapping Lakatos's historical examples of mathematical responses to exceptions, counterexamples, and other troubling phenomena to Mary Douglas's anthropology of purity and pollution. For Bloor, Lakatos's classification of responses to his titular refutations corresponded to Douglas's (1978) grid-group scheme for understanding how members of different social groups respond to disruptions to social order based on their degree of hierarchy and their concentration or mobility in or between social groups. This analysis exemplified the strong program's identification of classification as a central and constitutive activity of science, befitting Douglas's anthropological interest in the maintenance of categories as a central facet and manifestation of social order.

An appendix to Bloor's 1978 article took Lakatos's posthumous editors, John Worrall and Elie Zahar, to task for adding explanatory footnotes that toned down what Bloor saw as the more radical relativist implications of Lakatos's analysis. Worrall replied at length in the next issue of the journal, to which another shorter back and forth was appended (Worrall 1979). For Bloor, Lakatos's account showed the interminable revisability of mathematical concepts as well as mathematical proofs. For Worrall, Lakatos demonstrated that revisability was important in mathematics up to a point at which reasoning could be reduced to certain and rigorous

logical deduction. This, to Bloor, exemplified just the sort of response to a troubling phenomenon that had been analyzed with Douglas's scheme in Bloor's article, with Worrall answering a perceived threat to the integrity of logical deduction by reclassifying disagreements about proofs as unthreatening revisions to theorems (Bloor in Worrall 1979, p. 79).

The notion – preposterous to Bloor but dreadfully serious to a number of mathematicians, logicians, and mathematical scientists – that relativist and social constructivist accounts of mathematics were a threat to logic and mathematics itself (or at least to broader support for and trust in these) became a significant component of the so-called Science Wars of the 1990s (see Harris 1999). To a degree, social constructivist challenges to the obviousness or inevitability of extremal elementary mathematics invited the criticism that they were simply (perhaps ignorantly or perhaps malignly) attempting to deny the obvious. Mathematical physicist Alan Sokal gained notoriety by publishing a hoax article in the journal *Social Text*, which he claimed had demonstrated a willingness of cultural critics to endorse any anti-scientific nonsense. Cultural critics, including historians and sociologists of science, took the occasion to historicize and sociologize the animus of Sokal and his supporters, as well as their own disciplines, much as Bloor had done in reply to Worrall (e.g., Scharping 2001).

Bloor's commitment to controversies as stages for displaying the assumptions and stakes of scientific reasoning applied reflexively to his own scholarly practice, which featured exemplary exchanges with theorists who, unlike Worrall or Sokal, shared many of his relativist and constructivist principles and motivations. A 1992 back and forth with ethnomethodologist Mike Lynch in an edited volume dedicated in part to staging methodological debates in the sociology of science drew out the importance of social structure and authority for Bloor's understanding of agreement and closure in mathematics (Lynch 1992; Bloor 1992). The authors revisited a parable from Wittgenstein's *Philosophical Investigations* exploring the indeterminacy of rules of counting. A pupil counts successfully by twos until 1000, then continues the series 1004, 1008, and 1012, and is reprimanded for misunderstanding.

For Bloor, the pupil has shown that an infinite variety of logical rules would have been consistent with the sequence up to 1000, including an infinite number that would have the sequence continue as the pupil did. The correct interpretation was underdetermined and is settled only by an exercise of social authority over the pupil. For Lynch, by contrast, the lesson is that rules are not reducible to their formal logical representations. The pupil has indeed erred, and this shows not that the rule was underdetermined but that there was more to the rule than what Bloor's account supposed. Following a mathematical rule is not just following (or obeying) a logical dictate; it is recognizably conforming to a pattern of observable practice. Bloor returned to the problem of rule-following at several later points, including a rich 2004 debate in *Social Studies of Science* initiated by philosopher Martin Kusch that directly revisited the 1992 arguments (Kusch 2004; Bloor 2004; Kusch opens the exchange with a review of the rule-following problem in the philosophy of mathematics).

A 1999 exchange of journal articles with Bruno Latour, although not so directly concerned with mathematical knowledge, developed another dimension of Bloor's

commitment to social explanation (Bloor 1999a, b; Latour 1999). Latour had famously framed one of his most significant interventions in the sociology of agency and agreement in science as a generalization of Bloor's principle of symmetry. Where Bloor's symmetry principle emphasized the ultimate status of social causes to account for the stabilization of claims about nature, Latour argued that this principle of symmetry must be extended to account for the very opposition between social and natural causes. They thus differed over whether (per Bloor) symmetry was a matter of prioritizing human agency in the adjudication of claims about nature or (per Latour) a necessary next step to account for the constructed opposition of nature and society as part of the formation of scientific knowledge.

The stakes of symmetry weigh especially heavily in strong program interpretations of mathematics. For Worrall, for example, finite logical deductions are precisely the indubitable ground truths that cannot be denied (almost by definition) by any rational person. The formal underpinning of mathematical rigor may appear to be outside the nature-society binary that concerned both Bloor and Latour in their own ways, and little in their debate would seem to contradict this impression. Latour's intervention explicitly invoked the Science Wars context of their disagreement and the difficulty of articulating nature's role (such as it is) in a social theory of scientific knowledge. In Bloor's view, centering social explanation to the exclusion of nature was not to deny nature's relevance to scientific knowledge but to assert that, in the end, responsibility for scientific knowledge was exclusively human.

To insist on human, collective, social responsibility for the ground truths of mathematics was indeed a radical and challenging position. However, as the next section discusses, it was a position very much in keeping with a robust tradition of analysis of the social contexts and implications of mathematics, perhaps too easily missed in readings of strong program sociology outside of its formative intellectual and political contexts. Bloor built on a tradition of social study of mathematics that was based foremost in social critique rather than critique of mathematics as such, and the primacy of social responsibility was accordingly all the more visible.

5 Roots and Radicals

Restivo (1982, 1985 [1983]) identified the formative role for sociologies of mathematics of a train of Marxist interest in the social and material underpinnings of apparently idealist mathematical knowledge. These date to Boris Hessen's famous 1931 intervention as part of the exceptional Soviet delegation to that year's International Congress of the History of Science, in London, transgressively proposing a Marxist analysis of the hosts' national hero Isaac Newton and his *Principia Mathematica*. Varying forms of Marxist philosophy and politics were important commitments for Dirk Struik (e.g., 1942, 1986) and Joseph Needham (e.g., 1956, 1959), among those Restivo identifies as forerunners of the later sociology of mathematics. Restivo also includes in this list Ludwig Wittgenstein, whose later philosophy of mathematics has been argued to have been influenced by his engagement with Marxist thinkers (see Kitching and Pleasants 2002). This was not confined to the First and Second World of the Cold War; we might add Indian Marxist

mathematician and historian Damodar D. Kosambi as a contemporary deeply invested in materialist historiography.

At one level, the significance of Marxism to the social study of mathematics derives from the philosophical commitment to understand the development of ideas and production of knowledge in social, economic, and material terms, as part of the global capitalist system. But, especially for the postwar contributors to the Marxist tradition of social studies of mathematics, the dimensions of societal direction of and social responsibility for mathematics motivated inquiry into how mathematical knowledge responded to social contexts and conditions. Such "social and economic connections" and their relevance to questions of responsibility and the role of mathematics in modern society were central motivations for later work that somewhat critically recognized the influence of Marxist perspectives (e.g., Bos and Mehrtens 1977, pp. 10–12).

Restivo (1985 [1983], see also 1993, 2001) embraced the strong program as a foundation for continued materialist investigation of mathematics. Having engaged in detail with existing Marxist sociology of mathematics that he argued had been insufficiently grounded in materialism, Restivo took sociological inspiration from Spengler (1926) whom he identified as the first sociologist of knowledge to have expressed a strongly constructivist view of mathematics as a human activity inherently defined by its social, cultural, and historical milieu (1985 [1983], p. 86). Restivo recognized the "sociological imperative" to understand mathematics as a social pursuit (1993, p. 249) and responded by laying down an ambitious program for a radical sociology of mathematics "in which all talk about mathematics is social talk." This built upon the Marxist premise that all science is a collective and socially embedded activity and drawing on essential social elements of scientific activity, such as language, other "material," and the scientists' existence within multiple layers of social institutions. Restivo shifted the focus from the individual mathematician to the "math world": a collective representation of the activities and institutions inhabited by mathematicians. Restivo's key argument was that mathematics is not manufactured by individual mathematicians alone but is a product of "math worlds." These "math worlds" are embodied by mathematical objects and are inseparable from the social history of their own construction. Mathematical objects, in turn, are collective representations, sensibilities, and worldviews – and not merely symbols.

This, for Restivo, showed the necessity of "entering math ethnographically" (1993, p. 249). Restivo's later work defends the sociology of mathematics from "misguided" fears of its relativism when in his view it is in fact realist (2001). Restivo brushes aside the "unreasonable effectiveness of mathematics" conundrum as something that in fact demonstrates the realism of his sociology of mathematics: he declares the success of mathematical applications in the "real world" unsurprising, since mathematics is "'borrowed' from the real world to begin with" (1985, pp. 180–181, see also Hamming 1980). Restivo elevates mathematics, along with logics and religion, to the status of a "battleground" (2008) similar in scope to the Copernican revolution, and highlights the challenges math-fueled technological progress can bring if not accompanied by sociological understanding of mathematics as a social act and a complex of social institutions (see also 2016, 2017).

The critical and transformative potential of social accounts of mathematics has been especially visible in mathematics education and in studies motivated by that field. Restivo and Bloor both offer social constructivist foundations for philosopher of mathematics education Paul Ernest's examination of learning, language, and ethics in mathematics (e.g., Ernest 1991, 1994, 1998, 2000, 2018; see also Ernest's chapter in this handbook; cf. Pinxten 1994). Leone Burton (e.g., 1995) outlines a feminist epistemology of mathematics and mathematics education. The visible and large-scale social contexts and consequences of mathematics education, especially primary mathematics education, have motivated social studies of race and marginalization in and beyond mathematics classrooms (e.g., Gutiérrez 2013, 2018; Martin 2019). Sociological critiques have addressed institutional inequalities and injustices based on race, class, and gender, in various settings: elitism in academic journals (e.g., Chang and Fu 2021), racist political regimes (e.g., Mehrtens 1987), Hunsicker and Rittberg's discussion of the social injustices caused by unequal dissemination of mathematical ideas (2022), or the unequal effects of mathematization in different applied contexts (e.g., Ziols and Kirchgasler 2021), to name a few examples.

Critiques of school mathematics and the particular eurocentric model on which many hegemonic curricula have been based have underwritten programs of investigation into other forms of mathematical learning and activity under the rubric of ethnomathematics (e.g., d'Ambrosio 1997, 2000). Emphasizing the diversity, contextual specificity, and mathematical substance of knowledge and practice across different places and by historically underrepresented and underprivileged people, ethnomathematical research has ranged across non-Western cultures as well as examples of activity in Western settings that has not traditionally been understood to be mathematical but can be interpreted from a mathematical perspective (e.g., Lave 1988).

Gellert and Jablonka (2007) approach mathematics sociologically by locating mathematical knowledge in power relations and social inequality. They link and problematize the concurrent trends of mathematization and dematematization in contemporary society and warn of the resulting political "threat to democracy because of a widening gap of mathematical knowledge between constructors and consumers." Williams and Wake (2007), employing a microsociological and ethnographic approach to inform their analysis, similarly find that mathematical knowledge in workplace division of labor is increasingly "black-boxed, with groups of people excluded or insulated from mathematical knowledge and power" (2007). François (2007) argues that ethnomathematics has the potential to play a transformative role in Western school curricula, help undo some of the social stratification perpetuated by existing mathematics education, and demystify the "untouchable and frightening" public perception of mathematics. Drawing on Alan J. Bishop's "M-m distinction," she argues that school curricula reproduce social stratification by restricting vocational pupils' exposure to "small m mathematics" ("a purely deductive science, using a technique-oriented curriculum" and basic competencies such as counting, designing, explaining, locating, measuring), while far fewer pupils have easy access to mathematics (the Western scientific discipline, including ideas of mathematical beauty and creativity).

6 Watching and Listening to Mathematicians

Sociologists of mathematics have long struggled to catch mathematics in the act. Even when one can directly observe people at times and in places where they are doing mathematics, seeing the mathematics being done as such can be vexingly difficult. Mathematical communication combines visible, audible, and other sensory qualities with elements of invisible cognition that can be necessary to make sense of the more observable aspects of such communication. Proxies, inferences, and reconstructions can make up for some of this inaccessibility while bringing their own compromises and challenges.

At one extreme end of the scale of observation, cognitive scientists Lakoff and Núñez (2000) have contributed to materialist understandings of mathematics by correlating specific mathematical concepts and intuitions to features of embodied cognition. Their universalist conclusions about undifferentiated masses of human beings, and their corresponding generalizations about mathematical knowledge, are by design and supposition un-sociological while sharing sociologists' interest in the tangible human production of mathematics and resisting idealist notions of mathematics as entirely separate from the human condition.

Drawing on ethnomethodological approaches to scientific knowledge (e.g., Livingston 1986), Greiffenhagen (sometimes in collaboration with Sharrock) has expanded the scale of observation from that of individual cognition to small settings of interpersonal communication including university classroom lessons and supervisor-student meetings (Greiffenhagen 2014; Greiffenhagen and Sharrock 2011, 2019). Barany and MacKenzie (2014) combined ethnographic observations with interviews to characterize the material production of mathematical knowledge at the scale of small research communities, identifying the situated and dynamic roles of particular forms of writing (on blackboards, scrap paper, and other media) for mobilizing particular kinds of understanding and communication. Kremakova (2016a, b) used ethnography and autoethnography to uncover links between mathematicians' careers, professional identities, and mathematical activities such as teaching, learning, collaboration, proving, publishing, supervising graduate students, and so on, locating these in the context of ongoing marketization of higher education and the precarious scientific labor market.

Methods of institutional sociology permit indirect characterizations of mathematics and mathematicians at larger scales. Early examples can be found in Hagstrom (1965), Fischer (1973), and Startup (2021 [1979]), who investigated the composition of mathematics and other scientific university departments and the social characteristics and working lives of academic mathematicians, reflecting on (and analyzing mathematicians' own reflections of) aspects of mathematical labor and practice along the way. More recently, combining surveys, publication data, and other methods, a large international collaboration to study gender gaps in science has produced large-scale systemic findings about the contemporary mathematics profession (e.g., Mihaljević and Santamaría 2020; Steinfeldt and Mihaljević 2023).

Sociologists of mathematics have also used variations on historical methods, focusing on discursive and contextual analyses of written records to watch

mathematicians indirectly by viewing the records of their written communication with each other. Such studies have ranged in scale from Merz and Knorr Cetina's (1997) examination of the email records of a single collaboration to Rosental's (2008) analysis of usenet discussion posts about a contested theorem in mathematical logic, to Martin and Pease's (2013) study of online mathematics interactions including question and answer forums, and to Steingart's (2012) examination of the paper trail of a large-scale and long-running project in group theory. Written records can, through their references and omissions, show aspects of unwritten practice and indicate how mathematicians rely on combinations of formalized and unformalized reasoning to make sense of concepts and arguments and to establish confidence or articulate its absence whether or not these make sense.

A vital source of knowledge about mathematical knowledge has been mathematicians themselves. In the course of their mathematical practice and careers, mathematicians necessarily develop personal and shared understandings – to varying degrees of explicitness and formality – of relationships between mathematical institutions and mathematical knowledge. Many core activities of mathematical work, from evaluating research proposals and results to assessing peers and students to navigating careers and organizations, all depend on sociological knowledge. Restivo (1993, 248) calls this "[t]he everyday folk sociology of mathematicians." Beyond their personal navigation of their professional worlds, some mathematicians pursue (with varying degrees of rigor and amateurism) social and humanistic studies of their subjects. These can all be significant sources of insight and understanding in the context of social theories and more systematic sociological analyses of mathematics (e.g., Lane et al. 2019) while being susceptible to the biases and limitations of their informants (see Hanna and Larvor 2020). Socially and politically engaged mathematicians have employed sociological theories in support of both theoretical and political arguments against institutional racism and other sources of inequality and discrimination in academic mathematics departments (see, e.g., H[Harron] 2017).

Many academic mathematicians engage in sociologizing their "math worlds" and turn to insights from anthropology, sociology, history, and philosophy of mathematical practice to make sense of the mathematical experience. Mathematicians have produced social commentaries on various aspects of the mathematical existence, work, career, biography, professional identity, teaching, the nature of mathematics, and the way it is done (see, e.g., Hardy 2019 [1940]; Wilder 1981 [1950]; Davis and Hersh 1981; Hersh 1997; De Millo et al. 1979; Thurston 1994; Lockhart 2009 [2002]; Harris 2015). Mathematicians' endogenous sociologies proliferate in global public channels such as blogs and public talks, steering public discussions in the mathematics community and beyond (Terence Tao, Timothy Gowers, Edward Frenkel, Marcus du Sautoy, Günter M. Ziegler, Piper H[Harron], to name a few). These voices, often in dialogue or open disagreement with one another, serve the dual purpose of being an ongoing participatory sociology of mathematics produced by "mathematical natives" in its own right and forming a metacorpus of data for nonmathematician sociologists of mathematics to interrogate and draw upon. These "native sociologies" also provide a key point of tangency for sociologists and

mathematicians at which mutual assumptions, biases, and ignorances become exposed, helping to develop a more realistic mutual understanding of what it is that mathematicians, and sociologists, actually *do*.

Extensive conversations with mathematicians formed the basis for Bettina Heintz's (2000) study of contingency and flexibility in the "inner world of mathematics," which sets mathematicians' perceptions of their principles and practices in the context of sociological theories including those of Bloor and Restivo. Lane et al. (2019) apply Ingold's notion of crafting as wayfaring to analyze mathematical metaphors of landscape, journey, and craft to understand the way in which mathematicians' work is represented and contrast the realities of mathematicians' work with the myth of the lonely genius. In Kremakova's (unpublished) sociology of labor and career patterns of mathematicians, personal experiences and perceptions of career trajectories among mathematicians aggregate into a systemic view of patterns of research and mobility with significant social and intellectual consequences. Kremakova's approach of analyzing what mathematicians do as scientific labor reveals tight links between what happens inside mathematics as a scientific discipline (such as the content of emerging mathematics, the formation of university curricula, or the choice of research questions and methods related to available funding), on the one hand, and the workings of scientific institutions and the conditions of the scientific labor market, on the other. Analyzing the mathematician as his or her own "nomological machine," Kremakova discusses mathematicians as embodied, material, and social beings, working with material (or quasi-material) artifacts and constantly negotiating tensions and conflicts between individual career paths and constraints within mathematical institutions.

Important constraints that define and anchor the mathematical experience in socioeconomic realities include the precarity of early career work and contracts, the need to reconcile personal life and caring responsibilities with the challenges of standardized academic career trajectories tailored to an idealized atomized and disembodied rational agent, and other challenges posed by the "new spirit of academic capitalism" (Kremakova 2016b). Narratives of inspiration and passion prevalent in mathematics, as well as the challenges of quantifying and demonstrating mathematical "outputs," make mathematicians (similarly to other academics) prone to exploitation and self-exploitation as workers embedded in contemporary capitalism. The "technocratic time" of contemporary neoliberal academia allows little room for "playful exploration" and insufficient "thinking time" required to do *new* science, while the current "managerial turn" in universities (characterized, inter alia, with increased reliance on productivity metrics for research assessment) produces strategic responses by mathematicians who, in order to remain successful in their field, are led to turn more into managers and less into artists or adventurous explorers. As a result shorter-term research programs become prioritized, pragmatic publication strategies are adopted, and "safe choices" (e.g., of which open problems to pursue or of which candidates to hire in mathematics jobs) are made. The precarious conditions and increasing internationalization of the job market for early career mathematicians are hard to combine with a viable (gendered) work-life balance and also contribute to the prioritization of short, "doable" research projects – or

even projects that have already been carried out but are as yet unpublished – as mathematicians without stable positions find themselves unable to commit to and carry out ambitious long-term research agendas. The (perceived and real) lack of stability and the pressure to publish fast and move countries for work create insurmountable barriers for mathematicians with caring responsibilities (disproportionately female), disabilities, lower socioeconomic backgrounds, and results in perpetuating multiple social inequalities within the mathematics community.

At the same time, mathematics "in the making" also brings moments of joy and other intense emotions. Mathematicians experience and talk about playfulness, creativity, pure curiosity, intellectual satisfaction, the addictive reward of sudden "a-ha moments" following long periods of hard toil, the skill of making mistakes without embarrassment, the pride in developing your own mastery and seeing your students surpass you, and the joy of collaboration and belonging to a shared community of practice and of values (both in person and across space and time). Mathematics is made up of a wide variety of social roles valued by the community. Mathematicians devise, maintain, and reproduce collective institutions, roles, and myths in service of their own community: examples include the "troll on the bridge" (the head of department who craftily uses university leadership to shield department members from excessive bureaucratic demands), funding councils and journals that legitimize and uphold the importance of various strands of mathematics against the tyranny of fashionable buzzwords, or outreach programs targeting marginalized groups both within research communities (conferences such as "Women in Numbers" and organizations such as the Network of Minorities in Mathematical Sciences) and among the general public.

7 Conclusion: Making and Breaking Order

Sociologies of mathematics aim to comprehend the collective achievement of order and agreement in mathematical knowledge and mathematical institutions. The orderly production of theorems and the orderly production of mathematicians to produce, store, and share them are, in their own ways, distinctive social problems addressed by common social forms in the modern world. Sociologies of mathematics matter because these mathematical orders are never as stable as they seem, and that instability can present risks and opportunities that run through and beyond mathematics as a discipline, subject, and form of knowledge.

We note and anticipate productive scholarship in these directions at the intersections of traditional and new subfields, such as the sociology and anthropology of work and labor, of gender, of social inequality and of education, higher education and academia, feminist sociologies of care and social (re)production, and the sociology of technology and computing.

Institutional understandings of mathematics are especially important in the midst of a long era where the uses of mathematics and the processes of producing that mathematics have each grown more complex and more consequential across greater segments of the social world. Algorithmic and informatic mathematics draws a great

deal of attention in this regard, but even seemingly esoteric, isolated, or pure areas of mathematics participate in a political economy and social arrangement of making, knowing, and sharing mathematics that conditions its higher-profile manifestations. The institutions of mathematics span modern societies and involve many interconnected and hierarchical relationships of knowledge and power with profound implications for what mathematics may do in the world, how, by whom, for whom, and to whom.

It therefore bears emphasizing, to an extent not emphasized in the foregoing analysis, that mathematics in its institutions *and* as a form of knowledge is today (and long has been) in many respects a tremendous locus of inequality and injustice, a resource for the empowered, and a burden for the marginalized. What we have, rather, emphasized in the foregoing is the extent to which concern for the social ramifications of mathematics has motivated and guided investigations into its social and material foundations and the potential for such investigations to revise what is taken for granted about what may and must be in the world of mathematics and from mathematics in the world.

8 Cross-References

- ▶ Mathematics Education Research on Mathematical Practice
- ▶ Polymath as an Epistemic Community
- ▶ Practice Makes Perfectoid
- ▶ Structural Semiotics as an Ontology of Mathematics
- ▶ The Epistemological Subject(s) of Mathematics
- ▶ The Ethics of Mathematical Practice
- ▶ The Social Constitution of Mathematical Knowledge: Objectivity, Semantics, and Axiomatics
- ▶ The Social Epistemology of Mathematical Proof
- ▶ What Mathematicians Do: Mathematics as Process and Creative Rationality

References

Barany MJ (2014) Savage numbers and the evolution of civilization in Victorian prehistory. Br J Hist Sci 47(2):239–255

Barany MJ (2017) Of Polyhedra and Pyjamas: Platonism and induction in meaning-Finitist mathematics. In: De Freitas E, Sinclair N, Coles A (eds) What is a mathematical concept? Cambridge University Press, Cambridge, pp 19–35

Barany MJ (2021) "A young man's game": youth, gender, play, and power in the personae of mid-twentieth-century global mathematics. In: Niskanen K, Barany MJ (eds) Gender, embodiment, and the history of the scholarly persona: incarnations and contestations. Springer, Cham, pp 21–53

Barany MJ (2022) Mathematical institutions and the 'In' of the association for women in Mathematics. In: Beery JL, Greenwald SJ, Kessel C (eds) Association for women in mathematics: the first fifty years. Springer, Cham. pp 325–341

Barany MJ (2023a) "Some call it Arsmetrike, and some Awgryme": misprision and precision in algorithmic thinking and learning in 1543 and beyond. In: Ames MG, Mazzotti M (eds) Algorithmic modernity. Oxford University Press, Oxford, pp 31–44
Barany MJ (2023b) On remediation: media, repair, and the discipline of fantasy in the theory and practice of algorithmic modernity. Osiris 38:40–57
Barany MJ, MacKenzie D (2014) Chalk: materials and concepts in mathematics research. In: Coopmans C, Lynch M, Vertesi J, Woolgar S (eds) Representation in scientific practice revisited. MIT Press, Cambridge, MA, pp 107–129
Belhoste B (2003) La Formation d'une technocratie: L'École polytechnique et ses élèves de la Révolution au Second Empire. Belin, Paris
Bloor D (1973) Wittgenstein and Mannheim on the sociology of mathematics. Stud Hist Phil Sci 4(2):173–191
Bloor D (1976) Knowledge and social imagery. Routledge, London
Bloor D (1978) Polyhedra and the abominations of Leviticus. Br J Hist Sci 11(3):245–272
Bloor D (1983) Wittgenstein: a social theory of knowledge. Macmillan and Columbia University Press, New York
Bloor D (1992) Left and right Wittgensteinians. In: Pickering A (ed) Science as practice and culture. University of Chicago, Chicago, pp 266–282
Bloor D (1997) Wittgenstein, rules and institutions. Routledge, London
Bloor D (1999a) Anti-Latour. Stud Hist Philos Sci 30(1):81–112
Bloor D (1999b) Reply to Bruno Latour. Stud Hist Philos Sci 30(1):131–136
Bloor D (2004) Institutions and rule-Scepticism: a reply to Martin Kusch. Soc Stud Sci 34(4): 593–601
Bloor D, Barnes B, Henry J (1996) Scientific knowledge: a sociological analysis. Athlone/Chicago University Press, London
Bos HJM, Mehrtens H (eds) (1977) The interactions of mathematics and society in history. Some exploratory remarks. Hist Math 4:7–30
Burton L (1995) Moving towards a feminist epistemology of mathematics. Educ Stud Math 28(3): 275–291
Chang HCH, Fu F (2021) Elitism in mathematics and inequality. Hum Soc Sci Commun 8(1):7
Comte A (1842) Cours de Philosophie Positive. Bachelier, Paris
D'Ambrosio U (1997) Ethnomathematics and its place in the history and pedagogy of mathematics. In: Powell AB, Frankenstein M (eds) Ethnomathematics: challenging Euro- centrism in mathematics education. State University of New York Press, Albany, pp 13–24
D'Ambrosio U (2000) Historiographical proposal for non-Western mathematics. In: Selin H (ed) Mathematics across cultures. The history of non-Western mathematics. Kluwer Academic Publishers, Dordrecht, pp 79–92
Davis PJ, Hersh R (1981) The mathematical experience. Birkhäuser, Boston
De Millo RA, Lipton RJ, Perlis AJ (1979) Social processes and proofs of theorems and programs. Commun ACM 22(5):271–280
Desrosières A (1998) The politics of large numbers: a history of statistical reasoning (trans: Naish C). Harvard University Press, Cambridge, MA
Douglas M (1978) Cultural bias. Royal Anthropological Institute, London
Ernest P (1991) The philosophy of mathematics education. The Falmer Press, London
Ernest P (ed) (1994) Constructing mathematical knowledge: epistemology and mathematics education (Studies in mathematics education). Routledge, London
Ernest P (1998) Social constructivism as a philosophy of mathematics. SUNY, Albany
Ernest P (2000) Why teach mathematics? In: White J, Bramall S (eds) Why learn Maths? London University Institute of Education, London. http://webdoc.sub.gwdg.de/edoc/e/pome/why.htm. Accessed 2023
Ernest P (ed) (2018) The philosophy of mathematics education today. ICME-13 monographs. Springer, Cham
Ernest P (2023) Rejection, disagreement, controversy and acceptance in mathematical practice: episodes in the social construction of infinity. Global Philos 33:15

Fischer CS (1973) Some social characteristics of mathematicians and their work. Am J Sociol 78(5): 1094–1118
François K (2007) The untouchable and frightening status of mathematics. In: François K, van Bendegem JP (eds) Philosophical dimensions in mathematics education. Springer, New York
François K, van Bendegem JP (2010) Ethnomathematics and the philosophy of mathematics (education). In: Löwe B, Müller T (eds) PhiMSAMP. Philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London, pp 121–154
Gellert U, Jablonka E (eds) (2007) Mathematisation – demathematisation: social, philosophical and educational ramifications. Sense Publishers, Rotterdam
Greiffenhagen C (2014) The materiality of mathematics: presenting mathematics at the blackboard. Br J Sociol 65(3):502–528
Greiffenhagen C, Sharrock W (2011) Does mathematics look certain in the front, but fallible in the back? Soc Stud Sci 41(6):839–866
Greiffenhagen C, Sharrock W (2019) Tensions in Garfinkel's ethnomethodological studies of work programme discussed through Livingston's studies of mathematics. Hum Stud 42(2):253–279
Gutiérrez R (2013) The sociopolitical turn in mathematics education. J Res Math Educ 44(1):37–68
Gutiérrez R (2018) Introduction: the need to Rehumanize mathematics. In: Goffney GR, Boston M (eds) Rehumanizing mathematics for black, indigenous, and Latinx students. Annual perspectives in mathematics education, vol 2018. National Council of Teachers of Mathematics, Reston
H[H]arron P (2017) Get out the way, Inclusion/Exclusion American Mathematical Society Blog. https://blogs.ams.org/inclusionexclusion/2017/05/11/get-out-the-way/. Accessed 2023
Hacking I (1990) The taming of chance. Cambridge University Press, Cambridge
Hagstrom WO (1965) The scientific community. Basic Books, New York
Hamming RW (1980) The unreasonable effectiveness of mathematics. Am Math Mon 87(2):81–90
Hanna G, Larvor B (2020) As Thurston says? On using quotations from famous mathematicians to make points about philosophy and education. ZDM 52:1137–1147
Hardy GH (2019 [1940]) A mathematician's apology. Annotated by Cain AJ (2019) https://archive.org/details/hardy_annotated/. Accessed 2023
Harris M (1999) I know what you mean! https://webusers.imj-prg.fr/~michael.harris/Iknow.pdf. Accessed 2023
Harris M (2015) Mathematics without apologies: portrait of a problematic vocation. Princeton University Press, Princeton
Heintz B (2000) Die Innenwelt der Mathematik: Zur Kultur und Praxis einer beweisenden Disziplin. Springer, Berlin
Hersh R (1997) What is mathematics, really? vol 6. Oxford University Press, Oxford
Hersh R, Davis PJ (1981) The mathematical experience. Birkhäuser, Basel
Hessen B (1931) The socio-economic roots of Newton's Principia. Paper presented at Second International Congress of the History of Science, London
Hunsicker E, Rittberg CJ (2022) On the epistemological relevance of social power and justice in mathematics. Axiomathes 32(Suppl 3):1147–1168
Kaiser D (1998) A Mannheim for all seasons: Bloor, Merton, and the roots of the sociology of scientific knowledge. Sci Context 11(1):51–87
Kitching G, Pleasants N (eds) (2002) Marx and Wittgenstein: knowledge, morality and politics. Routledge, London
Kremakova M (2016a) Mathematicians against the clock: accelerated work and accelerated careers in the neoliberal university. LSE blog, 3/3/2016. https://blogs.lse.ac.uk/impactofsocialsciences/2016/03/03/mathematicians-against-the-clock-neoliberal-university/. Accessed 2023
Kremakova M (2016b) The "new spirit of academic capitalism": can scientists create generative critique from within? Teorie vědy/Theory Sci 38(1):27–51
Kusch M (2004) Rule-Scepticism and the sociology of scientific knowledge: the Bloor-Lynch debate revisited. Soc Stud Sci 34(4):571–591
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge

Lakatos I (author), Worrall J, Currie G (eds) (1980) Philosophical papers: vol 1: The methodology of scientific research programmes and vol 2: Mathematics, science and epistemology. Cambridge University Press, Cambridge
Lakoff G, Núñez RE (2000) Where mathematics comes from: how the embodied mind brings mathematics into being. Basic Books, New York
Lane L, Martin U, Murray-Rust D, Pease A, Tanswell F (2019) Journeys in mathematical landscapes: genius or craft? In: Hanna G, Reid D, de Villiers M (eds) Proof technology in mathematics research and teaching. Mathematics education in the digital era, vol 14. Springer, Cham
Latour B (1999) For David Bloor. . . and beyond: a reply to David Bloor's Anti-Latour. Stud Hist Philos Sci 30(1):113–129
Lave J (1988) Cognition in practice: mind, mathematics and culture in everyday life. Cambridge University Press, Cambridge
Livingston E (1986) Ethnomethodological foundations of mathematics. Routledge, London
Lockhart P (2009 [2002]) A Mathematician's lament: how school cheats us out of our most fascinating and imaginative art form. Bellevue Literary Press, New York
Löwe B, Müller T, Müller-Hill E (2010) Mathematical knowledge as a case study in empirical philosophy of mathematics. In: Van Kerkhove B, Van Bendegem JP, De Vuyst J (eds) Philosophical perspectives on mathematical practice. College Publications, London, pp 185–203
Lynch M (1992) Extending Wittgenstein: the pivotal move from epistemology to the sociology of science. In: Pickering A (ed) Science as practice and culture. University of Chicago Press, Chicago, pp 215–265
MacKenzie D (2001) Mechanizing proof. MIT Press, Cambridge, MA
Mannheim K (1979 [1936]) Ideology and utopia: an introduction to the sociology of knowledge. Routledge/Kegan Paul, London
Martin DB (2019) Equity, inclusion, and antiblackness in mathematics education. Race Ethn Educ 22(4):459–478
Martin U, Pease A (2013) Mathematical practice, crowdsourcing, and social machines. In: Carette J, Aspinall D, Lange C, Sojka P, Windsteiger W (eds) Intelligent computer mathematics. MKM, calculemus, DML, and systems and projects 2013, held as part of CICM 2013, Bath, UK, July 8–12, 2013, proceedings, vol 7961 of Lecture Notes in Computer Science. Springer, pp 98–119
Mehrtens H (1987) The social system of mathematics and National Socialism: a survey. Sociol Inq 57:159–187
Merz M, Knorr Cetina K (1997) Deconstruction in a 'Thinking Science': Theoretical Physicists at Work, Social Studies of Science 21(1):73–111
Mihaljević H, Santamaria L (2020) Authorship in top-ranked mathematical and physical journals: role of gender on self-perceptions and bibliographic evidence. Quant Sci Stud 1(4):1468–1492
Miller S (2019) Social institutions. Stanford Encyclopaedia of Philosophy (online). https://plato.stanford.edu/entries/social-institutions/. Accessed 2023
Needham J (1956) Science and civilization in China: volume 2, History of scientific thought. Cambridge University Press, Cambridge
Needham J (1959) Science and civilisation in China: volume 3, Mathematics and the sciences of the heavens and the earth. Cambridge University Press, Cambridge
Pinxten R (1994) Anthropology in the mathematics classroom? In: Lerman S (ed) Cultural perspectives on the mathematics classroom. Mathematics education library, vol 14. Springer, Dordrecht
Restivo S (1982) Mathematics and the sociology of knowledge. Knowl Creat Diff Util 4(1): 127–144
Restivo S (1985 [1983]) The sociology of mathematics. In: The social relations of physics, mysticism, and mathematics. Springer, Dordrecht, pp 161–175
Restivo S (1993) Conclusion. In: Restivo S, van Bendegem JP, Fischer R (eds) Math worlds: philosophical and social studies of mathematics and mathematics education. State University of New York Press, Albany
Restivo S (2001) Mathematics in society and history: sociological inquiries. Springer Science + business Media, Dordrecht

Restivo S (2008) Mathematics and science. In: Restivo S, Denton P (eds) Battleground science and technology. Greenwood, Westport/London, pp 278–286
Restivo S (2016) Red, black, and objective: science, sociology, and anarchism. Routledge, New York
Restivo S (2017) The social construction of mathematics. In: Sociology, science, and the end of philosophy: how society shapes brains, gods, Maths and logics. Palgrave Macmillan, London
Robson E (2008) Mathematics in ancient Iraq: a social history. Princeton University Press, Princeton
Rosental C, trans Porter C (2008) Weaving self-evidence: a sociology of logic. Princeton University Press, Princeton
Scharping M (2001) Wissenschaftsfeinde? 'Science Wars' und die Provokation der Wissenschaftsforschung. Westfälisches Dampfboot, Münster
Spengler O (1926) The decline of the west. Alfred A. Knopf, New York
Startup R (1983) Sociological approaches to mathematics. Int J Sociol Soc Policy 3(4):64–73
Startup R (2021 [1979]) The university teacher and his world: a sociological and educational study. Routledge, London
Steinfeldt C, Mihaljević H (2023) A machine learning approach to quantify gender bias in collaboration practices of mathematicians. Front Big Data 5:989469
Steingart A (2012) A group theory of group theory. Soc Stud Sci 42(2):185–213
Struik D (1942) On the sociology of mathematics. Sci Soc 6:58–70; reprinted in Schaaf WL (ed) (1948) Mathematics, our great heritage. Harper, York, pp 82–96
Struik D (1986) The sociology of mathematics revisited: a personal note. Sci Soc 50(3) Special 50th Anniversary Issue (Fall, 1986):280–299
Thurston WP (1994) On proof and progress in mathematics. Bull Am Math Soc 30(2):161–177
Wilder RL (1981 [1950]) Mathematics as a cultural system. Pergamon Press, Oxford
Williams J, Wake G (2007) Black boxes in workplace mathematics. Educ Stud Math 64:317–343
Wittgenstein L (1922) Tractatus logico-philosophicus (trans: Pears DF, McGuinness BF, 1974). Routledge/Kegan Paul, London
Wittgenstein L (1953) Philosophical investigations (trans: Anscombe GEM), 3rd edn. Blackwell, Oxford
Wittgenstein L (1956) Remarks on the foundations of mathematics (trans: Anscombe GEM). Blackwell, Oxford
Worrall J (1979) A reply to David Bloor. Br J Hist Sci 12(40):71–81
Ziols R, Kirchgasler K (2021) Health and pathology: a brief history of the biopolitics of US mathematics education. Educ Stud Math 108:1–20

Practice Makes Perfectoid

Pasts, Futures, and the Historiography of Contemporary Mathematics

Michael J. Barany

Contents

1 Prologue (with Apologies to Michael Harris) 2620
2 Making a Promising Theory 2621
3 History Repeating 2623
4 Time for Mathematics 2624
5 Temporalities in Counterpoint 2626
6 History and Contemporary Mathematics 2633
7 Cross-References 2635
References 2635

Abstract

Comparing my historical account of the early years of Laurent Schwartz's theory of distributions with number theorist Michael Harris's narrative of the early years of Peter Scholze's perfectoid theory, I develop a perspective on change and temporality in mathematics that emphasizes the relationships between concepts, expectations, and communities of practice. Contemporary mathematics, understood as mathematics imbued with temporality, reflects the dynamic relationship between the people, ideas, pasts, and prospects of mathematical knowledge. Studying these historically may offer critical perspectives on the social and political conditions and implications of mathematical research and the communities that practice it.

Keywords

Expectation · Horizon · Temporality · Confidence · Contemporary mathematics · Perfectoid theory · Distribution theory · Concepts · Fields Medal

M. J. Barany (✉)
Science, Technology & Innovation Studies, The University of Edinburgh, Edinburgh, UK
e-mail: michael@mbarany.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_83

1 Prologue (with Apologies to Michael Harris)

I first met number theorist Michael Harris when I was an undergraduate student, by way of a number of essays posted to his website. I had no notion at the time that Harris was a number theorist, but the essays made clear that he was a mathematician who was also a clever writer and an insightful reader of French critical theory, all things I too hoped to be. I was particularly inspired by Harris's (1999) remarks on the Science Wars, a cultural conflict of the 1990s that continued to reverberate in the academic contexts where I aspired to study both mathematics and the cultural theories that perhaps (allegedly, controversially) misunderstood or undermined it.

I first introduced myself to Harris some years later, in 2011, at one of the famous coffee breaks at the Institute for Advanced Study in Princeton, New Jersey. I was, by then, a new graduate student in history of science at the neighboring university, part way toward abandoning my ambition to be a mathematician and somewhat less focused on the particular French cultural theories that most excited me as an undergraduate. I had no notion that at that very moment a veritable revolution was unfolding at the Princetitute, that Harris was among the number-theoretic eminences who were in the process of witnessing the spectacular arrival of the perfectoid.

The perfectoid concept was developed and introduced by Peter Scholze, a German mathematician born in the same year I was. Scholze and I began our respective life courses toward Harris in the latter part of the decade in which the precocious Harris "acquired charisma" as a recognized mathematical leader, ratified with a tenured professorship at age 27 (Harris 2015a, Chap. 2). Still a graduate student at the time of the perfectoid debut, Scholze completed his doctorate by the end of 2011 and attained a full professorship the subsequent year at age 24.

I first learned of Scholze from Harris two years after our fleeting 2011 introduction, during our first proper conversation in Harris's Paris office. I had come to Harris's mathematics institute to work with the institute's renowned historians of modern mathematics on the planned subject of my doctoral research, the theory of distributions. Introduced by French mathematician Laurent Schwartz in the latter half of the 1940s, the theory of distributions was an exciting revision to the foundations of mathematical analysis and partial differential equations, earning Schwartz one of the second ever pair of Fields Medals in 1950.

Nearing the end of my Paris research, I asked Harris for an informal meeting. As luck would have it, Harris had just two days prior read a preprint of my report of my sociological research into the material dimensions of mathematical seminars and research practices, from the year prior to my arrival in Princeton (eventually published as Barany and MacKenzie 2014). We had a lot to talk about. I do not remember how the topic came up, but (as can happen in elite mathematics institutes) the conversation turned at one point to future Fields Medalists. I offered a naive guess, which Harris disabused before sharing with great confidence that, sooner or later, the most exciting young mathematician in the world would surely win the coveted prize. The mathematician, whose work Harris was following with great interest and was seeing his students' generation engaging with infectious zeal, was, of course, Scholze.

I caught up with Harris again at a 2014 conference in London (Larvor 2016), by which time I had some thesis results to share (Barany 2016) and Harris could share intimations of his forthcoming monograph, *Mathematics Without Apologies* (Harris 2015a). I was delighted to be invited the next year to review the book for *Mathematical Reviews* (Barany 2015b), my first book review for the venerable bibliographic enterprise whose history had just recently captured my interest as a significant element of mid-twentieth-century mathematics (see Barany 2021a). This occasioned a renewed email correspondence with Harris, during which he shared a draft of an essay on the perfectoid concept, which he presented as a "test case for an absent theory."

One thing led to another, and I found myself attempting to revise Harris's draft into a joint essay that would combine Harris's perspective on the emergence of the perfectoid concept with a sociological theory taken from my own work and my interpretation of Harris's recent book. The attempted collaboration fizzled, and Harris's chapter ultimately appeared in a version close to what Harris had generously shared initially (Harris 2017a). For me, the abortive collaboration was a tremendously fruitful chance to rethink my own understanding of what it means to debut an exciting new theory. As I continued to develop my account of the dramatic entry of distributions at mid-century, Harris had primed me to notice the striking convergences in the trajectories of distribution theory and perfectoid theory.

With apologies to the mathematician without apologies, I aim to share here an elaborated account of my theory for his test case, based on Harris's (2017a) narrative of perfectoid theory and my own collaborative account of the corresponding period of Schwartz's career (Barany et al. 2017; Barany 2018b). My theory proposes a social-historical perspective on contemporary and unfolding mathematics. I draw from historical theories of the narrative construction of temporality and from sociological theories of the dynamics of confidence and evaluation in mathematical communities. This represents an alternative to the rich and provocative account of the relationship between mathematics and history presented by Blåsjö (2021) in this handbook. The concerns of mathematicians and of critical social historians of mathematics have many significant interfaces, and the perspectives of the latter have much to offer the former.

2 Making a Promising Theory

Laurent Schwartz was born in 1915 into a world divided by a Great War and was brought up in a family of scientists and scholars. He came of age and embarked on advanced mathematical training as his world divided again in the Second World War, excelling in elite settings of French mathematics under mentors connected to some of the most forward-looking and internationally oriented currents of modern mathematical research. As an unestablished researcher in Grenoble, Paris, and Nancy, as the war wound to a close and reconstruction began, he developed the initial ideas that would become his theory of distributions, derived from recent developments in the theory of topological vector spaces.

Schwartz claimed that a radical shift in theoretical perspective could resolve a long-standing foundational concern about the meaning of differential equations involving non-differentiable functions, including functions that had become central in fields such as electrical engineering and quantum physics. He focused initially on potentially intricate technical problems, such as the role of convolution and Fourier duality in his theory, pursued in an intensive French context strongly shaped by the Bourbaki collaboration of pedagogical and foundational radicals. The prestigious 1946 Prix Peccot of the Collège de France gave him a platform in Paris to pursue theoretical questions and build a case and an audience for his theory's significance to modern mathematical analysis.

His theory was based on difficult ideas from recent mathematics, but their implications could be intuited and adopted into mathematical practices in a variety of registers without demanding special expertise in the subjects underpinning Schwartz's arguments. Indeed, for the first half-decade of the theory of distributions, most of the theory's systematic justification and exposition rested in a future textbook that was little more than a promissory note. Most users of the theory could realize its conceptual promise using familiar frames of calculation such as integration by parts, a basic calculus technique whose dramatic reinterpretation in Schwartz's theory demanded very little change to the representational or intuitive dimensions of reckoning about differential equations.

Several influential senior figures of international mathematics encountered Schwartz and his theory in the first years after World War 2, including Norbert Wiener, Marshall Stone, and (most importantly) Harald Bohr. The latter connection was cemented at a 1947 colloquium on harmonic analysis sponsored by the Rockefeller Foundation and Centre National de la Recherche Scientifique as part of an international effort to rebuild outward connections in French science. Bohr was particularly excited by the possibilities he saw in the theory as a bridge between the frontiers of modern pure mathematics and its applications, and he supported Schwartz in promoting the theory on those terms. With Bohr's help, Schwartz attracted a wide international interest between 1947 and 1949, including his first international and transoceanic speaking invitations.

Bohr's final intervention, shortly before his death, was to engineer Schwartz's recognition with a Fields Medal at the 1950 International Congress of Mathematicians, decided by a committee Bohr chaired and was not shy about guiding toward his preferred outcome (Barany 2018a). Sensing that his presentation to the 1950 congress might be his last major appearance on the international stage of mathematics, Bohr presented Schwartz's theory as the future of the discipline. He argued that distributions had already transformed basic calculus through the modernizing march of generalization in a way that he predicted would be central to future mathematics curricula.

Things did not go quite as Bohr predicted, but the promise of distributions made Schwartz a sought-after international visitor and interlocutor and helped him secure a stable platform in France on which to build an influential career that did not ultimately depend on the theoretical fruitfulness of early claims about distributions. Mathematicians on multiple continents believed they could build successful careers by engaging, adapting, and teaching the theory of distributions, and they acted on

that belief to create an efflorescence of expository texts and proposals for the theory's potential uses and implications. The theory did indeed become foundational in some branches of advanced mathematics, but not necessarily in a way that drove future research programs. Building on his recognition, international travels and connections, and scientific credibility, Schwartz became a public leader in mathematical research, science and education policy, and antiwar and human rights activism (see Journoud 2013; Coeuré 2022).

3 History Repeating

Peter Scholze was born in 1987 into a world divided by a Cold War and was brought up in a family of scientists and scholars. He came of age and embarked on advanced mathematical training in a world still geopolitically defined by war and by global (financial) crisis, albeit not of the sort that surrounded Schwartz, excelling in elite settings of German mathematics under conditions that similarly prepared him for an international orientation to widely recognized mathematical challenges. As a student in Bonn, he developed his perfectoid theory based on recent developments in the field of arithmetic geometry.

Like Schwartz, Scholze debuted his theory with a claim that it resolved a decades-old problem, in his case a special case of the weight-monodromy conjecture. That claim, made without a full demonstration at the Institute for Advanced Study in Princeton in 2011, was followed by a visit and series of six lectures to the Princetitute's Paris counterpart, the Institut des Hautes Études Scientifiques. The first public demonstration took place at these lectures at the IHES, a site of historical and symbolic importance for the problems in algebra, geometry, and number theory for which Scholze's approach appeared to offer a refreshing and spectacular breakthrough. In 2012, as Schwartz had in 1946, Scholze received the Prix Peccot the year after first presenting his theory to the public of mathematical elites.

Some things move quicker in the early 2010s than in the late 1940s. News, videos, and notes of Scholze's presentations were rapidly available across much of the world of arithmetic geometry research. Scholze travelled far and wide, published rapidly, and alone and in collaboration demonstrated a striking series of further uses of his theory. His virtuoso performances of proof and exposition convinced eager audiences, primed to spot the pitfalls of algebraic intuition and ready to be inspired by their circumvention in the resolution of previously stubborn constructions and deductions. As Harris puts it, the cascade of theoretical applications, together with the way the perfectoid concept made sense of relevant antecedent theories, made it clear that this was *the right* (Harris's emphasis) concept to bring together timely combinations of concepts and research agendas from arithmetic geometry. A professorship and associated institutional responsibilities accrued to Scholze much more swiftly than they had for Schwartz.

There were more prizes to win, and Scholze won quite a few of them. He declined one such prize for which Harris had nominated him, for reasons that he did not publicize but about which Harris sympathetically speculated (Harris 2015b; see also Harris 2014). (It will be significant below that the declined prize was called New

Horizons.) He did not win a Fields Medal right away, an effect of a major shift in the medal's criteria and perception since 1966 (Barany 2015a). Under the pre-1966 criteria, Scholze would have been unlikely to win at all. His University of Bonn professorial forebear Friedrich Hirzebruch was disqualified from the 1958 medal for having risen too fast in the discipline and consequently being too well recognized to qualify under the prevailing older criteria despite his young age (Barany 2018a).

In Harris's telling, Scholze rapidly nucleated a generation shift in his field, supplying new terms, representations, styles, and methods to young researchers who saw his difficult theory as a promising highway toward the vanguard of contemporary research. Graduate students organized study groups to dissect the flood of materials coming from a burgeoning research community. Harris's account recalls the exhilarating sense of motion from supervising a Masters thesis on perfectoid theory begun shortly after Scholze's IHES lectures. By following the theory as it spread and developed in those frenetic months, the student not only rapidly outpaced Harris's own sense of competence in the field in which Scholze intervened but could chase the theory as far as new results that were unimagined even to Scholze and his collaborators at the time Harris's student had started the project. It was hard to keep up, a difficulty it seems older generations felt more acutely, but those who managed to keep pace with the community of perfectoid researchers were quickly and thrillingly rewarded with access to a fast-moving forefront of mathematical theory.

Berkeley's Mathematical Sciences Research Institute hosted a staggeringly well-attended 2014 "Hot Topics" meeting on perfectoid spaces that brought together senior figures who saw their work carried forward in perfectoid research with younger researchers keen to see where the theory could take them. By then, Harris was aware of competent and confident users of perfectoid theory who had not personally closely collaborated with Scholze. This milestone for the reproduction and sustainability of a perfectoid research community corresponded to developments in Scholze's own research trajectory toward *next* concepts and challenges that would go beyond probing the limits of his 2011 formulation.

I spotted Scholze in the queue to enter the opening ceremony of the 2018 International Congress of Mathematicians in Rio de Janeiro. He was conspicuously one of only a very few, especially among the younger attendees, to be wearing something approaching formal attire, as though prepared to be in the spotlight. His was among the least surprising Fields Medal awards in history. It was evident in Scholze's plenary address later in the Congress that the perfectoid concept as such had become just a part of a multifaceted theoretical framework and research program.

4 Time for Mathematics

Setting Schwartz next to Scholze prepares us to analyze the history of *contemporary mathematics.* By contemporary, I mean mathematics as it is being produced, shared, interrogated, and evaluated. It is mathematics understood to be "now," whenever "now" is. "Now" is a moving target, and its mathematics is therefore always

precariously poised between particular pasts and futures, which can alike be mobilized in contemporary mathematical practices. It is thus essential to understand such mathematics as being contemporary in another sense, not just *together in time* (the usual meaning of contemporary) but *imbued with time.* Contemporary mathematics is mathematics pulsing with temporality.

There are two basic ways for mathematics to be *in time*: static or dynamic. Static mathematics exists on a fixed timeline, with ideas and arguments and events each having their place on a linear continuum stretching from past to present to future. The core questions of static mathematics are ontological and epistemological: what things are, what they mean, what they signify or imply, how they relate to each other logically or conceptually or genealogically. These relations can be fixed to identifiable durations: prime numbers in the seventeenth century, epsilon and delta arguments in the 1860s, homotopy type theory pondered over breakfast last week. They can exist outside of duration: the eternal congruences of triangles, the principle of induction, the propositions on page 214 of a textbook. The stuff of static mathematics is separately defined by when it is, where it is, what it means, and so on.

Dynamic mathematics, by contrast, exists at the always-shifting join of specific pasts and futures. There is not a single timeline of dynamic mathematics, because for each person in each place at each moment there may be a different constellation of experiences, understandings, and expectations that shape what mathematical concepts and practices mean (cf. Barany 2017). In the formulation of conceptual historian Reinhart Koselleck (1985), historical time is not a fixed backdrop of human events but a constantly recreated condition generated by the interface of ever-changing experience and ever-retreating horizons of expectation. Rather than what a concept signifies, a dynamic analysis asks (with a subtle but crucial shift in perspective) what its significance is, how it brings together and makes relevant particular experiences and understandings of what the concept relates to, where it comes from, and what it suggests, inspires, or disrupts. The stuff of dynamic mathematics is always on the move, always reconfiguring a multiplicity of pasts and futures (cf. Epple 2011).

Static and dynamic perspectives give two distinct ways to explain the parallel successes of perfectoid and distribution theory. They draw attention to different facets of what matters in the foregoing narratives, with correspondingly different implications for the historical understanding of mathematical practice (cf. Barany 2020). While both kinds of accounts are retrospective, I claim that the retrospection of a static past misrepresents essential features of the contemporary development, adoption, and dispersion of new mathematical ideas identified through a dynamic conception of this history.

Historical and contemporary mathematicians, however, often frame their own conceptions of their past and present in static terms. Customs of narration and of genealogical and conceptual accounting in mathematics can directly obscure the dynamism and instability that is not just intrinsic to contemporary mathematics in the making but is necessary to explain how such mathematics moves and transforms. That is, dynamic practices produce the conditions for the epistemic closure of mathematical understanding in static terms.

5 Temporalities in Counterpoint

In homage to a narrative device from Harris (2015a), let us consider a series of oppositions posed by a fictional Number Theorist (NT) and Mathematics Historian (MH), categorically not suggested to represent the biographical number theorist and mathematics historian from this chapter's prologue. Specifically, the biographical number theorist Michael Harris deliberately avoided making the kinds of claims in his perfectoid narrative that I shall here attribute to the figure of NT, and he in fact often appears (to my reading at least) to take a perspective closer to MH where he does theorize elsewhere. The number theorist will venture static explanations of the early histories of perfectoid and distribution theory, based on a commonsense view of the sociology of mathematics articulated for optimal contrast with the ensuing counterpoint, not on an empirically grounded summary of views of any particular number theorist(s). On the importance of such commonsense endogenous sociologies among mathematicians to the ordinary practice of mathematics, see my chapter with Kremakova (forthcoming) in this handbook. The mathematics historian will then reframe NT's static explanations in dynamic terms so that we may consider their different implications for the histories in question.

Our first counterpoint concerns the relationship between mathematical concepts and mathematical communities.

NT: Communities of mathematicians develop mathematical concepts and theories.
MH: Mathematicians use concepts and theories to organize into communities.

Mathematical concepts and theories are, from either perspective, fundamentally communal, developed through exchange, contestation, interpretation, and other communicative activities. For concepts and theories in their infancy, the relationship between communities and concepts runs importantly in both directions: communities change concepts, but concepts are critical loci and resources for changing communities in turn. Concepts and theories can be exciting, inspirational, even *hot,* and these evaluations and designations matter because they transform the communities that engage those concepts and theories. I have in fact elsewhere suggested that, from a social-historical perspective, theories may be *defined* in such terms, as "social systems built around relationships of partial mutual understanding," such that the "theory of distributions" can be defined as a "social system built around relationships of partial mutual understanding of distributions" (Barany 2018b, p. 264). (The significance of the "partial mutual understanding" aspect is a major theme in the rest of this section.) The early formulations of perfectoid spaces, distributional derivatives, and other elements of Schwartz's and Scholze's theories were most historically significant, for reasons we shall explore, for how they could be used to nucleate and expand rapidly growing and highly variegated communities of interested mathematicians with the social, material, technical, and theoretical capacities and motivations to promote the respective theories in a variety of capacities.

This reorientation of the relationship between communities and concepts changes how to interpret the role of those whose names become famous alongside

their theories. Both Schwartz and Scholze gained forms of celebrity status with their theories, and their celebrity drew from *and contributed to* the success of their theoretical interventions. Major awards and prizes, likewise, responded to, reshaped, and advanced the mathematicians' theoretical agendas as well as their careers and public images.

NT: Awards and career milestones follow from, and recognize, important results.
MH: Awards and career milestones certify and reinforce the importance of existing results, making them more important for more people, thereby promoting further results and certifications of importance.

The celebrity faces of distribution and perfectoid theory had to convince people in positions to recognize their theories' potential significance, and both succeeded in doing so from early stages in their research and careers. Not only did they convince those with more social and epistemic authority in the discipline of their theories' potential, they were both strikingly successful in establishing that they were the best people to bring their theories to wider and wider mathematics communities. Generous, charismatic, persuasive, and credible, Schwartz and Scholze became recognizable as mathematical celebrities because they were easily legible as potential celebrities.

Prestigious lecture invitations, their respective Prix Peccot and later Fields Medals, and other accolades signaled that these were mathematicians not only worth watching but worth joining. Particularly with Scholze, not just awarded but anticipated recognitions added fuel to the blaze. Conversations like my early one with Harris, speculating about major prizes and future distinctions, lent excitement and momentum, a sense of a theory going somewhere. Both theories and both mathematicians began to receive formal recognitions well in advance of having definitively and widely established the results that would come to be seen as the most significant indications of their importance, and well before there could be said to be a stable consensus about what their interventions really meant for their areas of mathematics. Both mathematicians used the occasions of their honors and associated opportunities to develop their research conversations, introduce further ideas, and project toward the future directions and implications of work in progress.

I have documented how Bohr's enthusiastic and arguably mostly exaggerated impressions of the joint potential of Schwartz and distribution theory motivated Bohr to engineer invitations and recognitions to help that potential be realized (Barany 2017). The Fields Medal and Bohr's highly visible characterization of Schwartz and distributions at the 1950 International Congress of Mathematicians, in particular, decisively set a celebrity trajectory that helped Schwartz establish what I suggest is the first mathematical theory from a previously unrecognized mathematician to establish an interconnected and multi-continental research community in its first decade (Barany 2018b). Schwartz charted a path and benefited from nascent institutions that Scholze could traverse and benefit from on a much faster timescale half a century later, helping mathematicians the world over to identify their often-challenging interventions as worth knowing and pursuing.

Public fanfare combined with information infrastructures (preprints, travel, rumor mills, letters, seminars) to build research communities at a pace that far exceeded that of the kind of individual interpersonal interaction that is the usual focus of accounts of collaboration and sense-making in mathematics. To take off or break through requires a conviction of going somewhere. Prizes and careers are simultaneously indications and engines of *going somewhere*, enabling the crucial shifts in perception and scale for theories and names to become big.

NT: Inspiring leaders of mathematical research communities assume positions of influence and prestige that reflect the leadership they offer.
MH: Potential leaders exercise influence and accrue prestige by assuming and representing the roles and expectations of leadership.

The personal and prospective natures of celebrity and leadership in mathematics mean that the visible, recognizable, attributable identities of mathematicians shape their theories and their disciplinary communities. One cannot lead if nobody is there to follow, if whoever is there does not expect one at least potentially to lead. Credibility and leadership are self-reinforcing. Leaders who successfully occupy, justify, and satisfy the optimistic expectations others have of their leadership gain more opportunities and expectations to do the same. Those who do not start with such optimism or who do not fit templates or roles for expectations of leadership, conversely, can do many of the things that would otherwise be associated with innovation and leadership without provoking the loop of expectation and recognition that defines the meteoric careers of others (Rossiter 1993).

From early stages in their careers, both Schwartz and Scholze converted smaller platforms into larger platforms by stepping into and satisfying escalating expectations of mathematical charisma and leadership. Harris (2015a, p. 143) perhaps even understates the case when observing that by age 24 Scholze had become "one of the youngest elder statesmen ever" in his branch of mathematics. As articulate young men, comfortable traveling and confident in settings of academic demonstration and socialization, they launched their careers from a presumption of belonging and potential (see Barany 2021b). Their leadership asked of them qualities and actions they were prepared to exhibit. Both were inarguably exceptional communicators, theorists, and much else, but their success was critically conditioned on, in many other respects, being visibly unexceptional in elite mathematical communities.

Expectations of celebrity and leadership make a big difference to the trajectories of the careers and theories of celebrity leaders, but they also quite profoundly overstate the difference that those celebrity leaders personally make. Like Bohr before him (but thankfully without the same sense of mortal urgency), Harris and a number of comparably positioned established mathematicians followed their somewhat circumstantial encounters with perfectoid theory at its earliest comings-out with later personal and professional activities and commitments to embrace and support the theory and its leading proposer and spokesperson. These established figures, like Bohr and Bohr's generational correspondents, attended lectures, spoke in private and wrote in public, connected students and colleagues, nominated for

prizes, and otherwise treated perfectoid and distribution theory, and Scholze and Schwartz, as theories and people of which and whom to take note.

In public and in private, Harris has expressed quite strongly that he was marginal to the early trajectory and success of the theory and mathematician he found so exciting. He makes a repeated point in his essay (Harris 2017a) and elsewhere (e.g., Harris undated) to disclaim any deep understanding or expertise in perfectoid theory, emphasizing that he has watched and appreciated the rise of perfectoid theory to a significant extent as an outsider. Without wishing to suggest an equivalence in their respective historical roles relative to their theories of interest, I quite suspect Bohr would have felt something similar. Bohr, for his part, was never a recognized mathematical contributor to the theory of distributions as such and was perceived by Schwartz as a somewhat peripheral (though valued) early supporter. Though their roles and situations have many important differences, I think Harris and Bohr each challenge us to reconsider what role such (sometimes self-consciously) theoretically marginal disciplinary elites have in the pivotal formative stages of a fast-rising theory or career.

Bohr's more-than-circumstantial centrality to the history of distributions was a somewhat surprising finding for my collaborators and me based on detailed archival research across many countries and languages (Barany et al. 2017). Without having done the same investigation of perfectoid theory, it is clear enough to me that there are significant differences in precisely how mathematicians in positions of established authority encouraged the latter theory's success. To the extent the comparison is informative here, Harris should be understood as an avatar for the subset of mathematicians who, in parallel to Bohr, recognized and encouraged Scholze's early theory from positions of established professional authority while being peripheral to the core perfectoid research community. In particular, within that community Harris did not appear to play the most Bohr-like role, which might instead be speculatively assigned to Harris's collaborator, the highly decorated number theorist Richard Taylor (an impression Harris has shared with me in more recent correspondence). It was Taylor, for example, who as Distinguished Visiting Professor at the Institute for Advanced Study in 2011 facilitated Scholze's spectacular IAS perfectoid debut. To what extent and to whom further comparisons may apply is a question that may await and surprise future historians.

NT: Mathematical theories and concepts start with those who best understand them and take hold as more people come to understand them better. Informal and imprecise understandings help motivate and support mathematicians' acquisition of more rigorous understandings.

MH: Mathematical theories and concepts are always simultaneously understood in a variety of ways, and those with middling understanding can be crucial for the theory's early success. Informal understandings, even arguable misunderstandings, often quite adequately replace apparently rigorous understandings, and can be more effective than normatively rigorous accounts at producing and promoting meaningful mathematical knowledge.

The experience of core users and active researchers of a concept and theory only begins to explain its significance and success. Theories and concepts and associated careers necessarily depend on convincing non-experts at different degrees' remove from the active research community. These non-experts must act on partial, provisional, approximate, or proxy understandings to form their evaluations and shape their advocacy.

The unmistakable lesson of the histories of perfectoid and distribution theory is that these non-expert understandings work for the social and intellectual purposes they serve. Among other purposes, such understandings supply a means of organizing and evaluating mathematicians' labor, histories, and values, and these uses are not in general reducible to the technical details that supply the concepts' notional content. It takes an extraordinary amount of work to understand the ideas and methods that have come to be associated with perfectoid spaces and still more work to adopt those ideas and methods in one's own research. One need not, however, be a perfectoid researcher to be a perfectoid advocate, mentor, sponsor, and (in some settings at least) expositor. Nor need one be a perfectoid expert to appreciate the kinds and extent of difficulty it takes to become such an expert and to support and reward it accordingly.

Some distance from the active research community can in fact better position an advocate for the social and intellectual roles they play in a theory's success. That distance demands and poises mediating elites to find terms and perspectives that bridge the central ideas and insights of a theory to those who are not (yet or ever) active researchers or experts but who may become experts, interact with them, celebrate or support them, or otherwise be in positions of finding meaning and value in the theory. A small remove from the day-to-day difficulties and frustrations of research with a theory also usefully obscures a sense of its shortcomings, a general phenomenon in the sociology of technical systems that Donald MacKenzie (e.g., 1990, pp. 370–372) has called the certainty trough. Those with the most confidence in a technical system, whether a mathematical research program or a nuclear missile guidance device, are not those with the most or the least expertise. Rather, they tend to be those who know enough to appreciate and be excited by such systems without knowing enough to mitigate that optimistic picture with the quotidian grind of subtle limitations and promises not yet fulfilled.

Even closer to the beating thrum of active research, the multiplicity of meanings matters. Mathematical concepts and theories are always polysemic, always understood in a variety of ways by people attempting to comprehend, use, teach, apply, and reimagine them for mathematical and other purposes. Indeed, many mathematical environments, technologies, and practices appear to be maintained specifically for the purpose of mobilizing unstable and polysemic representations (Barany and MacKenzie 2014). This gives a sociological explanation for the importance of multiple forms and modalities of communication for theories in the making, including venues for localizing and particularizing activities and habits of "reading down" mathematics as a converse to the formalizing work of "writing up" statements and conclusions in stable and circulable forms. I have argued (Barany 2018b) that the proliferation of sometimes mutually conflicting theoretical understandings of the

theory of distributions in different research contexts and communities was necessary for its rapid intercontinental trajectory. No one mathematician, Schwartz (or Scholze) included, is ever the comprehensive and exclusive master of everything that makes their theories meaningful, suggestive, useful, valid, or otherwise meritorious. The imagined future uses of a theory, the features that are most important for its adoption and adaptation in growing communities of research and practice, are perhaps most susceptible to and dependent on the constitutive polysemy of dynamic mathematics.

NT: A theory's antecedents culminate in a new theory, and that culmination is evidence of the theory's promising future. Together, the antecedents and expected future explain (or at least support) its successful adoption. Citations acknowledge the dependence of new mathematics on the past.

MH: A theory's antecedents and futures are negotiated (and continually renegotiated) as part of the theory's formation and successful adoption, supporting and shaping what the theory means to those developing and embracing it. Citations identify instrumental pasts that situate new mathematics for the future.

In Harris's and Bohr's respective accountings, there is a striking resemblance in the extent to which Scholze and Schwartz established the promise of their theoretical interventions by successfully mobilizing relevant and suggestive pasts for their theories. In both stories, the most significant antecedents emerge and become canonical as part of processes of explaining and developing a new theory. This frequently happens, and did happen for both of these theories, in dialogue and with the active participation of experts and practitioners in the traditions and theories that become adopted as canonical forerunners and underpinnings of the new theory. While these can be understood in the moment as processes of *recognizing* and *appreciating* those relevant pasts, in a vital dynamic sense, the pasts are re-narrated and created as part of the stories and trajectories of the theories for which they become crucial justifications and resources. The pasts link theories to communities of users, suggest conceptual and social affinities, and promise theoretical futures in the form of projections along a train of development. They show current and potential users and appreciators that the theory is (in Harris's term) *the right* one for them, for their goals, values, and interests. Thus canonized, antecedents establish that the theory brings together the pasts that matter to its audiences and shows them a viable way forward.

For Schwartz, and amplified in Bohr's telling, the most significant pasts and antecedents for the theory of distributions could be quite indirectly related to the genealogical configuration of theories and problems from which the theory of distributions first emerged. The theory's early meanings and audiences were thus created in significant part by the addition of temporal and conceptual antecedents that made the new theory meaningful to early interlocutors. In particular, Schwartz grafted distributions onto a conceptual and temporal trajectory of symbolic calculi reaching back through the Dirac calculus of formal derivatives in mathematical

physics to the Heaviside calculus in electrical engineering. In this framing, distributions were a culmination of a half-century of symbolic reckoning that motivated the utility and possibility of generalizing the process of differentiation to discontinuous or otherwise non-differentiable functions. Bohr, for his part, enrolled Schwartz in an even longer history of generalization and conceptual closure that included the development and algebraic stabilization of number systems.

These arrangements of pasts for distributions and their associated implicit (and sometimes, especially for Bohr, explicit) futures directly shaped early audiences for distributions and their understandings of the theory. In France, Denmark, and Canada, Schwartz met with crucial early audiences who proved willing to recognize distributions as an ideal solution to a foundational problem about the use of partial differential equations at the interface of pure and applied mathematics. Important advocates following Bohr, such as George Temple in England, domesticated the early framing of the theory as part of a progression of symbolic calculi associated with English theorists. Temple presented his version of Schwartz's theory as an import that reconnected the essential insights and essential utility of distributions to a local tradition of analysis and practical generalization.

The communal and programmatic orientation of the work of establishing canonical antecedents to new theories can be seen in the historical uses of abstracting, bibliography, and citation in modern and recent mathematics (Barany 2021a), as well as in mathematicians' historiographical practice (Paumier and Aubin 2016). Schwartz actively used reviews and citations to establish conceptual and programmatic relationships between his distribution-theoretic work and other bodies of research that may or may not otherwise have had much of anything to do with distributions. The history of citing and acknowledging prior work and of using those citations in ongoing work is thick with the ongoing labor of establishing research communities, reconciling understandings, and demonstrating significance by drawing the past through present projects toward a promising future.

When Scholze introduced perfectoid spaces, he could call upon the skills, techniques, terminologies, and intuitions of the fields at whose crossroads he situated his concept. He could expect researchers in those fields to be prepared to take notice of his intervention, if only he could interest them. These fields gave Scholze a ready means of establishing his work's significance and calling attention to his concept, by providing both a vocabulary for asserting relevance and a stable of open problems that could prove his concept's worth. Importantly, Scholze's early interlocutors saw themselves in the antecedents brought to bear on perfectoid theory and wrote the history of their own work, priorities, and engagements into the methodological and conceptual trajectories Scholze proposed.

By situating perfectoid spaces both genealogically and analogically in Grothendieck's algebraic geometry and Fontaine's constructions, Scholze supplied a ready basis for affirming their optimality (by analogy to affine and general schemes) and their fruitfulness (by extension of the fruitful genealogy of a tradition in arithmetic geometry). This promise of fruitful applications, based on successfully framing perfectoid spaces as the conceptually appropriate and most practically promising concept for a number of programs of research, helped Scholze recruit

(both directly and indirectly) a community of researchers who would begin to fulfill that promise and thereby reinforce the validity of the initial assessments. Despite the difficulty of perfectoid reckoning, its novelty meant that an enthusiastic young researcher could hope to study up to a meaningful research competency after a relatively short period of intensive study. With the goal of perfectoid geometry in view, such a student could take a short cut (if not a royal road) through only those aspects of the theory's background as seemed most relevant to a desired line of inquiry. Thus, Harris's Masters student could race in the space of a short project from ignorance of the theory almost directly into the community of perfectoid researchers.

Mathematics is, as Harris (2015a) put it, a problematic vocation. That is (in part), mathematics is organized around posing and solving problems (see Barany 2015b). Since fields of study are effectively defined by the programs of active research within them, by frameworks for articulating workable problems and meaningful solutions, the growing programmatic orientation of mathematical researchers toward perfectoid spaces as a unifying concept had the effect of embedding perfectoid geometry as a fundamental framework for a growing body of scholars. Above all, this reflected the self-fulfilling character of being *right*, worthwhile, or fruitful. Scholze's concept had to appear *right* long before its promise for arithmetic geometry began to be realized in great measure and in a wider community of scholars. He set this realization in motion by situating perfectoid spaces in a technical and communal framework that assured his audience of the ability to follow and appreciate his maneuvers and supplied a ready-made basis for establishing their significance. Casting backward to Hodge through Fontaine, Tate, and Grothendieck made it possible to cast forward to a perfectoid-theoretical future.

6 History and Contemporary Mathematics

I continue to treasure my sporadic exchanges with Harris, who continues to offer a distinctive and invaluable perspective on the contemporary discipline of mathematics. He was one of the first to see a draft of this essay. A recent theme of our conversations on which Harris has also had much to say in public[1] is the political economy of mathematical scholarship and its implications for the social and political responsibilities of mathematicians. Harris is a member of the Azat Miftakhov committee, a group modeled on Schwartz's Committee of Mathematicians to challenge human rights abuses and political imprisonment of mathematicians in our own time, focusing on Miftakhov's political imprisonment (ongoing as of this writing) in Russia. Scholze participated in the group's 2021 conference to raise awareness and show solidarity following Miftakhov's sentencing, an effort to convert some mathematical celebrity into political consciousness.

Harris has made a point of drawing attention to historical analyses, including some of my own, out of a conviction that mathematicians stand to benefit from

[1] See esp. Chap. 3 of Harris (2015a). See also Harris (2017b) and Harris and Jackson (2014).

awareness of the fruits of scholarly research on the social and political history of their field. Part of his exercise of professional and social responsibility is to use his standing as a mathematician and his profile as a cultural critic of mathematics to promote research that helps mathematicians critically interrogate their present and consider the implications of possible futures for the discipline. This kind of engagement and its political and intellectual motivations have a long tradition (see Barany and Kremakova forthcoming).

They are not, however, what Viktor Blåsjö (2021) refers to in this handbook with "Mathematicians used to be highly invested in the study of the history of their own field." Examining the purposes and uses of history "from the mathematician's point of view," Blåsjö advocates for the mathematical value of what he calls "critical internalist" histories of mathematics as a creative fount of mathematical understanding. Blåsjö sees a virtue in mathematicians' endeavor to use the values and methods of conceptual mathematical thinking to make sense of historical mathematics as a form of research experience, one that might build intuition or perspective on the conceptual dimensions of mathematics as a field.

I have suggested here that there are ways of thinking about the power of mathematical ideas that are good for producing mathematical ideas but bad for understanding mathematical communities and the different sorts of power that runs through them. In contrast to Blåsjö's critical internalist historiography, I think this is where a critical social historiography can and should speak to mathematics and mathematicians. An understanding of the inauguration and development of mathematical theories that focuses on static conceptual relations will necessarily miss the most significant dimensions of mathematical creativity and community. These dimensions, of a necessarily dynamic temporality, challenge familiar distinctions between internalist and externalist historiography. Community dynamics have tremendous conceptual consequences, and vice versa. Is hope internal? Is excitement? Is credibility? Is conviction?

From the perspective of the history of contemporary mathematics, of mathematics emphatically temporally in the making, theories and concepts must be understood as both products and tools of social organization, based on the temporalized production and mobilization of relevant pasts and promising horizons. Truth alone does not make a theory successful or even correct (at least in the broader sense I associate with Harris's invocation of being *right*) and is only a small part of its context. Insightfulness, significance, and value are negotiated in communities of practice.

These, in turn, give purchase on the critical potential and importance of the political economies and social contexts of contemporary research. A dynamic temporality of mathematics trains attention on the facets of mathematical knowledge work where privilege and inequality relate to the central features of mathematics as a system of knowledge and practice, and vice versa. Critical social histories can obscure or rationalize as well as identify or challenge the communal-conceptual dynamics of mathematical practice – we might signal the ambivalence of sociohistorical analyses by reference to the indecision of history narrated in the *past imperfectoid*, with apologies to Tony Judt's (1992) critical history of the political rationalizations and recriminations of Schwartz's generation. Critical potential is no

guarantee of critical purchase, but it is at least a way of interrupting the critical vacuity of static mathematical conceptions of mathematical change.

In his 1543 arithmetic primer, a long distant forerunner to the kind of arithmetic that made Scholze famous, Robert Recorde exhorted that "rules without practice, is but a light knowledge, and practice it is, that maketh men perfect and prompt in all things" (quoted and discussed in Barany 2023). The study of static mathematical rules without dynamic mathematical practice is indeed but a light knowledge. There is, rather, much to be gained from a historiography of contemporary mathematics in the sense sketched here, one in which practice makes not just perfect but also perfectoid.

7 Cross-References

► Fermat's Last Theorem
► What Mathematicians Do: Mathematics as Process and Creative Rationality

Acknowledgment Portions of this research and writing were supported by a UKRI Frontier Research grant (reference EP/X033961/1).

References

Barany MJ (2015a) The myth and the medal. Not Am Math Soc 62(1):15–20
Barany MJ (2015b) Review of Harris (2015). Math Rev, MathSciNet MR3289987
Barany MJ (2016) Remunerative combinatorics: mathematicians and their sponsors in the mid-twentieth century. In Larvor 2016:329–246
Barany MJ (2017) Of polyhedra and pyjamas: platonism and induction in meaning-finitist mathematics. In: De Freitas et al (eds) (2017), pp 19–35
Barany MJ (2018a) The Fields Medal should return to its roots. Nature 553(18 January):271–273
Barany MJ (2018b) Integration by parts: wordplay, abuses of language, and modern mathematical theory on the move. Hist Stud Nat Sci 48(3):259–299
Barany MJ (2020) Histories of mathematical practice: reconstruction, genealogy, and the unruly pasts of ruly knowledge. ZDM 52:1075–1086
Barany MJ (2021a) Abstract relations: bibliography and the infra-structures of modern mathematics. Synthese 198:6277–6290
Barany MJ (2021b) "A young man's game": youth, gender, play, and power in the personae of mid-twentieth century global mathematics. In: Niskanen and Barany (eds) Gender, embodiment, and the history of the scholarly persona: incarnations and contestations. Palgrave, pp 21–53
Barany MJ (2023) "Some call it Arsmetrike, and some Awgryme": misprision and precision in algorithmic thinking and learning in 1543 and beyond. In: Ames and Mazzotti (eds) Algorithmic modernity: mechanizing thought and action, 1500–2000. Oxford, pp 31–44
Barany MJ, Kremakova M (forthcoming) Inverting hierarchies: the sociology of mathematical practice. This handbook
Barany MJ, MacKenzie D (2014) Chalk: materials and concepts in mathematics research. In: Coopmans C, Lynch M, Vertesi J, Woolgar S (eds) Representation in scientific practice revisited. MIT, pp 107–129
Barany MJ, Paumier A-S, Lützen J (2017) From Nancy to Copenhagen to the world: the internationalization of Laurent Schwartz and his theory of distributions. Hist Math 44(4):367–394

Blåsjö V (2021) Historiography of mathematics from the mathematician's point of view. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer
Coeuré S (2022) East-West scientific collaboration face-to-face with dissidence: the logic of the International Committee of Mathematicians in defence of Plyushch and Shikhanovich. Cahiers du monde russe 63(1):103–122
De Freitas E, Sinclair N, and Coles A (eds) (2017) What is a mathematical concept? Cambridge
Epple M (2011) Between timelessness and historiality: on the dynamics of the epistemic objects of mathematics. Isis 102:481–493
Harris M (1999) I know what you mean! https://webusers.imj-prg.fr/~michael.harris/Iknow.pdf. Accessed 2023
Harris M (2014) Dispatch from the Oscars of science. Slate, November 19. https://slate.com/technology/2014/11/breakthrough-prize-ceremony-oscars-of-science-honor-mathematicians-life-scientists-physicists.html. Accessed 2023
Harris M (2015a) Mathematics without apologies: portrait of a problematic vocation. Princeton
Harris M (2015b) Working the red carpet, part 2. Mathematics Without Apologies blog. https://mathematicswithoutapologies.wordpress.com/2015/11/09/working-the-red-carpet-part-2/. Accessed 2023
Harris M (2017a) The perfectoid concept: test case for an absent theory. In: De Freitas et al (eds) (2017), pp 143–158
Harris M (2017b) Do mathematicians have responsibilities? In: Sriraman B (ed) Humanizing mathematics and its philosophy. Birkhäuser
Harris M (undated) The purpose of the article: a clarification. http://www.math.columbia.edu/~harris/otherarticles_files/Clarification.pdf. Accessed 2023
Harris M, Jackson A (2014) Mathematicians discuss the Snowden revelations. Not Am Math Soc 61(6):623–624
Journoud P (2013) Laurent Schwartz et le Vietnam: la « perte de l'innocence ». In: Durand M, Tazzioli (eds) Des mathématiciens et des guerres: Histoires de confrontations (XIX^{e}–XX^{e} siècle). CNRS Éditions
Judt T (1992) Past imperfect: French intellectuals, 1944–1956. University of California Press
Koselleck R (1985) "Space of experience" and "horizon of expectation": two historical categories. In: idem., Futures past: on the semantics of historical time, trans. Keith Tribe. Columbia, pp 255–275
Larvor B (ed) (2016) Mathematical cultures: the London meetings 2012–2014. Birkhauser
MacKenzie D (1990) Inventing accuracy: a historical sociology of nuclear missile guidance. MIT
Paumier A-S, Aubin D (2016) Polycephalic Euclid? Collective practices in Bourbaki's history of mathematics. In: Remmert VR, Schneider M, Kragh Sørensen H (eds) Historiography of mathematics in the 19th and 20th centuries. Birkhäuser, Basel, pp 185–218
Rossiter MW (1993) The ~~Matthew~~ Matilda Effect in science. Soc Stud Sci 23:325–341

Mathematics Education Research on Mathematical Practice

Keith Weber and Matthew Inglis

Contents

1 Introduction: The Relationship Between Mathematical Practice and Mathematics Education 2638
1.1 Authentic Mathematical Activity in the Mathematics Classroom 2638
1.2 Mathematics Educators' Need for Understanding Mathematical Practice 2639
2 The Relationship Between PMP and Mathematics Education Research on Mathematical Practice 2640
2.1 The Different Ways That Claims About Mathematical Practice Are Used in PMP and Mathematics Education 2640
2.2 How Mathematics Education Research on Mathematical Practice Can Contribute to PMP 2641
3 Research on Mathematicians' Understanding 2644
3.1 The Importance of Mathematical Understanding 2644
3.2 Open-Ended Interviews with Mathematicians About Their Understanding of Abstract Concepts 2645
3.3 Task-Based Interviews on How Mathematicians Develop an Understanding of Advanced Mathematics 2647
3.4 Discussion of Mathematics Education Research on Mathematicians' Understanding and Its Implications for PMP 2649
4 Mathematics Education Research on Mathematicians' Proof Reading 2650
4.1 The Importance of Proof Reading 2650
4.2 Hypothesis Generation Versus Hypothesis Testing 2650
4.3 Competent Performance in Proof Validation 2651
4.4 How Do Mathematicians Check Proofs for Correctness? 2653
4.5 Factors Mathematicians Consider When Assessing the Persuasiveness of Arguments 2656
4.6 Why Do Mathematicians Read Proofs? 2658

K. Weber
Rutgers University, New Brunswick, NJ, USA
e-mail: keith.weber@gse.rutgers.edu

M. Inglis (✉)
Department of Mathematics Education, Loughborough University, Loughborough, UK
e-mail: M.J.Inglis@lboro.ac.uk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_88

5 Concluding Remarks 2659
6 Cross-References 2660
References 2660

Abstract

In the mathematics education research literature, there is a growing body of scholarship on how mathematicians practice their craft. The purpose of this chapter is to survey some of this literature and explain how it can contribute to the philosophy of mathematical practice. We first describe how mathematics educators use empirical methodologies to investigate the behaviors of mathematicians and argue that findings from these studies can inform the philosophy of mathematical practice. We then illustrate this by summarizing research on mathematicians' understandings and mathematicians' proof reading.

Keywords

Mathematics education · Philosophy of mathematical practice · Proof · Proof-reading · Understanding

1 Introduction: The Relationship Between Mathematical Practice and Mathematics Education

Many mathematics educators believe that mathematicians' practices can and should inform the way that mathematics is taught to students. For this reason, mathematics educators regard empirical research into the psychology and sociology of professional mathematics as a legitimate topic for mathematics education research. In the last two decades, a number of research reports on mathematicians' practices have appeared in highly prestigious journals, on topics as diverse as the genre of writing in mathematics journals (Burton and Morgan 2000), how mathematicians choose research questions (Misfeldt and Johansen 2015), and whether mathematicians skim proofs before reading them closely (Inglis and Alcock 2013; Weber and Mejía-Ramos 2013). The purpose of this chapter is to survey some of the educational research on mathematical practice and to describe how this research can inform the wider literature on the philosophy of mathematical practice (from hereon, PMP).

1.1 Authentic Mathematical Activity in the Mathematics Classroom

A central reason that mathematics educators are interested in mathematical practice is because many believe that mathematics classroom communities should behave similarly to professional mathematical communities and that students should be engaged in activities that are reflective of professional mathematicians' activities.

The calls for classrooms to be reflective of disciplinary practices date back to Bruner (1960) who claimed that there should be "a continuity between what a scholar does on the forefront and what a child does in approaching it for the first time" (pp. 27–28) and Schwab (1978), who argued for a school curriculum "in which there is, from the start, a representation of the discipline" (p. 269). Applying this general theme to the specific subject of mathematics, Lampert (1992) wrote that "classroom discourse in authentic mathematics has to bounce back and forth between being authentic (that is, meaningful and important) to the immediate participants and being authentic in its reflection of a wider mathematical culture" (p. 310). In concert with Lampert, Schoenfeld (1992), Sfard (1998), and Rasmussen, Zandieh, Teppo, and King (2005) all argued that having students engage in authentic mathematical activity is a central goal of mathematics education.

Satisfying the calls from the previous paragraph requires an understanding of what it is that mathematicians do at the forefront of their discipline and what it would mean for a classroom community to be authentic in its reflection of mathematical practice. Following Stylianides et al. (in press), the situation can be framed as follows. Professional mathematicians engage in goal-directed activities. That is, mathematicians have goals (e.g., deciding whether a conjecture is true) and they engage in activities with the aim of satisfying those goals (e.g., seeing if the conjecture holds for particular examples or trying to prove the conjecture) (There can be a nested structure to mathematicians' goal-directed activities. For instance, they may engage in the activity of proving with the goals of demonstrating that a statement is true and understanding why it is true. But successfully engaging in proving (i.e., writing a proof) can thereby become a meaningful subgoal. They may then engage in the activity of reading the research literature to find techniques that will allow them to achieve their goal of writing a proof.). Mathematics classroom communities also engage in goal-directed activities. Activities can be said to satisfy Lampert's criterion of being "authentic in its reflection of a wider mathematical culture" if mathematicians have similar goals and engage in the same types of activities to attain them. Activities can be said to satisfy Lampert's criterion of being "authentic. . . to the immediate participants" if the students can perceive how these activities are enabling them to achieve these goals.

1.2 Mathematics Educators' Need for Understanding Mathematical Practice

If mathematics educators are to successfully engage students in authentic mathematical practice, they need to have an accurate understanding of what professional mathematical activity is. Our preceding description of authentic mathematical activity suggests three specific questions about mathematical practice that are of interest to mathematics educators:

- *What are mathematicians' activities and why do they engage in these activities?* For instance, it seems clear that mathematicians frequently read proofs.

Mathematics educators desire that students read proofs for the same reasons that mathematicians do (e.g., Hanna and Barbeau 2008). Why do mathematicians read proofs?

- *What are mathematicians' goals and how can mathematics educators decide when a goal is achieved?* For instance, it is widely accepted that a goal of mathematics education is to prepare students to solve mathematical problems and prove mathematical theorems in a way that would pass muster to a mathematician. How can a mathematics educator decide if a student has successfully resolved a problem or proven a theorem?
- *How, at the level of cognitive mechanism, do mathematicians engage in their goal-directed activities?* If we want students to successfully engage in a goal-directed activity, then we need to teach them how they should do so, which in turn requires mathematics educators to have a model of competent performance with regard to that activity. Mathematics educators frequently study mathematicians' behavior to build that model of competent performance (e.g., Schoenfeld 1985).

In short, it behooves mathematics educators to have an accurate understanding of what mathematicians do, why they do it, and how they do it. Accordingly, many mathematics educators conduct empirical studies addressing these questions.

2 The Relationship Between PMP and Mathematics Education Research on Mathematical Practice

2.1 The Different Ways That Claims About Mathematical Practice Are Used in PMP and Mathematics Education

There is an important commonality between how mathematics educators and philosophers of mathematical practice investigate mathematical practice. Both groups of scholars use empirical data to make descriptive claims as a basis for drawing implications. However, the *types of implications* each group draws are different. The implications that mathematics educators draw ultimately aim to be practical – tools for teachers and researchers to analyze student performance, and inform instruction. The implications that philosophers draw naturally pertain to philosophical matters, often concerning the epistemological and ontological status of mathematical claims. Because mathematics educators and philosophers aim to make different claims, they will seek out and value different empirical data about mathematical practice.

To mathematics education researchers, findings about mathematical practice are useful to the extent that they can inform the practical work of the teacher or researcher. This often leads to a desire for *precision* in how mathematicians' behavior in their practice is modeled. Consider, for instance, Polya's (1945) *How to solve it*. Mathematics educators want students to be able to solve challenging nonroutine problems; in the past, they have attempted to teach students the four-step problem solving process and the heuristics contained in Polya's volume. However, teaching students from Polya was not sufficient to improve students' problem-

solving performance. Schoenfeld's (1985) diagnosis was that the heuristics that Polya described were too *vague* to be useful for students. For instance, Schoenfeld considered a specific heuristic: To better understand an unfamiliar problem, exemplify the problem by considering various special cases. But identifying what serves as a suitable special case is a complex skill that depends on context. What example serves as a special case in one context may be unsuitable to another. Without being able to articulate the complex subskill of identifying a suitable special case, simply telling students to consider a special case, or illustrating how it is done for a particular problem, is not sufficient to help students solve problems. Some philosophers may find a description of the fine-grained cognitive processes needed to identify special cases to be psychological minutia that is not of philosophical interest, but this articulation is important for mathematics educators.

In the search for practicality, mathematics educators may also downplay issues that are of philosophical importance if they do not have implications for teaching or research. For instance, Dubinsky (2000) approved of Beth and Piaget's (1966) declaration that the status of platonism is an unreasonable question for the psychologist or the educator, as is the issue of whether mathematical ideas are perceived in a platonic realm or constructed by the learner as the resolution to these questions will not impact how a teacher helps students learn mathematics: "If we decide to adopt a Platonist point of view, then we still must determine the mechanism by which an individual gains access to the mathematical ideas that exist in some ideal world, and if we are thinking in constructivist terms, we still must work out the mechanisms by which the individual comes to construct these ideas" (Dubinsky 2000, p. 215). We do not present this passage to imply that Dubinsky speaks for all mathematics educators when he declared that the status of platonism is not important to mathematics education research (For instance, some mathematics educators, such as Paul Ernest (1985), explicitly call for rejecting platonism on the grounds that this philosophy will lead to problematic teaching practices.). Rather, we are illustrating how many mathematics educators can be indifferent to philosophical concerns if these concerns do not have practical implications for mathematics education research and teaching.

2.2 How Mathematics Education Research on Mathematical Practice Can Contribute to PMP

Even though mathematics educators and philosophers have different agendas, we nonetheless argue that mathematics education research on mathematical practice can contribute to PMP for three reasons. First, mathematics education research can corroborate or challenge claims about mathematical practice that philosophers have made. Second, mathematics educators' claims about mathematical practice can constrain philosophical accounts of mathematical practice. Third, surprising results from mathematics education studies about mathematical practice can give philosophers interesting puzzles to explain. We elaborate on each point further below.

Philosophers of mathematics often assert claims about how mathematicians think and behave and then use these claims to warrant philosophical conclusions. For instance, Fallis (1996) claimed that computer-assisted proofs are universally accepted by all mathematicians as unproblematic bona fide proofs; he used this claim to argue that proofs do not need to be a priori or transparent to gain mathematical acceptance and as part of an argument for why mathematicians should accept exotic proofs that use DNA evidence. In contrast, Rota (1997) averred that mathematicians are generally skeptical of computer-assisted proofs (proofs of the four-color theorem in particular), viewing such proofs as error-prone and not definitive. Rota used this observation as a basis for claiming that mathematicians find some proofs (or convincing verifications, if the mathematicians deny that computer-assisted proofs are proofs) to be inadequate. Aberdein (2009) believed that the mathematics community is split on the status of computer-assisted proofs. He cited a computer-assisted proof as an instance of proof*, accepted as a proof by some mathematicians but rejected by others. Aberdein used the existence of proof*s to highlight the role that argumentation that extends beyond pure deduction plays in mathematics discourse.

Whether mathematicians *ought* to accept computer-assisted proofs and the theoretical import of the publication of the proof of the four-color theorem are legitimate philosophical issues. However, the *actual* stance adopted by mathematicians about the status of computer-assisted proofs is an empirical question and therefore should be addressed by empirical research. Further, this empirical question has philosophical consequences. If Fallis, Rota, or Aberdein were found to be wrong about their beliefs about the status of computer-assisted proofs, the argumentative support for their philosophical positions would be undermined. The difficulty with describing an empirical question about mathematical practice is that there is rarely a single piece of evidence that definitively answers the question, especially when the claims are general in nature – that is, the claims do not concern the views of a single mathematician or the development of a particular mathematical concept, but rather pertain to the collective viewpoint of the mathematical community or general trends about how mathematical concepts develop. As Löwe and Van Kerkhove (2019) emphasized, philosophers of mathematical practice should seek *triangulation* to buttress their empirical claims about mathematical practice. Different types of evidence should be sought to support the same empirical claim.

The value that mathematics education research can provide to PMP is that mathematics educators tend to collect different data to support their claims. Philosophers often warrant their claims about mathematical practice with naturally occurring preexisting data, such as proofs that appear in journals, private correspondence between mathematicians, and memoirs and reflections from famous mathematicians. These data have the strength of mathematical authenticity; they come from real mathematicians engaging in real mathematical activity. However, these data can be limited in terms of generalizability. In contrast, if mathematics educators are interested in a particular issue, they will often create laboratory settings with multiple mathematicians in which data are *elicited* that bear on the specific phenomenon. To avoid misinterpretation, we are not claiming that mathematics educators' methods

for investigating mathematical practice are *better* than the methods employed by philosophers. To the contrary, our methods too have significant limitations, especially with regard to our data's authenticity to mathematical practice. We advance the more modest argument that our methods can complement philosophers' methods by addressing limitations inherent in philosophers' methods (and vice versa).

As an example of how findings from mathematics education research bear on mathematical practice, Weber and Czocher (2019) conducted a survey in which 94 mathematicians were shown a computer-assisted proof that had been published in the *American Mathematical Monthly* and were asked a series of questions about it, including whether they would judge the proof to be valid. The key findings were that 39% of the participants said the argument was a proof (61% said it was not), but 90% of the participants believed that at least 10% of mathematicians would disagree with their judgment. Further, 75% of participants felt their judgment of the proof's validity would depend on context, and 85% of the participants felt that the computer-assisted proof would qualify as a proof in at least some contexts. These findings thus support Aberdein's (2009) assertion that computer-assisted proofs are regarded as controversial within the mathematics community. These results also reveal that *some* mathematicians are suspicious of computer-assisted proofs and are thus consistent with Rota's claims, although the data neither corroborate nor refute Rota's claim that mathematicians never regard a computer-assisted proof as being definitive. These results challenge Fallis's (1996) claim that there is a consensus among mathematicians that computer-assisted proofs are not problematic.

A second way in which mathematics education research on mathematical practice can contribute to PMP is by offering a constraint on the accounts that philosophers provide. For instance, Weber and Czocher's (2019) results show that mathematicians' acceptance of a proof depends upon context and they will not agree on the acceptability of some proofs (We should be careful not to overstate our case about mathematicians' disagreement. Weber and Czocher also found that nearly all mathematicians accepted proofs using routine techniques from number theory and claimed their acceptance was independent of context. It may be the case that mathematicians only disagree on the correctness of a small subset of the proofs that they encounter.). Any philosophical account of how proofs guarantee the truth of assertions for mathematicians should account for the fact that the acceptability of some proofs depends on context and any account of how proofs may foster consensus on which theorems are true should account for the fact that mathematicians may disagree on the acceptability of some proofs. Finally, mathematics educators sometimes unearth surprising results about mathematical practice that are in need of philosophical explanation. The finding that mathematicians *disagree* on their judgments, including on the acceptability of inferences within a proof, has occurred often in studies on mathematical practice (see Hanna and Larvor 2020, for a survey of some of this research). It is an interesting question about how the mathematical community reaches consensus on the status of most theorems yet they disagree on the status of some types of inference (Or at least the common *perception* among mathematicians is that such a consensus exists. That itself is a perception worth verifying empirically.).

In the next two sections, we will summarize mathematics education research on what it means for mathematicians to understand a concept and on how mathematicians read mathematical proof. For each topic, we discuss how the results that mathematics educators have found might inform PMP.

3 Research on Mathematicians' Understanding

3.1 The Importance of Mathematical Understanding

There is universal agreement among mathematics educators that students should not just learn procedures. Students should also develop an *understanding* of the concepts that they are studying. Unfortunately, mathematics educators have struggled to define precisely what it means to understand a mathematical concept (see Hiebert 2013, for a review of some of this literature). Indicative of this struggle is the plethora of articles claiming that what mathematics educators lack is an understanding of understanding (Davis (1992), Goldenberg (1995), Michener (1978), and Nickerson (1985) have each written articles where the wordplay "understanding understanding" has appeared in the title, and Thompson (2013) dedicated a book chapter to the "meaning of meaning" in mathematics education.

The difficulty is especially pronounced when it comes to understanding the types of abstract concepts that occur at the university level. For instance, what does it mean for a student or mathematician to understand a group, a limit, a complex function, or a topological space? With respect to abstract algebra in particular, mathematics education researchers have proposed very different models of instruction where the differences are due, at least in part, to very different ideas about what it means to understand algebraic concepts (Burn 1996; Dubinsky et al. 1994, 1997; Larsen 2004, 2013). In summary, there is an undesirable state of affairs in mathematics education research: Teaching for understanding is accepted as a primary goal in mathematics instruction, but it is not clear what understanding is or how it should be measured, especially with the abstract mathematics taught at the university level.

Mathematical understanding is also important for PMP. For many mathematicians, their primary reason for engaging in mathematics is to understand mathematics better (e.g., Gowers 2000; Thurston 1994). As the mathematician and statistician David Blackwell once said, "I'm not interested in doing research and I never have been. I'm interested in understanding, which is quite a different thing" (as cited in Albers and Alexanderson 1985, p. 19). Because understanding is a primary motive for many mathematicians, it would seem that any adequate philosophical account on the growth of mathematical knowledge would need some account of what mathematical understanding is. Yet, perhaps surprisingly, there are few accounts in PMP of what constitutes mathematical understanding (We say "surprisingly" because there is an emerging literature in the philosophy of science on what comprises scientific understanding (e.g., Grimm, Baumberger, and Ammon 2016), but the literature on mathematical understanding is comparatively sparse (but see, e.g., Avigad 2008, in

press; Delarivière and Van Kerkhove in press; Folina 2018; Inglis and Mejía Ramos in press).)

Some mathematics education researchers have conducted empirical studies with mathematicians to provide a better account of what it means to understand advanced mathematics and how mathematical understanding develops. In this section, we review some of this literature. We then discuss some potential implications of this research for PMP.

3.2 Open-Ended Interviews with Mathematicians About Their Understanding of Abstract Concepts

Sfard (1994) and Sinclair and Tabaghi (2010) each conducted *open-ended semi-structured interviews* (Hammer and Wildavsky 1993) with mathematicians about what it means for them to understand a concept. In an open-ended interview, a small number of mathematicians meet individually with a mathematics education researcher, who asks them a series of questions about a particular phenomenon of interest. The individual mathematician's responses is typically recorded in some way, transcribed, and subjected to analysis. The "open-ended" modifier signals that at least some of the questions are designed to elicit open-ended responses that allow for the mathematician to elaborate on her viewpoints. The "semi-structured" modifier means that the interviewer has the right to use her professional judgment to ask unplanned follow-up questions, for purposes such as clarification or elaboration. The way that mathematics educators use data obtained from semi-structured interviews is in some ways similar to the ways that philosophers use the memoirs and reflections of famous mathematicians (e.g., Hardy 1940). In both cases, researchers are using mathematicians' accounts of their mathematical beliefs and activities to provide models of mathematical practice. There are several advantages that semi-structured interviews have over relying on the reflections of famous mathematicians. Because the interviewer chooses the question and steers the discussion, the interviewer can target a phenomenon of interest, rather than hoping that a mathematician will happen to write about it. Because the interviewer can ask for clarification and elaboration, this will reduce ambiguity in the mathematicians' comments about the targeted phenomenon, although ambiguity can never be eliminated entirely. Interviewing multiple mathematicians allows the researcher to compare and contrast their responses, which offers a richer and more textured perspective on the phenomenon in question (The memoirs of famous mathematicians have advantages over semi-structured interviews. For instance, these memoirs represent the most mature reflection rather than the commentary from mathematicians who are put on the spot. As always, our point is not that the data that mathematics educators typically rely on is *better* than the data that philosophers are more familiar with. Rather, it is that the data have different strengths and weaknesses.).

Sfard (1994) reported a study based on extended (3+ h each) semi-structured interviews with three mathematicians about what it meant to understand abstract mathematics. Sfard analyzed her data using Lakoff's (1987) linguistic theory based

on the premise that our understandings are based on metaphors which can ultimately be traced down to our physically embodied experience. Sfard reported data illustrating the following claims. First, the mathematicians she interviewed claimed they understood a concept only if they could view the concept metaphorically. For instance, one mathematician said, "To understand a concept, I must create an appropriate metaphor. A personification. Or a spatial metaphor. A metaphor of structure. Only then can I answer questions, solve problems" (p. 48). Second, mathematicians provided different types of metaphors, including personification metaphors where the mathematical concept is viewed as taking up physical space, performing actions, and/or having a personality. Indeed, the mathematicians claimed that they would become intimately familiar with these concepts, and this familiarity was critical for reasoning about them. Third, mathematicians' visual representations of concepts were vague frameworks for reasoning and action, rather than a concrete instantiation of the concept itself.

Sinclair and Tabaghi (2010) conducted open-ended semi-structured interviews with six mathematicians about their understanding of eigenvectors. In their analysis, the authors used a theoretical frame in which they distinguished between conceptual mathematics (CM) and ideational mathematics (IM) (as introduced by Sinclair and Schiralli 2003). CM are public representations of mathematical ideas that exist in public space and have shared meanings in the mathematical community; IM are the personal representations that individuals use to represent mathematical ideas to themselves.

Sinclair and Tabaghi's key findings were that when their participants were asked to describe eigenvectors, all used metaphors involving motion. Although the formal CM definition of the eigenvector of a matrix involved equality (i.e., $Ax = \lambda x$), participants were inclined to say that a matrix A acted on an eigenvector x by transforming it in a manner that stretched or shrank x. For instance, one mathematician described eigenvectors as follows:

> I think of a matrix I think of applying a vector to, no applying a matrix [lifts left hand] to the vector [moves left hand toward right hand and then back again] and then what you get out is another vector [moves his right hand, bended fingers], that is in the same direction but either stretched [clenches his hands and extends his arms] or shrunk [brings his clenched hands and extended arms together], but you also have to be aware of like [moves both hands up and down in an alternate fashion] flip direction [hands moving toward each other and index fingers pointed to each other] or it could just sent to zero [moves left index finger down] and that will still count as an eigenvector (p. 229).

Sinclair and Tabaghi (2010) emphasized two things from this excerpt and others (Sinclair and Tabaghi (2010) included photographs of this participant (and other participants) illustrating their gestures that we do not include here for the sake of brevity.). First, the mathematicians' descriptions of eigenvectors had a *temporal* quality that was not captured by static equality. The mathematicians treated a matrix as a transformation that you "apply" to a vector and you compare what you "get out" with the original vector. Second, in all their descriptions, participants used gesturing to illustrate their ideas, particularly how vectors were being transformed, which

suggested to Sinclair and Tabaghi that these mathematicians were using enactive metaphors based on their embodied cognition.

Sinclair and Tabaghi also observed that individual mathematician's understanding of metaphors was idiosyncratic and involved personal metaphors, such as thinking of eigenvectors in stretching or shrinking a rubber band, the motions of hands on a distorted clock, or stressors acting on a plate. Finally, participants sometimes used diagrams to describe their understanding of eigenvectors. When diagrams were introduced, the participants associated motion with the diagrams that they often represented using gestures (e.g., using a sweeping motion of the hands to describe a mapping represented in a diagram).

In another study, Soto-Johnson et al. (2016) asked five mathematicians to describe what representations or explanations might be useful to understand the continuity of complex valued functions (Soto-Johnson et al. (2016) also asked their participants to determine whether two specific complex functions, $f(x) = Re(z)/|z|$ and $g(z) = zRe(z)/|z|$, could be extended to $z = 0$ in a way that would preserve their continuity. The latter task is known as a task-based interview, which we will say more about in the next subsection.). Like Sinclair and Tabaghi (2010), Soto-Johnson et al. (2016) used Sinclair and Schiralli's (2003) constructs of conceptual and ideational mathematics (CM and IM) to analyze participants' responses. Soto-Johnson et al.'s findings were consistent with Sfard's (1994) and Sinclair and Tabaghi (2010) in that all participants used IM comprised of idiosyncratic personalized metaphors for understanding the continuity of complex functions. Soto-Johnson et al. further observed that the metaphors that participants used were often based on their previous embodied experiences in nonmathematical contexts. For instance, one mathematician who hiked as a hobby used the hiking metaphor of falling off a cliff to describe function discontinuities.

3.3 Task-Based Interviews on How Mathematicians Develop an Understanding of Advanced Mathematics

We now consider the issue of how mathematicians may develop their understanding of an advanced mathematical context. In both popular culture and mathematicians' memoirs, some mathematicians describe their process of coming to understand a piece of mathematics as comprised of small increments of local understanding followed by sudden jumps. The mathematician Andrew Wiles offered the following metaphor to explain his experience doing mathematics:

> Perhaps I could best describe my experience of doing mathematics in terms of entering a dark mansion. One goes into the first room, and it's dark, completely dark. One stumbles around bumping into the furniture, and gradually, you learn where each piece of furniture is, and finally, after six months or so, you find the light switch. You turn it on, and suddenly, it's all illuminated. You can see exactly where you were. (Andrew Wiles, as quoted by NOVA 1997)

Later, Wiles offered some elucidation on what it might mean to stumble around in a dark mansion:

> I sometimes might scribble. I do doodles. I start trying to find patterns, really, so I'm doing calculations which try to explain some little piece of mathematics, and I'm trying to fit it in with some previous broad conceptual understanding of some branch of mathematics. (Andrew Wiles, as quoted by NOVA 1997)

Poincaré also described a flash of mathematical insight as emerging after a series of more modest connections were formed:

> A host of ideas kept surging in my head; I could almost feel them jostling one another, until two of them coalesced, so to speak, to form a stable combination. When morning came, I had established the existence of one class of Fuchsian functions. . . I then only had to verify the results, which took only a few hours. (Poincaré 1952, 52–53)

For the purposes of mathematics education, we view these quotes as descriptive rather than prescriptive. While they provide important insights into how Wiles and Poincaré do mathematics, they are not specific enough to guide instruction. To gain more precision, mathematics educators sometimes conduct *semi-structured task-based interviews* (Clement 2000; Goldin 2000; Mejía-Ramos and Weber 2020). In these interviews, a modest number of mathematicians meet individually with a researcher who asks them to think aloud as they complete predetermined tasks that are meant to approximate mathematical practice. The benefits of task-based interviews are that they offer a more fine-grained account of how mathematicians complete tasks than can be gained from interviewing mathematicians or reading their reflections.

To "understand how mathematicians make sense of ideas that are unfamiliar to them" (p. 22), Wilkerson-Jerde and Wilensky (2011) conducted task-based interviews in which ten mathematicians were observed reading a paper from geometric topology about "links." The paper was outside of the participants' areas of research so they would be learning the topic for the first time. However, the paper was accessible to nonexperts. This task was meant to approximate how mathematicians might learn a new mathematical topic in their professional practice.

As Wilkerson-Jerde and Wilensky's participants read the paper, they continually engaged in two metaprocesses. In the first metaprocess, which the authors labeled as deconstruction, the mathematicians sought to understand how individual ideas fit together. In the second metaprocess, which the authors labeled as coordination, the mathematicians would fit ideas together in a coherent framework which they can use to predict mathematical relationships that they had not yet considered. In both cases, mathematicians engaged in a cycle in which they asked questions, answered their questions, and provided explanations. For instance, when the mathematicians encountered a new definition, they may ask how it applies to a particular example with which they were familiar or they may form a hypothesis about the meaning of an idea in the definition. After answering their question, they may offer a more high-level explanation about what the definition means.

Wilkerson-Jerde and Wilensky's (2011) findings are useful to mathematics educators because they offer potential prescriptive accounts of how abstract mathematics can be learned. Their work offers the hypothesis that mathematicians' understanding is based on a rich set of connections of basic ideas in a mathematical theory that are developed by mathematicians constantly questioning the meaning of certain terms and how they relate to other ideas and examples. Students can, in fact, be trained to ask similar questions when they are reading mathematical text. When students are trained to do so, their comprehension of the text improves dramatically (Alcock et al. 2015; Hodds et al. 2014).

3.4 Discussion of Mathematics Education Research on Mathematicians' Understanding and Its Implications for PMP

In this subsection, we offer several cross-cutting findings about professional mathematicians' understanding of abstract mathematics. Before doing so, we offer the following caution. Each of the studies cited above was *exploratory studies* that were used for *hypothesis generation*. These studies were exploratory in the sense that the researchers were not conducting the studies to test the viability of any prespecified theoretical claim. Rather, the researchers' aim was to put mathematicians in situations that would elicit interesting data about a phenomenon of interest from which the researchers could construct plausible models of mathematical practice that would be consistent with their data. Consequently, the findings presented here should be viewed as tentative hypotheses that are empirically grounded in the data that were collected rather than definitive claims about mathematical practice.

A first finding is that mathematicians viewed dynamic metaphors as essential to their understanding. The primary salience of dynamic metaphors in mathematicians' reports on their understanding would seem to us to put a demand on any philosophical account of mathematical understanding. A viable account should either explicate the role that dynamic metaphors play in understanding or at least explain why mathematicians hold the belief that dynamic metaphors are important for understanding.

A second finding is that mathematicians' metaphors are idiosyncratic and based on individual mathematician's personalized experience. If we view "mathematical understanding" of a concept to consist of mathematicians' productive ways of thinking about that concept that successfully guide their work, then there is no single mathematical understanding. Alternatively, if we claim that there is a specific way of thinking about a concept that constitutes true understanding, that account would need to explain why mathematicians so often deviate from that understanding in their own idiosyncratic ways.

Finally, as Sinclair and Schiralli (2003) argued, the metaphors used by mathematicians to create concepts are distinct from the metaphors that contemporary mathematicians use to understand the concepts. For instance, surely the concept of the continuity of complex functions was not created with hiking in mind, but this is how

one mathematician understood continuity. Lakoff and Núñez (2000) offered a theory by which all mathematical concepts can be traced via metaphor to our embodied cognition in our physical world. Even if Lakoff and Núñez have offered a viable theory about where mathematics comes from, this should not be mistaken for a theory of how mathematics is understood or what mathematics is today.

4 Mathematics Education Research on Mathematicians' Proof Reading

4.1 The Importance of Proof Reading

As with conceptual understanding, the issue of proof reading is of both educational and philosophical importance. In university mathematics courses, proofs are common both in lectures (e.g., Mills 2014) and in textbooks (e.g., Raman 2004). Indeed, much of students' time in university mathematics courses is spent studying proof. Unfortunately, there is evidence that students do not understand much of what they read. Numerous studies have found that university mathematics students with substantial training in proof are unable to distinguish between valid and invalid arguments (e.g., Alcock and Weber 2005; Inglis and Alcock 2012; Ko and Knuth 2013; Selden and Selden 2003; Weber 2010). Understanding a proof should extend beyond checking a proof for correctness. For instance, a proof may be used to provide explanation or illustrate a new mathematical method. However, there is also evidence that students do not recognize these broader purposes of proof (e.g., Healy and Hoyles 2000; Weber and Mejía-Ramos 2014). When students observe a proof in their lectures, they may miss the broader points that the professor is trying to convey (Cowen 1991; Lew et al. 2016). Improving students' ability to read proofs can increase their understanding of mathematics, especially in the proof-oriented courses taught at the university level.

The ways in which mathematicians read proofs have philosophical consequence as well. Proof is a primary means by which mathematicians persuade each other about the veracity of theorems and, more generally, how they communicate mathematical knowledge to one another. Philosophers sometimes cite mathematicians' practices with proof to account for the fact that mathematics is reliable and objective and that mathematicians generally agree on what statements are true and what theorems are established (e.g., Azzouni 1994; Berry 2016). To understand the role proof can play in these regards, it is important not only to know what proof is but how the mathematical community reads these proofs.

4.2 Hypothesis Generation Versus Hypothesis Testing

In the previous section, we claimed that the mathematics education research on concept understanding has involved exploratory studies focused on hypothesis generation. Here we describe an alternative approach to conducting empirical

research which involves *hypothesis testing* (Inglis and Aberdein 2020). Researchers who study mathematical practice often form hypotheses about how mathematicians behave. These hypotheses are often generated from the types of exploratory studies described in the previous section, but they sometimes emerge from other sources as well, such as the researcher's informal conversations with mathematicians or reflections on her own experience doing mathematics. As a concrete example that we will discuss in this section, a researcher may hypothesize that mathematicians will find arguments more persuasive if they come from an authoritative source. To test the viability of her hypothesis, the researcher uses deductive reasoning to form an observational prediction that she would expect to see if her hypothesis is true. For instance, in our hypotheses on the role of authoritative sources, a researcher might predict that mathematicians would evaluate a specific argument as more persuasive if they knew that the argument came from a famous mathematician than if they were not aware of who wrote the argument. The researcher then conducts an empirical study to determine if her observational prediction occurs. If the prediction is realized, this is strong support for the researcher's hypothesis, especially if the prediction is surprising or is inconsistent with what other popular competing hypotheses would predict. If the prediction is not observed, this casts doubt upon the hypothesis. While hypotheses testing studies often do not offer conclusive proof that a hypothesis about mathematical practice is true or false, they can offer much stronger empirical support for or against hypotheses than can be obtained by exploratory studies (Hypothesis generation studies allow the researcher a greater number of degrees of freedom to interpret their data than hypothesis testing studies. For further discussion, see Lindsay et al. (2016).) In general, hypothesis testing studies are uncommon in both PMP and mathematics education, but they have been used regularly in mathematics education research on mathematicians' proof reading.

4.3 Competent Performance in Proof Validation

Mathematics education research into university students' reading of mathematics began in earnest with Selden and Selden's (2003) study on whether students could reliably distinguish between correct proofs and invalid arguments. In their paper, they presented eight students with four purported proofs of the claim, "If n^2 is divisible by 3, then n is divisible by 3." One argument contained the inference that because $3x = nn$, then 3 divides n. Selden and Selden declared the argument was invalid because this inference contained a gap that could not be filled. However, on their first reading of the proof, only one of the eight students judged this ostensibly invalid proof to be incorrect (although students' performance improved due to the interviewer asking the students questions that prompted reflection).

Weber (2008) and Inglis and Alcock (2012) investigated the reasoning processes that mathematicians used to determine if a proof was valid. In both studies, one of the items that the mathematicians considered was the aforementioned argument from the Selden and Selden study. All eight of the mathematicians participating in Weber's study noted that the gap identified by Selden and Selden could be bridged with the

theorem, "if a prime p divides ab, then p divides a or p divides b." All were hesitant with their judgments about the validity of the proof and only five of the eight mathematicians judged the argument to be invalid. Similarly, only 7 of the 12 mathematicians in Inglis and Alcock's study judged the argument to be invalid. Based on this finding, the authors of these studies hypothesized that mathematicians may disagree about the validity of a proof when it contains large gaps.

To test this hypothesis, Inglis, Mejía-Ramos, Alcock, and Weber (2013) presented 109 mathematicians with an argument justifying why the antiderivative of x^{-1} was ln x (This proof was similar to another proof in Inglis and Alcock's (2012) study where mathematicians disagreed about the validity of the proof.) The last line of the proof commuted the limit and integral in an algebraic expression without justification for why such a step was permissible. This gap in the proof could be bridged – the commutation of the limit and integral was with respect to a uniformly continuous function – but the proof contained no indication that there was a gap in the proof at all. The participants were told that this proof was submitted to be published in the *Mathematical Gazette* and were asked if the argument was a proof. Inglis et al. predicted that mathematicians would disagree on the validity of this proof.

Inglis et al.'s (2013) prediction was confirmed: 27% of their participants judged the argument to be a valid proof and 73% disagreed. Participants were asked what percentage of mathematicians would agree with their evaluation. The average response was 75%. Among mathematicians who thought the proof was valid, this average was 65%. For those who thought the proof was invalid, the average was 78%. Hence, while participants collectively believed that the majority of mathematicians would agree with their judgment, they were generally aware that a substantial portion of the mathematical community would disagree with them. Finally, Inglis et al. wanted to rule out that the possibility that the disagreement was merely the result of a performance error – that is, that those who declared the proof to be valid simply overlooked what was problematic in the last line of the proof. To address this, participants were told that another mathematician had said the proof was invalid since one cannot assume that the integral and limit commute without considering the type of convergence involved. Participants were asked if this critique was enough to render the proof invalid. Only one participant who judged the proof to be valid agreed with this, suggesting that the judgment of validity was not simply a matter of oversight.

The results from this study suggest that mathematicians do not always agree on whether an argument constitutes a proof. These findings offer a challenge to philosophical theories that are based on the premise that mathematicians agree on what constitutes a proof. For instance, Azzouni (2004) claimed that "mathematicians are so good at agreeing with one another whether some proof convincingly establishes a theorem" (p. 84) and used this claim to support the conclusion that the correctness of a proof can be checked by a mechanical process, and Berry (2016) argued that the norm that published proofs are readily available provides mathematical practice with the virtues of autonomy (i.e., an individual mathematician can verify the truth of any theorem herself) and consensus (i.e., mathematicians agree on what are or are not theorems). Berry's argument relies on his claim that "any mathematician can

(without much effort) find his or her own explicit reasons for believing any Publically Accepted Claim . . . journals rightly believe the results they publish are Reliable and *expect the arguments for them to be found convincing by all*" (p. 117, italics are our emphasis). The aforementioned findings about mathematicians' disagreement do not disprove Azzouni's or Berry's conclusions; however, they do challenge their premises. Mathematicians do not *always* agree on whether a proof convincingly establishes a theorem. It may be the case that they agree *often enough* with respect to the published literature, but justifying such a premise requires further argumentation.

Our results place constraints on philosophical accounts of how proof can work in mathematical practice. For instance, philosophical accounts that posit that mathematicians use similar processes for checking a proof would need to account for the fact that mathematicians sometimes reach different conclusions on the proof status of some arguments and, further, are aware that mathematicians would disagree with their judgments.

4.4 How Do Mathematicians Check Proofs for Correctness?

Earlier in the section, we noted that there are many studies in mathematics education that document how university students cannot distinguish between correct proofs and invalid arguments (e.g., Ko and Knuth 2013; Selden and Selden 2003). These studies identify a mathematics education research problem in need of remediation: how can we help students learn how to check proofs for correctness? To address this issue, mathematics educators need to know how students should be reading proofs to assess their correctness. Some mathematics educators have studied how mathematicians read proofs to delineate the processes needed to read proofs competently.

Weber and Alcock (2005) argued that checking a proof for correctness involves a process that they called "inferring warrants." To illustrate, consider the statement, "Since 7 is prime, 1007 is prime." Weber and Alcock observed that when mathematicians checked the correctness of this statement, they did not consider whether 1007 is a prime number. Rather, they rejected the inference as invalid because it appeared to be based upon the general principle that adding 1,000 to a prime number produced a prime number and this general principle is false. Of course, the general principle was not explicitly stated in the statement that the mathematicians considered; mathematicians inferred the person making this assertion must have had this general principle in mind. Weber and Alcock then hypothesized that when mathematicians read a new statement of the proof, checking its validity sometimes involves identifying previous statements in a proof and a general inferential scheme from which the new statement can be derived.

To test Weber and Alcock's (2005) hypothesis, Inglis and Alcock (2012) conducted a task-based eye-tracking experiment, which recorded those parts of the text that 12 mathematicians and 18 university mathematics students fixated upon as they checked the validity of 12 purported proofs. Inglis and Alcock identified particular inferences in a proof that required inferring a warrant. For instance, one proof began as follows:

1. Suppose there are finitely many primes of the form $4k + 1$.
2. Then these primes can be listed $p_1, p_2, p_3, \ldots, p_n$.
3. Define a number a as follows: $a = p_1 p_2 p_3 \ldots p_n + 4$.

Inglis and Alcock judged line 2 as requiring an implicit warrant – namely that finite lists can be enumerated in a bounded list (i.e., we would not ordinarily be able to assert that any arbitrary class of integers could be enumerated as a list of n elements). In contrast, Inglis and Alcock said that line 3 did not require a warrant as stipulating the meaning of a could always be done, regardless of how the previous variables in the proof have been defined. Based on their assumption that inferring warrants for line x would involve looking at line x, then line x-1, then line x again, Inglis and Alcock predicted that mathematicians would be more likely to show an $x \rightarrow x\text{-}1 \rightarrow x$ eye-movement pattern than students when line x required inferring a warrant compared to when it did not. This prediction was borne out. Especially interesting was the fifth line in the proof described above: "Every number that leaves remainder 1 when divided by 4 is divisible by a prime that also leaves remainder 1 when divided by 4." This statement is simply false (9 is a counterexample) and all 12 mathematicians believed the falsity of the statement was sufficient to render the proof invalid. (In contrast, only 10 of the 18 students judged the proof to be invalid, once again highlighting students' limited ability to identify errors in proofs). The mathematicians in the study spent a great deal of time oscillating their attention between lines 4 or 5 in this proof, apparently searching in vain for an implicit warrant to justify this false statement. Students, in contrast, rarely moved their attention back and forth between lines 4 and 5. The main implication that Inglis and Alcock drew for pedagogical purposes is that students needed to devote more effort in inferring warrants as they read a proof, but this also supports the conclusion that mathematicians infer warrants when reading proofs in their practice.

The next topic we consider is the role of examples in proof validation. Weber (2008) conducted a task-based study in which eight mathematicians were asked to think aloud as they determined if eight purported proofs were correct. Weber was especially curious how mathematicians determined whether new statements in a proof followed from previous statements when the correctness of a new statement was not clear. In most cases, mathematicians' justifications were based on deductive argumentation, where the mathematicians wrote a subproof or gave an informal argument for why the new statement was
a logical consequence of previous statements. However, there were also instances in which mathematicians used empirical evidence to justify the new assertion – that is, they concluded the new assertion was acceptable because it held for some examples that the mathematicians checked. For instance, one mathematician encountered the following consecutive lines in the proof: n is congruent to 3 (mod 4). Note that n is not a perfect square. Below is the transcript of how the mathematician concluded the obvious implicit warrant that numbers congruent to 3 modulo 4 are not perfect squares:

> I'm using examples to see what, where the proof is coming from. So 5^2 is 25 and that's 1 mod 4. 36 is 0 mod 4. 49 is 1 mod 4, 64 is 0 mod 4. I'm thinking that, ah! So it is. . . 24 times 24 is 0 mod 4. So a perfect square has to be 1 mod 4, doesn't it? n^2 equals 1 mod 4 or 0 mod 4. Alright. (Weber 2008, p. 443)

When asked about his reasoning later in the interview, this mathematician concluded that reasoning from patterns in this case was reliable since these patterns usually held and could be verified by induction, while acknowledging that such reasoning often would not work in topology, which was his area of research.

To test the hypothesis that mathematicians would sometimes accept statements in a proof because they held for particular examples, Mejía-Ramos and Weber (2014) conducted a survey with 118 mathematicians. In the survey, the participants were asked questions about how and why they read proofs. Two of the questions from the survey were as follows:

Q1. When I read a proof in a respected journal and I am not immediately sure that a statement in a proof is true, it is not uncommon for me to increase my confidence in the statement by checking it with one or more carefully chosen examples.
Q2. When I read a proof in a respected journal and I am not immediately sure that a statement in a proof is true, it is not uncommon for me to gain a sufficiently high level of confidence in the statement by checking it with one or more carefully chosen examples to assume the claim is correct and continue reading the proof.

Mejía-Ramos and Weber's hypotheses were that more mathematicians would agree than disagree in responding to Q1 and Q2. For Q1, 79% of the respondents agreed with the statement and 13% disagreed. For Q2, 56% of the respondents agreed and 30% disagreed. The remainder, in each case, responded they were neutral with regards to the statement. In each case, there were significantly more (in the sense of statistical inference) mathematicians who agreed with the statement than who disagreed. Of Mejía-Ramos and Weber's 118 respondents, 54 reported having substantial experience refereeing papers. They were also asked Q1 and Q2 with respect to refereeing. For Q1, the results were similar. For Q2, 43% of the mathematicians agreed and 30% of the mathematicians disagreed. The difference between those who agreed and disagreed was not statistically significant for Q2, but the results indicate that a nontrivial number of mathematicians will accept statements in a proof based on empirical evidence, even when they are refereeing papers.

The findings about mathematicians' use of empirical evidence provide a challenge to philosophical accounts for how proof is essential to ensuring the reliability of mathematics. A common account is that empirical evidence in support of an assertion, no matter how seemingly persuasive, is unreliable. Yet proof, because of the notion of logical necessity, provides an extremely high level of reliability (see Berry 2016, for one elaboration of such a viewpoint). The difficulty is that mathematicians sometimes use empirical evidence to check specific assertions within a proof. From our perspective, this suggests that an account of proof that posits that (i) mathematicians read proofs to gain certainty that theorems are true, (ii) certainty

is gained by checking that each line of the proof is valid, and (iii) empirical evidence cannot provide this certainty, seems inconsistent with the data presented in this subsection. If empirical evidence cannot provide certainty that a proof is correct, then either some mathematicians are not reading proofs to gain certainty (even when they referee) or their certainty is coming from another source beyond establishing the validity of each line in the proof (We believe that (i), (ii), and (iii) may all be false. Later in this section, we present evidence that mathematicians do not always check proofs for correctness (see Mejía-Ramos and Weber 2014) and mathematicians may use the overarching method used in the proof to conclude that it is correct rather than a line-by-line check. Further, there appear to be *some* cases in which mathematicians gain certainty or near certainty in a mathematical claim based on empirical evidence (see Paseau 2011, for a discussion).)

4.5 Factors Mathematicians Consider When Assessing the Persuasiveness of Arguments

An interesting debate in mathematics education concerns what evidence should be used by students to convince themselves that mathematical claims are true (see Stylianides et al. 2017, or Weber et al. 2014, for a review of this literature). The general premise among mathematics educators is that students should be convinced by the same types of evidence as mathematicians (e.g., Harel and Sowder 1998). A particularly interesting question concerns the role authority play in mathematicians' acceptance of a proof and what role authority should play in students' acceptance of a proof.

Based on an open-ended interview study, Müller-Hill (2010) found that mathematicians would be more accepting of an argument if it came from an authoritative source, even in the context of refereeing. For instance, one mathematician in her study said:

> Let's say a famous mathematician comes up with a paper and I have to referee it. Then I am preoccupied with the fact that he is a famous mathematician and so that it will probably be correct. And then you say, 'yes, this really seems plausible, but I'm not really sure if it's true' and then you end up with the question, 'is this because I don't have enough knowledge'. (Müller-Hill 2010, as cited and translated in Geist et al. 2010, p. 162)

To avoid misinterpretation, there are two points worth noting. First, the role that authority plays in the Müller-Hill quote above concerns the correctness of an argument, not the truth of a claim. The issue at stake in this question is not whether mathematicians will believe a claim because an authority says it is so, but whether they will believe a proof is correct because it came from an authoritative source. Second, the excerpt from Müller-Hill does not say that mathematicians would accept *any* argument from an authoritative source, even one that contains a clear mistake. Rather, what is claimed is that the identity of the author of the argument influences how the argument is read and a mathematician may be more inclined to grant the

benefit of the doubt if the argument came from an authoritative source. Other mathematics educators have made claims consistent with the excerpt above. For instance, Hanna (1983) asserted that to mathematicians, the author having an unimpeachable reputation as an expert matters more than the rigor of an argument in deciding its acceptability. Like Müller-Hill (2010), Weber (2008) quoted mathematicians stating that the author of an argument would influence how mathematicians would read it.

To test the hypothesis that authority affects the persuasiveness of an argument to mathematicians, Inglis and Mejía-Ramos (2009) conducted a randomized controlled experiment. As Inglis and Aberdein (2020) discussed, experiments offer a powerful means of testing hypotheses about mathematical practice, especially when these hypotheses predict that a particular factor influences mathematicians' judgments or performance. To see whether a particular factor *F* (in this case, the authority of the author) influences performance, mathematicians who participate in an experiment are randomly divided into a control group and an experimental group, where each is asked to complete a task. The tasks that the control group and experimental group receive are identical in all respects, *except* factor *F* is present in the experimental group's task but not the control group's task. If there is an observed difference between the control group and the experimental group's performance, then the difference can only be attributable to the presence of factor *F* or due to chance, as *F* is the only difference between the tasks that the two groups completed. If the performance difference between the groups is statistically significant, this controls the probability that the difference was due to chance and it is reasonable for a researcher to conclude that it was due to factor *F*.

Inglis and Mejía-Ramos (2009) asked 190 mathematicians to evaluate (on a scale of 0 through 100) how persuasive they found a visual proof of the Fixed Point Theorem. The control group and experimental groups received the same tasks, except the experimental group was told that J.E. Littlewood wrote the argument and the control group was not. The experimental group rated the argument as 17 points more persuasive than the control group, a statistically significant difference. Because the only difference in the tasks between the groups is that the experimental group was told that Littlewood wrote the argument, it is reasonable to conclude that the mathematicians knowing that the argument came from an authoritative source led them to find it more persuasive. Similar results were reported by Inglis and Mejía-Ramos when mathematicians were asked to evaluate a heuristic (nondeductive) argument that was written by Timothy Gowers.

In Mejía-Ramos and Weber's (2014) survey, the large majority of mathematicians indicated that they would find proofs to be more persuasive if they were published in an academic journal, especially one that was of high quality, and if it came from an authoritative source that they trusted. The broad implication for these studies is that factors concerning social context influence mathematicians' acceptance of proofs. This supports philosophical accounts of practice such as Auslander's (2008) that describe the role of trust and testimony in the mathematical community's acceptance of claims and the growth of knowledge.

4.6 Why Do Mathematicians Read Proofs?

Weber and Mejía-Ramos (2011) conducted open-ended interviews on how and why mathematicians read proofs that were published in journals. Three interesting findings emerged: First, some mathematicians claimed they did not read proofs to check them for correctness; they assumed the proofs that they were reading were probably correct by dint of the fact that they were published. Second, all participants claimed that the primary reason they read the published literature was to learn about new techniques that they could use to solve problems in their own research. Mejía-Ramos and Weber's (2014) subsequent survey found that their respondents overwhelmingly agreed that they sometimes did not check the proofs that they read for correctness and their primary reason for reading proofs was to identify new techniques that they could use. Third, their participants collectively viewed their understanding of the proof as consisting of four components: being able to offer a high-level summary of the proof, apply the methods of the proof to a new context, break the proof into subproofs and understand how these subproofs or modules related to each other, and relate the ideas of the proof to specific examples (Mejía-Ramos et al. 2012). These results corroborate claims from several philosophers on the nature of proof, such as Rav's (1999) claim that proofs to mathematicians are primarily about revealing methods and Avigad's (2020) recent claims on the role of modularizing and examples in the ways that proofs are written.

Broadly speaking, Weber and Mejía-Ramos' (2011) findings are consistent with the claims in the philosophical literature that for mathematicians, fully understanding a proof goes beyond being able to logically justify each of the steps in a proof (e.g., Feferman 2012; Thurston 1994). However, if this body of research presents an accurate account of mathematical practice with regard to proof reading, we would expect to find evidence for their claims in how mathematicians read proofs. Eye-tracking studies to date have failed to find this evidence. For instance, if mathematicians sought to understand proofs in terms of the global methods that were employed, we might expect mathematicians reading a proof to go globally through the proof (evidence of viewing the proof as a unified whole) rather than focusing only on consecutive lines within the proof (evidence of viewing the proof as a chain of specific inferences). However, when Inglis and Alcock (2012) conducted their eye-tracking study, they found evidence of the latter behavior but not the former. Further, if understanding a proof is different than checking a proof for correctness, we would expect mathematicians who were reading a proof for understanding would exhibit different behaviors than mathematicians who were reading a proof with the aim of checking for correctness. Yet when Panse et al. (2018) conducted an eye-tracking study in which an experimental group of mathematicians was asked to read a proof with the aim of understanding and a control group was asked to read a proof with the aim of checking for correctness, the authors could find no difference between the two groups' eye movements.

These results do not necessarily imply that checking a proof for correctness is tantamount to understanding a proof, or even that mathematicians will not engage in different behaviors when reading the proofs for different purposes. But these results

do pose a philosophical puzzle. What type of behavior do mathematicians exhibit when reading a proof? Does this behavior constitute evidence that mathematicians try to view a proof holistically, rather than as a series of individual assertions? If checking a proof for correctness is different than checking a proof for understanding, what types of different behaviors would be consistent with this hypothesis? And why weren't these behaviors documented in the eye-tracking studies that were conducted?

5 Concluding Remarks

Research on mathematicians' practices has been flourishing in the mathematics education research community for some time. As this chapter has shown, there is a growing literature focused on what mathematicians do, why they do it, and what mechanisms they use while doing it. The rationale for this interest is a belief that a detailed understanding of the answers to these questions should be of use to teachers and curriculum designers. Specifically, an understanding of expert mathematical practice should help to inform (i) the goals of the curriculum, (ii) teachers' judgments about whether students have achieved these goals, and (iii) pedagogical methods that are likely to be most effective at helping students succeed.

Because of this belief that an understanding of expert mathematical practice is educationally relevant, mathematics education researchers have conducted empirical research on a number of topics that are of interest to PMP researchers. Here we have reviewed recent research on mathematicians' understanding of mathematical topics, as well as work on mathematicians' reading of proofs. But these topics are merely illustrative: The mathematics education literature also includes work on mathematical language (Burton and Morgan 2000) and aesthetic judgments (Inglis and Aberdein 2020).

Although mathematics education researchers are interested in understanding mathematical practice, we should not lose sight of the differences between educational and philosophical research on the topic. We have noted differences in terms of both goals and typical research methods. Because of their desire to (eventually) have a positive impact on educational practice, there may be differences between the issues that attract mathematics education researchers' and philosophers' attention. Nevertheless, there are also important overlaps.

As befitting academics working in an empirical social science, mathematics education researchers typically warrant their claims using empirical evidence collected in laboratory settings. Here we have reviewed a variety of examples, including studies which used task-based interviews, which tracked mathematicians' eye movements, and which deployed controlled experimental techniques. These methods can be criticized for lacking authenticity and not taking mathematical context seriously. In contrast, in the PMP literature, there is a much greater reliance upon conceptual analysis, introspective accounts, and historical case studies. While these methods have the advantages of exploring authentic mathematics and giving proper weight to

mathematical context, they are limited in terms of generalizability and run the risk of confirmation bias.

If our observations about the research methods typically used by mathematics education researchers and philosophers are right, then there is the potential for significant cross-fertilization between the two disciplines. More specifically, the research methods used by mathematics education research, particularly those which use large samples and hypothesis testing, can address some of the limitations inherent in introspective accounts and historical case studies. Equally, the authenticity gained by studying mathematicians' accounts – both historical and contemporary – of their own practices can help to address the inevitable artificiality introduced by empirical lab-based methods.

The value of mathematics education research to PMP goes beyond an expansion of research methods. We have suggested that it can contribute in at least three ways: by corroborating or challenging claims made by the PMP literature, by constraining future PMP theorizing, and by producing surprising empirical findings that provoke interesting philosophical questions. In sum, deeper connections between the two research communities are likely to benefit both.

6 Cross-References

- ▶ Definitions (and Concepts) in Mathematical Practice
- ▶ Heuristics and Mathematical Practice
- ▶ The Design of Mathematical Language
- ▶ The Ethics of Mathematical Practice

References

Aberdein A (2009) Mathematics and argumentation. Found Sci 14(1–2):1–8

Albers D, Alexanderson GL (1985) Mathematical people: profiles and interviews. CRC Press, Boca Raton

Alcock L, Weber K (2005) Proof validation in real analysis: inferring and checking warrants. J Math Behav 24(2):125–134

Alcock L, Hodds M, Roy S, Inglis M (2015) Investigating and improving undergraduate proof comprehension. Not Am Math Soc 62:742–752

Auslander J (2008) On the roles of proof in mathematics. In: Gold B, Simons RA (eds) Proofs and other dilemmas: mathematics and philosophy. Mathematical Association of America, Washington, DC, pp 61–77

Avigad J (2008) Understanding proofs. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 317–353

Avigad J (2020) Modularity in mathematics. Rev Symbol Logic 13(1):47–79

Avigad J (in press) Varieties of mathematical understanding. Bull Am Math Soc

Azzouni J (1994) Metaphysical myths, mathematical practice: the ontology and epistemology of the exact sciences. Cambridge University Press, Cambridge

Azzouni J (2004) The derivation-indicator view of mathematical practice. Philos Math 12(2):81–106

Berry D (2016) Proof and the virtues of shared enquiry. Philos Math 26(1):112–130

Beth EW, Piaget J (1966) Mathematical epistemology and psychology. Reidel, Dordrecht
Bruner JS (1960) On learning mathematics. Math Teacher 53(8):610–619
Burn B (1996) What are the fundamental concepts of group theory? Educ Stud Math 31(4):371–377
Burton L, Morgan C (2000) Mathematicians writing. J Res Math Educ 31(4):429–453
Clement J (2000) Analysis of clinical interviews: foundation and model viability. In: Kelly AE, Lesh R (eds) Handbook of research design in mathematics and science education. Lawrence Erlbaum, Mahwah, pp 547–589
Cowen CC (1991) Teaching and testing mathematics reading. Am Math Mon 98(1):50–53
Davis RB (1992) Understanding "understanding." J Math Behav 11(3):225–241
Delarivière S, Van Kerkhove B (in press) The mark of understanding: in defense of an ability account. Axiomathes
Dubinsky E (2000) Meaning and formalism in mathematics. Int J Comput Math Learn 5(3):211–240
Dubinsky E, Dautermann J, Leron U, Zazkis R (1994) On learning fundamental concepts of group theory. Educ Stud Math 27(3):267–305
Dubinsky E, Dautermann J, Leron U, Zazkis R (1997) A reaction to Burn's "what are the fundamental concepts of group theory?". Educ Stud Math 34(3):249–253
Ernest P (1985) The philosophy of mathematics and mathematics education. Int J Math Educ Sci Technol 16(5):603–612
Fallis D (1996) Mathematical proof and the reliability of DNA evidence. Am Math Mon 103(6): 491–497
Feferman S (2012) And so on...: reasoning with infinite diagrams. Synthese 186(1):371–386
Folina J (2018) Towards a better understanding of mathematical understanding. In: Piazza M, Pulcini G (eds) Truth, existence and explanation. Boston studies in the philosophy and history of science, vol 334. Springer, Cham, pp 121–146
Geist C, Löwe B, Van Kerkhove B (2010) Peer review and knowledge by testimony in mathematics. In: Löwe B, Müller T (eds) PhiMSAMP, philosophy of mathematics: sociological aspects and mathematical practice, volume 11 of texts in philosophy. College Publications, London, pp 155–178
Goldenberg EP (1995) Multiple representations: a vehicle for understanding understanding. Software goes to school: teaching for understanding with new technologies, 155–171
Goldin GA (2000) Affective pathways and representation in mathematical problem solving. Math Think Learn 2(3):209–219
Gowers T (2000) Two different cultures of mathematics. In: Arnold V, Atiyah M, Lax P, Mazur B (eds) Mathematics: Frontiers and perspectives. AMS Publications, New York
Grimm SR, Baumberger C, Ammon S (eds) (2016) Explaining understanding: new perspectives from epistemology and philosophy of science. Taylor & Francis, New York
Hammer D, Wildavsky A (1993) The open-ended, semistructured interview: an (almost) operational guide. In: Craftways: on the organization of scholarly work. Transaction Publishers, New Brunswick, NJ, pp 57–101
Hanna G (1983) Rigorous proof in mathematics education. OISE Press, Toronto
Hanna G, Barbeau E (2008) Proofs as bearers of mathematical knowledge. In: Explanation and proof in mathematics. Springer, Boston, pp 85–100
Hanna G, Larvor B (2020) As Thurston says? On using quotations from famous mathematicians to make points about philosophy and education. ZDM 52(6):1137–1147
Hardy GH (1940) Ramanujan, twelve lectures suggested by his life and work. Cambridge University Press, Cambridge
Harel G, Sowder L (1998) Students' proof schemes. CBMS Issues Math Educ III:234–283
Healy L, Hoyles C (2000) A study of proof conceptions in algebra. J Res Math Educ 31(4):396–428
Hiebert J (2013) The constantly underestimated challenge of improving mathematics instruction. In: Vital directions for mathematics education research. Springer, New York, pp 45–56
Hodds M, Alcock L, Inglis M (2014) Self-explanation training improves proof comprehension. J Res Math Educ 45(1):62–101
Inglis M, Aberdein A (2020) Are aesthetic judgements purely aesthetic? Testing the social conformity account. ZDM 52:1127–1136

Inglis M, Alcock L (2012) Expert and novice approaches to reading mathematical proofs. J Res Math Educ 43(4):358–390
Inglis M, Alcock L (2013) Skimming: a response to Weber and Mejía-Ramos. J Res Math Educ 44 (2):472–474
Inglis M, Mejía Ramos JP (in press) Functional explanation in mathematics. Synthese
Inglis M, Mejía-Ramos JP (2009) The effect of authority on the persuasiveness of mathematical arguments. Cogn Instr 27(1):25–50
Inglis M, Mejía-Ramos JP, Weber K, Alcock L (2013) On mathematicians' different standards when evaluating elementary proofs. Top Cogn Sci 5(2):270–282
Ko YY, Knuth EJ (2013) Validating proofs and counterexamples across content domains: practices of importance for mathematics majors. J Math Behav 32(1):20–35
Lakoff G (1987) The death of dead metaphor. Metaphor Symb 2(2):143–147
Lakoff G, Núñez R (2000) Where mathematics comes from. Basic Books, New York
Lampert M (1992) Practices and problems in teaching authentic mathematics. In: Oser F, Dick A, Patry J-L (eds) Effective and responsible teaching: the new synthesis. Jossey-Bass, San Francisco, pp 295–314
Larsen S (2004) Supporting the guided reinvention of groups and isomorphism: a developmental research project. Doctoral dissertation, Arizona State University
Larsen SP (2013) A local instructional theory for the guided reinvention of the group and isomorphism concepts. J Math Behav 32(4):712–725
Lew K, Fukawa-Connelly TP, Mejía-Ramos JP, Weber K (2016) Lectures in advanced mathematics: why students might not understand what the mathematics professor is trying to convey. J Res Math Educ 47(2):162–198
Lindsay DS, Simons DJ, Lilienfeld SO (2016) Research preregistration 101. APS Observer 29 (10) https://www.psychologicalscience.org/observer/research-preregistration-101
Löwe B, Van Kerkhove B (2019) Methodological triangulation in empirical philosophy. In: Inglis M, Aberdein A (eds) Advances in experimental philosophy in logic and mathematics. London, Bloombury Academic, pp 15–37
Mejia-Ramos JP, Fuller E, Weber K, Rhoads K, Samkoff A (2012) An assessment model for proof comprehension in undergraduate mathematics. Educ Stud Math 79(1):3–18
Mejía-Ramos JP, Weber K (2014) Why and how mathematicians read proofs: further evidence from a survey study. Educ Stud Math 85(2):161–173
Mejía-Ramos JP, Weber K (2020) Using task-based interviews to generate hypotheses about mathematical practice: mathematics education research on mathematicians' use of examples in proof-related activities. ZDM 52:1099–1112
Michener ER (1978) Understanding understanding mathematics. Cogn Sci 2(4):361–383
Mills M (2014) A framework for example usage in proof presentations. J Math Behav 33:106–118
Misfeldt M, Johansen MW (2015) Research mathematicians' practices in selecting mathematical problems. Educ Stud Math 89(3):357–373
Müller-Hill E (2010) Die epistemische Rolle formalisierbarer mathematischer Beweise. Formalisierbarkeitsbasierte Konzeptionen mathemaischen Wissens und mathematischen Rechtfertigung innerhalb einer sozioempirisch informierten Erkenntnistheorie der Mathematik. Unpublished doctoral dissertation. Rheinische Friedrich-Wilhelms-Universitat, Bonn
Nickerson RS (1985) Understanding understanding. Am J Educ 93(2):201–239
Nova (1997). The proof. Transcript downloaded from: http://www.pbs.org/wgbh/nova/transcripts/2414proof.html. Last downloaded May 14, 2021
Panse A, Alcock L, Inglis M (2018) Reading proofs for validation and comprehension: an expert-novice eye-movement study. Int J Res Undergraduate Math Educ 4(3):357–375
Paseau A (2015) Knowledge of mathematics without proof. Br J Philos Sci 66(4):775–799
Poincaré H (1952) Oeuvres de Henri Poincaré: Tome VIII. Gauthier-Villars, Paris
Polya G (1945) How to solve it: a new aspect of mathematical method. Princeton University Press, Princeton

Raman M (2004) Epistemological messages conveyed by three high-school and college mathematics textbooks. J Math Behav 23(4):389–404
Rasmussen C, Zandieh M, King K, Teppo A (2005) Advancing mathematical activity: a practice-oriented view of advanced mathematical thinking. Math Think Learn 7(1):51–73
Rav Y (1999) Why do we prove theorems? Philos Math 7(1):5–41
Rota GC (1997) The phenomenology of mathematical beauty. Synthese 111(2):171–182
Schoenfeld AH (1985) Mathematical problem solving. Elsevier, Orlando, FL
Schoenfeld AH (1992) Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In: Grouws DA (ed) Handbook of research in mathematics teaching and learning. Macmillan, New York, pp 334–370
Schwab JJ (1978) Education and the structure of the disciplines. In: Westbury J, Wilkof NJ (eds) Science, curriculum, and liberal education: selected essays. The University of Chicago Press, Chicago, London, pp 229–272
Selden A, Selden J (2003) Validations of proofs considered as texts: can undergraduates tell whether an argument proves a theorem? J Res Math Educ 34(1):4–36
Sfard A (1994) Reification as the birth of metaphor. Learn Math 14(1):44–55
Sfard A (1998) On two metaphors for learning and the dangers of choosing just one. Educ Res 27 (2):4–13
Sinclair N, Schiralli M (2003) A constructive response to "where mathematics comes from". Educ Stud Math 52(1):79–91
Sinclair N, Tabaghi SG (2010) Drawing space: mathematicians' kinetic conceptions of eigenvectors. Educ Stud Math 74(3):223–240
Soto-Johnson H, Hancock B, Oehrtman M (2016) The interplay between mathematicians' conceptual and ideational mathematics about continuity of complex-valued functions. Int J Res Undergraduate Math Educ 2(3):362–389
Styilianides AJ, Komatsu K, Weber K, Stylianidies GJ (in press) Teaching and learning authentic mathematics: the case of proving. To appear. In: Danesi M (ed) Handbook of cognitive mathematics. Springer Nature
Stylianides GJ, Stylianides AJ, Weber K (2017) Research on the teaching and learning of proof: taking stock and moving forward. In: Cai J (ed) Compendium for research in mathematics education. National Council of Teachers of Mathematics, Reston, pp 237–266
Thompson PW (2013) In the absence of meaning…. In: Leatham K (ed) Vital directions for mathematics education research. Springer, New York, NY, pp 57–93
Thurston WP (1994) On proof and progress in mathematics. Bull Am Math Soc 30:161–177
Weber K (2008) How mathematicians determine if an argument is a valid proof. J Res Math Educ 39 (4):431–459
Weber K (2010) Mathematics majors' perceptions of conviction, validity, and proof. Math Think Learn 12(4):306–336
Weber K, Alcock L (2005) Using warranted implications to understand and validate proofs. Learn Math 25(1):34–51
Weber K, Czocher J (2019) On mathematicians' disagreements on what constitutes a proof. Res Math Educ 21(3):251–270
Weber K, Mejía-Ramos JP (2011) Why and how mathematicians read proofs: an exploratory study. Educ Stud Math 76(3):329–344
Weber K, Mejía-Ramos JP (2013) On mathematicians' proof skimming: a reply to Inglis and Alcock. J Res Math Educ 44(2):464–471
Weber K, Mejía-Ramos JP (2014) Mathematics majors' beliefs about proof reading. Int J Math Educ Sci Technol 45(1):89–103
Weber K, Inglis M, Mejía-Ramos JP (2014) How mathematicians obtain conviction: implications for mathematics instruction and research on epistemic cognition. Educ Psychol 49(1):36–58
Wilkerson-Jerde MH, Wilensky UJ (2011) How do mathematicians learn math?: resources and acts for constructing and understanding mathematics. Educ Stud Math 78(1):21–43.

Argumentation in Mathematical Practice

Andrew Aberdein and Zoe Ashton

Contents

1 Introduction 2666
2 What Is an Argument? 2667
3 Proof as Argumentation 2672
3.1 Toulmin Layouts 2672
3.2 Argumentation Schemes 2674
4 Mathematical Reasoning as Argumentative 2678
4.1 Mathematics and Rhetoric 2678
4.2 Problem Choice 2679
4.3 Reasoning About Refutations 2680
4.4 Presenting Reasoning 2681
5 Mathematical Communities as Argumentative 2681
References 2684

Abstract

Formal logic has often been seen as uniquely placed to analyze mathematical argumentation. While formal logic is certainly necessary for a complete understanding of mathematical practice, it is not sufficient. Important aspects of mathematical reasoning closely resemble patterns of reasoning in non-mathematical domains. Hence the tools developed to understand informal reasoning, collectively known as argumentation theory, are also applicable to much mathematical argumentation. This chapter investigates some of the details of that application. Consideration is given to the many contrasting meanings of the word "argument"; to some of the specific argumentation-theoretic tools that

A. Aberdein (✉)
School of Arts and Communication, Florida Institute of Technology, Melbourne, FL, USA
e-mail: aberdein@fit.edu

Z. Ashton
Department of Philosophy, Ohio State University, Columbus, OH, USA

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_12

have been applied to mathematics, notably Toulmin layouts and argumentation schemes; to some of the different ways that argumentation is implicated in mathematical practices; and to the social aspects of mathematical argumentation.

Keywords

Argumentation schemes · Argumentation theory · Mathematical argumentation · Mathematical reasoning · Mathematical rhetoric · Problem choice · Toulmin layouts

1 Introduction

Since logic developed the tools to adequately represent formal derivations, many philosophers of mathematics have been tempted to conclude that formal derivation suffices to account for all interesting features of mathematical practice. However, there have always been other philosophers who perceived the shortcomings of such a reduction. Here, for example, is Henri Poincaré:

> If you are present at a game of chess, it will not suffice, for the understanding of the game, to know the rules for moving the pieces. That will only enable you to recognize that each move has been made conformably to these rules, and this knowledge will truly have very little value. Yet this is what the reader of a book on mathematics would do if he were a logician only. To understand the game is wholly another matter; it is to know why the player moves this piece rather than that other which he could have moved without breaking the rules of the game. It is to perceive the inward reason which makes of this series of successive moves a sort of organized whole. This faculty is still more necessary for the player himself, that is, for the inventor (Poincaré 1913, 218).

One response to this limitation of formal logic is to recognize an analogy with a similar limitation in another domain: formal logic is also an imperfect tool for understanding everyday reasoning. Solutions have been proposed for that problem: systems of informal logic, argumentation theory, or dialectic have been devised since antiquity to address ordinary reasoning (van Eemeren et al. 2014). Hence some philosophers of mathematical practice have reasoned that these theories might also lend themselves to the understanding of mathematical reasoning (Aberdein 2009).

This chapter surveys the uses to which argumentation theory has been put in order to understand mathematical practice. Section 2 addresses the many ambiguities implicit in the word "argument" – and how they are specifically related to mathematics. Section 3 discusses two prominent proposals for the application of particular tools from argumentation theory to mathematics: Toulmin layouts and argumentation schemes. Section 4 considers the argumentative aspects of mathematical practices beyond proof and Sect. 5 focuses on the contribution of communities of mathematical practice to such argumentation.

2 What Is an Argument?

Several overlapping distinctions may be drawn in our understanding of arguments.

- **Argument-that/Argument-about:** A sequence of statements whereby premises offer support for a conclusion is an argument. But, an exchange of conflicting views held by different people is also an argument. We may refer to the former as an *argument-that* and to the latter as an *argument-about*. Arguments-that are also known as arguments$_1$ and arguments-about as arguments$_2$ (O'Keefe 1977). As Michael Gilbert helpfully glosses the distinction, "one person makes an argument$_1$ and [at least] two people have an argument$_2$" (Gilbert 2014, 21). Although the most familiar mathematical arguments tend to be arguments-that, arguments-about arise in mathematics too, such as priority disputes, contested axioms or principles, or debates over the legitimacy of a technique or the admissibility of a proof. A salient recent example is the contested status of Shinichi Mochizuki's claimed proof of the *abc* conjecture (Aberdein 2023).
- **Process/Product:** Arguments-that are sometimes represented as products of the argument-about process, but this is arguably a mischaracterization: "If, as part of organizing the domain of argumentation theory, we merely want to distinguish acts of arguing from arguments-as-objects, we should not use the misleading process/product labels to do so. At the very least such labels imply a relationship that does not exist and so distort our perceptions of the domain of study" (Goddu 2011, 87). Nonetheless, we may usefully distinguish between an argument understood as an act of arguing and the informational trace that act leaves behind (a transcript, a recording, etc.), also often called an argument (Sundholm 2012, 948). This distinction straightforwardly applies to mathematics. (For further reflection on proofs as acts of proving or proof-events, see Goguen 2001; Stefaneas and Vandoulakis 2014, 2015.)
- **Monologue/Dialogue/Polylogue:** The number of participants in an argument may vary considerably. Most attention has traditionally been paid to monologues and dialogues: arguments-that are characteristically presented as monologues, arguments-about as dialogues. If dialogues are understood loosely, as also covering argumentation with more than two participants, that distinction would exhaust the options. However, some authors have made a case for differentiating the two-participant dialogue from the many-participant polylogue (Lewiński 2014). The Polymath Project, a series of crowd-sourced proofs of open conjectures, is a rich source for research on mathematical polylogues (Allo et al. 2021).
- **Small scale/Large scale:** Arguments can vary significantly in scale. The scale of an argument may be measured in several conceptually distinct ways. The duration of an argument may range from arguments that take seconds to arguments which last for hundreds of years. The size of an argument may range from a few inferences expressed in short sentences about a simple issue to inferentially complex structures involving a great many very long sentences. Likewise,

mathematical proofs can vary in length from a few lines to tens or even hundreds of pages. In exceptional cases, proofs can be so long as to defy the capacity of any single mathematician to survey the whole (Coleman 2009). This can arise in proofs achieved by traditional means, such as that of the classification of finite simple groups, the components of which comprised thousands of pages in several hundred articles by dozens of mathematicians (Steingart 2012). It is even more acute in the case of computer-assisted proofs, such as that of the four-color theorem or the Kepler conjecture, or indeed the subsequent computer verification of these proofs (Gonthier 2008; Hales et al. 2017).

- **Static/Dynamic:** Static arguments have achieved a final and definite form; dynamic arguments are fluid and ongoing. In general, the evolution of knowledge may be understood as the product of dynamic argumentation. Dynamic arguments are common in everyday life – but they are also central to the development of scientific thought. A well-known and influential analysis of a dynamic argument in mathematics is Imre Lakatos's *Proofs and Refutations*. Lakatos shows how the protracted search for a proof of the Descartes–Euler conjecture, which relates the quantities of vertices, edges, and faces of convex polyhedra, involved significant redefinition of most of the concepts used in that conjecture (Lakatos 1976). (For further discussion, see Sect. 4.3 below or Reyes 2021.)
- **Centralized/Distributed:** Arguments may be either centralized or distributed with respect to several factors including time, people, space, and media. A highly centralized argument may be restricted to what one person communicates in one place at one time in one mode of expression. But arguments may involve a varying cast of arguers and be drawn out over long periods, in multiple locations and media. Large-scale arguments are characteristically distributed: As a mathematical example, the classification of finite simple groups involved many years of work by a large, geographically widespread collective of mathematicians (Steingart 2012).
- **Sequential/Parallel:** Logic, whether formal or informal, tends to reconstruct arguments as a linear sequence of premises from which intermediate statements are derived, culminating in a final conclusion. (Some logical systems prefer to invert this sequence.) However, argumentation in natural contexts often occurs in a more parallel fashion: Several strands of argument may be developed simultaneously, the conclusion may be derived from initial premises and then reinforced by subsequent subarguments, and so forth. Parallel arguments are also more likely to arise in projects with many participants, whether by accident or design: "To permit a large collaboration, . . . long proofs have been broken up into series of shorter lemmas" (Hales et al. 2017, 11). The educationists Christine Knipping and David Reid have proposed more fine-grained subdivisions of parallel arguments in mathematics (Knipping and Reid 2019; for further discussion, see Sect. 3.1 below.)

The seven overlapping distinctions addressed so far have received unequal interest in the philosophy of mathematics. The formal logical approach mentioned in the introduction best coheres with a small-scale, static, centralized, sequential product like a published proof. Much recent work in the philosophy of mathematical

practice focuses on mathematical arguments along the other dimensions. Some further distinctions arise from consideration of the goals of the arguers:

- **Persuasive/Directive/Polemic:** That arguments may be distinguished by their objective is an ancient idea: Aristotle distinguished forensic arguments (concerned with past acts), display arguments (concerned with present circumstances), and deliberative arguments (concerned with future acts) (Aristotle 1991, 1358b). Erik C. W. Krabbe and Jan Albert van Laar propose an updated distinction between three different functions of reasoning: persuasive (to convince the other party), directive (to get the other party to act), and polemic (to intimidate the other party) (Krabbe and van Laar 2007, 29 f.). They contrast these "inherently argumentative" functions with three further functions of reasoning: explanatory (to enhance understanding), explorative (to investigate connections between statements), and probative (to establish new knowledge).
- **Adversarial/Nonadversarial:** On a strict interpretation, all arguments begin in conflict: a difference in belief, or concerning how to act, or of some other kind. On a broader interpretation, arguments need not be strictly adversarial, hence they may proceed from other situations, including shared uncertainty or one party knowing what another does not.

These nine different dimensions of comparison interact in important ways. Firstly, they are not pairwise independent. For example, as the scale of an argument increases, we may tentatively expect both the number of participants and the number of different objectives to increase, but the likelihood of the argument being either static, centralized, or sequential to tend to zero. Secondly, the last two dimensions combine to produce what Douglas Walton calls *dialogue types* (see Fig. 1, adapted from Walton and Krabbe 1995, 81). Further complicating this picture, Walton and Krabbe observe that dialogues can shift from one type to another (for example, an inquiry might turn into a persuasion dialogue if one inquirer becomes an advocate for a particular result) or be embedded in a dialogue of a different type (a deliberation over which course of action to pursue might contain an inquiry into the merits of one action, say). Walton maintains that arguments may arise in any dialogue type (Walton 1998); Krabbe is more conservative and regards argumentation as restricted to adversarial contexts (the lefthand fork of Fig. 1) (Krabbe and van Laar 2007, 33).

How do these distinctions apply to argumentation in mathematics? The most widely discussed case is that of mathematical proof, which many authors have maintained is intrinsically argumentative. But even here, we may distinguish multiple distinct activities which give rise to arguments of different kinds. For example, Krabbe observes the following distinct stages:

1. Thinking up a proof to convince oneself of the truth of some theorem
2. Thinking up a proof in dialogue with other people (inquiry dialogue; probative functions of reasoning)
3. Presenting a proof to one's fellow discussants in an inquiry dialogue (persuasion dialogue embedded in inquiry dialogue; persuasive and probative functions of reasoning)

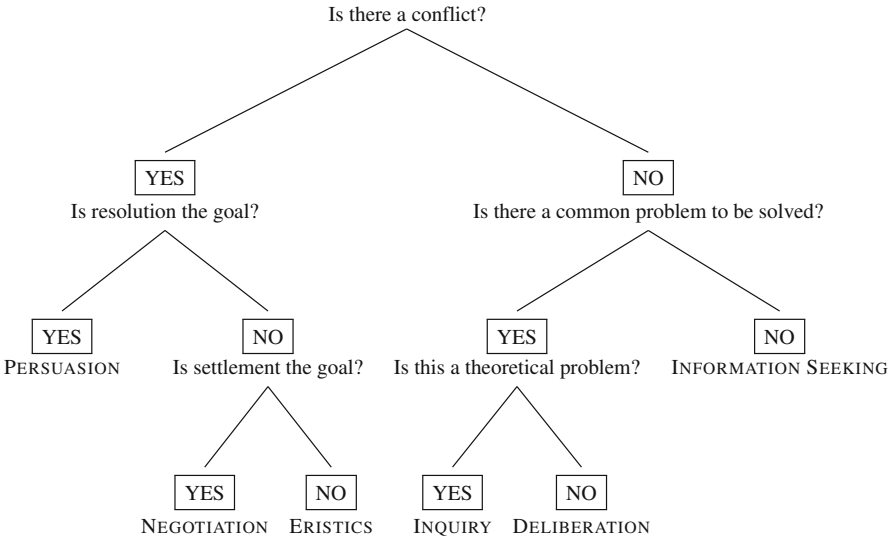


Fig. 1 Determining the type of dialogue (after Walton and Krabbe 1995, 81)

4. Presenting a proof to other mathematicians, e.g., by publishing it in a journal (persuasion dialogue; persuasive and probative functions of reasoning)
5. Presenting a proof when teaching (information-seeking and persuasion dialogue; explanatory, persuasive, and probative functions of reasoning) (Krabbe 2008, 457)

This sequence is familiar from many mathematicians' descriptions of the proving process, although in actual examples some steps may be repeated as proof attempts come unstuck.

In deference to his intrinsically adversarial conception of argument, Krabbe only considers three dialogue types as hospitable to proofs. In other work, one of us has suggested that proofs (or other mathematical arguments) may be found in other dialogue types too (see Table 1, from Aberdein 2021, 165; see also Aberdein 2007b, 148). "Oracular" information-seeking owes its inspiration to an influential aside of Alan Turing concerning a machine "supplied with some unspecified means of solving number-theoretic problems; a kind of oracle as it were" (Turing 1939, 172). An oracle is a "black box" – it supplies answers but not explanations. For some skeptics of computer-assisted proofs, this is a compelling analogy for the role that the computer plays in such proofs (e.g., Tymoczko 1979). Deliberation differs from inquiry in seeking only a provisional conclusion. Mathematicians aspire to more permanent stability for their results. Nonetheless, there are circumstances where they are obliged to settle for less than they would wish, despite the rigor of their arguments. These include the "architectural conjectures" upon which many mathematical research programs depend (Mazur 1997, 198). Negotiation characteristically adds resource sensitivity to the provisional outcome typical of deliberation.

Table 1 Some mathematical dialogue types

Dialogue type	Initial situation	Main goal	Goal of proponent	Goal of respondent
Inquiry	Open-mindedness	Prove or disprove conjecture	Contribute to main goal	Obtain knowledge
Persuasion	Difference of opinion	Resolve difference of opinion with rigor	Persuade respondent	Persuade proponent
Pedagogical information-seeking	Respondent lacks information	Transfer of knowledge	Disseminate knowledge of results and methods	Obtain knowledge
Oracular information-seeking	Proponent lacks information	Transfer of knowledge	Obtain information	Inscrutable
Deliberation	Open-mindedness	Reach a provisional conclusion	Contribute to main goal	Obtain warranted belief
Negotiation	Difference of opinion	Exchange resources for a provisional conclusion	Contribute to main goal	Maximize value of exchange
Eristic	Personal conflict	Reveal deeper conflict	Win in the eyes of onlookers	Win in the eyes of onlookers

While idealized accounts of mathematical practice disregard such factors, they are unavoidable in some contexts, especially in applied mathematics. (And it has been controversially suggested that "semi-rigorous" proofs might come with price tickets, costing the computational resources necessary for certainty [Zeilberger 1993].) Even eristic dialogues can be the context for mathematical reasoning, as demonstrated in the mathematically inventive quarrels of early modern mathematicians such as Girolamo Cardano and Niccolò Tartaglia (Toscano 2020).

A final distinction among the different senses of argument cuts across most of those discussed above and arises not from argumentation theory but directly from mathematical practice. Here it is presented by Joel David Hamkins:

- **Hard/Soft:** "A hard argument is one that is technically difficult; perhaps it involves a laborious construction or a difficult calculation; perhaps it involves bringing disparate fine details together in just the right combination in order to succeed; or perhaps it involves proving various specific facts about a comparatively abstract construction, perhaps relating disparate levels of abstraction. A soft argument, in contrast, is one that appeals only to very general abstract features of the situation, and one needs hardly to construct or compute anything at all" (Hamkins 2020, 166).

Hamkins's use of "hard" and "soft" echoes G. H. Hardy's division of analysis into "the 'hard, sharp, narrow' kind as opposed to the 'soft, large, vague' kind" (Hardy 1929, 64). It also owes something to Alexander Grothendieck's celebrated analogy between two strategies for solving mathematical problems and two ways of opening

a nut: cracking it with a hammer or softening it in water until it opens with light pressure (McLarty 2007, 301). Grothendieck favored the latter strategy, of immersing problems in a much wider theory from which a solution could (eventually) be readily inferred. This in turn is suggestive of Freeman Dyson's division of mathematicians into birds and frogs: "Birds fly high in the air and survey broad vistas of mathematics out to the far horizon. They delight in concepts that unify our thinking and bring together diverse problems from different parts of the landscape. Frogs live in the mud below and see only the flowers that grow nearby. They delight in the details of particular objects, and they solve problems one at a time" (Dyson 2009, 212). Hard arguments play to the tightly focused strengths of the frogs; soft arguments require the birds' sweeping perspective.

3 Proof as Argumentation

The view mentioned in Sect. 1, that formal derivation suffices to account for all interesting features of mathematical practice, squarely focuses on the role of mathematical proof. Pure mathematicians trade in proofs. But, even in the domain of mathematical proof, formal logic does not capture all there is of interest. Toulmin layouts and argumentation schemes, as we will see below, are useful methods of examining mathematical proofs. But, notably, such techniques tease apart the descriptive and normative aspects of (apparent) proof. Whereas a formal derivation is either correct or not, strictly speaking, a derivation at all, the methods of argumentation theory provide the means to describe proofs independently of whether they succeed as proofs.

3.1 Toulmin Layouts

One of the most influential treatments of informal argumentation is that of Stephen Toulmin (1958). His "layout" can represent deductive inference, but encompasses many other species of argument besides. In its simplest form, shown in Fig. 2a, the

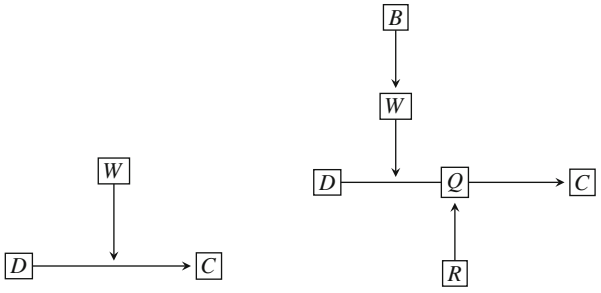


Fig. 2 Toulmin layouts

(a) Basic Layout (b) Full Layout

layout represents the derivation of a claim (*C*), from data (*D*), in accordance with a warrant (*W*). This DWC pattern resembles a deductive inference rule, such as *modus ponens*, but it can be used to represent looser inferential steps. The differences between the types of inference which the layout may represent are made explicit by the additional elements of the full layout shown in Fig. 2b. The warrant is justified by its dependence on backing (*B*), possible exceptions, or rebuttals (*R*), are indicated, and the resultant force of the argument is stated in the qualifier (*Q*). Hence the full layout may be understood as "Given that *D*, we can *Q* claim that *C*, since *W* (on account of *B*), unless *R*." In a frequently cited example (derived from Toulmin 1958, 104), "Given that HARRY WAS BORN IN BERMUDA, we can PRESUMABLY claim that HE IS BRITISH, since ANYONE BORN IN BERMUDA WILL GENERALLY BE BRITISH (on account of VARIOUS STATUTES . . .), unless HIS PARENTS WERE ALIENS, SAY."

Toulmin wrote *The Uses of Argument* in England in the 1950s as a critique of what he perceived as a formalizing trend in contemporary philosophy; the Toulmin layout was subsequently adopted by communication theorists in America in the 1960s and 1970s; from there it seems to have passed to mathematics educationists in Germany in the 1990s. (Although Toulmin himself briefly considers a mathematical example, Theaetetus's proof that there are exactly five platonic solids [Toulmin et al. 1979, 89]. For discussion, see [Aberdein 2005, 290 ff.].) In particular, Götz Krummheuer is usually credited with the first sustained application of the Toulmin layout to mathematical argumentation (Krummheuer 1995; for a recent survey, see Krummheuer 2015).

Toulmin draws a distinction between analytic and substantial arguments depending on whether the claim is already at least implicit in the backing. This may suggest that only analytic arguments occur in mathematics. But that overlooks the emphasis Toulmin places on how out of the ordinary analytic arguments are: "it begins to be a little doubtful whether any genuine, practical argument could ever be properly analytic" (Toulmin 1958, 127). Granted, immediately after this passage, Toulmin places arguments in (pure) mathematics among the analytic arguments. However, the philosophers and educationists who have applied Toulmin's work to mathematics endorsed his account of argument, not his philosophy of mathematics. They have typically treated (at least some) mathematical arguments as substantial rather than analytic. For example, "It is the substantial argumentation that is seen here as more adequate for the reconstruction" of mathematics classroom situations (Krummheuer 1995, 236). In other words, the part of Toulmin's work that we should apply to many mathematical arguments is what he has to say about nonmathematical arguments.

Much subsequent work applying the Toulmin layout to mathematical reasoning has concerned ways in which it may be extended to cover cases that he does not directly address – including many of the less favored sides of the distinctions drawn in Sect. 2 above. Toulmin's own later work discusses how more than one layout may be chained together to represent a multistep argument (Toulmin et al. 1979, 79). Other authors have shown how more complicated structures, including linked and convergent arguments, may be represented by combinations of layouts (Aberdein

2006, 214). Christine Knipping and David Reid have paid particular attention to larger-scale structures of parallel argument, distinguishing source-like argumentation structure, where "arguments and ideas arise from a variety of origins, like water welling up from many springs," from "reservoir structure," wherein arguments "flow towards intermediate target-conclusions that structure the whole argumentation into parts that are distinct and self-contained," and spiral argumentation structure in which "the final conclusion is proven in many ways" (Knipping and Reid 2019, 18 ff.). Matthew Inglis and colleagues have argued persuasively for the relevance of the full Toulmin model to mathematical arguments, rather than the simplified DWC or DWBC versions that have found widest application among mathematics educationists (Inglis et al. 2007). In particular, they observe that nondeductive warrants can play an essential role in mathematical argumentation, just so long as this is signaled by the use of appropriate qualifiers. Work has also been done to explore the connections between the Toulmin layout and other models of argument applicable to mathematics (Pease and Aberdein 2011), or to link it to broader conceptual analyses of mathematical cognition, such as the "ck¢-enriched" Toulmin model (Pedemonte and Balacheff 2016).

3.2 Argumentation Schemes

Argumentation schemes are stereotypical patterns of reasoning. Although their origins lie in the topoi of classical rhetoric, they have lately found extensive application in the analysis and evaluation of argumentation. This revival is substantially due to the work of the argumentation theorist Douglas Walton. Most attention has been paid to defeasible schemes typical of informal reasoning, although deductive inference rules can also be considered special cases of argumentation schemes. The defeasible nature of the reasoning is not made explicit among the premises, but captured by an additional device, critical questions, which point to possible exceptions. Many of the defeasible schemes may ultimately be understood as more or less specialized instances of the very general scheme of *Defeasible Modus Ponens* (Walton et al. 2008, 366). In Scheme 1, we have presented it in a way designed to bring out its similarities to the Toulmin layout:

Argumentation Scheme 1 *Defeasible Modus Ponens*
Data P.
Warrant As a rule, if *P*, then *Q*.
Therefore,
Qualifier presumably,
Conclusion . . . *Q*.

Critical Questions

1. *Backing:* What reason is there to accept that, as a rule, if *P*, then *Q*?
2. *Rebuttal:* Is the present case an exception to the rule that if *P*, then *Q*?

The strength of the argumentation scheme approach lies in its heterogeneity: An influential (but not exhaustive) survey identifies over one hundred different schemes (Walton et al. 2008, 308 ff.). Hence schemes are typically presented with much greater specificity than Scheme 1. For example, Scheme 2 is a scheme for Argument from Analogy:

Argumentation Scheme 2 *Argument from Analogy*
Similarity Premise Generally, case C_1 is similar to case C_2.
Base Premise A is true (false) in case C_1.
Conclusion A is true (false) in case C_2.

Critical Questions

1. Are there differences between C_1 and C_2 that would tend to undermine the force of the similarity cited?
2. Is A true (false) in C_1?
3. Is there some other case C_3 that is also similar to C_1, but in which A is false (true)? (Walton et al. 2008, 315).

Analogies in mathematics can be formal and thereby capable of rigorous proof (for a specific example and further discussion of this scheme, see Aberdein 2013b, 373). They can also be informal heuristics, for example, the "strong analogy between the pluralist nature of set theory and what has emerged as an established plurality in the foundations of geometry" (Hamkins 2020, 296; for further discussion of this analogy, see Berry 2020). Analogical reasoning has been a topic of wide interest and mathematical analogies in particular have been the subject of focused discussion (Schlimm 2008; Bartha 2013; Priestley 2013).

In more recent work, Walton proposed a partial taxonomy of schemes (Walton and Macagno 2015, 22), although he acknowledged that some schemes remained outside this classification. Many of these schemes have been applied to mathematical arguments: Table 2 is based on (Walton and Macagno 2015, 22, Table 1), but adds references to prior work in which such applications have been developed. Just as Walton's classification of schemes is incomplete, so is their application to mathematics. Some of these schemes may be of limited usefulness in the analysis of specifically mathematical argumentation, but others have direct application.

By varying which schemes are treated as admissible, it is possible to capture different conceptions of mathematical rigor. To this end, one of us has proposed a threefold distinction among the ways schemes may be employed in mathematical reasoning (Aberdein 2013b, 366 f.):

- **A-schemes** correspond directly to derivation rules. (Equivalently, we could think in terms of a single A-scheme, the "pointing scheme" which picks out a derivation whose premises and conclusion are formal counterparts of its data and claim.)
- **B-schemes** are exclusively mathematical arguments: high-level algorithms or macros. Their instantiations correspond to substructures of derivations rather

Table 2 Walton and Macagno's partial classification of schemes (adapted from Walton and Macagno 2015, 22), with prior applications to mathematical argumentation indicated

Discovery arguments	Applying rules to cases	Practical reasoning	Source-dependent arguments
1. Arguments establishing rules	1. Arguments based on cases	1. Instrumental argument from practical reasoning	1. Arguments from position to know
Argument from a random sample to a population	*Argument from an established rule*[f]	Argument from action to motive	a. *Argument from expert opinion*[a,c,i]
Argument from best explanation	*Argument from verbal classification*[b,i,j]	2. Argument from values	b. *Argument from position to know*[c]
2. Arguments finding entities	Argument from cause to effect	Argument from fairness	Argument from witness testimony
Argument from sign[f,h,i,k]	2. Defeasible rule-based arguments	3. *Value-based argument from practical reasoning*[g]	2. Ad hominem arguments
Argument from ignorance[c]	*Argument from example*[c,e,i,j]	a. *Argument from positive or negative consequences*[c,i,j]	a. Direct ad hominem
	Argument from analogy[d,j]	Argument from waste	b. Circumstantial ad hominem
	Argument from precedent[k]	Argument from threat	Argument from inconsistent commitment
	3. Chained arguments connecting rules and cases	Argument from sunk costs	Arguments attacking personal credibility
	Argument from gradualism[c]		i. Arguments from allegation of bias
	Precedent slippery slope argument		ii. Poisoning the well by alleging group bias
	Sorites slippery slope argument		3. Arguments from popular acceptance
			Argument from popular opinion[b]
			Argument from popular practice[b]

[a]Aberdein (2007a)
[b]Aberdein (2010)
[c]Aberdein (2013a)
[d]Aberdein (2013b)
[e]Aberdein (2019)
[f]Aberdein (2021)
[g]Cantù (2013)
[h]Dove (2009)
[i]Metaxas (2015)
[j]Metaxas et al. (2016)
[k]Pease and Aberdein (2011)

than individual derivations (and they may appeal to additional formally verified propositions).

- **C-schemes** are even looser in their relationship to derivations, since the link between their data and claim need not be deductive. Specific instantiations may still correspond to derivations, but there will be no guarantee that this is so and no procedure that will always yield the required structure even when it exists. Thus, where the qualifier of A- and B-schemes will always indicate deductive certainty, the qualifiers of C-schemes may exhibit more diversity. Indeed, different instantiations of the same scheme may have different qualifiers.

B-schemes are essentially what Saunders Mac Lane calls "processes of proof" or general rules, algorithmic procedures that are ultimately reducible to elementary logical inferences although not necessarily so analyzed by the mathematicians who routinely employ them: "whenever a group of elementary processes of proof occurs repeatedly in the course of many proofs, it is desirable to formulate this group of steps once for all as a new process" (Mac Lane 1935, 123). Much more recently, Yacin Hamami has used B-schemes, which he terms "hl-rules" or higher-level rules of inference, to defend the "standard view" of mathematical rigor, that rigorous proofs are those for which there is a routine translation into a formal derivation (Hamami 2022).

Hamami's account of rigor has three components: a descriptive thesis, a normative thesis, and a philosophical thesis asserting the conformity of the other two theses. Relative to some mathematical practice $\mathcal{M}$, these theses may be stated as follows. The descriptive thesis states that a mathematical proof P is rigorous$_D$ if and only if for every mathematical inference I in P, there exist $D \in \mathcal{D}^*$ and $V_1, \ldots, V_n \in \mathcal{V}^*$ such that (1) $D(I) = \langle I_1, \ldots, I_n \rangle$ and (2) $Vi(Ii) = \texttt{valid}$ for all $i \in [\![1, n]\!]$, where the set $\mathcal{D}^*$ consists of decomposition (or proof search) processes, whereby a mathematical inference may be rewritten as a sequence of immediate mathematical inferences, and the set $\mathcal{V}^*$ consists of hl-rules, whereby immediate mathematical inferences are judged valid if they correspond to instances of the hl-rules. The normative thesis states that a mathematical proof P is rigorous$_N$ if and only if P can be routinely translated into a formal proof. Hamami defines "routine translation" as the composition of three successive translations between proofs understood at four levels of granularity: vernacular level proofs, comprised of inferences presented at the level of formality normal to $\mathcal{M}$; higher-level proofs, comprised of inferences instantiating hl-rules in $\mathcal{M}$; intermediate-level proofs, comprised of inferences instantiating primitive rules of inference in $\mathcal{M}$; and lower-level proofs, comprised of inferences instantiating rules of inference in a purely formal system. The conformity thesis states that if P is rigorous$_D$ then P is rigorous$_N$. Hamami's account of rigor corresponds to one of four alternatives that one of us has discussed elsewhere (Aberdein 2013b, 369). It may be contrasted with the more conservative policy of only admitting A-schemes and thereby treating only formalized mathematics as truly rigorous and more liberal options in which C-schemes are also admitted, thereby tolerating a greater diversity of innovation in mathematical proof.

4 Mathematical Reasoning as Argumentative

In Sect. 3, we saw that to understand proof through the lens of argumentation theory is to see it in the context of a greater diversity of types of mathematical argument. While proofs are a vital component of modern mathematical practice, it is not the only aspect of mathematical reasoning to which argumentation theory may be applied. We now turn to these other aspects. Our discussion begins with general claims about mathematical reasoning and rhetoric. We then turn to a number of issues which surround proof including problem choice, reasoning about refutations, and presentation of mathematical information.

4.1 Mathematics and Rhetoric

A first connection between mathematical reasoning and argumentation theory involves rhetoric. Rhetoric, the study of the art of persuasive argument, has long been set in opposition to mathematics. It was thought that mathematics, as an objective, rational, and atemporal field, has little to do with the study of persuasion. But a number of authors have challenged this idea. An early paper by Philip J. Davis and Reuben Hersh identified two areas where mathematics and rhetoric intersect (Davis and Hersh 1986). The first involves importing or applying mathematics to theories in the social sciences, such as psychodynamics and economics. Such appeals to mathematiziation are rhetorical and argumentative moves, but they are not argumentation within mathematical reasoning. However, as Davis and Hersh point out, there is rhetoric in mathematics too, since "all proofs are incomplete, from the viewpoint of formal logic" (Davis and Hersh 1986, 66). Each proof requires rhetorical elements to convince the intended audience that the result is true. The mathematician relies on the audience's background knowledge or intuition to patch up gaps in the proof and understand the intentions of the prover.

Like Davis and Hersh, Edward Schiappa discusses multiple ways in which mathematical reasoning can be rhetorical. The first intersection, again, is the rhetorical use of mathematics. Mathematical methods can be used to persuade in a variety of arguments. Schiappa cites examples ranging from mundane advertisement – "four out of five dentists agree" – to the discovery of Neptune to argue that mathematical reasoning plays a role in lending credibility to arguments outside its purview (Schiappa 2021). The second intersection is the role of rhetoric within mathematics: the argumentative and stylistic modes of persuasion in mathematical arguments. Each aspect of mathematical practice historically required a social and persuasive component. Acceptance of axioms and stipulated definitions depends on the audience one aims to persuade. Even the available concepts which mathematicians reason about can be the result of rhetoric. For example, G. Mitchell Reyes argues that, since there was no empirical or geometric verification for infinitesimals, their substance was found in the rhetorical argumentation which surrounded them (Reyes 2004). Schiappa also argues that the language of mathematics is rhetorical since it is human-made. Symbols and concepts like the infinitesimal or the number zero were

additions where "social acceptance was not assured, meaning was contested, and alternatives competed" (Schiappa 2021, 49). Relatedly, in this collection, Reyes examines the relationship between rhetoric and mathematics in Lakatos's *Proofs and Refutations* (Reyes 2021).

4.2 Problem Choice

In addition to the rhetorical components pervasive in nonproof, there is room for argumentation in the process around proofs. Perhaps the most fundamental part of solving a problem, and of proving a theorem, is selecting a problem. Problem choice in mathematics has frequently been associated with the beauty and intrinsic worth of problems. Under this view, there is a special sensibility mathematicians employ when choosing a subject. According to Jacques Hadamard, mathematicians "feel that such a direction of investigation is worth following; [they] feel that the question in itself deserves interest . . . everybody is free to call or not to call that a feeling of beauty" (Hadamard 1945, 127). This approach to problem choice has been supplemented in recent years. Morten Misfeldt and Mikkel Willum Johansen interviewed research mathematicians about the factors which influence problem choice (Misfeldt and Johansen 2015). In line with Hadamard's claims, certain external factors, like funding, were not very influential on problem choice. Misfeldt and Johansen found that problem choice was largely motivated by personal interest, perceived ability to solve the problem, and the values of the community. It was not enough for mathematicians to be interested in the problem, they had to be assured that the mathematical community would be interested. Mathematicians expressed "the need to have an audience—the right audience—for their work" (Misfeldt and Johansen 2015, 368). This connection between an audience and an arguer is an area which argumentation is primed to explore.

Elsewhere, one of us has looked at problem choice through an argumentative lens by applying Chaïm Perelman and Lucie Olbrechts-Tyteca's notion of the contact of minds to problem choice in mathematics (Ashton 2018). The contact of minds is a set of conditions which must be met before argumentation can occur. Contact of minds requires four things: the arguer must attach importance to the audience, the speaker must not be beyond question, the audience must be willing to consider being convinced, and they must share a common language. Contact of minds is a prerequisite of any argumentation and mathematical arguments are no different. Ashton argues that it is an audience-based factor that influences problem choice alongside traditional considerations of beauty, intrinsic worth, and practical benefits. But the contact of minds was not originally meant to be applied to mathematics: Perelman and Olbrechts-Tyteca specifically oppose their study of argumentation to the mathematical sciences. Argumentation, they claimed, was distinct from mathematics since argumentation was social and its conclusions were probable (Perelman and Olbrechts-Tyteca 1969). But barring mathematical practice from the domain of argumentation is inapproporiately limiting (Dufour 2013; Ashton 2018, 2021). Misfeldt and Johansen's interviews indicate that mathematicians consider the

interest of other mathematicians while choosing problems. Given that problem choice is broader than merely the structure of proof, a choice about what to research is a choice of what to argue about. In this way, the process of choosing a problem to research is distinctly social and related to the mathematical community (Ashton 2018; discussed further in Sect. 5).

4.3 Reasoning About Refutations

The influential informal logician Ralph Johnson (2000) follows the rhetoricians Perelman and Olbrechts-Tyteca (1969) in denying that proofs can be arguments because there are features of proof, such as necessity, that are incompatible with the social dimension of arguments. For Perelman and Olbrechts-Tyteca, mathematics deals in demonstrations which are mechanically checkable and result in certainty, regardless of the audience. Mathematical reasoning, for them, does not involve uncertainty or controversy. Likewise, Johnson viewed proofs as lacking a dialectical tier. For Johnson, arguments have two components. The first component of argument is the illative core which is a discursive structure where reasons support the conclusion. But this logical aspect alone is not a full argument. A dialectical tier must supplement the illative core. Dialectical tiers are where arguers discharge their dialectical obligations by responding to objections, criticisms, or implications of their view. According to Johnson, proofs do not have dialectical tiers since they are conclusive and anyone trained in the field must recognize that they are conclusive. But refutations are a natural part of mathematical reasoning. Reasoning about purported refutations may best be understood in terms of argumentation.

In an idealized view of proof, one in which proof cannot be an argument, the reader of a proof is an expert who "needs nobody to grasp a proof, otherwise she is not an expert" (Dufour 2013, 69). Such a view, according to Dufour, ignores the important role of checking a proof for correctness. There may still be legitimate room for refutations at this stage. As Fallis (2003) notes, proofs have many intentional gaps. The gaps are purportedly something a reader could fill in, with enough time and background knowledge. Verifying that a proof is correct requires checking these gaps. If the gaps cannot be traversed by either the reader or author, the proof itself may be refuted. Sometimes large gaps may result in the apparent refutation of otherwise good proofs. According to Dufour, Galois experienced such problems of communication (Dufour 2013, 71). Galois's exaggerated brevity led his audience to believe that the mathematics itself was insufficiently developed. Proofs may be refuted because of unintelligible or untraversable gaps. These gaps, and their relationships to audience understanding, are evidence that reasoning about the correctness of proofs, and their refutations, involves argumentation.

Contra Johnson, Ian Dove has argued that proofs do have dialectical tiers (Dove 2007). In particular, Dove argues that the method of monster barring seen in Lakatos's *Proofs and Refutations* is part of a dialectical tier (Lakatos 1976). Lakatos reconstructs a series of purported proofs of Euler's formula. Cauchy's proof of Euler's formula for polyhedra is considered and then counterexamples are raised

which are not convex and not simply connected. Dove argues that there is a dialectical tier for Cauchy's proof of Euler's formula since (a) objections are raised to the proof and (b) the objections receive responses in the proof. The method of response is to bar exceptions to the formula so that Euler's formula holds for simply connected, convex polyhedra. In other words, monster barring to improve a conjecture is one way a proof can have a dialectical tier.

Both Dove and Dufour found that reasoning about refutations is an argumentative practice. Purported proofs are not always uncontroversial and incorporating refutations involves filling out a dialectical tier or considering a relevant audience.

4.4 Presenting Reasoning

After a problem has been chosen and its solution has been verified, mathematical reasoning must be disseminated to a larger public. There is much work to be done to discover how mathematicians incorporate audience consideration while presenting their solutions. But Line Edslev Andersen's interviews with working mathematicians indicate that audience consideration does feature into how papers are revised. Andersen's interviews provide insight into how mathematicians write for mathematicians (Andersen et al. 2021), how peer reviewers receive and evaluate papers (Andersen 2017), and how mathematicians choose which gaps to include in their papers (Andersen 2020). Additionally, mathematical results must be translated to students in a pedagogical setting. Results also often need to be communicated to scientists or mathematicians in other fields. In each of these cases, it is common to reword or recast a proof to aid in communication. All "strategic rewordings belong to the field of mathematical argumentation" (Dufour 2013, 73). Some of these rewordings may even be usefully cast within Johnson's concept of a dialectical tier. This is a rich area for further research, but, as Dufour points out, one must first admit mathematical proofs into the realm of argumentation. Section 5 examines how communities play a role in mathematical practice broadly, not just in presenting reasoning.

5 Mathematical Communities as Argumentative

Throughout this chapter, we have looked at the different aspects of mathematical practice that involve argumentation. Mathematicians present and produce argumentations. Their choice of problems and reasoning while solving problems also involved argumentation. But arguments are presented to, or developed with, an audience in mind. We now turn to mathematical audiences and communities. We argue that mathematical audiences influence which investigations are undertaken and how they are undertaken.

Given that the interest of this section is mathematical audiences, the receivers of mathematical arguments, one might begin by asking what kind of interaction mathematicians have with their audiences. Mathematics is often portrayed as an

isolated activity. Mathematicians lock themselves in an attic and spend days proving complicated conjectures alone. Indeed, stories of famous mathematicians seem to support such an idea. Andrew Wiles did work on his proof of Fermat's last theorem in an attic. And mathematical advances can be so particular, as in the case of Mochizuki's purported proof of the *abc* conjecture, that only a handful of people can verify or understand them. But this isolated view is not an accurate portrayal of mathematical practice.

Without an appropriate community, or engagement with that community, mathematical investigations can falter. Besides the obvious importance of mentors or coauthors, there needs to be a certain level of engagement with a larger community of mathematicians. Take, for example, William Thurston's discussion in *On Proof and Progress in Mathematics* (Thurston 1994). According to Thurston, mathematicians tend to follow fads. Fad-following is in line with his claim that there is a vital social component to mathematical progress. He drew this conclusion from his own experiences. When Thurston first entered the field of foliations, he rapidly solved a number of open problems. But the field seemed to empty out within a few years of his entrance and Thurston lost interest soon after. In response to this experience, he tried to develop and present infrastructure instead of proving theorems. He writes that:

> I have put a lot of effort into non-credit-producing activities that I value just as I value proving theorems: mathematical politics, revision of my notes into a book with a high standard of communication, exploration of computing in mathematics, mathematical education, development of new forms for communication of mathematics . . . I think that what I have done has not maximized my "credits" . . . I do think that my actions have done well in stimulating mathematics. (Thurston 1994, 177)

We can see that, for Thurston, the health of the mathematical community is invaluable to retaining mathematicians who are interested in that area. This is part of their tendency to "follow fads" to newly exciting communities.

Audience and community involvement seem important, but what exactly does argumentation theory have to offer in this area? As one of us has argued, Thurston's story exemplifies a broader issue, namely that problem choice in mathematics rests on the assurance that there exists a "contact of minds" between the audience and the mathematician (Ashton 2018). As we can see, the interest and activity of a mathematical community is vital to successful mathematical practice.

In addition to the role that communities play in problem choice, they play an important role in the ongoing dialectic involved in problem-solving. The method of proofs and refutations described by Lakatos relies heavily on community involvement in problem-solving. A first "proof" is suggested to show that, for polyhedra, $V - E + F = 2$, where V is the number of vertices, E the number of edges, and F the number of faces of the polyhedron. The proof faces a number of global and local counterexamples, that is, counterexamples to the conjecture and to specific proof steps, respectively. Accommodating these counterexamples requires a reconceptualization of the definition of polyhedron. In Section 4, we saw that reasoning about those refutations involved filling out a dialectical tier. It is also

important that community involvement plays an essential role in the reasoning. The history of the proof is recounted using the story of a class attempting to solve the problem. But it is clear that the refinements of definition, the counterexamples, and the resulting proofs are all products of dialectical engagement between the initial prover and an audience.

This community involvement plays out on a historical scale, as in (Lakatos 1976), but also in other collaborative problem-solving. For example, communities play a key role in crowdsourced mathematics such as Mathoverflow and (Mini-)Polymath. These activities have been studied by Ursula Martin and Alison Pease who connect the activities undertaken to the method of proofs and refutations (Martin and Pease 2013a, b). Building on these datasets, in later work with Joseph Corneli and other colleagues, they model mathematical arguments by analyzing how the discourse unfolds (Corneli et al. 2019). They use what they call Inference Anchoring Theory + Content to further understand how these dialogues introduce and track salient features of mathematical progress. This is done by examining the dialogue within a community of mathematicians.

So far we have considered the role that an external audience plays in problem choice. We have also looked at how we could identify communities in terms of the argument schemes that they allow or even disavow. But the final role of mathematical audiences is best described as an internal one. The core idea is that mathematical audiences play an important role in the development of proofs and other mathematical arguments. Valentin Bazhanov, for example, has argued that proofs are an appeal to a scientific community (Bazhanov 2012). A new result is offered to the community as a purported proof. To become a proof, the community must be persuaded that the argument is reliable and reproducible. This decision is borne out in dialogue.

Catarina Dutilh Novaes argues for a dialogical conception of proof (Dutilh Novaes 2021). According to the dialogical conception, the concept of proof is an embodiment of a semiadversarial dialogue between two people: Prover and Skeptic. Skeptic grants certain premises to Prover. Prover then puts forward statements claimed to follow from the premises. At each move, or inference, Skeptic has three potential moves. He brings up objections, counterexamples, and requests for clarification. If all of the steps are indefeasible, that is, without counterexamples, the proof is a winning strategy for Prover. Of course proofs are not usually dialogues between two interlocutors following Prover–Skeptic rules. Dutilh Novaes argues that the Skeptic is actually internalized into the method itself. In this sense, Skeptic, who is a particular kind of audience, plays a vital role in what constitutes proof.

In addition to the internalized skeptic, one of us has argued that the audience under consideration in a proof is actually a universal audience (Ashton 2021). This is a concept drawn from Perelman and Olbrechts-Tyteca: An argument to the universal audience is one that is meant to convince all people. The universal audience itself is an imagined audience – an arguer can never stand before all people and ask whether or not they assent to his argument. Rather, the universal audience is an abstraction from experiences with real audiences that the arguer has encountered. The account applies also to mathematics: mathematicians encounter real audiences in their education and practice. In addition, they learn what certain groups of real audiences

react to – knot theorists accept inferential moves involving diagrams that certain algebraists might not. By considering all these different, real audiences, mathematicians construct an internalized audience which reflects the standards of reasonableness found within each audience. In this way, mathematicians construct their own universal audience.

According to both Dutilh Novaes and Ashton, there is an internalized audience which, through experience with real audiences, is vital to proof development. In one case, a proof is an argument to an internalized skeptic. In the other, it is an argument to an internalized standard of reasonableness. Whether the mathematician aims to convince the adversarial skeptic or "all reasonable people," there is clearly a role for the audience in standards of proof. In other words, proofs should be examined in terms of the audiences who could have influenced them.

Communities, conceived of as audiences to mathematical arguments, are a vital component of the practice on a number of levels. The assurance of an interested audience and the contact of minds can influence research programs and problem choice. The ongoing dialogue between provers and the mathematical community helps to verify results, introduce new methods, and clear hidden assumptions. In addition to these explicit, external roles, the internalized audience plays an important role in proof development and the associated normativity. Given all this, it's clear that no account of argumentation in the philosophy of mathematics could be complete without a thorough discussion of the role of these audiences.

References

Aberdein A (2005) The uses of argument in mathematics. Argumentation 19(3):287–301
Aberdein A (2006) Managing informal mathematical knowledge: techniques from informal logic. In: Borwein JM, Farmer WM (eds) MKM 2006, vol 4108. Springer, Berlin, LNAI, pp 208–221
Aberdein A (2007a) Fallacies in mathematics. Proceedings of the British Society for Research into Learning Mathematics 27(3):1–6
Aberdein A (2007b) The informal logic of mathematical proof. In: Van Kerkhove B, Van Bendegem JP (eds) Perspectives on mathematical practices: bringing together philosophy of mathematics, sociology of mathematics, and mathematics education. Springer, Dordrecht, pp 135–151
Aberdein A (2009) Mathematics and argumentation. Found Sci 14(1–2):1–8
Aberdein A (2010) Observations on sick mathematics. In: Van Kerkhove B, Van Bendegem JP, De Vuyst J (eds) Philosophical perspectives on mathematical practice. College Publications, London, pp 269–300
Aberdein A (2013a) Mathematical wit and mathematical cognition. Top Cogn Sci 5(2):231–250
Aberdein A (2013b) The parallel structure of mathematical reasoning. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, Dordrecht, pp 361–380
Aberdein A (2019) Evidence, proofs, and derivations. ZDM 51(5):825–834
Aberdein A (2021) Dialogue types, argumentation schemes, and mathematical practice: Douglas Walton and mathematics. J Appl Logics 8(1):159–182
Aberdein A (2023) Deep disagreement in mathematics. Global Philosophy 33(1):17
Allo P, Van Bendegem JP, Van Kerkhove B (2021) Polymath as an epistemic community: analyzing the digital traces of the polymath project. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer, Cham
Andersen LE (2017) On the nature and role of peer review in mathematics. Account Res 24(3): 177–192

Andersen LE (2020) Acceptable gaps in mathematical proofs. Synthese 197(1):233–247
Andersen LE, Johansen MW, Sørensen HK (2021) Mathematicians writing for mathematicians. Synthese 198(Suppl 26):6233–6250
Aristotle (1991) The art of rhetoric. Penguin, London, translated by H. Lawson-Tancred
Ashton Z (2018) Mathematical problem choice and the contact of minds. In: Zack M, Schlimm D (eds) Research in history and philosophy of mathematics: the CSHPM 2017 annual meeting in Toronto, Ontario. Birkhäuser, Cham, pp 191–203
Ashton Z (2021) Audience role in mathematical proof development. Synthese 198(Suppl 26):6251–6275
Bartha P (2013) Analogical arguments in mathematics. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, Dordrecht, pp 199–237
Bazhanov VA (2012) Mathematical proof as a form of appeal to a scientific community. Russ Stud Philos 50(4):56–72
Berry S (2020) Taking the analogy between set theory and geometry seriously, online at https://seberry.org/Hamkins.pdf
Cantù P (2013) An argumentative approach to ideal elements in mathematics. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, Dordrecht, pp 79–99
Coleman E (2009) The surveyability of long proofs. Found Sci 14(1–2):27–43
Corneli J, Martin U, Murray-Rust D, Nesin GR, Pease A (2019) Argumentation theory for mathematical argument. Argumentation 33(2):173–214
Davis PJ, Hersh R (1986) Mathematics and rhetoric. In: Davis PJ, Hersh R (eds) Descartes' dream: the world according to mathematics. Penguin, London, pp 57–73
Dove IJ (2007) On mathematical proofs and arguments: Johnson and Lakatos. In: Van Eemeren FH, Garssen B (eds) Proceedings of the Sixth Conference of the International Society for the Study of Argumentation, vol 1. Sic Sat, Amsterdam, pp 346–351
Dove IJ (2009) Towards a theory of mathematical argument. Found Sci 14(1–2):137–152
Dufour M (2013) Arguing around mathematical proofs. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, Dordrecht, pp 61–76
Dutilh Novaes C (2021) The dialogical roots of deduction: historical, cognitive, and philosophical perspectives on reasoning. Cambridge University Press, Cambridge
Dyson F (2009) Birds and frogs. Not Am Math Soc 56(2):212–223
van Eemeren FH, Garssen B, Krabbe ECW, Henkemans AFS, Verheij B, Wagemans JHM (2014) Handbook of argumentation theory. Springer, Dordrecht
Fallis D (2003) Intentional gaps in mathematical proofs. Synthese 134(1–2):45–69
Gilbert MA (2014) Arguing with people. Broadview Press, Peterborough, ON
Goddu GC (2011) Is 'argument' subject to the product/process ambiguity? Informal Logic 31(2):75–88
Goguen J (2001) What is a proof? Online at http://cseweb.ucsd.edu/~goguen/papers/proof.html
Gonthier G (2008) Formal proof—the four color theorem. Not Am Math Soc 55(11):1382–1393
Hadamard J (1945) An essay on the psychology of invention in the mathematical field. Princeton University Press, Princeton, NJ
Hales TC, Adams M, Bauer G, Dang DT, Harrison J, Hoang TL, Kaliszyk C, Magron V, McLaughlin S, Nguyen TT, Nguyen TQ, Nipkow T, Obua S, Pleso J, Rute JM, Solovyev A, Ta AHT, Tran TN, Trieu DT, Urban J, Vu KK, Zumkeller R (2017) A formal proof of the Kepler conjecture. Forum Math Pi 5(e2):1–29
Hamami Y (2022) Mathematical rigor and proof. Rev Symb Log 15(2):409–449
Hamkins JD (2020) Lectures on the philosophy of mathematics. The MIT Press, Cambridge, MA
Hardy GH (1929) Prolegomena to a chapter on inequalities. J Lond Math Soc 4(1):61–78
Inglis M, Mejía-Ramos JP, Simpson A (2007) Modelling mathematical argumentation: the importance of qualification. Educ Stud Math 66(1):3–21
Johnson RH (2000) Manifest rationality: a pragmatic theory of argument. Lawrence Erlbaum Associates, Mahwah
Knipping C, Reid DA (2019) Argumentation analysis for early career researchers. In: Kaiser G, Presmeg N (eds) Compendium for early career researchers in mathematics education. Springer, Cham, pp 3–31

Krabbe ECW (2008) Strategic maneuvering in mathematical proofs. Argumentation 22(3):453–468
Krabbe ECW, van Laar JA (2007) About old and new dialectic: dialogues, fallacies, and strategies. Informal Logic 27(1):27–58
Krummheuer G (1995) The ethnography of argumentation. In: Cobb P, Bauersfeld H (eds) The emergence of mathematical meaning: interaction in classroom cultures. Lawrence Erlbaum Associates, Hillsdale, pp 229–269
Krummheuer G (2015) Methods for reconstructing processes of argumentation and participation in primary mathematics classroom interaction. In: Bikner-Ahsbahs A, Knipping C, Presmeg N (eds) Approaches to qualitative research in mathematics education: examples of methodology and methods. Springer, Dordrecht, pp 51–74
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge
Lewiński M (2014) Argumentative polylogues: beyond dialectical understanding of fallacies. Stud Log Gramm Rhetoric 36(1):193–218
Mac Lane S (1935) A logical analysis of mathematical structure. Monist 45(1):118–130
Martin U, Pease A (2013a) Mathematical practice, crowdsourcing, and social machines. In: Carette J, Aspinall D, Lange C, Sojka P, Windsteiger W (eds) Intelligent computer mathematics, LNAI, vol 7961. Springer, Berlin, pp 98–119
Martin U, Pease A (2013b) What does mathoverflow tell us about the production of mathematics?, presented at SoHuman, 2nd International Workshop on Social Media for Crowdsourcing and Human Computation. ACM Web Science, Paris. Online at http://arxiv.org/abs/1305.0904
Mazur B (1997) Conjecture. Synthese 111(2):197–210
McLarty C (2007) The rising sea: Grothendieck on simplicity and generality. In: Gray JJ, Parshall KH (eds) Episodes in the history of modern algebra (1800–1950). American Mathematical Society, Providence, RI, pp 301–325
Metaxas N (2015) Mathematical argumentation of students participating in a mathematics–information technology project. Int Res Educ 3(1):82–92
Metaxas N, Potari D, Zachariades T (2016) Analysis of a teacher's pedagogical arguments using Toulmin's model and argumentation schemes. Educ Stud Math 93(3):383–397
Misfeldt M, Johansen MW (2015) Research mathematicians' practices in selecting mathematical problems. Educ Stud Math 89(3):357–373
O'Keefe D (1977) Two concepts of argument. J Am Forensic Assoc 13(3):121–128
Pease A, Aberdein A (2011) Five theories of reasoning: interconnections and applications to mathematics. Log Log Philos 20(1–2):7–57
Pedemonte B, Balacheff N (2016) Establishing links between conceptions, argumentation and proof through the ck¢-enriched Toulmin model. J Math Behav 41:104–122
Perelman C, Olbrechts-Tyteca L (1969) The new rhetoric: a treatise on argumentation. University of Notre Dame Press, Notre Dame
Poincaré H (1913) The foundations of science. The Science Press, New York, translated by G. B. Halstead
Priestley WM (2013) Wandering about: analogy, ambiguity and humanistic mathematics. J Humanist Math 3(1):115–135
Reyes GM (2004) The rhetoric in mathematics: Newton, Leibniz, the calculus, and the rhetorical force of the infinitesimal. Q J Speech 90(2):163–188
Reyes GM (2021) Rhetorical approaches to the study of mathematical practice. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer
Schiappa E (2021) In what ways shall we describe mathematics as rhetorical? In: Wynn J, Reyes GM (eds) Arguing with numbers: the intersections of rhetoric and mathematics. Penn State Press, Philadelphia, PA, pp 33–52
Schlimm D (2008) Two ways of analogy: extending the study of analogies to mathematical domains. Philos Sci 75(2):178–200
Stefaneas P, Vandoulakis IM (2014) Proofs as spatio-temporal processes. Philos Sci 18(3):111–125
Stefaneas P, Vandoulakis IM (2015) On mathematical proving. J Artif Gen Intell 6(1):130–149

Steingart A (2012) A group theory of group theory: collaborative mathematics and the 'uninvention' of a 1000-page proof. Soc Stud Sci 42(2):185–213
Sundholm G (2012) "Inference versus consequence" revisited: inference, consequence, conditional, implication. Synthese 187(3):943–956
Thurston WP (1994) On proof and progress in mathematics. Bull Am Math Soc 30(2):161–177
Toscano F (2020) The secret formula: how a mathematical duel inflamed renaissance Italy and uncovered the cubic equation. Princeton University Press, Princeton, NJ
Toulmin S (1958) The uses of argument. Cambridge University Press, Cambridge
Toulmin S, Rieke R, Janik A (1979) An introduction to reasoning. Macmillan, London
Turing AM (1939) Systems of logic based on ordinals. Proc Lond Math Soc 45(1):161–228
Tymoczko T (1979) The four-color problem and its philosophical significance. J Philos 76(2):57–83
Walton DN (1998) The new dialectic: conversational contexts of argument. University of Toronto Press, Toronto
Walton DN, Krabbe ECW (1995) Commitment in dialogue: basic concepts of interpersonal reasoning. State University of New York Press, Albany, NY
Walton DN, Macagno F (2015) A classification system for argumentation schemes. Argum Comput 6(3):219–245
Walton DN, Reed C, Macagno F (2008) Argumentation Schemes. Cambridge University Press, Cambridge
Zeilberger D (1993) Theorems for a price: tomorrow's semi-rigorous mathematical culture. Not Am Math Soc 40(8):978–981

Rhetorical Approaches to the Study of Mathematical Practice

G. Mitchell Reyes

Contents

1 Introduction 2690
2 The Problems with Formalism 2691
3 A Dialectical Heuristic for Informal Mathematics 2694
4 The Principle of Fallibility 2697
5 Problem Situations and Historical Contexts 2699
6 Natality: The Emergence of Rhetorical Study of Mathematical Discourse 2700
7 Extension: Mathematics as a Translative Rhetorical Force 2705
References 2707

Abstract

What does rhetoric, the study of argument and persuasion, have to do with mathematics, the study of patterns of relations (in number, geometry, topology, . . .) and their necessary entailments? The following chapter explores a growing body of research that seeks to address this question – scholarship that pursues the counterintuitive idea that rhetoric and mathematics might just have more to do with each other than we have been taught to believe. It argues that many of the seeds for contemporary study of rhetoric and mathematics emerge from Imre Lakatos's *tour de force*, *Proofs and Refutations*. There we find a rejection of mathematical formalism and the *telos* of absolute truth, an emphasis on the arts of mathematical practice, and a model of informal mathematics that positions rhetorical argument as an engine of mathematical innovation. The book thus opens the way, however unintentionally, for understanding informal mathematics as intimately intertwined with rhetorical practice. The chapter concludes with discussion of how contemporary scholars have built upon Lakatos's work and, in recent years, expanded beyond it.

G. M. Reyes (✉)
Lewis & Clark College, Portland, OR, USA
e-mail: mreyes@lclark.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_85

Keywords

Rhetoric · Mathematics · Lakatos · Argument · Informal mathematics · Mathematical argument · Mathematics education

1 Introduction

Since the rise of modern science in the sixteenth century, rhetoric and mathematics have been estranged – the former aligned with persuasion and manipulation, the latter with rationalism and pure reason (see Cifoletti 2006). (Giovana Cifoletti's work shows that prior to the sixteenth century rhetoric and mathematics were not diametrically opposed – that instead they were often profitably interconnected – suggesting that the current opposition is both historically particular and mutable.) Only recently has this perspective started to break down, with increasing numbers of scholars becoming interested in how rhetoric and mathematics intertwine. In this essay, I suggest that one way to understand the story of the slow emergence of research on rhetoric and mathematics – and the themes and contours therein – is to consider closely the groundbreaking work of Imre Lakatos's *Proofs and Refutations* (1976). There we will find many of the seeds of what has grown into a lively interdisciplinary study of the myriad rhetorical-mathematical entanglements that mark our age.

Soon after its original publication, *Proofs and Refutations* was recognized as a *tour de force* in the study of reasoning and concept formation in mathematics. It was, in Thomas Kuhn's words, "a brilliant analysis of the permissible rejoinders to mathematical refutations [that] provides the most telling arguments I know against a naïve falsificationist position." (Kuhn 1970, pp. 13–14). Philip Kitcher, another luminary in the philosophy of science and mathematics, opined, "*Proofs and Refutations* presents a program that needs to be taken seriously by anyone who is interested in the nature of mathematics. We shall understand mathematics better when Lakatos's ideas are developed and we gain a clear idea of their merits and shortcomings" (1977, pp. 782–783). In the last two decades *Proofs and Refutations* has begun to garner the attention it deserves, specifically among scholars interested in the history and philosophy of mathematical practice, yet in the midst of that literature little attention has been paid to the role of rhetoric in Lakatos's treatise. Attending to *Proofs and Refutations* more closely, I argue, reveals many of the ways that rhetoric and mathematics intertwine, granting us an opportunity to see how contemporary research into rhetoric and mathematics has both built upon and extended Lakatos's work.

What made *Proofs and Refutations* a *tour de force*? First, in focusing on informal rather than formal mathematics the book challenged the orthodox formalist model dominant within mathematics at the time – a model that emphasized the proof procedures of formal mathematics over the practices of innovation characteristic of informal mathematics (this captures Lakatos's sense of formalism, and he appears to be targeting Hilbert's (1925) work as a classic example; on formalism as a system of

logical rules that one "plays" as one might play chess see Rotman 2000, pp. 5–7). Second, it was unique in its fallibilist approach, rejecting as pathological the *telos* of absolute truth, which according to Lakatos could never explain the growth of mathematics or the evolution of the field. Third, it established the importance of historical context not only for understanding but, more crucially, for practicing mathematics. And finally, it argued for a dialectical heuristic that rendered issues of mathematical criticism, conceptual polysemy, and conceptual translation vital to the growth of mathematical knowledge. In so doing, *Proofs and Refutations* not only offers a counterstatement to any prevailing emphasis on proof, justification, and infallibility in studies of mathematics, it also illustrates the role of argumentation and symbolic action in mathematical concept formation. The book thus opens the way, however unintentionally, for understanding informal mathematics as intimately intertwined with rhetorical practice.

To develop these points, the chapter begins with close textual analysis of Lakatos's treatise so that we might see how rhetoric (as the study and practice of argument) emerges as crucial to the practices of mathematical innovation. From that vantage, we can then trace the paths that scholars have taken to extend Lakatos's work, expand our grasp of mathematical practice, and begin to develop an understanding of mathematics not as the language of truth but as a powerful translation machine that is rapidly reshaping the social-material world.

2 The Problems with Formalism

On the first page of his introduction, Lakatos positions his book as a counterstatement to mathematical formalism. Formalism – the notion that in order for mathematics to be irrefutably true it must be "formalized" into an internally consistent language governed by a set of purely logical rules – dominated the philosophy of mathematics in the mid-twentieth century. Lakatos viewed formalism as a kind of mathematical dogma that fundamentally misrepresented mathematical practice:

> Formalism disconnects the history of mathematics from the philosophy of mathematics, since, according to the formalist concept of mathematics, there is no history of mathematics proper. . .. Formalism denies the status of mathematics to most of what has been commonly understood to be mathematics, and can say nothing about its growth. None of the 'creative' periods and hardly any of the 'critical' periods of mathematical theories would be admitted into the formalist heaven, where mathematical theories dwell like the seraphim, purged of all the impurities of earthly uncertainty (Lakatos 1976, pp. 1–2).

This passage establishes Lakatos's critiques of the formalist model of mathematics and gives a sense of his authorial voice (sharp, irreverent, sometimes biting; these characteristics of voice matched, by all reports, Lakatos's scholarly personality – see Larvor 1998; Kampis et al. 2002; Lakatos et al. 1999). The formalist desire for internal consistency and validity leads one to believe that the most important question to ask of a mathematical argument is whether there is a series of valid

inferences from the premises to the conclusion. If there is not – if the argument is somehow logically inconsistent due, for example, to conceptual development – then the argument is deemed irrational. This is why Lakatos argues formalism cannot account for the growth of mathematical knowledge: formalists separate creation from justification – or informal mathematics from formal mathematics – and then argue that only the latter can be studied and established as purely rational. The bifurcation is misguided for Lakatos because it implies that informal mathematics is an irrational or quasi-rational practice.

Lakatos uses both style and content to develop his critiques of formalism. Instead of a series of statements, deduced from accepted axioms, leading to a conclusion (a common argumentative form in mathematics), the reader enters a dialogue between a "Teacher" and his students wherein each student loosely corresponds with one or several actual mathematicians whose work was important to the discussion at hand. The purpose of the class is to determine the mathematical relationship between the number of "edges," (E) "vertices," (V) and "faces" (F) of polyhedra. The dialogue begins with the Teacher reminding the class that in the previous lesson they "arrived at a conjecture concerning polyhedra, namely, that for all polyhedra $V - E + F = 2$. We tested it by various methods. But we haven't yet proved it. Has anybody found a proof?" Lakatos then has his first student character, Sigma, respond, "'I for one have to admit that I have not yet been able to devise a strict proof of this theorem. As however the truth of it has been established in so many cases, there can be no doubt that it holds good for any solid. Thus the proposition seems to be satisfactorily demonstrated.' But if you have a proof, please do present it" (1976, p. 7). As the dialog unfolds, Lakatos weaves in historical footnotes pointing to the historical specificity of particular statements or arguments. In this case, he is actually quoting Euler, who was the first mathematician to come up with the polyhedra conjecture. The remainder of the book proceeds accordingly, with different students – each representing a host of different mathematicians who worked on polyhedra – raising counterarguments to the original conjecture, revising the original, or proposing a novel conjecture – all of which traces the historical process of growth of mathematical knowledge of polyhedra. In so doing Lakatos dramatizes the deliberative process on the Euler conjecture that took place within mathematical communities over several centuries.

If such stylistic choices (use of dialogic form; emphasis on historical context) were not enough to clarify Lakatos's views of formalism, he makes his position explicit in a way only Lakatos can. After several counterexamples to Euler's original conjecture, Lakatos has Delta play the formalist: "I turn in disgust from your lamentable 'polyhedra,' for which Euler's beautiful theorem doesn't hold: I look for order and harmony in mathematics, but you only propagate anarchy and chaos" (1976, p. 19). Lakatos then clarifies that throughout history mathematics has been plagued by such thinking, leading many a great mathematician to reject counterexamples to particularly "elegant" theorems as "monsters," aberrations better ignored than seriously addressed. These "monster-barrers," as Lakatos calls them, dwell in a "formalist heaven" in which "order" and "harmony" are more important than the growth of mathematical knowledge. With each counterexample that emerges the

monster-barrer simply retreats to a safe haven, reducing the purview of the original theorem and labeling the exceptions monsters. Monster-barrers proceed thus because they believe the purpose of proofs is to "prove" a theorem irrefutably true. As such, they view the deductive style of proof that moves from proven premises through valid logical inference to an irrefutable conclusion (theorem) as the only legitimate form.

From the formalist perspective counterexamples are not simply problems to solve, they are dangerous – destructive of the order and harmony mathematics promises. Lakatos summarizes the consequences of formalist dogma near the end of *Proofs and Refutations*:

> As long as a counterexample was a blemish not only to a theorem but to the mathematician who advocated it, as long as there were only proofs or non-proofs, but no sound proofs with weak spots, mathematical criticism was barred. It was the infallibilist philosophical background of Euclidean method that bred the authoritarian traditional patterns in mathematics, that prevented publication and discussion of conjectures, that made impossible the rise of mathematical criticism. . . . Mathematical or scientific criticism cannot exist while we only appreciate a mathematical or scientific result if it yields perfect truth. . . . The idea – expressed so clearly by Seidel – that a proof can be respectable without being flawless, was a revolutionary one in 1847, and, unfortunately, still sounds revolutionary today (1976, p. 139).

Here Lakatos underscores how the formalist desire for certainty can function ideologically to quell criticism and obstruct mathematical growth. For Lakatos formalism is more about *ethos* than *logos* – more about capturing the *ethos* of authority of absolute truth than capturing the way mathematics actually develops. "Deductivist style hides the struggle, hides the adventure. The whole story vanishes, the successive tentative formulations of the theorem in the course of the proof-procedure are doomed to oblivion while the end result is exalted into sacred infallibility" Lakatos goes on to explain that this style "while claiming objectivity, in fact fosters a private guild-language . . . [and] suffocates criticism." "Deductivist style," Lakatos concludes, "tears the proof-generated definitions off their 'proof-ancestors,' presents them out of the blue, in an artificial and authoritarian way. It hides the global counterexamples which led to their discovery" (1976, pp. 142–144).

Ultimately, formalism – and the strict deductivist style it promotes – is a problem for Lakatos because it renders mathematics nothing more than a formal abstraction, ignoring any consideration of mathematical heuristic, plausible reasoning, and the evaluative language of mathematicians, who often judge certain results deep and some proofs more worthy of celebration than others. Formalism contends that static deductive rationality is *the* form of rationality proper, thus neglecting study of conceptual formation, conceptual change, and knowledge growth in mathematics. Formalists want to formalize mathematics so as to identify it with infallible truth but they can give no account of the content they formalize. In their narrow view of rationality and mathematical proof they leave little room for interaction between mathematical conjectures, proofs, and refutations that for Lakatos are absolutely

crucial to the growth of mathematical knowledge. Against this formalist backdrop he offers a dialectical heuristic that seeks to account for these processes.

3 A Dialectical Heuristic for Informal Mathematics

For Lakatos the dialectical unity of proofs and refutations is fundamental to mathematical innovation. Proofs do not merely prove, they allow the mathematician to deconstruct a conjecture, to establish its assumptions (lemmas), and thus to aid in the process of criticism through counterexample. Those counterexamples act as refutations that then lead the mathematician back to examination of the proof for hidden assumptions. Here one begins to see why Lakatos ultimately rejected Karl Popper's criterion of falsification for differentiating between science and nonscience (see Popper 1966). When one is confronted by a mathematical refutation, one does not simply cast out the conjecture as false and move on (as strict falsificationists would have it). Equally important, one does not simply retreat into a safe haven of validity (as exception-barring formalists would have it). One returns to the proof and attempts to improve it and the conjecture simultaneously. This is why the Teacher says early in the dialogue that, though it may sound counterintuitive due to prevailing formalist dogma, he is just as interested in "proving" false conjectures as he is true ones: many mathematicians, the Teacher laments, think

> that a proof necessarily proves what it has set out to prove. My interpretation of proof will allow for a false conjecture to be 'proved,' i.e. to be decomposed into subconjectures. If the conjecture is false, I certainly expect at least one of the subconjectures to be false. But the decomposition might still be interesting! I am not perturbed at finding a counterexample to a 'proved' conjecture; I am even willing to set out to 'prove' a false conjecture! (1976, p. 23).

In contrast to the simplistic and impoverished formalist notion of proof, Lakatos's sense of proof harkens back to the ancient Greeks before the rise of the Euclidean axiomatic method: "I propose," says the Teacher, "to retain the time-honoured technical term 'proof' for a *thought-experiment – or 'quasi-experiment' – which suggests a decomposition of the original conjecture into subconjectures or lemmas*, thus *embedding it* in a possibly quite distant body of knowledge" (1976, p. 9).

The consequences of defining proofs as "thought-experiments" are many: one immediately sees, for instance, that for Lakatos the purpose of proof is not to prove a conjecture but to open the way for analysis – the lifeblood of mathematical innovation. And this analysis, as it emerges in the book, is decidedly dialectical, tacking back and forth between proofs and refutations and in the process transforming the concepts involved. As Alpha points out, "Discovery does not go up or down [as deductivist style would lead one to believe], but follows a zig-zag path: prodded by counterexamples, it moves from the naïve conjecture to the premises and then turns back again to delete the naïve conjecture and replace it by the theorem. Naïve conjectures and counterexamples do not appear in the fully fledged deductive structure: the zig-zag of discovery cannot be discerned in the end-product" (1976, p. 42). Indeed, much of

Lakatos's book is designed to tear away the mask of deductivism to reveal the dialectical argumentative structure within the doing of mathematics. The dialogue dramatizes this dialectical argumentative structure with the concept "polyhedron," showing how the dialectical process, which took place over several centuries, changed and improved the concept.

Consider, for example, how the proof-deconstruction of the original Euler conjecture immediately reveals its hidden assumptions. On the first page of the dialogue the Teacher, in response to Sigma's (Euler's) request for a proof, states the following:

> In fact I have one. It consists of the following thought-experiment. *Step 1*: Let us imagine the polyhedron to be hollow, with a surface made of thin rubber. If we cut out one of the faces, we can stretch the remaining surface flat on the blackboard, without tearing it. The faces and edges will be deformed, the edges may become curved, but V and E will not alter, so that if and only if $V - E + F = 2$ for the original polyhedron, $V - E + F = 1$ for this flat network – remember that we have removed one face. . . . *Step 2*: Now we triangulate our map – it does indeed look like a geographical map. We draw (possibly curvilinear) diagonals in those (possibly curvilinear) polygons which are not already (possibly curvilinear) triangles. By drawing each diagonal we increase both E and F by one, so that the total $V - E + F$ will not be altered. *Step 3*: From the triangulated network we now remove the triangles one by one. To remove a triangle we either remove an edge—upon which one face and one edge disappear, or we remove two edges and a vertex—upon which one face, two edges and one vertex disappear. Thus if $V - E + F = 1$ before a triangle is removed, it remains so after the triangle is removed. At the end of this procedure we get a single triangle. For this $V - E + F = 1$ holds true. Thus we have proved our conjecture (1976, pp. 7–8). (In the book Lakatos provides illustrations to help the reader follow this proof procedure.)

I cite the dialogue at length here because it reveals much about Lakatos's dialectical heuristic. Note that in the very first step we see the Teacher imbedding Euler's naïve conjecture in a "quite distant body of knowledge," namely, topology. (Some might find suspicious the fact that Lakatos begins in the middle of the story, so to speak, but the purpose of the dialogue is not to present an historical account but rather a rational reconstruction. I will deal with this distinction and its implications later in the chapter.) Euler conceived of polyhedra as solid three-dimensional objects. The first major innovation was to redefine polyhedra as systems of faces, edges, and vertices with two dimensions, one dimension, and zero dimensions, respectively. The turn to topology to prove Euler's conjecture was Cauchy's breakthrough 55 years after Euler, in 1813. This move is a crucial part of mathematical invention because it stretches the original Eulerian concepts beyond their initial conceptualization. In so doing, it introduces the change and growth in mathematical concepts that calls for a dialectical understanding of mathematical practice in the first place.

Where does that change and growth come from? It emerges out of the interaction between seemingly unrelated bodies of knowledge. Consider, for example, how the turn to topology shifted the intellectual terrain: instead of imagining polyhedra as solid three-dimensional objects, one can imagine them as skins stretched into two-dimensional topological surfaces. This reimagining of the problem enabled a novel mathematical proof of the Euler conjecture to emerge that revealed several hidden

presuppositions. For when one follows the Teacher's proof-procedure one can more easily detect counterexamples that expose assumptions in the argument. All one need do is envision instances in which the proof-procedure fails: imagine, for instance, a smaller cube set *inside* a larger cube (called a "nested cube"). This counterexample reveals a significant hidden assumption in the original conjecture, namely, that "*a polyhedron is a solid whose surface consists of polygonal faces*" (1976, p. 14). If this definition stands then the nested cube is a polyhedron and Euler's conjecture must fall; that is, unless one is willing to return to the proof and refine it. In Lakatos's dialogue this is a dynamic process, for mathematicians do not respond uniformly to counterexamples. Some (perhaps most according to Lakatos) respond simply by redefining the concepts in question to quarantine the counterexamples (this is the monster-barring technique). Delta, impersonating Jonquières's response to the nested cube counterexample in the late 1800s, redefines a polyhedron as "*a surface consisting of a system of polygons,"* effectively transforming the nested cube into "*two* polyhedra – two surfaces, one completely inside the other" (1976, pp. 14–15). Against this new definition of polyhedra, however, new counterexamples arise, to which Delta (this time impersonating Möbius) responds "of course I did not mean that any system of polygons is a polyhedron. By polyhedron I meant *a system of polygons arranged in such a way that (1) exactly two polygons meet at every edge and (2) it is possible to get from the inside of any polygon to the inside of any other polygon by a route which never crosses any edge at a vertex*" (1976, p. 15). Thus we see – even in the most conservative monster-barring responses – the growth of our knowledge of polyhedra through the interactions and translations between geometry and topology. Those processes stretch the original conception of polyhedra; they are at once moments of innovation and enablers of knowledge growth.

For Lakatos, what is also important here is that one sees a dialectical heuristic for informal mathematics begin to emerge, in this case between definition and counterexample. Although monster-barring is the most simplistic form of this dialectic (a kind of negative dialectic), it nevertheless shows a rational pattern of call and response, and in the process the concept of polyhedra *becomes more precise and sophisticated*. The purpose of *Proofs and Refutations* is to develop an understanding of the complex dialectical patterns of argument that animate informal mathematics. Lakatos eventually codifies this pattern (or heuristic) of informal mathematical reasoning in his first appendix:

> *(1) Primitive conjecture. (2) Proof (a rough thought-experiment or argument, decomposing the primitive conjecture into subconjectures or lemmas). (3) 'Global' counterexamples (counterexamples to the primitive conjecture) emerge. (4) Proof re-examined: the 'guilty lemma' to which the global counterexample is a 'local' counterexample is spotted. This guilty lemma may have previously remained 'hidden' or may have been misidentified. Now it is made explicit, and built into the primitive conjecture as a condition. The theorem – the improved conjecture – supersedes the primitive conjecture with the new proof-generated concept as its paramount new feature* (1976, pp. 127–128).

Lakatos then notes that while these four steps are the most common in informal mathematical reasoning they are sometimes augmented by the following three:

"*(5) Proofs of other theorems are examined to see if the newly found lemma or the new proof-generated concept occurs in them: this concept may be found lying at cross-roads of different proofs, and thus emerge as of basic importance. (6) The hitherto accepted consequences of the original and now refuted conjecture are checked. (7) Counterexamples are turned into new examples – new fields of inquiry open up*" (1976, pp. 127–128). These seven capture Lakatos's understanding of the broad patterns that inform the practice of doing informal mathematics, yet they are not immutable and they do not compose a method. No, for Lakatos anything that evolves has within it the capacity for its own self-overcoming.

4 The Principle of Fallibility

It would be easy to interpret Lakatos's seven steps as *the method* of reasoned argument in informal mathematics. To do so, however, would be to miss the principle of fallibility at the heart of his thinking. A crucial dimension of his heuristic, in other words, is that it remains open to the new and the unimagined. The purpose of the dialectic of proofs and refutations is not – in contrast to formal mathematics – infallible truth; the purpose is to understand the arts of mathematical invention and the processes of mathematical knowledge growth. Unlike Hegelian dialectic, Lakatos's dialectical thinking never ends in an ultimate synthesis (for discussion of Hegelian and Lakatosian dialectics, see Dusek 2015; Larvor 1998). Every "proved" theorem is always provisionally proven and, if truly innovative, always opens up "new fields of inquiry." Moreover, the set of procedures he offers are not laws of reasoning; they are, according to Lakatos, a list of the most common patterns of arguing and thinking in informal mathematics. As he stated in his original thesis, "Nothing bores me more than the hopeless effort to build up a perfect *system* of heuristic rules. One can point out some tentative rules which may help us to avoid some deeply entrenched wrong heuristic habits. But attempting to turn heuristics into a system of rules which claim to take account of the art of discovery seems to me pathological." (Lakatos 1961, p. 75).

At the end of the dialectical process, then, one might have a better theorem, more sophisticated concepts, and increased knowledge of the mathematical objects under consideration, but one will not have absolute truth. The process increases one's knowledge and, for Lakatos, helps one approach the truth; but one must remain ever open to the possibility of the new and the unanticipated. One must avoid the *hubris* of certainty. Lakatos conveys this clearly in an early encounter between Beta and the Teacher. Overcome by this *hubris*, Beta rejects the idea of naïve conjectures as "guesswork," calling his conjecture "insight," to which the Teacher responds, "I abhor your pretentious 'insight.' I respect conscious guessing, because it comes from the best human qualities: courage and modesty" (1976, p. 30). Here we see a kind of Lakatosian mathematical identity emerging in which the dogma of the absolute is sternly rejected for the courage and humility of fallibilism. Beta plays the part well: by the end of the dialogue he has transformed: instead of espousing a formalist's *hubris*, he turns full circle, observing that the belief that a "'proof' indubitably

establishes the truth of the naïve conjecture," that it produces a "certainty . . . beyond doubt . . . is only a symptom of lack of imagination, of conceptual poverty. It produces smug satisfaction and prevents the growth of knowledge" (1976, p. 87).

A keen reader might think Lakatos wants it both ways: he wants to reject absolute truth and champion knowledge growth simultaneously. But how does knowledge grow except by the accumulation of true statements about mathematical objects? Lakatos would not necessarily disagree, given one adds an important qualification: mathematical knowledge grows as we accumulate more and more sophisticated provisionally true statements – provisional precisely because the possibility of a new counterexample emerging to the latest iteration of a theorem always exists; more sophisticated because the new statement has incorporated the previous counterexamples, transforming them into examples, and thus expanded the theorem's domain of validity. It is only one's "lack of imagination," as Beta put it, which closes one off from the possibility of future counterexamples and allows one to be filled with dogmatic certainty. Lakatos's concern is not with certainty per se but with its dogmatic obstruction of knowledge growth. As he clarifies at the end of *Proofs and Refutations*,

> The dominance of the exception-barring method shows how the Euclidean method can, in certain problem situations, have deleterious effects on the development of mathematics. Most of these problem situations occur in growing mathematical theories, where growing concepts are the vehicles of progress, where the most exciting developments come from exploring the boundary regions of concepts, from stretching them, and from differentiating formerly undifferentiated concepts. In these growing theories intuition is inexperienced, it stumbles and errs. There is no theory which has not passed through such a period of growth; moreover, this period is the most exciting from the historical point of view and should be the most important from the teaching point of view. These periods cannot be properly understood without understanding the method of proofs and refutations, without adopting a fallibilist approach" (1976, p. 140).

Thus, the principle of fallibility embeds in Lakatos's dialectical philosophy a skepticism (perhaps even a fear) of all absolutist claims. There is no single logic of mathematical invention. Mathematical invention is a practice that must be allowed to evolve in response to new problems. For Lakatos, however, this does not mean it is an irrational practice. To the contrary, he finds the need for a rational heuristic all the more essential precisely because of his mathematical fallibilism, which left unbalanced could render informal mathematics nothing more than a subjective guessing game. Part of Lakatos's dialectical task, then, is to walk a tightrope between rationalism and fallibilism, between reason and possibility. (Unfortunately, Lakatos loses this balance in his later work on scientific research programs, where he attempts to establish a unifying logic of science that can be used to demarcate between science and nonscience.) One can see the delicacy of this task best in the dialectic Lakatos constructs between problem situations and historical contexts.

5 Problem Situations and Historical Contexts

When one considers the treatment of time in *Proofs and Refutations*, one begins to see how cunningly its form reinforces its content. For just as the art of mathematical invention "follows a zig-zag path," as Alpha notes, so the reader's path tacks back and forth between the dialogue and the historical footnotes. This method of presentation humanizes mathematics even as it reinforces Lakatos's balance of rationalism and fallibilism; it also constructs a useful dialectical tension between rational reconstruction of the problem situations relevant to polyhedra and the particularities of historical contexts out of which these problem situations emerged.

One of the great rhetorical benefits of Lakatos's temporal style is the way in which it rehumanizes mathematics. We have already observed how the deductivist style of presentation can dehumanize mathematics, shrouding it in a fog of mysticism for the uninitiated. Both the dialogue and the historical footnotes challenge this tendency. Instead of starting with the end product (the latest theorem of polyhedra) and then being asked to use it to solve similar or related problems, Lakatos begins with Euler's original conjecture and traces the dialectical story from there; but his is a story feathered with surprisingly flawed characters – at once intelligent and small-minded, curious and arrogant, generous and impatient. Lakatos's students – meant to represent many of the greatest mathematicians of the eighteenth and nineteenth centuries – are a far cry from the heroic image of the individual mathematical genius one often finds in histories of mathematics and in mathematics textbooks. Flawed as they are, these characters subtly encourage a questioning, irreverent, skeptical habit of mind that – as a resource of criticism and counterexample – is crucial for mathematical innovation. Equally important, Lakatos's characters allow the reader to identify with the mathematical "geniuses" of the past and view their accomplishments as incremental, collaborative, and, most importantly, within the realm of the possible (see Marchi (1976) for a similar argument about the humanizing effects of Lakatos's work).

In addition to humanizing mathematics, playing the dialogue against the historical footnotes allows Lakatos to negotiate –without attempting to resolve – the dialectic between fallibilism and rationalism. To understand Lakatos's position here we need to distinguish between rationalism in individuals and rationalism in historical processes. Lakatos accepted that individuals were both rational and irrational, and his dialogue accentuates both, but he also thought that if one accepts the premise that mathematics does in fact progress toward mathematical truths, then as a historical process mathematics must have a rational heuristic guiding it. That is why, even as he has individual characters animate his dialogue, the dialogue itself is not interested in individual rationalism so much as it is in developing a rational reconstruction of the problem situation related to polyhedra and teasing out the rational heuristic guiding informal mathematics. (This is not to say that individual rationalism plays no role in *Proofs and Refutations*; to the contrary, Lakatos positions individual rationalism in dialectical tension with the rationalism of mathematics itself, which results in an interesting de-centering of individual agency. I address this issue in the next section.) In isolation, the dialogue's rational reconstruction

would take Lakatos toward the philosophical dead-end of static rationality (see Lakatos 1976, p. 93); the historical footnotes help him avoid this, introducing the fragmented, irregular, discontinuous nature of historical specificity and thus maintaining the dialectical tension between rationalism and fallibilism. (Lakatos argues against the static rationality of logical positivism – itself an outgrowth of the merger of formal logic and mathematics in the nineteenth century – because it treats conceptual evolution as an irrational process and insists on fixed conceptual meanings.) Lakatos follows Popper in this respect, rejecting the notion of historical necessity that Popper and many others saw in Hegel's dialectic (see Popper 1966). For to accept historical necessity would be to reduce "free human beings to the status of mere conduits for world historical forces" (Larvor 1998, p. 78), erasing the openness and possibility that Lakatos positions as fundamental to mathematical innovation.

One might say in the end that Lakatos uses the dialogic form to bend and mold the fragmentary, irrational material of history into a rational, coherent narrative. However, he does not conceal that history, nor its role as constitutive of possibility and thus, for Lakatos, of mathematical innovation itself.

6 Natality: The Emergence of Rhetorical Study of Mathematical Discourse

Having traced the development of Lakatos's dialectical heuristic for informal mathematics and some of its philosophical implications, what now can be said of the relationship between rhetoric and mathematics? Wherein lie the Lakatosian seeds for study of their imbrication? Perhaps the best way to address these questions is to examine the earliest scholarly efforts to explicitly link the two fields.

While *Proofs and Refutations* advanced a paradigmatic shift in focus in philosophical circles from formal to informal mathematics, the field of rhetorical studies also underwent a paradigm shift in the mid-twentieth century. This is why, when contemporary scholars claim mathematics has rhetorical dimensions *they are not claiming math is merely another form of persuasion, deception, or manipulation*. Surely mathematics can be put to such ends, as many scholars have shown (see Gould 1996; Rotman 2000; O'Neil 2016), but in examining the connections between rhetoric and math, contemporary scholars often have something different in mind. Indeed, the old senses of rhetoric as oratory, or persuasion, or, in Aristotle's words, "the faculty of discovering the possible means of persuasion" (1924, p. 15; 1.2, 1355a) must be temporarily put aside if we are to understand the deeper entanglements between rhetoric and math. The possibility for study of those entanglements emerged in the twentieth century out of a recalibration that shifted the field of rhetoric from a narrow focus on public oratory to a broader examination of argumentation and other forms of symbolic inducement. Out of the work of scholars like Kenneth Burke (1945, 1957, 1969), Chaim Perelman (1982), and Lucie Olbrechts-Tyteca (Perelman and Olbrechts-Tyteca 1969), a novel theoretical platform was established for beginning to think from a rhetorical perspective about the role of

argument in mathematical concept formation as well as the role of mathematics in political discourse (its influence on public policy, for instance; see Schiappa 2021 for a survey of this research). That potentiality, however, lay fallow for several decades, not taken up in earnest until Philip Davis's and Reuben Hersh's pioneering essay "Rhetoric and Mathematics," which not by chance was published 11 years after *Proofs and Refutations* in 1987 (in fact, they published a different version of this essay in their book *Descartes' Dream* (1986), and Hersh's indebtedness to Lakatos can be traced back to Hersh (1979)).

Philip Davis and Reuben Hersh were both trained mathematicians, but they had read Lakatos's work and they were intrigued by Perelman's and Olbrechts-Tyteca's new theory of rhetoric as the study of argument. Their real interest, however, lay in how the study of mathematical argument both within and outside the field might advance scholarly understanding of mathematical innovation, mathematics pedagogy, and the influence of mathematics in matters political. Their essay was the first to establish substantive connections between rhetoric and mathematics, and they showed "that mathematics is not really the antithesis of rhetoric, but rather that rhetoric may sometimes be mathematical, and that mathematics may sometimes be rhetorical" (Davis and Hersh 1987, p. 54). In their essay one gains the double sense in which Davis and Hersh mean to discuss rhetoric and mathematics: in the first sense they consider how mathematics is *used* rhetorically in other fields (in economics, for example). In the second sense Davis and Hersh maintain that there are "rhetorical modes of argument and persuasion" (p. 54) at play in the *doing* of mathematics. Here, like Lakatos, they are interested in showing the non-foundational nature of mathematical proof.

Already in the first beginnings of contemporary study of rhetoric and math, then, one can detect affinities with Lakatos's work. In calling for study of the rhetorical modes of argument within mathematical communities, Davis and Hersh are clearly extending Lakatosian principles, especially the embrace of fallibilism and rejection of formalism. But they also go beyond Lakatos, who thought of rhetoric as oratory and persuasion and so would have rejected the centrality of "rhetorical modes of argument" in mathematical practice. Yet if we shift to a more contemporary notion of rhetoric as the study of particular arguments and why they become compelling to particular audiences, then the idea of analyzing practices of mathematical argument from a rhetorical perspective begins to make more sense. Add to that Davis's and Hersh's call to study the ways mathematical discourse increasingly influences public culture, and one has a compelling case for the study of both the force of rhetorical argument in mathematics and the force of mathematics in the traditional rhetorical domains of public deliberation.

Much of the work on rhetoric and mathematics since the 1980s has followed the two-fold approach of Davis and Hersh, either investigating argumentation among mathematicians or the mobilization of mathematics within other contexts (see Ashmore 1997; Bernard 2003; Hunt 1991; McCloskey 1998; Mercer 2006; Merriam 1990; Miller 2003; Potter et. al. 2016; Wynn 2009). Deirdre McCloskey's (1998) work on the rhetoric of economics, for instance, examines the rhetorical force of mathematics among economists – taking up Lakatos's critiques of formalism and

extending them to criticize economic positivism, which McCloskey shows is often more about reputation and authority (e.g., ethos) and using math as a cudgel than it is about expanding our knowledge or understanding of economic processes. Likewise, James Wynn's (2009) study of how mathematization changed the field of Biology traces the strategies of argument for and against mathematical approaches to biological inquiry, teasing out the topoi of mathematical arguments and how they slowly transformed the concept of evolution and rewrote biological epistemology.

A number of scholars have also taken up the other strand of Davis's and Hersh's original study, examining everything from how mathematicians argue with one another to how rhetorical style influences student perception of and identification (or dis-identification) with mathematics (Ernest 1999; Aberdein 2009; Aberdein and Dove 2013; Carrascal 2015). (A broad collective of researchers is now embracing Andrew Aberdein's description of informal mathematics as a "species of argument" that has more in common with "ordinary" or "informal" argument than a finished axiomatic-deductive proof. Begoña Carrascal shows how the study of the practices of mathematical invention can enhance the reasoning skills of student mathematicians and thereby improve math pedagogy.) Numerous studies of mathematics education, for instance, link formalism with an axiomatic rhetorical approach in the classroom, which tends to alienate students while communicating that mathematics is an abstract language of truth primarily about rule-driven problem-solving (Boaler 2000; Rav 2007; Ho and Hedberg 2005).

These studies of rhetoric and mathematics have taught us much, yet the notion of rhetoric they deploy remains fairly traditional in flavor. By that I mean to say that these studies are primarily interested in issues of "suasion" and how particular arguments either persuade or dissuade particular audiences (whether of mathematicians, biologists, economists, or lay citizens). There is, however, a different sense of "argument" operative in *Proofs and Refutations* that a few rhetorical scholars have built upon in the twenty-first century. To see this different sense of argument, consider how much analytic work the concept "argument" does in Lakatos's analysis. Consider, for instance, how the *performance* of argument in *Proofs and Refutations* demystifies and humanizes mathematical practice. By dramatizing mathematical arguments over polyhedra Lakatos shows that, contrary to popular belief, the practices of arguing and reasoning in informal mathematics are strikingly similar to forms of argument in other fields of inquiry. Most serious inquiry goes through a process of making claims (conjectures), seeking evidence to support those claims (confirming examples), engaging with counterexamples, attending to the warrants (hidden assumptions) those counterexamples reveal, making those warrants explicit (lemma incorporation), and either rebutting the counterexamples or qualifying one's argument in response to the counterexamples. This process of arguing and thinking will ring familiar to any student of Stephen Toulmin's *The Uses of Argument* (1958). It is a generic form of "proof construction" practiced in myriad disciplines.

But argument provides an even more fundamental service in Lakatos's work. For it is dialogic argument that enables him to map the topography of invention found in particular problem situations and historical contexts and thus begin to reveal the patterns of inventional practice in informal mathematics. Argument is the means by

which Lakatos fashions his rational heuristic. Argument is not necessarily constitutive of that rational heuristic for Lakatos – he considers himself to be *discovering* the patterns of thinking in the history of mathematics – but it is nonetheless an essential vehicle for *revealing* and *understanding* those patterns and for distinguishing between those that advance and those that obstruct mathematical knowledge growth. Productive patterns include an agonistic model of intellectual engagement, proof-deconstruction via thought-experiment, and translation procedures between seemingly unrelated bodies of knowledge; regressive patterns appear in the form of monster-barring and formalist commitments to static rationality. Argument thus plays an essential role in *understanding* mathematical innovation (a topic I examine extensively in Reyes 2020); it is simultaneously the mathematician's means for exploring uncharted territory and Lakatos's cartographic tool – his proverbial pen and ink. (One implication here is that argument plays a role in both mathematical innovation and mathematics pedagogy, for it is in the mode of *understanding – how* rather than *knowing* that a rhetorical-argumentation approach to mathematics is especially useful.)

In *Proofs and Refutations*, then, argument is not merely a means of communication, persuasion, and critique; it is also an analytic tool for discovering something new about mathematical practice. Equally important, for Lakatos the practice of argument in mathematics is positively linked with mathematical knowledge growth, such that argument becomes an engine rather than an obstacle to the growth of mathematics.

Over the last two decades rhetorical scholars have taken up and extended the notion of argument – and thus rhetoric – as an engine of mathematical invention. My earliest work on rhetoric and mathematics, for example, examined the invention of the Calculus and the rise of the concept of infinitesimals (Reyes 2004). In that study I wanted to show that rhetoric (as argument) was not merely a means to communicate or critique Newton's and Leibniz's ideas but, instead, actually *constituted* the idea of infinitesimals – one of the more important concepts in the history of mathematics. In this sense, we are no longer talking about the *rhetoric of* mathematics – how mathematicians argue and persuade one another, how math is deployed in other domains – but the *rhetoric in* mathematics – how argument and symbolic action function as constitutive agents within mathematical practice.

Two projects in particular exemplify how rhetoric plays constitutive roles in the arts of mathematical innovation. The first is Brian Rotman's semiotic approach to mathematics (see Rotman 1987, 1993, 2000), which culminates in his book *Mathematics as Sign*. The second is George Lakoff and Rafael Núñez's *Where Mathematics Comes From* (2000), which offers a cognitive-metaphorical paradigm for understanding mathematical practice (for extended analysis of these works see Reyes 2014). Rotman – another trained mathematician – shows perhaps better than any other how the dominant perspectives among mathematicians (either realism or formalism) conceal the materiality of mathematical discourse. One consequence, Rotman contends, is that mathematicians tend to accept metaphysical explanations of math (its a priori existence, for example) and ignore the embodied practices of inscription elemental to both formal and informal mathematics. As a result, the

means of production of mathematical knowledge are hidden and the constitutive work of the mathematical sign is systematically denied. Close study of the semiotics of mathematics, Rotman shows, leads to the conclusion that the meaningfulness of a mathematical statement does not come from a formal mathematical code but from the ideas that motivate those inscriptions, which emerge from outside of mathematics. To the keen reader, this will sound conspicuously similar to Lakatos's notion of concept stretching in which a particular thinker either stretches a concept outside of mathematics into a mathematical domain or vice versa. In either case, such practices, which both Rotman and Lakatos show are vital to mathematical innovation, remain invisible to the staunch formalist, for whom only the mathematical code matters.

Against Rotman's semiotic analysis, Lakoff and Núñez's approach appears comparatively more direct (though perhaps less philosophically rich). Absent from all the research on mathematics, Lakoff and Núñez contend, is analysis of the ideas implicit in equations (what Rotman called the "meta-code"), the ways mathematical ideas are weaved together in mathematical statements, and why the veracity of a mathematical statement comes from those ideas and their interdependence. "Many of the confusions, enigmas, and seeming paradoxes of mathematics," they explain, "arise because conceptual metaphors that are part of mathematics are not recognized as metaphors but are taken as literal" (2000, p. 6). Deliteralizing those metaphors will thus help clarify some of the conflicts within mathematics, but, they caution, doing so will not eliminate mathematics' metaphorical content: "Conceptual metaphor . . . [enables] us to reason about one kind of thing as if it were another. This means that metaphor is . . . a *grounded, inference-preserving cross-domain mapping* – a neural mechanism that allows us to use the inferential structure of one conceptual domain (say, geometry) to reason about another (say, arithmetic)" (2000, p. 6). Sophisticated mathematical ideas, Lakoff and Núñez argue, often weave several conceptual metaphors together; only through metaphoric analysis, then, can we begin to reveal their highly complex conceptual content (research building on this approach includes Fauconnier and Lakoff 2013; Confalonieri et al. 2018).

Coursing through all these disparate research projects, then, is an abiding interest in the ideas that animate mathematical statements, which for these authors is what makes math both meaningful and dynamic. More importantly, these books and articles have expanded the space for researchers to explore connections between rhetoric and math, whether through semiotic analysis, cognitive-metaphor analysis, or rhetorical analyses of mathematical invention. Scholars like Joseph Mazur, for example, have examined how the semiotics of math function in many ways like words, rendering tropological thoughts concrete through metaphor, analogy, and mathematical translation. Such translation practices simultaneously constrain and enable the ideas of the practicing mathematician, and Mazur contends that even conventional notions of mathematical beauty can be understood in semiotic terms: "Cultural predispositions of aesthetically persuasive symbols may play into our emotional appreciation of beauty." And in mathematics, he contends, beauty is at least partially "attributable to the illuminating efficiency of smart and tidy symbols" (2014, p. 220; along these lines also see Csiszar 2003).

Other scholars of the rhetoric *in* math have been, like Lakatos, more focused on invention than aesthetic convention. Joseph Little, for instance, has examined the importance of analogical inference in Hantaro Nagaoka's 1904 Saturnian account of atomic spectra. Nagaoka's rhetorical analogies "provided the underlying conceptual scheme for Nagaoka's discursive problem space, and in so doing licensed the axioms and subsequent lines of mathematical reasoning that would be seen by his audience as logically permissible in undertaking a solution" (Little 2021). Likewise, Lynn Stalling's (2000) and Kay O'Halloran's (2005) research examines the historical evolution of mathematical discourse, though in a more diachronic mode than Little. In study of the change of mathematical symbolics over time, O'Halloran shows that algebra's "mathematical symbolism developed as a semantic resource with a grammar through which meaning was *unambiguously* coded in ways which involved *maximal economy* and *condensation*" (2005, p. 97).

Taken together, these studies of the rhetoric *in* mathematics have helped break down the walls between the two fields, contributing to our understanding of the evolution of mathematics and helping scholars better grasp the arts of mathematical innovation, many of which are intimately connected (as Lakatos and others have now shown) with traditional rhetorical strategies of argumentation, analogical reasoning, metaphorical enfolding, and semiotic translation.

7 Extension: Mathematics as a Translative Rhetorical Force

My latest research extends many of the insights of the above literature while simultaneously pushing the boundaries of that literature. For coursing through the great diversity of research projects on rhetoric and mathematics is a general acceptance (my own early work included Reyes 2004) of the twentieth-century theory of rhetoric as the study of argument and symbolic action. As such, this research often finds itself tethered to "the symbolic," whether in the context of professional mathematicians or lay publics. And while that tethering enables a certain kind of scholarly focus on the discourses of mathematics, it also often tacitly reinforces a "two-worlds" philosophy that tends to separate "words" and "things." The problem here is that mathematics often begins with words and symbols and diagrams but ends with things – that is, with hardened mathematical objects that are themselves hybrids of human and nonhuman agency. Those mathematical objects are often combined into complex propositions that end up *mattering* precisely because, as Karen Barad eloquently put it, "matter is . . . not a thing but a doing, a congealing of agency" (2007, p. 184). Indeed, as any practicing mathematician can attest, when one does math one is not alone; instead, one interacts, thinks with, and thinks through different mathematical codes. And it is precisely the interplay of agencies (human, discursive, nonhuman) that give rise to a mathematical proposition that *matters* – one that can function as a *principle of composition* with the potential to transform the world. This is how mathematics underwrites the worlds of capital and technoscience – not because it is true in some transcendental sense but because it is a powerful form of discursive translation, collapsing distance by mobilizing traces, transforming

particularities of difference into stable similarities of form, and, because of that formal coherence, encouraging associations and combinations between phenomena that, left untranslated in the world, have no relation whatsoever. Those relations do not exist a priori, just waiting to be discovered; they are fabricated extensions of the existing collective. From this perspective, mathematical statements are not representations of reality, they are multipliers of reality.

To understand mathematical discourse and practice in this way, however, requires a new understanding of rhetoric not merely as the study of argument and symbolic action, but also (and perhaps more fundamentally) as the study of symbolic-material relations, how those relations emerge and become materially manifest, and how those manifestations ultimately transform and expand reality (humans included). In short, the goal is to understand how math functions as a translative rhetorical force. Only in this way, I argue, can we collapse the artificial distinction between the technical spheres of mathematical innovation and the public spheres of mathematical application. Only in this way do we explicitly begin to reject the bifurcation between words and things and begin to think of both in terms of relations and networks of relations – a basis from which we can begin to understand the traffic mathematics enables between symbolicity and materiality. From this relational perspective, the ultimate goal is to comprehend how the symbolic-material worlds of technical mathematical discourse become materially manifest and, in those manifestations, accelerate the becoming (the ontological transformation) of the collective of relations we call reality (for a more developed version of this position see Reyes 2019).

To describe math as a translative rhetorical force, then, suggests a certain kind of shift in our thinking about both rhetoric and mathematics. Associating mathematics primarily with the study of timeless, a-historical, a priori objects (as mathematical realists do) has alienated it from rhetoric, wherein history, culture, and context are crucial and must be taken seriously (Perelman and Olbrechts-Tyteca (1969) offer a classic critique of Descartes' mathematical ontology from this perspective). Likewise, associating rhetoric primarily with argument and persuasion has long alienated it from the mathematical arts of demonstrative proof and refutation. None of these associations, however, are permanent or essential, and one might say the liveliness of both rhetoric and math (as fields of study) can be measured by their capacities for evolution and diversification. If, however, we can begin to think of both in terms of the relations that form the networks that compose the social-material world, then we can escape their estrangement and begin to take their increasingly consequential relations seriously.

Thinking of both rhetoric and mathematics through the lens of relationality has real potential. Not only can it offer scholars an avenue of escape from the pieties of mathematical realism and formalism that Lakatos so effectively exposed, more importantly it promises a better explanation and understanding of both how mathematics evolves and how it has become the meta-discourse of our current social-material world. The ideologies of purification in mathematics – which can be found in most arguments for mathematical infallibility – come at a high cost, blinding us from the very practices that enable the emergence and articulation of new mathematical relations with enough agential force to unleash not only whole new realms of

math but also whole new worlds of social-material possibility (in Reyes 2020 I explore imaginary numbers as exemplars of this process). This is why I argue that, from a rhetorical-relational perspective, mathematics becomes both better understood and more interesting. From this perspective, math is not a truth-production machine; it is a collection of powerful translation machines. Or, put differently, a collection of languages and ways of thinking (for every language is a way of thinking) that translate those that learn to "speak" them just as much as they do the world that such "speaking" re-articulates. From this vantage point, one might ask not just "What problems can mathematics solve" but also "Who do we become when we think mathematically," "What becomes of the world through a particular mathematical lens," and "How do particular mathematical networks function as principles of composition." (By "principle of composition" I do not mean a principle of writing in a composition classroom but instead a principle out of which the networks of relations that compose our world emerge.)

Ultimately, mathematics can and should be understood in light of the networks and relations it renders, the ways it simultaneously concentrates and conceals power, the means by which it translates the objects of its gaze, the productive metaphorical structures it uses, and who we become when we do math. In short, it can and should be understood in its myriad rhetorical forms. Lakatos's work opened the way for such understanding and perhaps now one can see how more recent scholarship has built upon and extended his original insights. And while we have come some distance in grasping the links between rhetoric and mathematics since Lakatos, much work remains if we are going to come to grips with the increasing translative rhetorical force of mathematical discourse in the twenty-first century.

References

Aberdein A (2009) Mathematics and argumentation. Found Sci 14(1):1–8
Aberdein A, Dove IJ (2013) The argument of mathematics. Springer, Dordrecht
Aristotle (1924) The "art" of rhetoric (trans: Freese J). G. P. Putnam's Sons, New York
Ashmore M (1997) Fraud by numbers: quantitative rhetoric in the Piltdown forger discovery. In: Smith B, Plotnitsky A (eds) Mathematics, science, and postclassical theory. Duke University Press, Durham
Barad K (2007) Meeting the universe halfway: quantum physics and the entanglement of matter and meaning. Duke University Press, Durham
Bernard A (2003) Ancient rhetoric and Greek mathematics: a response to a modern historiographical dilemma. Sci Context 16(3):391–412
Boaler J (2000) Mathematics from another world: traditional communities and the alienation of learners. J Math Behav 18:379–397
Burke K (1945) A grammar of motives. Prentice-Hall, New York
Burke K (1957) The philosophy of literary form; studies in symbolic action (Rev. ed., abridged by the author). Vintage Books, New York
Burke K (1969) A rhetoric of motives. University of California Press, Berkeley
Carrascal B (2015) Proofs, mathematical practice and argumentation. Argumentation 29(3):305–324
Cifoletti G (2006) Mathematics and rhetoric. Introduction. Early Sci Med 11(4):369–389
Confalonieri R, al e (2018) Concept invention: foundations, implementation, social aspects and applications. Springer, Cham

Csiszar A (2003) Stylizing rigor: or, why mathematicians write so well. Configurations 11(2):239–268
Davis P, Hersh R (1986) Descartes' dream: the world according to mathematics, 1st edn. Harcourt Brace Jovanovich, San Diego
Davis P, Hersh R (1987) Rhetoric and mathematics. In: Nelson J, Megill A, McCloskey D (eds) The rhetoric of the human sciences: language and argument in scholarship and public affairs. University of Wisconsin Press, Madison
Dusek V (2015) Lakatos between Marxism and the Hungarian heuristic tradition. Stud East Eur Thought 67(1–2):61–73
Ernest P (1999) Forms of knowledge in mathematics and mathematics education: philosophical and rhetorical perspectives. Educ Stud Math 38(1/3):67–83
Fauconnier G, Lakoff G (2013) On metaphor and blending. J Cogn Semiotics 5(1–2):393–399
Gould S (1996) The mismeasure of man (Rev. and expanded. ed). Norton, New York
Hersh R (1979) Some proposals for reviving the philosophy of mathematics. Adv Math 31(1):31–50
Hilbert D (1925) On the infinite. In: Benacerraf P, Putnam H (eds) Philosophy of mathematics: selected readings, 2nd edn. Cambridge University Press, Cambridge, pp 183–201
Ho KF, Hedberg JG (2005) Teachers' pedagogies and their impact on students' mathematical problem solving. J Math Behav 24:238–252
Hunt B (1991) Rigorous discipline: Oliver Heaviside versus the mathematicians. In: Dear P (ed) The literary structure of scientific argument: historical studies. University of Pennsylvania Press, Philadelphia
Kampis G, Kvasz L, Stöltzner M (2002) Appraising Lakatos: mathematics, methodology, and the man (Vienna Circle Institute Library; 1). Kluwer Academic, Dordrecht
Kitcher P (1977) On the uses of rigorous proof. Science 196(4291):782–783
Kuhn TS (1970) Logic of discovery or psychology of research? In: Lakatos I, Musgrave A (eds) Criticism and the growth of knowledge. Cambridge University Press, Cambridge, pp 1–23
Lakatos I (1961) Essays in the logic of mathematical discovery. Unpublished PhD dissertation, Cambridge University
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. Cambridge University Press, Cambridge
Lakatos I, Feyerabend P, Motterlini M (1999) For and against method: including Lakatos's lectures on scientific method and the Lakatos–Feyerabend correspondence. University of Chicago Press, Chicago
Lakoff G, Núñez RE (2000) Where mathematics comes from: how the embodied mind brings mathematics into being. Basic Books, New York
Larvor B (1998) Lakatos: an introduction. Routledge, London
Little J (2021) Rhetoric and mathematics in the Saturnian account of atomic spectra. In: Wynn J, Reyes GM (eds) Arguing with numbers: essays at the intersections of rhetoric and mathematics. University Park, Pennsylvania State University Press
Marchi P (1976) Mathematics as a critical enterprise. In: Lakatos I, Cohen RS, Feyerabend P, Wartofsky MW (eds) Essays in memory of Imre Lakatos. D. Reidel Pub, Dordrecht/Boston
Mazur J (2014) Enlightening symbols: a short history of mathematical notation and its hidden powers. Princeton University Press, Princeton
McCloskey D (1998) The rhetoric of economics, 2nd edn. University of Wisconsin Press, Madison
Mercer C (2006) Leibniz on mathematics, methodology, and the good: a reconsideration of the place of mathematics in Leibniz's philosophy. Early Sci Med 11(4):424–454
Merriam AH (1990) Words and numbers: mathematical dimensions of rhetoric. South Commun J 55(4):337–354
Miller CR (2003) The presumptions of expertise: the role of ethos in risk analysis. Configurations 11(2):163–202
O'Halloran K (2005) Mathematical discourse: language, symbolism and visual images. Continuum, London
O'Neil C (2016) Weapons of math destruction: how big data increases inequality and threatens democracy, 1st edn. Crown, New York

Perelman C (1982) The realm of rhetoric. University of Notre Dame Press, Notre Dame
Perelman C, Olbrechts-Tyteca L (1969) The new rhetoric: a treatise on argumentation (trans: Wilkinson J, Weaver P). University of Notre Dame Press, Notre Dame
Popper K (1966) The open society and its enemies, 5th edn. Princeton University Press, Princeton
Potter J, Wetherell M, Chitty A (2016) Quantification rhetoric – cancer on television. Discourse Soc 2(3):333–365
Rav Y (2007) A critique of a formalist-mechanist version of the justification of arguments in mathematicians' proof practices. Philos Math 15:291–320
Reyes GM (2004) The rhetoric in mathematics: Newton, Leibniz, the calculus, and the rhetorical force of the infinitesimal. Q J Speech 90(2):163–188
Reyes GM (2014) Stranger relations: the case for rebuilding commonplaces between rhetoric and mathematics. Rhetor Soc Q 44(5):470–491
Reyes GM (2019) Algorithms and rhetorical inquiry: the case of the 2008 financial collapse. Rhetoric Public Affairs 22(4):569–614
Reyes GM (2020) Mathematics and rhetoric. In: Sriraman B (ed) Handbook of the mathematics of the arts and sciences. Springer, Cham, pp 1–21
Rotman B (1987) Signifying nothing: the semiotics of zero. St. Martin's Press, New York
Rotman B (1993) Ad infinitum – the ghost in Turing's machine: an essay in corporeal semiotics. Stanford University Press, Stanford
Rotman B (2000) Mathematics as sign: writing, imagining, counting. Stanford University Press, Stanford
Schiappa E (2021) In what ways shall we describe mathematics as rhetorical? In: Wynn J, Reyes GM (eds) Arguing with numbers. Pennsylvania State University Press, University Park
Toulmin S (1958) The uses of argument. Cambridge University Press, Cambridge, UK
Wynn J (2009) Arithmetic of the species: Darwin and the role of mathematics in his argumentation. Rhetorica 27(1):76–97

Bayesian Perspectives on Mathematical Practice

James Franklin

Contents

1 Introduction 2712
2 The Relation of Conjectures to Proof 2713
3 Applied Mathematics and Statistics: Understanding the Behavior of Complex Models 2714
4 The Objective Bayesian Perspective on Evidence 2716
5 Evidence for and Against the Riemann Hypothesis 2718
6 Probabilistic Relations Between Necessary Truths 2721
7 The Problem of Induction in Mathematics 2722
8 Conclusion 2724
9 Cross-References 2724
References 2724

Abstract

Mathematicians often speak of conjectures as being confirmed by evidence that falls short of proof. For their own conjectures, evidence justifies further work in looking for a proof. Those conjectures of mathematics that have long resisted proof, such as the Riemann hypothesis, have had to be considered in terms of the evidence for and against them. In recent decades, massive increases in computer power have permitted the gathering of huge amounts of numerical evidence, both for conjectures in pure mathematics and for the behavior of complex applied mathematical models and statistical algorithms. Mathematics has therefore become (among other things) an experimental science (though that has not diminished the importance of proof in the traditional style). We examine how the evaluation of evidence for conjectures works in mathematical practice. We explain the (objective) Bayesian view of probability, which gives a theoretical framework for unifying evidence evaluation in science and law as well as in

J. Franklin (✉)
School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, Australia
e-mail: j.franklin@unsw.edu.au

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_84

mathematics. Numerical evidence in mathematics is related to the problem of induction; the occurrence of straightforward inductive reasoning in the purely logical material of pure mathematics casts light on the nature of induction as well as of mathematical reasoning.

Keywords

Mathematical conjectures · Bayesian · Logical probability · Problem of induction · Riemann hypothesis · Numerical evidence

1 Introduction

Mathematical practice cannot consist solely of someone's writing down mathematical proofs and someone else's publishing and reading them. That may be the final output (in formal pure mathematics), but there are many preliminary activities that must take place on the way. Much of it involves evaluating inconclusive evidence as to whether a hoped-for theorem is true and whether one's attempts at proving it are likely to succeed.

It is similar in the art of navigation (prior to google maps), and for the same reasons: navigation cannot consist just of paths from A to B. Back of house (in the metaphor of Reuben Hersh 1991), there must be work to plan how to find them, to choose among alternatives, and to estimate whether a developing path is on the right track. In mathematical research too, one must navigate a path to the final result, a polished proof, and there are probabilistic skills needed to evaluate progress.

Especially in the early stages of work on a problem, mathematicians need not merely hunches and beliefs, but ones that are as a matter of fact well supported by the evidence. The working mathematician must begin by asking questions like: Which of the many conjectures that could be generated are worth investigating? Which ones are relevant to the problem at hand? Which can be confirmed or refuted in some easy cases, so that there will be some indication of their truth in a reasonable time? Which might be capable of proof by a method in the mathematician's repertoire? Which might follow from someone else's theorem? Which are unlikely to yield an answer until after the next review of tenure? The mathematician must answer these questions to allocate time and effort. But not all answers to these questions are equally good. To stay employed as a mathematician, a proportion of these questions must be answered *well* (Franklin 1987).

The area where a mathematician must make the finest discriminations of this kind – and where he might, in theory, be guilty of professional negligence if he makes poor decisions – is as a supervisor advising a prospective Ph.D. student. It is usual for a student beginning a Ph.D. to choose some general field of mathematics and then to approach an expert in the field as a supervisor. The supervisor (after estimating that the student is capable of a Ph.D.) then selects a problem in that field for the student to investigate. In mathematics, more than in any other discipline, the initial choice of problem is the crucial event in the Ph.D.-gathering process. The problem must be unsolved at present, not being worked on by someone who is likely to solve it soon, but most importantly, tractable, that is, probably solvable, or at least partially

solvable, by three years' work at the Ph.D. level. It is recognized that of the enormous number of unsolved problems that have been or could be thought of, the tractable ones form a small proportion, and that it is difficult to discern which they are. The skill in rational evaluation of conjectures required of a supervisor is high.

The next stage in successful mathematical practice is applying for a grant. While it is not unknown to ask for money for already completed work, a grant application normally talks up the importance of an unsolved problem or general area, its "impact," and the applicant's prospects of making progress on it. Expert referees for grants will usually have rather fixed ideas on what problems in their fields are important, and as fellow pure mathematicians they will ignore the applicant's claims that the research will eventually lead to a cure for cancer and world peace. Thus much of the applicant's effort must go into arguing persuasively that the problem is at least largely solvable with the resources the grant will provide and that the applicant and team are the perfect people to do it. For established researchers, the correct strategy is to exhibit their track records and speak confidently of their expected timeline of results (several examples in Zeilberger 2012). It is more difficult in the case of early-career researchers, both for the applicant and the evaluators, as there is less track record to go on so the evaluators need to consider whether the applicant's approach is realistic in the light of both the problem and the applicant's background. They may be tempted to fall back on proxy indicators of "promise" such as "having been supervised by a Fields medallist."

None of that argument involves presenting proofs. It is probabilistic in nature but bears on the provability of theorems – and their provability with given resources, not their provability in principle.

With the grant won, or not, the next task is to work on the actual prospective theorems, developing insights and talking to fellow experts on the problem, and constantly evaluating the evidence that progress towards a solution is being made. We look at examples of evidence for several conjectures below.

It is not adequate to describe the relation of evidence to hypothesis as "subjective," "heuristic," or "pragmatic": although those descriptions may be true as far as they go. For the practice of evaluating conjectures to be successful there must be an element of what it is rational to believe on the evidence. The Bayesian perspective on evidence, to be explained below, gives a philosophical account of that.

2 The Relation of Conjectures to Proof

Although it is not the only possibility, in the practice of pure mathematical research consideration of the probability of conjectures is usually allied to the traditional view that the main aim of pure mathematics is to construct semiformal but rigorous proofs of theorems.

As one researcher describes the process:

> The computer is often used as an exploratory tool in problem solving. This exploratory analysis is very important in gaining insight into problems and theorems and often leads to

> new conjectures. Heuristic reasoning may then lead from the conjecture to a formal theorem whose validity leads itself to further numerical investigation. When these numerical investigations present overwhelming numerical evidence that the proposed theorem is true, the researcher sets upon a search for a mathematical proof. (Schuster 1985); a more extensive account in (Wilf 2008), many examples in (Borwein and Bailey 2004).

That is how mathematicians speak in cases such as the Riemann hypothesis, considered below, where evidence has been collected and published over a long timespan because a famous conjecture has resisted proof.

Although purely computational evidence is not generally publishable by itself, it is possible to prove some restricted cases of a conjecture and exhibit computational evidence that it holds more generally (e.g., Boij 1999).

However, there are some other possibilities for how probabilistic reasoning can work in mathematical practice, other than simply providing evidence towards a traditional formal or semiformal proof.

One is that there may be probabilistic proofs, that is, proofs or "proofs" that in their nature provide only a high probability that the result is true. A well-known case is probabilistic testing of primality, widely used in cryptography, where an algorithm delivers only a high probability that a number is a prime (Rabin 1980).

In the light of such examples, there has been some discussion of whether mathematicians should abandon their model of formal deductive proof and rest content with some probabilistic substitute (e.g., (Fallis 1997), discussion in (Easwaran 2009); related arguments in (Paseau 2015; Sørensen 2016)). A popular article on the topic titled *The death of proof* by the journalist John Horgan (1993) attracted some heated responses (e.g., (Krantz 2011), but some support in (Zeilberger 1993)) and by and large mathematical practice has remained committed to the deductive model of proof as the ideal. That ideal is likely to gain support from current projects on fully automated theorem provers (a progress report in (Ornes 2020)). The results may be hard for humans to survey, but they are fully deductive proofs which do not contain human errors and sloppiness.

We should also not forget the role of arguments from authority in mathematics. (Inglis and Mejía-Ramos 2009). While mathematics in principle offers the possibility of understanding any result by following its proof, it is impossible in practice to do that for all the results one must know. Mathematicians, like other people, rationally attribute high probability to the theorems signed off by respected experts. Refereeing depends on it. So does learning mathematics. (Consequences for pedagogy considered in (Aberdein 2019), Sect. 5.)

3 Applied Mathematics and Statistics: Understanding the Behavior of Complex Models

Even more than in pure mathematics, numerical evidence has come to the fore in the large areas of applied mathematics and computational statistics that have been opened up by the vast increases in computational power in recent decades. In such

fields, one is typically dealing with a very complex mathematical object such as a simulation of climate, whose mathematical properties are well beyond the reach of analytic methods or proof. Understanding how it works – that is, its internal behavior rather than its match with the modeled outside reality – can only be approached by experimenting with its output and applying normal (probabilistic) scientific methods for drawing conclusions from experiments.

A typical area for such work is computational fluid dynamics. One begins with a complex physical phenomenon which one hopes to understand, such as turbulence generated by nonbreaking surface waves in the ocean (Tsai et al. 2015). One replaces the physical phenomenon with some standard numerical simulation of it, which one can run on a powerful computer with various choices of parameters. The simulation is a strictly mathematical object whose relation to the physical phenomenon being modeled needs understanding; for one thing, typically the model is discrete and the physical phenomenon continuous (at least at the scale being considered). But experimental evidence is available that the model and phenomenon behave quite similarly, or there would be no point in the simulation. One can therefore hope to understand the modeled reality by experimenting with and analyzing the model.

The behavior of the simulation (now considered in abstraction from its role in imitating the phenomenon) is a purely mathematical question. In principle it might be addressed by solving equations and writing proofs, but almost always the simulation is too complex for that to be possible. One can only perform (numerical) experiments, that is, run the simulation with different choices of parameters (such as total energy in the system and various initial conditions), and see what happens. If the phenomenon to be explained sometimes occurs and sometimes does not, progress has been made in understanding its causes. But the different simulation runs are only samples of the simulation's mathematical behavior, so there is an essentially probabilistic inference from the runs to the overall behavior, just as there is for any conclusion from finite experimental evidence to a general law.

Something similar occurs in "big data" applications where one sets up an algorithm intended to solve some statistical problem such as fraud detection or handwritten character recognition. Before testing the system "in the wild" on real data, one often seeks first to understand its performance in principle – does it have the inherent capacity to find certain structures in data? (The question is particularly urgent if there is a shortage of real data; example in (Hoffmann et al. 2019).) That is a strictly mathematical question and so in theory subject to proof, but with a complex system of the sort typically used, proof is not attainable. So one tests it on synthetic data, that is, computer-generated data into which one has inserted structures of the kind one hopes the system will be able to detect. If the system detects them, something is known about its performance. A run with purely synthetic data is an experiment that gives rise to purely mathematical knowledge.

Or one may want to compare different available mathematical methods for a type of task, such as spatial prediction where again there is a shortage of real data. Simulated data give a sufficient quantity of data to evaluate the capacities of the different methods (e.g., Harris et al. 2010). Again, the numerical experiments result in a purely mathematical conclusion.

4 The Objective Bayesian Perspective on Evidence

There is no agreed account of what Bayesianism means, but here we present a version that is suitable as a theoretical framework for the evaluation of evidence for conjectures in pure mathematics (Taken from (Franklin 2011, 2016) but based on (Keynes 1921; Jaynes 2003)). It is the version called (fully) objective Bayesianism or logical probability.

It holds that the relation of evidence to conclusion is a matter of strict logic, like the relation of axioms to theorems in mathematics but less conclusive. Given a fixed body of evidence – say in a trial in court, or in a dispute about a scientific theory, or in evaluating numerical evidence for a mathematical conjecture – and given a conclusion, there is a fixed degree to which the evidence supports the conclusion. If we could agree just what the standard of "proof beyond reasonable doubt" is, then, in a given trial, it is an objective matter of logical fact whether the evidence presented does or does not meet that standard, and so a jury is either right or wrong in its verdict on the evidence.

It is not essential to the Bayesian perspective that the relation of evidence to conclusion should be given a precise number. It is unlikely that in typical legal, scientific or mathematical cases a numerical probability could be calculated, even if it existed in principle. It is sufficient for objective Bayesianism that it is sometimes intuitively evident that some hypotheses, on some bodies of evidence, are highly likely, almost certain, or virtually impossible.

The most central theses of Bayesianism do not concern numbers but are certain qualitative principles of evidence. The first is the simplest principle of logical probability, called by Pólya ((1954), 4; further in (Mazur 2014)) "the fundamental inductive pattern" or "verification of a consequence." It is as follows:

q is a (nontrivial) consequence of hypothesis p
q is found to be true
So, p is more likely to be true than before

(In short, "Theories are confirmed by their consequences or predictions.") That is, taken for granted as rational in scientific experiment or historical inquiry (except by strict Popperian falsificationists). The point of drawing predictions from theories and checking if they are true is normally taken to be that confirmation of predictions supports a theory. That is as applicable to mathematical conjectures as to scientific theories: the simplest possible consequences of a general mathematical statement such as "All odd numbers have odd squares" are its instances such as "The square of 131 is odd," which can be checked by calculation. Finding the consequences to be true gives some degree of support for a generalization and can justify further work.

It is true that in some clearly defined circumstances, it is natural to associate a number with the degree to which evidence supports a conclusion. Those circumstances are cases of the "proportional syllogism" or "statistical syllogism," where the sole relevant evidence is a proportion in a set. If the sole evidence bearing on

whether a certain patient will be cured of disease A by drug B is that 89% of patients with disease A are cured by drug B, then it is natural to assign a number to the probability, on that evidence, that the patient will be cured; namely 0.89. (Even with the proportional syllogism, however, arguments with imprecise numbers are often the most applicable, such as "The vast majority of flights arrive safely, therefore I can relax on takeoff.")

When numbers are applicable, it is usual to model the relation of evidence e to conclusion h by a number $P(h \mid e)$, between 0 and 1 inclusive, "the (logical) probability of h given e," which satisfies the usual axioms of conditional probability:

$$P(\text{not} - h \mid e) = 1 - P(h \mid e)$$

$$P(h_1 \text{ and } h_2 \mid e) = P(h_1 \mid e) \times P(h_2 \mid h_1 \text{ and } e)$$

The theorem of probability that is most useful for updating evidence is the celebrated Bayes' theorem which expresses how the probability of evidence given a hypothesis relates to the probability of the hypothesis given the evidence:

$$P(h \mid e\&b) = P(e \mid h\&b) \times P(h \mid b)/P(e \mid b)$$

(where h is read as hypothesis, e as (new) evidence, and b as the background knowledge that is taken for granted in the context of the problem). Thus, new evidence supports a hypothesis which makes the evidence likely (but only if the evidence is surprising).

While Bayes' theorem gives its name to Bayesianism, we emphasize that for the purpose of theory evaluation in general, the numerical theorem is much less important than qualitative consequences of it such as Pólya's "fundamental inductive pattern."

Another intuitively appealing such consequence, also easily derivable, is that the more surprising a consequence of h (that is, the lower its "prior" probability $P(e \mid b)$ on the background knowledge b), the more it increases the probability of h. (Thus Einstein's surprising prediction of the gravitational bending of light gave strong support to the theory of general relativity when it was observed to be true.)

If one takes those principles to be logical, they are as applicable to evidence for conjectures in pure mathematics as they are for theories in science or allegations in courts.

The appearance in Bayes' theorem of the "prior" probabilities $P(h \mid b)$ and $P(e \mid b)$ calls attention to the need to have some knowledge of them. For example, if one takes a "dogmatic" prior according to which a hypothesis is impossible, $P(h \mid b) = 0$, then no amount of evidence will dig one out of that hole: according to Bayes' theorem, the "posterior" probability $P(h \mid e \,\&\, b)$ will still be zero after any evidence e. Some logical intuition will be needed to evaluate whether a mathematical proposition is inherently unlikely or not, but such intuition seems to be available in simple cases; for example, it seems intuitively unlikely that the square of every odd number should be odd unless there was some good mathematical reason for that.

Other versions of Bayesianism are possible, such as subjective Bayesianism which requires only the axioms and does not constrain prior probabilities (Howson and Urbach 2006) and a version of objective Bayesianism that falls short of claiming that the probabilities involved in evidence evaluation are matters of strict logic (Williamson 2010). These are less attractive in the very abstract reasoning of pure mathematics where it seems that any considerations that genuinely count can only be logical; but in any case those versions agree on the qualitative principles of probability such as Pólya's inductive pattern, which are the main concern in explaining rational methods of evaluating the strength of conjectures.

5 Evidence for and Against the Riemann Hypothesis

Much of what has been said so far – both about the process of evaluating evidence for conjectures and the general principles of Bayesian reasoning – can be illustrated by the evidence that has accumulated for and against the Riemann hypothesis, currently the most famous unsolved problem of mathematics. Because the hypothesis has resisted proof for so long, the mathematical community has been much more explicit than usual about the evidential situation.

Expert opinion is that the Riemann hypothesis is almost certainly true, and that a proof is not exactly imminent but may not be far off. Naturally, some of the considerations are only comprehensible to true experts, but some of them are comparatively shallow and can be widely appreciated.

Very relevant in this case are the shallowest arguments, those from authority. How people think about arguments from authority in mathematics is evident in reactions to people who claim to have proved the Riemann hypothesis. If someone unknown presents his "result" on BBC news and at a conference run by someone who has accepted fake papers, he will not be taken seriously (Steckles and Lawson-Perfect 2018). If the aged Fields Medallist Sir Michael Atiyah presents a sketch of a proof, it is still met with skepticism, because of the difficulty of the problem, the minimal amount of information provided, and the author's track record of making unestablished proof claims in his old age (Amit 2018). On the other hand, when Alain Connes (2016) learnedly explains the comparative strengths of different strategies that may lead to a solution, mathematicians in general will nod in agreement, whether they understand it or not. He is the expert.

The Riemann hypothesis states that "With certain trivial exceptions, all the (infinitely many) roots of the Riemann zeta function have real part equal to ½." For the present purpose an understanding of the Riemann zeta function as a function in the complex plane is not necessary, it is only important that the hypothesis is a simple universal proposition like "all ravens are black." It is also true that the infinitely many nontrivial roots of the Riemann zeta function have a natural order, so that one can speak of "the first million roots." (One of many popular but informed accounts is (Derbyshire 2003).)

As a straightforward generalization, the Hypothesis is susceptible to purely inductive evidence, that is, just calculating many zeros starting from the first one

Table 1 Number of roots of the Riemann zeta found to have real part ½

Worker	Number of roots found to have real part ½
Gram (1903) Backlund (1914) Hutchinson (1925) Titchmarch (1935/6)	15 79 138 1041
Lehmer (1956)	25,000
Meller (1958)	35,337
Lehman (1966)	250,000
Rosser, Yohe, and Schoenfeld (1968)	3,500,000
Brent (1979)	81,000,001
Te Riele, van de Lune et al. (1986)	1,500,000,001
Gourdon (2004)	10^{13}
Platt and Trudgian (2020)	3×10^{12} with more rigor

and checking that they have real part ½. (We consider inductive evidence in mathematics in Sect. 7 below.) That process was begun by Riemann after he made the conjecture in 1859 and has of course been much assisted by increasing computer power. The results have been, as they developed (Table 1):

Inductions in mathematics can certainly dwarf those in science.

While this inductive evidence is generally taken to be quite strong (to some indeterminate degree), there is also a reason to deny it is close to certainty. That comes from the close connection of the Hypothesis with the prime number theorem. This theorem states that the distribution of primes thins out logarithmically, or more exactly, the number of primes less than x is (for large x) approximately equal to the integral $\int_2^x \frac{dt}{\log t}$.

If tables are drawn up for the number of primes less than x and the values of this integral, for x as far as calculations can reach, then it is always found that the number of primes less than x is actually *less than* the integral. On this evidence, it was thought for many years that this was true for all x. Nevertheless, Littlewood proved that this is false. While he did not produce an actual number for which it is false, it appears that the first such number is extremely large – well beyond the range of computer calculations. That gives some reason – it is very hard to say how strong – for believing that there might be a very large counterexample to the Riemann Hypothesis even though there are no small ones.

The reasons why most mathematicians' confidence in the Riemann hypothesis does approach certainty are not taken from further numerical evidence but from its connections with other theses that are or "ought to be" right. The first of these was the prime number theorem just mentioned. It is a consequence of the Riemann Hypothesis and was proved independently in 1896. The support given by this proof is thus an instance of Pólya's schema, mentioned above, "theories are confirmed by their consequences or predictions." However, the prime number theorem is not a very strong consequence of the Riemann hypothesis: the Riemann Hypothesis "says a lot more," so the support given by the prime number theorem is only moderate.

Two other considerations (among those accessible to ordinary mathematicians) are much stronger. One is the remarkable "Denjoy's probabilistic interpretation of the Riemann Hypothesis," an easily understood thesis in number theory which is equivalent to the Riemann Hypothesis and is independently intuitively plausible. We will not review it here (for a very brief account see (Franklin 2014), Chap. 15). The other is again a result of its connection with the prime number theorem. That theorem gives a crude estimate of the density of primes (namely, the density of primes around a large number N is $1/\log N$). The Riemann Hypothesis implies a much more fine-grained knowledge of the distribution of primes (namely, that the difference between the number of primes up to x and the integral above is $O(\sqrt{x}\log x)$, the best possible). If the Riemann Hypothesis is true, that fine-grained distribution is the simplest and most natural one, but if the Hypothesis is false, some more complicated and unnatural phenomenon must be present. The type of probabilistic reasoning involved is what Mazur (2014) calls "reasoning from randomness": that after known constraints have been taken into account, the rest should be random. (To take a simpler example, given that all primes except 2 and 5 have last digits 1, 3, 7, or 9, one can initially expect these digits to occur equally often; that is, uniformly distributed as they would be if chosen at random.)

The Riemann Hypothesis is also supported by an argument from analogy (identified by Mazur as another probabilistic argument form common in mathematics). In some famous and difficult work, André Weil proved that the analogue of Riemann's Hypothesis is true for certain other zeta functions, and his related conjectures for an even more general class of zeta functions were proved to widespread applause in the 1970s. One expert says "It seems that they provide some of the best reasons for believing that the Riemann hypothesis is true – for believing, in other words, that there is a profound and as yet uncomprehended number-theoretic phenomenon, one facet of which is that the roots ρ all lie on Re $s = ½$" (Edwards 1974, 298). However, not all experts are convinced of the closeness of the analogy and outsiders are in no position to judge.

The Riemann Hypothesis, then, is a remarkable testbed for displaying the diversity of probabilistic evidence that can bear on a mathematical conjecture.

Similar stories could be told with the other famous unsolved problems of mathematics. For example, Goldbach's conjecture is generally strongly believed to be true (on the basis of both simple numerical and other evidence) but also to be very unlikely to be proved soon (Review in (Baker 2009), Sect. 3.2; evidence for some other currently unsolved conjectures listed in (Aberdein 2019), Sect. 1). It is also instructive to follow the historical development of evidence for now-proved conjectures, like Fermat's last theorem and the classification of finite simple groups. In those cases, the growing confidence of experts both that the conjecture was true and that proof was imminent turned out to be justified (Franklin 2014, Chap. 15). It has indeed been argued that there is a certain observable regularity in the time-to-proof of mathematical conjectures in general (Hisano and Sornette 2013). It would be hard to deny that probabilistic evidence in mathematics is worthwhile, in the face of evidence that it has proved in the past to be a reliable guide to the (provable) truth.

6 Probabilistic Relations Between Necessary Truths

There is one difficult question that needs consideration in applying Bayesian theory or nondeductive logic in *mathematics* in particular. P($h \mid e$) is intended to be a measure of the support that evidence e gives to hypothesis h. If e entails h, then P($h \mid e$) is 1, since h is certain given e. But in mathematics, the typical case is that e *does* entail h, though that is perhaps as yet unknown. If, however, P($h \mid e$) is really 1, how is it possible in the meantime to discuss the (nondeductive or probabilistic) support that e may give to h, that is, to treat P($h \mid e$) as less than 1? In other words, if h and e are necessarily true or false, how can P($h \mid e$) be other than 0 or 1?

The answer is that, in both deductive and nondeductive logic, there can be many logical relations between two propositions. Some may be known and some not. To take an artificially simple example in deductive logic, consider the following argument:

If all men are mortal, then this man is mortal
All men are mortal

Therefore, this man is mortal.

The premises entail the conclusion, certainly, but there is more to it than that. They entail the conclusion in two ways: firstly, by modus ponens, and secondly by instantiation from the second premise alone. That is, there are two logical paths from the premises to the conclusion. More realistic cases are common in mathematics, when different mathematicians produce different proofs of the same theorem, that is, different logical paths from the axioms to the theorem.

Now just as there can be two deductive paths between premises and conclusion, so there can be a deductive and nondeductive path, with only the latter known. Before the Greeks' development of deductive geometry, it was possible to argue the following:

All equilateral (plane) triangles so far measured have been found to be equiangular
This triangle is equilateral

Therefore, this triangle is equiangular

There is a nondeductive logical relation between the premises and the conclusion: the premises inductively support the conclusion. But when deductive geometry appeared, it was found that there was also a deductive relation, since the second premise alone entails the conclusion. This discovery in no way vitiates the correctness of the previous nondeductive reasoning or casts doubt on the existence of the nondeductive relation. That relation cannot be affected by discoveries about any other relation.

So the answer to the question "How can there be probabilistic relations between necessary truths?" is simply that those relations are additional to any deductive relations. They may be known independently of them (and are often easier to know).

Once that is established, it is possible to understand Pólya's remark that non-deductive logic is better appreciated in mathematics than in the natural sciences (1954, vol II, 24). In mathematics there can be no confusion over natural laws, the uniformity of nature, propensities, the theory-ladenness of observation, pragmatics, scientific revolutions, the social relations of science, or any other red herrings. There are only the hypothesis, the evidence, and the logical relations between them.

7 The Problem of Induction in Mathematics

As we have seen, induction, or inference from the observed to the unobserved, is found in pure mathematics as well as in the natural and social sciences. That has significance for understanding the nature of induction as well as for mathematics. If induction works in just the same way in pure mathematics as in science, its rationality and justification would seem to be independent of any contingent facts about the world, such as the obtaining of natural laws or the uniformity of nature.

Let us take an easy example, even simpler than the ones above concerning the roots of the Riemann zeta function.

The first million digits of π are random
Therefore, the second million digits of π are random

The first few digits of π are
3.14159265358979323846264338327950288419716939937510582097494 4
. . .

It can be seen that they are random in the sense of lacking pattern. (That notion is formalized in several ways, as in "passes standard statistical tests for randomness" or "incompressible," not "probabilistically generated," "stochastic" (Eagle 2018); of course the digits of π are generated by a deterministic formula. For present purposes, "looking patternless" is sufficient.) It is easily checked that the first million digits of π do both look random and pass statistical tests for randomness (such as having about the same number of each digit). (Calculations now stretch to 4 trillion digits in hexadecimal (Bailey et al. 2012).

Intuitively, that is good evidence that the patternlessness will continue into the second million digits, just in the same way that observing a million black ravens is good inductive evidence that further observed ravens will be black. It is widely believed that the digits continue to be random indefinitely, although it is not proved even that there are about equal numbers of each digit (Marsaglia 2005). (Some claims that the digits of π fail a more subtle statistical test for randomness in (Ganz 2014) with debate in (Bailey et al. 2017) and (Ganz 2017).)

Gronau and Wagenmakers (2018) argue that this is a perfect case for appreciating the Bayesian perspective in pure mathematics. They apply a formal Bayesian analysis to the hypothesis that π is normal (in base 10), that is, that the ten digits appear equally often in its decimal expansion. They conclude that "After all [the first]100 million digits [of π] have been taken into account, the observed data are 1.86×10^{30} times more likely to occur under [the hypothesis that π is normal] than under [the alternative hypothesis that is it not]. The extent of this support is overwhelming." As with any Bayesian analysis with exact numbers, the final conclusion depends on a choice of prior probabilities, and the authors discuss possible choices.

It is true, as argued by Baker (2007) that there is a special problem with inductive arguments in number theory in that all the observed cases are of small numbers. Any number that can be calculated with is very small, compared to numbers in general. That bias in the evidence could raise a question as to whether any induction of the form "All observed numbers have property X, therefore all numbers have property X" could have high probability. That does not imply, however, that inductive arguments in mathematics are generally poor. Firstly, a bias in the evidence towards small numbers does not affect inductive arguments with more modest conclusions, such as "All observed numbers have property X, so the next number calculated will have property X." (For example, the argument above about the randomness of the digits of π only extrapolated a finite distance, thus keeping to small numbers.) Secondly, many other inductive arguments have a bias in the evidence, without thereby becoming worthless (though they may become less secure). For example, extrapolative inductive inference like "All observed European swans are white, therefore all swans in the world are white" is a worthwhile inductive argument, even though the extrapolation beyond the observed range weakens it.

It is nevertheless true that purely inductive evidence is less credible in mathematics than it might be elsewhere, because of the possibility that some mathematical reason against a generalization might apply only for huge numbers. As we saw, the purely inductive evidence for the Riemann hypothesis was a possible instance of that, and it is possible to have "overwhelming" numerical evidence for a conjecture, where the reasons for it being false also reveal why there should be such good evidence for it (Schuster 1985). That is not the same, however, as saying that the evidence was not good in the first place or that a blanket skepticism is justified.

The occurrence of inductive arguments in pure mathematics suggests something about their nature as well as the nature of mathematics. If inductive arguments work in the necessary matter of pure mathematics, apparently on the same basis as they work in science, it seems they cannot depend on any contingent principles, such as the uniformity of nature or the holding of natural laws. The digits of π are surely the same in all possible worlds, and the rationality of arguments about them independent of any differences between possible worlds. That is more compatible with logical justifications of induction (e.g., Stove 1986) than ones that rely on either contingent facts like natural laws or sociological facts about what reasoners do. Again, it is easier to appreciate the force of probabilistic reasoning in mathematical cases, free of distractions.

8 Conclusion

The Bayesian perspective brings some order into part of the vast backstory of mathematical practice that underpins published research output. It allows us to understand the nature of the variety of methods of evaluating evidence – for pure mathematical conjectures, for the behavior of mathematical models, and for the performance of statistical algorithms. Experimental evidence, checking of consequences, weighing of intuitions, arguments from analogy, and expectations as to uniform distributions work the same way in mathematics as they do in natural and social sciences. The Bayesian perspective explains the unity of those methods and why they have rational force. The fact that they do work in the necessary matter of mathematics suggests that they are matters of pure logic.

9 Cross-References

- ▶ Experiments in Mathematics: Fact, Fiction, or the Future?
- ▶ Heuristics and Mathematical Practice
- ▶ Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice

References

Aberdein A (2019) Evidence, proofs, and derivations. ZDM 51:825–834

Amit G (2018) Riemann hypothesis likely remains unsolved despite claimed proof. New Scientist, 24 Sept. https://www.newscientist.com/article/2180504-riemann-hypothesis-likely-remains-unsolved-despite-claimed-proof/

Bailey DH, Borwein JM, Calude CS, Dinneen MJ, Dumitrescu M, Yee A (2012) An empirical approach to the normality of π. Exp Math 21(4):375–384

Bailey DH, Borwein JM, Brent RP, Reisi M (2017) Reproducibility in computational science: a case study: randomness of the digits of pi. Exp Math 26(3):298–305

Baker A (2007) Is there a problem of induction for mathematics? In: Leng M, Paseau A, Potter M (eds) Mathematical knowledge. Oxford University Press, Oxford, pp 59–73

Baker A (2009) Non-deductive methods in mathematics. Stanford Encyclopedia of Philosophy. http://plato.standord.edu/entries/mathematics-nondeductive

Boij M (1999) Betti numbers of compressed level algebras. J Pure Appl Algebra 134(2):111–131

Borwein J, Bailey D (2004) Mathematics by experiment: plausible reasoning in the 21st century. AK Peters, Natick

Connes A (2016) An essay on the Riemann hypothesis. In: Nash JF, Rassias MT (eds) Open problems in mathematics. Springer, Cham, pp 225–257

Derbyshire J (2003) Prime obsession: Bernhard Riemann and the greatest unsolved problem in mathematics. Joseph Henry Press, Washington, DC

Eagle A (2018) Chance versus randomness. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/chance-randomness/

Easwaran K (2009) Probabilistic proofs and transferability. Philosophia Mathematica 17(3): 341–362

Edwards HM (1974) Riemann's zeta function. Academic, New York

Fallis D (1997) The epistemic status of probabilistic proof. J Philos 94(4):165–186

Franklin J (1987) Non-deductive logic in mathematics. Br J Philos Sci 38(1):1–18
Franklin J (2011) The objective Bayesian conceptualisation of proof and reference class problems. Sydney Law Rev 33(3):545–561
Franklin J (2014) An Aristotelian realist philosophy of mathematics: mathematics as the science of quantity and structure. Palgrave Macmillan, Basingstoke
Franklin J (2016) Logical probability and the strength of mathematical conjectures. Math Intell 38 (3):14–19
Ganz RE (2014) The decimal expansion of π is not statistically random. Exp Math 23(2):99–104
Ganz RE (2017) Reply to "reproducibility in computational science: a case study: randomness of the digits of Pi" [Bailey et al. 17]. Exp Math 26(3):306–307
Gronau QF, Wagenmakers E-J (2018) Bayesian evidence accumulation in experimental mathematics: a case study of four irrational numbers. Exp Math 27(3):277–286
Harris P, Fotheringham AS, Crespo R, Charlton M (2010) The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. Math Geosci 42(6):657–680
Hersh R (1991) Mathematics has a front and a back. Synthese 88(2):127–133
Hisano R, Sornette D (2013) Challenges to the assessment of time-to-proof of mathematical conjectures. Math Intell 35(4):10–17
Hoffmann J, Bar-Sinai Y, Lee LM, Andrejevic J, Mishra S, Rubinstein SM, Rycroft CH (2019) Machine learning in a data-limited regime: augmenting experiments with synthetic data uncovers order in crumpled sheets. Sci Adv 5(4):eaau6792
Horgan J (1993) The death of proof. Sci Am 269(4):92–103
Howson C, Urbach P (2006) Scientific reasoning: the Bayesian approach, 3rd edn. Open Court, Chicago
Inglis M, Mejía-Ramos JP (2009) The effect of authority on the persuasiveness of mathematical arguments. Cogn Instr 27(1):25–50
Jaynes ET (2003) Probability theory: the logic of science. Cambridge University Press, Cambridge
Keynes JM (1921) A treatise on probability. Macmillan, London
Krantz SG (2011) John Horgan and "the death of proof?". In: The proof is in the pudding. Springer, New York, pp 219–222
Marsaglia G (2005) On the randomness of pi and other decimal expansions. http://www.yaroslavvb.com/papers/marsaglia-on.pdf
Mazur B (2014) Is it plausible? Math Intell 36(1):24–33
Ornes S (2020) How close are computers to automating mathematical reasoning? Quanta Magazine, Aug 27. https://www.quantamagazine.org/how-close-are-computers-to-automating-mathematical-reasoning-20200827/
Paseau A (2015) Knowledge of mathematics without proof. Br J Philos Sci 66:775–799
Platt D, Trudgian T (2020) The Riemann hypothesis is true up to $3 \cdot 10^{12}$, arXiv:2004.09765
Pólya G (1954) Mathematics and plausible reasoning (vol. I, Induction and analogy in mathematics, and vol. II, Patterns of plausible inference). Princeton University Press, Princeton
Rabin MO (1980) Probabilistic algorithm for testing primality. J Number Theory 12(1):128–138
Schuster EF (1985) On overwhelming numerical evidence in the settling of Kinney's waiting-time conjecture. SIAM J Sci Stat Comput 6(4):977–982
Sørensen HK (2016) 'The end of proof'? The integration of different mathematical cultures as experimental mathematics comes of age. In: Larvor B (ed) Mathematical cultures. Birkhauser, Cham, pp 139–160
Steckles K, Lawson-Perfect C (2018) Atiyah Riemann Hypothesis proof: final thoughts, aperiodical.com, Sept 28. https://aperiodical.com/2018/09/atiyah-riemann-hypothesis-proof-final-thoughts/
Stove D (1986) The rationality of induction. Clarendon, Oxford
Tsai W, Chen S, Lu G (2015) Numerical evidence of turbulence generated by nonbreaking surface waves. J Phys Oceanogr 45(1):174–180

Wilf HS (2008) Mathematics: an experimental science. In: Gowers T (ed) Princeton companion to mathematics. Princeton University Press, Princeton, pp 991–999
Williamson J (2010) In defence of objective Bayesianism. Oxford University Press, Oxford
Zeilberger D (1993) Theorems for a price: tomorrow's semi-rigorous mathematical culture. Notices Am Math Soc 40(8):978–981
Zeilberger D (2012) Appendix to Doron Zeilberger's Opinion 117: Links to posted Grant Proposals. https://sites.math.rutgers.edu/~zeilberg/Opinion117Appendix.html

Polymath as an Epistemic Community

Analyzing the Digital Traces of the Polymath Project

Patrick Allo, Jean Paul Van Bendegem, and Bart Van Kerkhove

Contents

1 Introduction 2728
2 Philosophical Import 2729
3 Perspectives on Polymath 2731
4 Polymath as E-Research? 2734
5 Description of the Data 2737
6 Global Participation Data 2740
7 Affiliations and Joint Affiliations 2743
8 The Relative Weight of Participants 2746
9 Specific Project Communities 2750
10 Conclusion 2753
11 Cross-References 2754
References 2754

Abstract

The Polymath Project is an online collaborative enterprise that was initiated in 2009, when Timothy Gowers asked whether and how groups could work together to solve mathematical problems that "do not naturally split up into a vast number of subtasks." Gowers proposed to answer this question himself by actually trying to set up such a collaboration, based on interactions taking place in the comment-threads of a series of posts on a WordPress blog. Hence, the first project officially started in early 2009, to be proclaimed successful only 6 weeks later (Gowers and Nielsen 2009). From its inception until April 2018, 15 more Polymath problems (and a handful of smaller or related ones) have been launched. These projects have attracted attention from different scholarly communities, including the philosophy of mathematical practices, from the perspective of which the Polymath Project can be seen as a vast repository of mathematics in action. This chapter continues previous work by its authors on the topic in question and topics

P. Allo · J. P. Van Bendegem · B. Van Kerkhove (✉)
Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium
e-mail: patrick.allo@vub.be; jean-paul.van.bendegem@vub.be; bart.van.kerkhove@vub.be

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_86

related. More specifically, for the purposes of this volume, it is our aim to both summarize and expand upon these earlier contributions. The starting point is the above observation that in the past decade, the issue of "massively collaborative mathematics" (to be qualified below) has drawn quite some attention. However, does it also warrant intensive *philosophical* attention in particular? For, as exciting as these developments might be from mathematical and other points of view, the enhanced possibilities that have come with it do not per se give rise to any philosophical import. In Van Bendegem (2011), one of us has indeed tentatively argued for the philosophical relevance of these new dynamics of proof construction flowing from the wide availability and efficiency of Internet technology. We shall below briefly rehearse and update the argument given there.

Keywords

Polymath Project · Massively collaborative mathematics · Dynamics of proof construction · Network analysis

1 Introduction

The Polymath Project is an online collaborative enterprise that was initiated in 2009, when Timothy Gowers asked whether and how groups could work together to solve mathematical problems that "do not naturally split up into a vast number of subtasks." Gowers proposed to answer this question himself by actually trying to set up such a collaboration, based on interactions taking place in the comment-threads of a series of posts on a WordPress blog. Hence, the first project officially started in early 2009, to be proclaimed successful only 6 weeks later (Gowers and Nielsen 2009). From its inception until April 2018, 15 more Polymath problems (and a handful of smaller or related ones) have been launched. These projects have attracted attention from different scholarly communities, including the philosophy of mathematical practices, from the perspective of which the Polymath Project can be seen as a vast repository of mathematics in action. This chapter continues previous work by its authors on the topic in question and topics related. More specifically, for the purposes of this volume, it is our aim to both summarize and expand upon these earlier contributions. The starting point is the above observation that in the past decade, the issue of "massively collaborative mathematics" (to be qualified below) has drawn quite some attention. However, does it also warrant intensive *philosophical* attention in particular? For, as exciting as these developments might be from mathematical and other points of view, the enhanced possibilities that have come with it do not per se give rise to any philosophical import. In Van Bendegem (2011), one of us has indeed tentatively argued for the philosophical relevance of these new dynamics of proof construction flowing from the wide availability and efficiency of Internet technology. We shall below briefly rehearse and update the argument given there.

In Allo et al. (2013), we have already proposed a preliminary theoretical model, based on epistemic logics, aimed at capturing (and understanding) how different

pieces of mathematical knowledge can be distributed within a larger community. As it happens, we shall also be borrowing arguments for the philosophical relevance of the phenomenon in question from that paper. The original contribution of the present chapter is of a more empirical nature, in that we study various Polymath projects in terms of the scientific communities themselves that emerged within and across them. We believe that this more social perspective can be particularly fruitful, as it allows us to address the two central goals of the Polymath Project jointly, namely enhanced efficiency and transparancy through collective proof-search processes. Although we admit that quantitative and network-oriented studies of the Polymath Project have been carried out before, this has never been executed at the present scale, and certainly not as an explicit contribution to the philosophy of mathematical practices.

Here is an overview of the chapter. We begin, in Sect. 2, by arguing in favor of the (potential) epistemological significance of considerations about social phenomena such as the Polymath Project. Section 3 continues with an overview of different (disciplinary) perspectives on said phenomena, in which we situate our own contributions. After that, the entire Sect. 4 is devoted to assessing the possible status of the Polymath Project as an instance of e-research. This concludes the first part of the chapter, which is rather surveying and reflective in kind. In the second part, we present the methodology and conclusions of our empirical work. Section 5 describes the nature and structure of the source material used. In Sect. 6, we make a global evaluation of various data concerning participation or affiliation, both to projects in general and to discussion-threads in particular. Next, Sect. 6 refines this analysis by also considering joint affiliations, giving rise to the tentative identification of collaborative networks per project. In a following step, Sect. 7 more specifically examines the role of core participants in the global project by means of the distribution of network links. Finally, in Sect. 8, we study the communities associated with single projects. We summarize our findings in Sect. 9.

2 Philosophical Import

In the years following the start of the Polymath Project, bold claims were made as to its repercussions. For example, Aron (2011) predicted that it would "democratise the process of mathematical discovery," while Nielsen (2012) firmly stated: "The Polymath Project is a small part of a much bigger story, a story about how online tools are transforming the way scientists make discoveries." Although we do not immediately share these very ambitious and drastic expectations, at the very least we would like to make a case for due philosophical attention to the developments in question.

Oddly enough, this requires a philosophical preparedness or openness to begin with. This attitude has been associated with accepting a degree of naturalism with respect to epistemological questions, making this type of philosophical questions about (the reliability of) mathematical results to an extent depend on matters of fact about mathematical inquiry as a process. As this chapter is part of a comprehensive handbook on the philosophy of mathematical practice, where this basic attitude is the

common point of departure, we shall not spend much time on arguing this general point. Once this is granted, clearly, the road to also take the Polymath Project as a legitimate topic of philosophical scrutiny (any significant results pending) is open. The reasons for opposing such a move are connected to the stubborn opposite philosophical apriority, namely that mathematical truths are cut loose from any empirical considerations related to their discovery or justification, and are thus independent of anything remotely connected to proof construction as a human affair. Neither individual nor collective beliefs have any bearing upon the acceptance of mathematical results. Only formal proof, proceeding by necessity from premises to conclusion, has the authority of establishing a mathematical knowledge claim.

There is a fundamental problem with the characterization just given, though. How does one establish that a particular formal proof is a or the correct one? Proofs do not tend to fall from the sky but are invariably conceived by concrete mathematicians (typically an individual or a small group of them) and subsequently have to be validated by others, a (mostly small) selection of peers, in order to have their validity substantiated. (Admittedly, the (growing) role of automated verification and other "proof technology" should also be acknowledged in this respect, a matter we can and will however not elaborate on here.) In other words, proofs have to be *recognized* to be correct, which typically involves any proofreaders to be *convinced* by the arguments on offer. Allow us to point to the fact that already at this stage, which is still to be associated with a rather traditional position within the epistemology of mathematics, an embryonic community of participating mathematicians is involved in sanctioning alleged results. This directly points to an irreducible reliance on what in the epistemological literature is called testimony, giving rise to fallibility. In the ideal world, that is, referees of proposed papers are supposed to establish whether the claims to absolute justification made by their authors are correct. However, we do not live in this ideal world, and the level of scrutiny of the refereeing process in reality varies greatly (Geist et al. 2010). Another aspect concerns the fact that published proofs are not just chains of formal derivations, but mostly also contain portions of motivational prose or drawings, sketchy (thus incomplete) arguments, and appeals to implicit knowledge. Ironically, these very aspects are often what makes proofs surveyable and robust to begin with (Van Bendegem and Van Kerkhove 2009). (We should also like to refer here to the chapter "The Social Epistemology of Mathematical Proof" by Line Edslev Andersen in the section "Proof" of this handbook for a more systematic discussion of various positions with respect to social proof dynamics.)

Let us now, however, get back to the more specific question that is facing us here: Is the Polymath Project an instance of an entirely new procedure of proof search in mathematics, or does it simply amount to the same old methods being implemented through new media, most notably the Internet? We are convinced the former is rather the case than the latter. In Allo et al. (2013), we have already made a first effort to clarify the matter from the perspective of epistemic logic, by suggesting a formal way to study and understand the construction of mathematical proof *within* communities, theoretically enabling one to model proofs-as-arguments from a perspective that emphasizes interaction and collaboration. In a similar vein, Van Bendegem (2011: Sect. 3) has

defended the idea that the way mathematical communities are organized does exert an influence on their problem-solving capacities and suggested the use of logics for multiagent systems (such as Dunin-Keplicz and Verbrugge (2010)) and dialogue logic to represent varying constellations of collaboration, whether or not including team leaders (as has been the case in the Polymath Project, see below).

Without being able to provide a knockdown argument allow us to add a philosophical rationale for at least exploring this possibility, inspired by Van Bendegem (2011: Sect. 2) (to which we refer for the complete, refined argument). The main point, we would like to make in this respect, exploits the idea introduced above that, given the finite world we live in, the problem-solving capacities of the mathematical community are limited. Hence, in mathematics, as elsewhere in human practice, we are dealing with an issue of resources. This gives rise to practical questions concerning the selection of problems that will be tackled and the efforts that will be made to do so (in terms of how many mathematicians will invest how much of their time in order to try and bring cases to a close). This in itself clearly does not preclude any of the possible problems to be tackled with a sufficient amount of research power.

Still, there is more to this. For in some cases, there appear to be absolute limits anyway, as problems present themselves the solution of which escapes the regular kind of individual-based, mathematical practice. Outside the area of massive or open collaboration, some proofs have indeed appeared to be well beyond the reach of individuals or smaller groups of them. A famous case in point is the Classification Theorem of Finite Simple Groups, also known as the Enormous Theorem (Van Bendegem and Van Kerkhove 2009: Sect. 4). This defies the view that, as the whole enterprise of mathematics is traditionally seen as the cumulation of individual achievements, no real interaction is needed beyond the obvious reliance on prior results. Yet, this description no longer covers the totality of mathematical inquiry.

In the present context, the crucial point is that the Polymath Project appears to have been conceived with this very same idea in mind. Here is what Timothy Gowers wrote on his blog in this respect: "I'm interested in the question of whether it is possible for lots of people to solve one single problem rather than lots of people to solve one problem each." (https://gowers.wordpress.com/2009/02/01/why-this-particular-problem/ (accessed 19 Jul 2020)) Cranshaw and Kittur (2011) confirm this, specifying that "Gowers was instead proposing to collaboratively solve a single problem that does not naturally split up into a vast number of subtasks." Given what was developed above, the philosophical relevance of studying the research dynamics that have been the product of this particular type of collaboration should be clear. As we develop our network analysis below (from Sect. 5 onward), we shall have the opportunity to point to various specific aspects of such philosophical interest.

3 Perspectives on Polymath

Several perspectives on the study of the Polymath Project can be adopted, corresponding to different categories of scholarly literature devoted to the topic. We are distinguishing three of them here (with a fourth appended to that).

First, one may think of Polymath as a specific (though atypical) example of science online (e-research, crowd / citizen science). This is the perspective adopted by scholars who study social and scientific practices online. Polymath indeed figures in some overviews of scientific collaborations online such as Franzoni and Sauermann (2014), and Michael Nielsen's "Reinventing Discovery" (Nielsen 2012). In Nielsen (2010), the latter described himself as an interested outsider participant of the first Polymath project (he later also contributed to the discussions in Polymath 4 and 8) and decisively contributed to the fame of the Polymath Project with his writings (Gowers and Nielsen 2009). Since Polymath indeed conforms to the definitions of e-science and crowd science, and undoubtedly shares some of the motivations of citizen science, this approach is well motivated, often highlighting features that remain absent from the small-scale analyses devoted to the fine-grained structure of mathematical arguments. Yet, as we shall argue in Sect. 4, since Polymath remains a rather unusual example of science online, this particular perspective cannot always do justice to the epistemic goals and achievements of the Polymath Project.

Second, one may think of Polymath as a rich and detailed repository of real mathematics in action. This is the traditional and most obvious perspective to adopt on the Polymath Project, when interested in the tiny and often messy details of mathematical practices. It is moreover a very natural perspective to adopt, since Polymath does effectively provide access to highly valuable data that normally remain hidden. Indeed, the entire project has been motivated by the desire to make the inner practice of research mathematics visible to outsiders to begin with, and most scholars of mathematical practice are to be considered a highly specialized group of such outsiders. From this perspective, the mathematical content of the collaboration (such as the types of arguments used by mathematicians) is the focus of attention. Other features are less salient, including social dynamics and the role of ICTs in the collaboration. As the targets are fine-grained analyses, and given the need to take into account the precise content of individual contributions, this kind of research has typically focused on smaller fragments of Polymath projects, leaving less room for overarching analyses. This approach also raises an important question: Can we really use Polymath as a source for the study of normal mathematical practices, or should we take the claim that it constitutes a "new kind of mathematics" seriously and therefore accept that we actually study a novel practice?

This second category notably includes work by Ursula Martin and Allison Pease, with its emphasis on problem-solving and argumentation in mathematics (as influenced by Lakatos, by Pólya, and by literature in formal argumentation), and with the stated intent to improve how large groups collaborate by making argumentative structures explicit. The focus is thereby on the use of ICTs in mathematics, relating to the metaphor of social machines (Lane et al. 2019: Sect. 5). Specifically, the third "mini" Polymath project was analyzed this way (these "mini" projects having been dedicated to Maths Olympiad-finals problems), identifying different categories of comments and the mathematical concepts these were relying on, and developing a typology of questions based on samples from MathOverflow, the discussion site for professional mathematicians on the StackExchange platform (Pease and Martin 2012; Martin and Pease 2013).

Incidentally, our own earlier work also belongs to this category (Allo et al. (2013), Van Bendegem (2011): Sect. 3).

The third category is dedicated to how specific information technologies are used for collaboration, and can naturally be related to hypertext research (in particular the use of blogs and wiki's) and humanistic computing. This category includes the work of Barany (2010), Cranshaw and Kittur (2011), and Varshney (2012). The authors of these papers try to identify specific features of how collaboration in the various Polymath projects has proceeded (e.g., remarks on noise-to-signal ratio, or the importance of leadership), propose criteria for success, and seek to evaluate to what extent collaboration in Polymath is successful. Unlike the present chapter, the focus in this work is limited to the first Polymath project. However, it does rely on social science methodology, network-analysis in particular, as will be done here. More recently, Kloumann et al. (2016) have specifically investigated the contrast between research-level Polymath projects (1–9) and "mini" Polymath projects, trying to disclose patterns of collaboration and linguistic features that can be used to distinguish collaboration on hard problems from collaboration on feasible ones.

The previously mentioned studies have inspired our current approach to Polymath as an epistemic community. This new dimension to the philosophical perspective arises naturally when taking into consideration the following features of the Polymath Project:

(i) The diversity of the epistemic goals and the value accorded to social-epistemic virtues like the openness of the enterprise, both in the sense of openness to potential contributors, and the disclosure of intermediate inputs. (Gowers 2010; Franzoni and Sauermann 2014)
(ii) The fact that Polymath is not governed by a technical platform, but by a set of social norms or rules that were initially put forward by Timothy Gowers (Van Bendegem 2011) and that still govern the interactions during the projects.
(iii) The presence of deliberative processes where questions concerning the research-agenda, or concerning the organisation and structure of the interactions, are collaboratively resolved. Some examples: (1) During the first Polymath project, there were several discussion-threads in which organizational and practical issues were discussed. This process led to further amendments to the rules Gowers initially proposed; (2) In September 2015, Gil Kalai posted a question on MathOverflow to crowdsource proposals for new Polymath projects. This question remains active, with regular input and occasional summaries and rankings of the proposals (https://mathoverflow.net/questions/219638/proposals-for-polymath-projects); (3) Projects typically start with a discussion-thread to gauge the interest before the project is officially launched.

Methodologically, this approach is continuous with some of the earlier quantitative and network-theoretic analyses of Polymath, including Cranshaw and Kittur (2011), Sarvate et al. (2011), and Kloumann et al. (2016). We do, however, subscribe to a more encompassing perspective on scientific communities, hoping to thereby lay the foundations for more substantial analyses of the socio-epistemological

dimensions of Polymath (or similar) projects. One type of questions, which can be explored along these lines, concerns the social norms or rules that govern this type of research, and especially the role of the "small contribution rule" that prevents the emergence of information-asymetries by asking for small and quick contributions, and discouraging longer periods of individual work (Wagenknecht 2014). Our research falls squarely within the category of "quantitative" approaches to the use of crowdsourcing in mathematical practices as recently identified by Pease et al. (2020: Sect. 4), in being aimed at developing and employing quantitative techniques (such as network-analysis) to study and test intuitions about said practices.

Before continuing with our own work, we shall in the upcoming section investigate in some more detail the perspective from e-science. But as announced, we cannot leave the current section without pointing to an obvious fourth category in the scholarly literature, broadly conceived. In this fourth category, we find insider and popular science (press) reports on the Polymath Project, which were the decisive element in how it gained notoriety, including in the first place Gowers and Nielsen's opinion piece in *Nature* (Gowers and Nielsen 2009) and Nielsen's book *Reinventing Discovery: the New Era of Networked Science*, in which Polymath featured prominently (Nielsen 2012). Additional inside reports include Nielsen (2010) and Gowers (2010), which covered the problem solved in the first Polymath project, the views about collaboration in mathematics that motivated the project, and lessons learned from organizing a large collaborative project. Finally, articles in *New Scientist* (Aron 2011), *Scientific American* (Castelvecchi 2010), *The Mathematical Intelligencer* (Nathanson 2010), and a short piece in *Nature* (Ball 2014) can be mentioned, as well as various blog posts and online publications such as *Quanta Magazine* (Klarreich 2013, 2015).

4 Polymath as E-Research?

The Polymath projects clearly fit the definition of e-research as research based on the use of digital tools and data – and the Internet in particular – for the distributed and collaborative production of knowledge (Meyer and Schroeder 2015: Chapt. 1). Because it leaves project participation and access to intermediate inputs entirely open, it also qualifies as a type of crowd science, as conceived in Franzoni and Sauermann (2014). Finally, because it is in principle open to amateur mathematicians, it is also a kind of citizen science. At the same time, the Polymath Project remains in several respects an outlier, and this presumably accounts for its relative absence from the literature in this field (with the exceptions noted in the previous section). We review some of the reasons for this exceptional status.

First, unless we compare it to paper and pencil mathematics, the Polymath Project is by contemporary standards a fairly low-tech enterprise. It relies on off-the-shelf tools, namely freely available blogging-platforms (and to a lesser extent Wiki software), and became possible through the availability of a LaTeX-plugin for WordPress. (See Gowers (2010). A richer discussion of this technological aspect was given by Jon Udell, a prominent voice on ICT research-infrastructure (https://blog.

jonudell.net/2009/07/31/polymath-equals-user-innovatio/.) Continuing on this theme, it is worth pointing out that the write-up process in Polymath 16 relied on the LaTeX-centric collaborative tool Overleaf (https://dustingmixon.wordpress.com/2019/08/05/polymath16-fourteenth-thread-automated-graph-minimization/#comment-24584).)
This is very different from most other types of e-research, which typically rely on the development of dedicated tools for collaboration and have in the case of Zooniverse even led to the creation of a reusable tool or platform for online collaboration (https://www.zooniverse.org/). Polymath, by contrast, is primarily defined by a set of rules, initially formulated by Gowers, which constitute a recipe for collaboration that can be easily reimplemented, and is only to a lesser extent characterized by the specific technology it relies on (blogs, but also wikis). Its appeal also remains limited to the formal sciences, with proposals to implement the same kind of open collaboration in AI (Zadrozny et al. 2015) and as a means to create research opportunities for mathematics undergraduates (Parsley and Rusinko 2017).

Second, Polymath is clearly a small-scale enterprise when compared to many collaborative online projects. The largest Polymath project (Polymath 8) attracted 180 participants, whereas the smaller projects (2 and 6) did not even reach 20 participants. Over a period of more than 9 years, 16 projects were initiated, and 622 distinct participants can be identified (see the detailed analysis below). By comparison, the seven Zooniverse projects reviewed in Sauermann and Franzoni (2015) reached between 3186 and 28,828 participants in their first 180 days. This makes the description of Polymath as "massive online collaboration" somewhat excessive when it is purely judged by the numbers. This is not really surprising. The scale of collaboration within mathematics is generally smaller than in other sciences (Grossman 2002), and this substantially lowers the threshold for what we may consider large, or even massive, collaborations. (Additionally, one could make the claim that there is a gap between Polymath-like "small science" and Zooniverse-like "Big Science.")

More importantly, Polymath does not fit the mantra of citizen science that "anyone can be a researcher." Even though participation is open to amateurs, and the choice for problems in combinatorics is at least partly motivated by the fact that knowledge of a small number of techniques can be sufficient to tackle open problems (Gowers 2000), active contribution is in reality limited to those with the required expertise (this being one dimension of the taxonomy of crowd science in (Franzoni and Sauermann 2014)). Polymath is meant to succeed in virtue of the fact that "different people know different things, so the knowledge that a large group can bring to bear on a problem is significantly greater than the knowledge that one or two individuals will have" (Gowers' initial post), or by leveraging so-called micro-expertise (Nielsen 2012: Chapt. 4). One may therefore hesitate to see this type of approach as one that relies on the illustrious Wisdom of Crowds, and even doubt that there is a clear kind of epistemic contribution that the Polymath projects are crowdsourcing. At this point, we can only indicate that the intended individual contributions are neither complete proofs (cfr. The one correct answer that is typically the goal on the StackExchange pages) nor clearly identifiable or even just potential fragments of such a proof.

Third, in contrast to both Zooniverse-like endeavors and large-scale collaborations in the sciences, the problem statement in Polymath projects is clear and explicit (Polymath 2 is probably an exception), but individual tasks are rarely clearly individuated and are never explicitly attributed to specific participants. Mainstream citizen science is typically based on a huge amount of similar tasks that are explicitly attributed to participants in such a way that all tasks are not only completed, but are also completed multiple times to allow for error correction. In Watson and Floridi (2018), large numbers of participants and hard-coded error correction strategies are explicitly related to the higher reliability of crowdsourced science. Similarly, outside the scope of citizen science, e-research and other large collaborative enterprises exhibit "a high degree of mutual dependence and task certainty" (Meyer and Schroeder (2015), building on the work of Whitley (2000)) that has no clear counterpart in the Polymath projects. This leads Franzoni and Sauermann (2014) to classify the Polymath projects as ill-structured projects that are characterized by the lack of independent subtasks (this being the other dimension of the taxonomy of crowd science referred to above).

The absence of hard-coded or other explicit error-correction procedures is equally important and should not entirely be traced to the lack of explicit subtasks (if the subtasks are unclear, the criteria for a successful microcontribution are equally unclear). Instead, it should also be linked to the fact that the aggregation of contributions is itself distributed instead of centralized. Because all contributions or comments are accessible to anyone, they are in principle open to scrutiny by every participant as well. The absence of procedures governing this has two potential drawbacks. First, it creates uncertainty concerning whether and how error-correction takes place. And second, it makes it hard for individual participants to identify the current state of the project (this being understood as what everyone should at a certain moment be agreeing on).

The open-ended and informal character of Polymath-contributions makes it harder to understand and explain the precise sense in which Polymath-like projects lead to more reliable or potentially more successful models for mathematical inquiry. Specifically, the highlighted contrasts with other forms of e-research and crowdsourced science reveal that epistemic success in Polymath cannot be retraced to good planning and reliable procedures. Other factors must instead be at play, and some of these factors may be grounded in the structure of the community or the specific shape of the patterns of interaction between the participants. As already highlighted, it would be a mistake to identify the epistemic value of Polymath-style collaborations to their potential to lead to mathematical results more quickly or efficiently. One explicit goal of the Polymath Project is precisely to reveal the inner workings of proof-search. Nielsen (2010: p.655) relates Polymath to "open science," but Gowers (2010: Sect. 2) is even more explicit, referring to the inefficiency of published proofs in conveying how they are discovered. If we take this goal into account it is immediately clear that a (hidden) centralized aggregation-procedure is of little use. We will need alternative criteria to understand why and how Polymath is valuable and potentially successful.

Fourth and last, it seems that Polymath-style research also diverges from mainstream e-research in the sense that, despite its high pace, Polymath can be understood as a form of slow science:

> I think with a large collaboration of this kind we can aim not only for success in settling the problems (which is, of course, the central goal) but also for higher standards in explaining the proof, improving the argument, making sure it is locally correct and detailed (hyperlinks can even allow to give details that are necessary for some while obvious for others), and in the exposition level of the paper. Since everything is in the open there is no pressure to rush on these matters. (Gil Kalai, comment of March 15, 2009 during Polymath 1) (https://gowers.wordpress.com/2009/03/10/polymath1-and-open-collaborative-mathematics/)

Thus, Polymath relies on "unused epistemic resources" in only one of two ways. It certainly does so in the sense that it seeks to reap the benefits of microexpertises that one would otherwise not have had access to, but it does not in terms of saving epistemic efforts (or attention). The latter point is underscored by several participants, pointing to the substantial investment of taking part, thereby denying that Polymath is mainly driven by reasons of economy, tapping into otherwise unused intellectual resources (Nielsen in Polymath 2014).

5 Description of the Data

After the previous, more encyclopedic, and philosophical sections, we now turn to the original, more technical contribution of this chapter, which is to offer a network-oriented analysis of the Polymath Project. We begin this endeavor in the present section, by describing the data that was used to this effect. Said data were obtained by scraping the discussion-threads of all the Polymath projects on the following blogs:

- Timothy Gowers' blog
- Tererence Tao's blog
- Gil Kalai's blog
- The Polymath-blog
- The Secret Blogging Seminar (only for some threads in Polymath 8)
- The Short, Fat Matrices blog run by Dustin G. Mixon (only for some threads in Polymath 16)

Lists of all the comment-threads were obtained from the Polymath-wiki at http://michaelnielsen.org/polymath1/index.php?title=Main_Page. The content and structure of the interactions was obtained by scraping and parsing the web pages of all comment-threads.

The structure of these data is related to:

Projects: There are 16 Polymath projects, each devoted to one research question.

Threads: Each project consists of a series of **blog-posts** and associated **discussion-threads**. Blog-posts are authored by the leader of the project. This is typically only one person. Polymath 1 is the main exception with parallel threads hosted on separate blogs.

Comments: Each discussion-thread is a succession of comments.

- With an author (which we refer to as **participants** in order to distinguish them from the authors of the blog-posts)

- With a content (of a given length or wordcount)
- Which are temporally ordered by their **timestamps**
- Which are organized in a tree-like structure, with:
  - Level 1 comments as child-nodes of the blog-post
  - Level $n + 1$ comments (for $n \geq 1$) as child-nodes of level n comments

In Fig. 1, the tree-like structure as well as the temporal ordering is depicted. The color of the nodes corresponds to the identity or at least the supplied name of the authors of the comments.

Where $n \geq 1$, the child-parent relation between comments is an explicit *reply-to* relation between the author of the child-comment and the author of the parent-comment, while the parent-child relation between these comments is an implicit *has reached* relation between the author of the parent-comment and the author of the child-comment. As said, we shall refer to the authors of comments as *participants*.

Participants can be identified within and across threads and projects by their supplied author-name. All these where compared for string-similarity. Similar names that could be retraced to the same commenter were reidentified, but ambiguous cases such as slight variants of the same name without surname were kept distinct. Various versions of "anonymous" are present in the data. Comments authored by an anonymous participant have not been attributed to a single participant but instead have been removed from the detailed analyses, with one exception. As we noticed an exceptionally high number of anonymous comments in Polymath 15 (more than in all other projects together), we decided that there were sufficient indications to

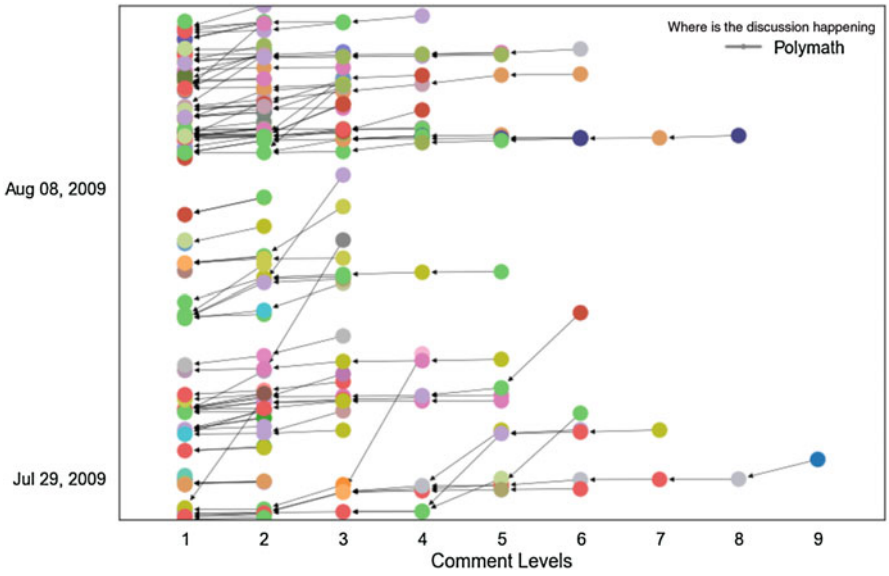


Fig. 1 Fragment of a discussion-thread in Polymath 4

identify these contributions as coming from a single participant, whom we renamed as "PM15 Anonymous" and treated as a regular identifiable participant.

The structure described captures the quantitative and relational information that can be obtained by scraping all the posts and discussion-threads.

While textual data could have been obtained from the content of the posts and comments, these have been left out of the analysis. First, this is not our focus here, but second, it does also avoid a further complication related to the fact that all LaTeX-content is rendered as figures, and hence unavailable for scraping. This makes the bulk of the mathematical content practically inaccessible to automated processing with standard tools for pulling data out of HTML such as Beautiful Soup, which naturally forms a huge obstacle to (comprehensive) detailed scrutiny. The codebase for our analysis can be found on GitHub at https://github.com/patrickallo/mathpracmod.

As we study the Polymath Project as a social network, we have placed most of our emphasis on the analysis of **relational data**. We can distinguish between the following types:

1. Relations that encode direct interactions
2. Relations that encode affiliations

   Relations derived from direct interactions are only defined between entities of the same kind. In our case, mainly participants. Affiliation-networks are, by contrast, bipartitite networks with relations only defined between entities of different kinds (Borgatti et al. 2018: Chapt. 13).

   When we study the Polymath community as a whole, we consider affiliations between participants and projects, and between participants and discusssion-threads. When we study the communities associated with a single project (particularly in Sect. 8), we consider affiliations between participants and threads, and between participants and episodes of intense interaction. Such episodes are identified by means of a clustering-algorithm.

   Given an affiliation-network, we can construct a so-called projected-graph that connects entities of one kind based on their joint affiliations. This gives rise to a third type of (derived) relational data:
3. Relations that encode joint affiliations

   It is thus natural to posit connections between participants based on their joint participation to projects, but the converse makes sense as well: We can connect projects based on the participants they have in common.

   When we describe the Polymath community as a whole, we consider coarse-grained affiliations. In such cases, the joint-affiliation relations we study are rather loose and do not even indicate that participants actually interacted. When we describe specific project-communities, joint affiliations to episodes do indicate actual interactions.

   Overall, we have found that the analysis of joint affiliations to episodes is a more reliable guide to the structure of interactions within Polymath than the analysis of direct interactions based on reply-to relations. We describe and motivate this procedure in more detail in Sect. 6.

When we consider affiliations, between participants and projects, and set the strength of an affiliation (the weight we assign to the edge) to a project to be the number of comments contributed by a participant to a project, then: (1) The degree of each participant in this bipartite network corresponds to the number of projects (the unweighted case) or the number of comments (the weighted case) of that participant. (2) The degree of each project in this bipartite network corresponds to the number of participants (the weighted case) or the number of comments (the unweighted case) in that project.

An important remark. Several fundamental metrics that are used to characterize relational data should be used with some caution in the present context, as they provide no insight beyond the analysis of the more basic scalar data used to deduce the existence of relations between participants and projects or threads. It is therefore useful to, first, make these connections explicit, and, second, to start our analysis with a direct exploration of these scalar data.

6 Global Participation Data

Let us begin by counting the comments and the participants in each project to get a first idea of their size as well as of the size of the global community involved (Fig. 2).

There have been 16 research-level Polymath projects over the 9 years between early 2009, when the first project was initiated, and early 2018, when the last two were. From 2009 to 2010, the first five projects were initiated, and from 2017 to 2018

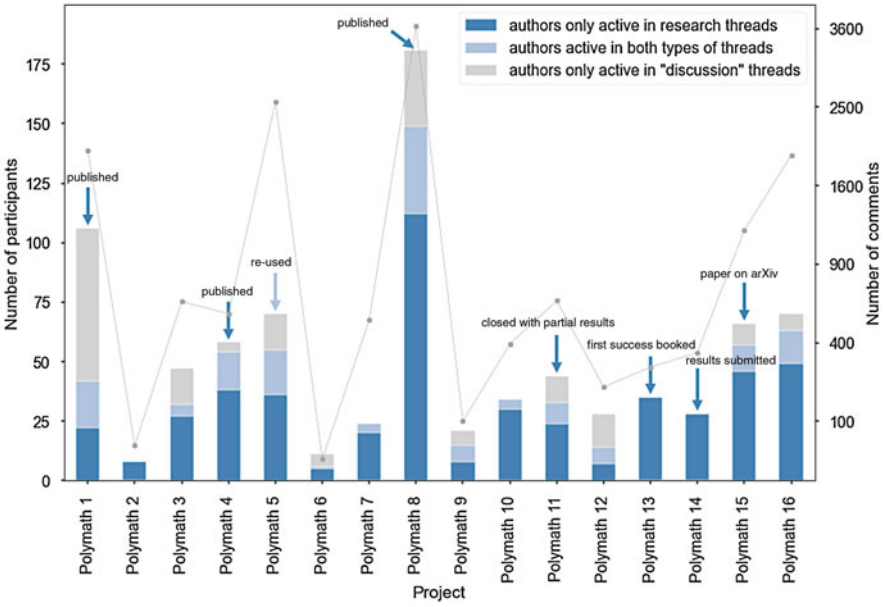


Fig. 2 Number of participants/comments for all projects

the five most recent projects were. The intermediate 6-year period led to fewer projects but was dominated by the largest Polymath project (to date), Polymath 8. For this chapter, we processed all commenting activity up to May 2020. Note, again, that we here do not consider the "mini" Polymath projects, dedicated to Maths Olympiad-final problems.

Polymath projects have had between 8 (Polymath 2) and 181 participants (Polymath 8). In terms of the number of interventions, the smallest project (Polymath 6) only had 27 comments, whereas the largest project (Polymath 8) had 3633. Across all projects, 622 distinct participants were identified (623 if "Anonymous" is included), and more than 14,000 comments were posted.

The distribution of the number of comments per participant appears to be extremely unequal. Across all projects, the average number of comments per participants is just below 24, with only 29 participants (less than 5%) commenting more than 100 times. Most notably, over 40% of all participants commented only once, and the two top contributors (Timothy Gowers and Terence Tao) wrote more than 25% of all comments. Data from the latest projects suggest, however, that the relative influence of Gowers and Tao has decreased in time.

To compare the inequality with results about Zooniverse from Franzoni and Sauermann (2014), we have plotted Lorenz-curves (not depicted to save space) and computed Gini-indices to quantify how unequal the Polymath projects are.

$$\text{Gini} = \frac{2\sum_i^n i \times y_i}{n\sum_i^n y_i} - \frac{n+1}{n}$$

with $[y_1, \ldots, y_n]$ a sorted vector of y_i 's. The range of this index is [0, 1], where the minimum value 0 corresponds to absolute distribution, while the maximum value 1 means absolute centralization. Hence, a high value of this index, approaching 1, means that a small proportion of participants is responsible for a high proportion of contributions.

The Gini-index for all comments across all projects is 0.88. This result is comparable to what was found for seven Zooniverse-projects (see Table S2 of the supplementary material from Franzoni and Sauermann (2014) at http://www.pnas.org/content/pnas/suppl/2015/01/02/1408907112.DCSupplemental/pnas.201408907SI.pdf), where a Gini-index of 0.85 is given as an average across all projects. In that sense, the value we find here is very high, but not exceptional. When we separate all projects, we find more variation. In general, larger projects tend to be more unequal in the sense that attracting more participants primarily means attracting more less active participants. This is to be expected since as projects go on, the relative impact of the most prominent participants increases.

In the specific context of Polymath, such a level of inequality may be more problematic than it is for other forms of crowd science. First, it may signal an extreme imbalance in the distribution of the cognitive labor that is at odds with the goal to engage in genuine collaborative problem-solving. Second, it may suggest that Polymath turns out to be less open than it has aspired to be. When we only count comments, we cannot settle this issue. We should therefore, minimally, try to gain a

better understanding of the nature of the inequality, by considering how it emerges, and what it may tell us about the role of prominent participants. We believe that the adoption of a network-perspective, as currently developed, is a fruitful approach to that end.

The distribution of comments can be described by a "heavy-tailed" distribution (Fig. 3). We have used the library described in Alstott et al. (2014) to fit the distribution and plot the results. The best fit is a Truncated Power Law (red line), but we cannot conclusively exclude alternative heavy-tailed distributions such as a Stretched Exponential or LogNormal distribution.

This distribution should not be seen as evidence for a *rich gets richer* dynamics. High values are the result of time and energy invested. This is obvious in this context, where we only look at the number of comments, but it is crucial to also keep this in mind when we study relational data below. Being affiliated to many projects, threads and episodes require substantial efforts; consequently, joint affiliations with many participants do also. Moreover, when we simultaneously take into account the numbers of comments and the numbers of affiliations (not depicted to save space), we can understand how this distribution emerges. The key insight is that participants who contribute just a few comments almost exclusively do so in only one project, while participants who contribute more comments are also more likely to be returning participants in other projects. As a consequence, the growth of the Polymath community has been such that the number of comments by the most active participants and the number of participants with very few comments have increased hand in hand.

In addition to the already highlighted fact that a very small fraction of the participants is responsible for the majority of the contributions, we also note that

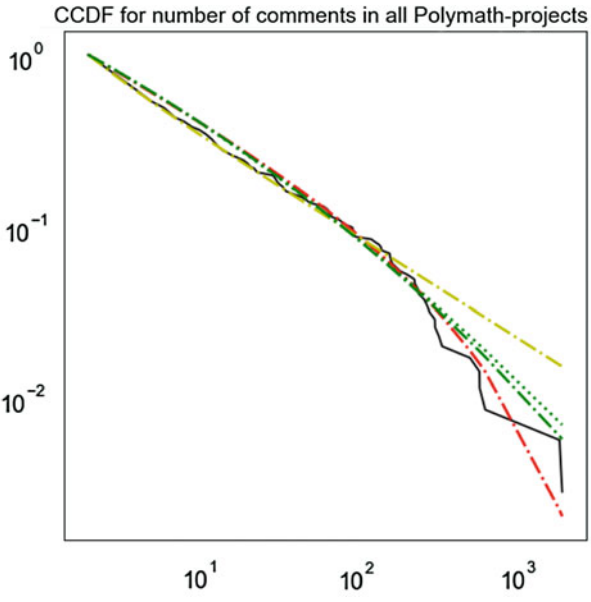


Fig. 3 Cumulative distribution of number of comments per participant

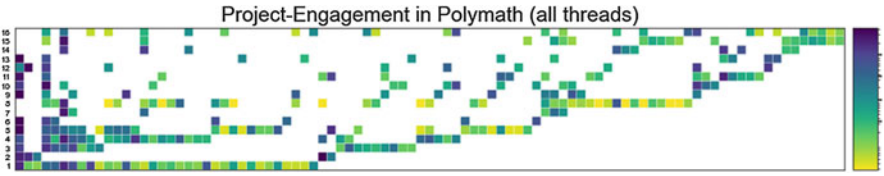


Fig. 4 Project affiliations of core participants based on the relative contribution to each project. This plot uses a logarithmic scale. The darkest blue depicts a relative contribution of over 40%; yellow and green depict relative contributions below 1%

only few participants have become long-term contributors: 529 of the 622 participants contributed to only one project, and as few as 11 contributed to at least five. We will henceforth refer to the 529 single-project participants as *peripheral* participants, and to the remaining 93 as *core* participants (Fig. 4). This terminology slightly diverges from the standard one which uses "core actors" (resp. "peripheral actors") to refer to actors affiliated to "core events" (resp. "peripheral events"), which would result in a larger core and smaller periphery (Borgatti et al. 2018: §13.6). A similar pattern, with the exception of Polymath 8 (marked by the presence of highly active single-project contributors), is found at the level of thread-affiliations: 90% of all participants participated in at most seven threads, and contributed at most six comments to each thread they participated in. Further analysis of the affiliation data moreover reveals that the weighted contribution to previous projects is a good (partial) predictor of future participation. This increases the plausibility of the hypothesised mechanism that produces the extreme inequality of the total sizes of contributions.

7 Affiliations and Joint Affiliations

We consider the affiliations of participants to projects and to threads by constructing bipartite networks with connections between participant-nodes and project- or thread-nodes. From these networks, we deduce joint-affiliation networks. Recall that the vast majority of Polymath participants we refer to as *peripheral* participants are only affiliated to a single project, and therefore only have joint affiliations that are based on a single project. As a consequence, joint-affiliation networks are characterized by the overwhelming presence of clusters of maximally connected participant-nodes that are all affiliated to the same project (Fig. 5 (left)). The connections between core participants, with affiliations to several projects, represent only a small fragment of all connections. Networks based on joint affiliations to threads (Fig. 5 (right)) share most of these features but can in addition be used to study more fine-grained patterns within projects.

With these caveats in mind, we can deduce a number of readily discernible features of the relevant networks. If we restrict our attention to the core participants, we find that a large proportion of them, about a quarter, were "early joiners," in that they joined already during the first project, and almost all had participated in more than one project by Polymath 5. Additionally, since the union of the affiliations of the two most active

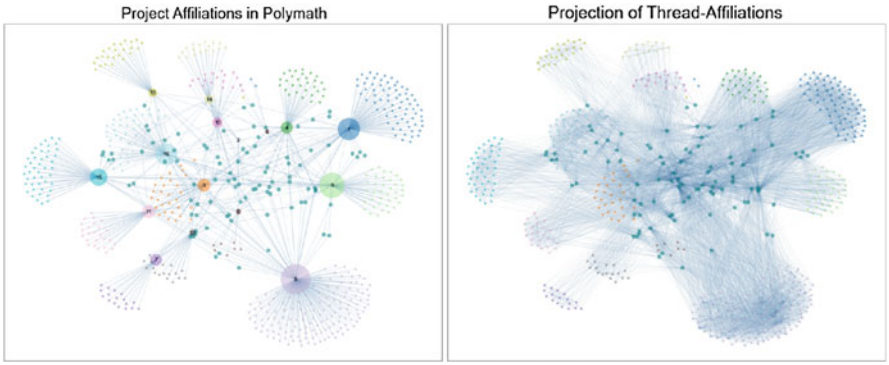


Fig. 5 Project affiliations (left) and joint affiliations to threads (right)

participants (Gowers and Tao) covers all projects *and* since their affiliations also overlap (with five common projects), we can conclude that the distance between the participants in the network based on joint project-affiliations is at most 3.

To further develop this analysis, we were in need of relying on a more systematic approach. We first give a precise description of the method that has been used to construct the joint-affiliation network and to assign *weights* to their connections between participants. The most common approach taken for generating coaffiliation networks is indeed to generate a weighted network from an unweighted bipartite network. One consideration is that joint affiliations to large entities (such as events that many people have attended) create weaker ties than joint affiliations to smaller entities. Another consideration is that having a joint affiliation to someone with many affiliations creates a weaker tie than having a joint affiliation with someone with fewer affiliations. Several approaches based on these types of considerations are reviewed in Borgatti and Halgin (2011). So if we want to take the strength of affiliations into account as well, we need a more elaborate approach, one that can be based on the following principles (explained relative to project-affiliations, but readily generalizable to thread-affiliations).

1. For every pair of participants l and k, let the vectors of weights $w_l = \left[w_l^1, \ldots, w_l^{16}\right]$, $w_k = \left[w_k^1, \ldots, w_k^{16}\right]$ correspond to the number of comments contributed to each project. This corresponds to a list of the weights between l (resp. k) and the neighbors of l (k) in the project-affiliation network; with zeros filled in to have fixed-length vectors and ensuring that the n th element of w_l is the number of comments of l contributed to the n th Polymath project. The sum of a vector w_l corresponds to the weighted degree of l in the project-affiliation network.
2. Let the vector $w = [\widehat{w}_1, \ldots, \widehat{w}_{16}]$ be the average number of comments contributed to each project.

We define the strength or weight of the tie beteween l and k as $\sum_{i=1}^{16} \frac{w_l^i \times w_k^i}{\widehat{w}^i}$

By multiplying weights, we ensure that the tie between two participants with few contributions is weak, and becomes stronger as either or both have more contributions. The normalizing factor ensures that the weights remain within the range of the original weights (based on absolute counts of comments). The results capture the idea that two highly active participants of a given project are more likely to have directly interacted. Moreover, the choice for absolute values (edge-weights as number of comments rather than project-relative proportions of comments) implies that stronger ties can in principle be created during larger projects.

Given what is known about the project- and thread-affiliations, and taking into account how coaffiliation networks are constructed, there are a number of notable properties of the resulting networks that we found unsurprising. They could also be deduced from our earlier considerations about project affiliations, and by observing the structural properties of coaffiliation relations. (Note that joint affiliations are symmetric or reciprocal, and affiliations to a specific project or thread are also transitive.)

1. Both project-level and thread-level joint affiliation networks are connected: There is a path between any two participants.
2. At the level of joint project-affiliations, path-lengths are very short: The distance between any two participants is at most 3.
3. The average clustering and transitivity of the network is very high. For joint project-affiliations, we found that the average clustering and the transitivity are twice as high as for a randomized network with the same degree-distribution.

   As such, these properties do not require further analysis or explanation. We can complement this with a few more specific observations.
4. When we consider the path-lengths for joint affiliations to projects, we see that 16 participants are within a distance of at most 1 of the majority of all participants. Given the variation in the project-sizes, these results do not entirely coincide with bare numbers in project-participation: Affiliations to a few of the larger projects place a participant close to many others.
5. When we make the same comparison for joint thread-affiliations, we only find two participants that are within a distance of at most 1 of the majority of all participants, namely Terence Tao and Gil Kalai. These are not the two original central figures of Polymath (and the two most active commenters), as it notably does not include Timothy Gowers. (In view of results obtained at earlier stages, where Gowers figured as a third, and indeed more prominent, central figure, it should be pointed out that Gowers' absence in Polymaths 14 through 16 tilted the balance.) This reminds us that joint affiliations to projects are rather loose ties. In this case, the central participants are exactly those who (by a substantial margin) participated in most threads.
6. Even when we look at joint affiliations to threads, we see that all but one participant are within a distance of at most 2 of the majority of other participants.

Jointly, these observations indicate that, given the specific position of the top participants, the size of the network is relatively small and that even single-project

participants (those we refer to as peripheral) are not peripheral in the more standard uses of the term. To better understand the role of the core participants, we will now engage in a still more sophisticated analysis.

8 The Relative Weight of Participants

Recall that the *degree* of a node is the number of links it has to other nodes, and that the *weighted degree* of a node is the sum of the weight of its links to other nodes. Note also that in the unweighted case the maximum degree of a node is bounded by the number of nodes in the network. When we consider the distribution of degrees, we compare the number of nodes with small degrees to the number of nodes with high or very high degrees. This type of analysis complements what we said about the distribution of the numbers of comments per participant.

We take the (unweighted) degrees of participants in the network based on joint thread-affiliations as our starting point. This is the most fine-grained and most extensive data set we have: There is a large number of participants and edges (≈9400), but there are also many (180) stages in the growth of this network. The purpose of this analysis is to characterize the emergence of nodes with very high degrees (so-called hubs), and to decide whether we have a regular (small- or ultrasmall-world) network or a network that is in the anomalous regime, where one central hub remains directly connected to all other nodes. We consider three aspects of the degree-distribution: (i) We fit the empirical distribution to a theoretical distribution to find whether it is heavy-tailed, (ii) we consider the evolution of the average shortest-path between nodes relative to the growth of the network, and (iii) we consider the evolution of the maximum degree relative to the growth of the network. The result can be found in Fig. 6.

As before, we can rule out that the degree-distribution is exponential. The best fit is a Log-Normal Distribution (blue line). As before, other heavy-tailed distributions such as the stretched exponential cannot be excluded. The dotted green line is the exponential distribution. (We refer again to Alstott et al. (2014) for a discussion of the fitting-method and the comparison of two competing distributions.) As we prefer to remain cautious, we postpone answering the question whether we have a small- or ultrasmall-world network, since we cannot straightforwardly consider the exponent of a fitted Power Law. (See Barabási (2016: Chap. 4) for an extended discussion.)

The analysis of the evolution turns out to be more informative (Fig. 7). We can see that the average shortest path-length grows linearly in ln ln (N), and that the maximum-degree almost keeps track with the number of nodes. From this, we may conclude that we have an ultrasmall network, which means that the presence of hubs substantially reduces the path-lengths in the network. However, note that with smaller network-sizes like these the difference between small-world and ultrasmall-world networks is not extremely well marked (Barabási 2016: p.132).

It is useful to note that this network is not in the so-called anomalous regime, where the degree of the largest hub is identical with the size of the network. The direct examination of the affiliations already revealed that this was not the case: No

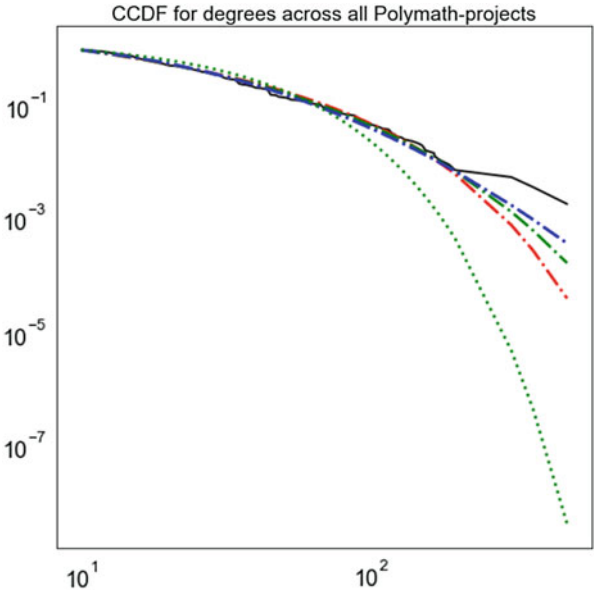


Fig. 6 Cumulative distribution of the degrees

participant is affiliated to all projects (and *a fortiori* to all threads), although the two core participants are sufficient to tie the entire network together.

It is insightful, at this point, to reflect on the mechanism that generates a specific degree-distribution. This can tell us how hubs emerge, and perhaps also shed light on the social role of the central participants. Traditionally, the presence of Power Law (or other heavy-tailed) degree-distributions is seen as the consequence of *preferential attachment*, which means that if new nodes are more likely to link to already well-connected nodes, then even a small preference for better-connected nodes will lead to the emergence of hubs (Barabási 2016: §5.3). Since we study a network based on joint affiliations, we cannot straightforwardly transpose this analysis to the present case. There are however at least two salient differences. First, new as well as already present nodes create links (or reinforce links). And second, participants do not directly link to other participants but link instead to projects or threads. In that sense, the emergence of hubs cannot be the direct result of preferential attachment.

But there is an additional complication, as participants cannot even freely choose their affiliations. At most stages, the number of available choices remains limited. With only one recent exception during Polymath 15 and 16 (which have proceeded simultaneously), there has always been at most one project to affiliate to. Additionally, during most projects (with Polymath 1 as a notable exception), there has only been one active thread to affiliate to. As such, the "choice" faced by a participant is binary: whether or not to affiliate to the current option. Moreover, every project or thread to affiliate to comes with preexisting affiliations that have to be taken as a given. These general features are slightly relaxed when we look at more fine-grained affiliations

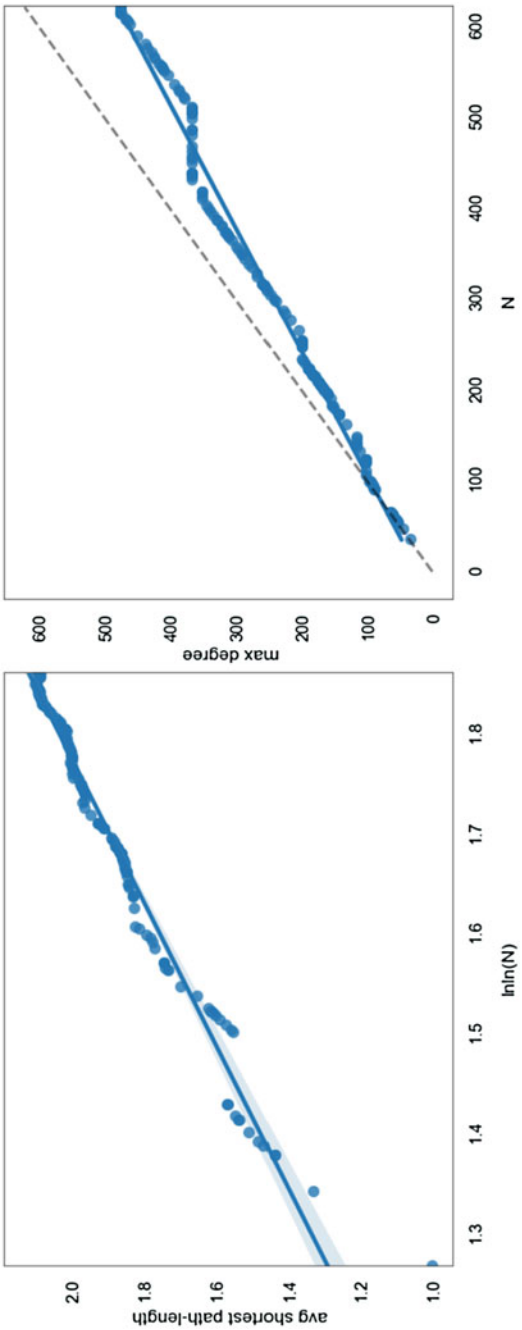


Fig. 7 Evolution of joint affiliations to threads

such as episodes, but they do not disappear. The discussion is always happening in a very small number of places.

This allows us to highlight two features of highly connected nodes:

1. They are omnipresent, in that they are always or almost always affiliated to the current option. This corresponds to a role for highly connected participants as *available* moderators or leaders.
2. They make the events they are affiliated to more attractive and thus increase the chance that other participants will affiliate rather than not affiliate. This corresponds to a role for highly connected participants as *attractive* moderators or leaders.

The first feature appears uncontroversial and squares with an earlier point we made about central participants: There is no lazy road to becoming a highly central participant, but centrality is due to extensive and sustained effort. The second feature is potentially problematic, for it may indicate that not every highly active participant can actually become a central participant, in the sense that only previously attractive participants can create a context in which they can become highly active in the presence of others. Contributing the majority of all comments does not, after all, make one highly connected. This only becomes possible if a sufficiently large pool of other contributors are willing to contribute the rest of the comments.

This worry connects to broader concern regarding the ability of Polymath to create the open environment it has aspired to. We cannot do full justice to this issue here but only reiterate the consequences of our two earlier insights. First, since centrality comes with effort, it is only within reach of those who are able and willing to spare that effort. And second, since the potential to spare the effort alone is presumably not sufficient to reach centrality, it may also not be within reach of those who are not already sufficiently attractive. Based on the self-reporting of the participants to Polymath 8, we should also stress that much effort remains hidden in the present analysis. Even just following without actively contributing requires a substantial investment given the high pace and level of most of the discussions (Polymath 2014).

The present analysis can unfortunately not shed any further light on the practical and institutional obstacles to reaching Polymath-prominence and cannot directly reveal anything about hidden investments to (or, indeed, the hidden returns of) the Polymath Project either. Nevertheless, the displayed method of precisely quantifiying the centrality of the most prominent participants can be of additional value, in that it will point to the relative centrality of less visible participants (something that cannot be inferred by merely reading the threads, but that surfaces by inspecting path-lengths), and can even give indications about the potential of newly emerging hubs (indeed, the largely inconclusive attempts to find subcommunities in the entire Polymath-community revealed that at any point there were apparent "runners-up" to the two major centers of attraction). In sum, by reasoning explicitly about the mechanisms that generate connections, nuance is added to how we evaluate the significance of centrality.

Granted, this type of highly abstracted analysis leaves many questions unanswered. Despite several attempts to decompose the entire Polymath community into

smaller subcommunities, and relate these communities to either the presence and influence of certain participants or to common interests and/or competences in certain types of mathematical problems, we could not draw any stable and convincing conclusions at this level. An examination of the types of mathematical problems that are proposed and actually accepted as Polymath projects did reveal that there are "typical" Polymath-problems, and newly proposed problems are often similar to earlier projects. Important differences between the types of problems that are tackled can be found as well though. One noteworthy instance are those projects that, such as Polymath 8 and more recently Polymath 16, depend on computational methods. Indeed, Polymath 16 quickly revealed close connections with Polymath 8 participants. What remains out of reach for now is the discovery of comparable, but less readily recognizable or predictable connections.

9 Specific Project Communities

In this final section, we take a closer look at the fine structure of the networks associated with specific projects. We shall do this by considering three types of networks for individual projects:

1. Directed (weighted) networks based on direct interactions
2. Undirected (weighted) networks based on joint affiliations to episodes
3. Directed (weighted) networks based on joint affiliations to episodes

Networks of the first type are directly extracted from the shape of the discussion-threads. In a directed weighted network, there is a directed edge with weight n from one participant to another *iff* there are n comments authored by the first participant that are a direct reply to comments of the second participant.

Networks of the second and third type are extracted from an intermediate affiliation-network between participants and episodes of intense interaction. Episodes are ordered collections of comments that, in view of their timestamps, supposedly belong together. As they are intuitively part of the same conversation, it is reasonable to presume that participants to a given episode have read all the comments belonging to it. We identified the episodes with the help of a clustering algorithm. A word on the latter. After initial tests with the `MeanShift` algorithm (Comaniciu and Meer 2002), we settled on the use of `DBSCAN` (Ester et al. 1996). Both are unsupervised algorithms that have the benefit of not requiring one to set a number of clusters in advance, but the latter algorithm also allows one to exclude single isolated data-points (or small numbers of points) from the clustering. This possibility is essential if we want to avoid the unpredictable and uninformative attribution of isolated comments to the closest cluster.

Undirected weighted networks based on the joint affiliation to episodes are defined as follows: There is an undirected edge with weight n between two participants *iff* the sum of the number of comments of the episodes both are jointly affiliated to is equal to n. Here, the individual contributions of the respective

participants do not play a role: only the sizes of the common episodes (a rough measure of how much information was exchanged in those episodes) matter.

Directed weighted networks based on the joint affiliation to episodes, in their turn, are defined as follows: There is a directed edge with weight n between one participant and another *iff* the sum of the number of comments contributed by the first participant in the common episodes is equal to n. Here, it is the size of the contribution in the common episodes of the source node that determines the weight; we look at how much information flowed from the first participant to the second.

Note that these approaches are distinct from those used at the level of project- and thread-affiliations in Sect. 6. There, we devised a method where only pairs of highly active participants could create strong ties, since the joint affiliation of two barely active participants to a large collection of comments could only indicate a weak connection. In the case of episodes, effective interaction is the norm, and it is the strength of the interaction (the amount of information exchanged or the number of actual interactions) that has to be quantified.

When weights are ignored, the second type of network is just an undirected variant of the third type. As such, it does not encode additional information and can therefore safely be left out of our analysis. Additionally, even if the third type of network is directed, all its relations are reciprocal (though potentially of a different weight).

In the remainder, we focus our attention on the third type, and provide a brief comparison of what can be learned from the first and third type to support this choice. Here are some general considerations in favor of the third type:

- The presence or absence of direct reply-to relations between comments is in general unreliable. Such relations were absent during the first few threads of Polymath 1 and exhibit substantially different patterns throughout later projects.
  - In Polymath 1, for instance, a separate ad hoc numbering was used for some time to explicitly and manually flag relations between comments. These relations were used by Cranshaw and Kittur (2011) to construct a comment reference-graph. The absence of such references during the subsequent projects rules out the extrapolation of that approach.
  - A comparison of the shape of the discussion-threads in Polymath 4 and 5 reveals a glaring discrepancy. Polymath 4 has comments of many levels (as shown in Fig. 1), but Polymath 5 has only level-1 and level-2 comments.
- As a means to track the flow of information between participants, a network extracted from explicit reply-to relations is exceedingly conservative. Taking into account the authors of "nearby comments" turns out to be a reasonable middle-ground between this conservative option and the liberal one of assuming that information is automatically shared with all participants. The latter option can readily be dismissed in view of the (reported) effort it takes to keep track of the discussion.
- Modeling information-flow relative to joint affiliations allows one to model comments as one-to-many communications. Correspondingly, we have limited our analysis to small, temporally dense episodes of interaction. There are good

reasons to believe that the discussion actually proceeds this way. There is not only the reported fact that the speed of the discussion is often high, but we also found indications of a diurnal pattern in the timestamps of the comments.
- More generally, the network obtained on the basis of affiliations to episodes appears to be a reliable guide to the structure of the community of a given project. We provide some evidence in favor of this below.

We now review a number of characteristic features of a selection of projects to compare them to the overall community. (Note that we had to exclude Polymath 1 from this analysis since the parallel development on two blogs turned out to be an obstacle to this type of direct comparison.) First a general remark: A comparison of the networks of type 1 and 3 for the Polymath 8 project reveals that when we only take into consideration the direct and explicit interactions between participants, the largest weakly connected component of this network only includes 93 from the total of 180 participants to the entire project, whereas when we look at joint affiliations to episodes, the largest weakly connected component includes 167 of these participants. As a consequence, the analysis of the interaction-graph will not convey any information about almost half of the participants. This seems unwarranted given that 30 of these isolated participants contributed more than 1 comment, and 6 of them contributed more than 5. It also turns out that some of these isolated participants are core participants of the global Polymath community. These observations suggests that, first, an exclusive focus on the interactions leaves out some potentially valuable information about the network, and, second, that it may also misrepresent how information is shared between participants.

A closer analysis of Polymath 4, 5, 8, and the first 6 threads of Polymath 16 (see Figs. 8 and 9 for illustration, others not depicted to save space) has revealed that:

- Distances in larger projects can be bigger, such as in Polymath 8, where the longest path within the main weakly connected component is 4, a very small number (6) of participants are within a distance of 1 of most other participants, but

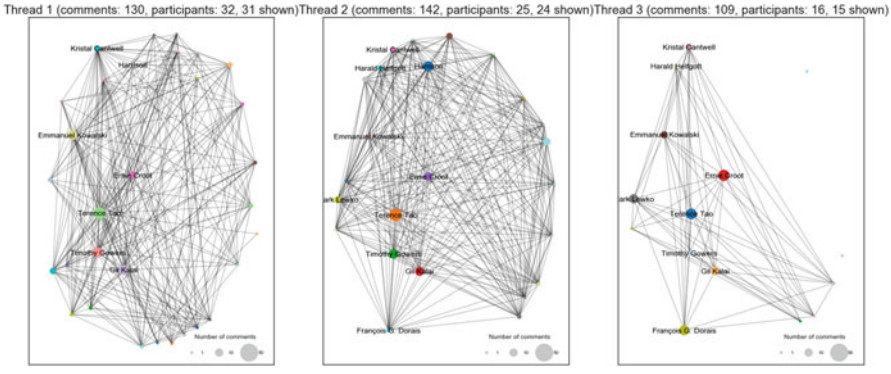


Fig. 8 Active network in Polymath 4 discussion-threads

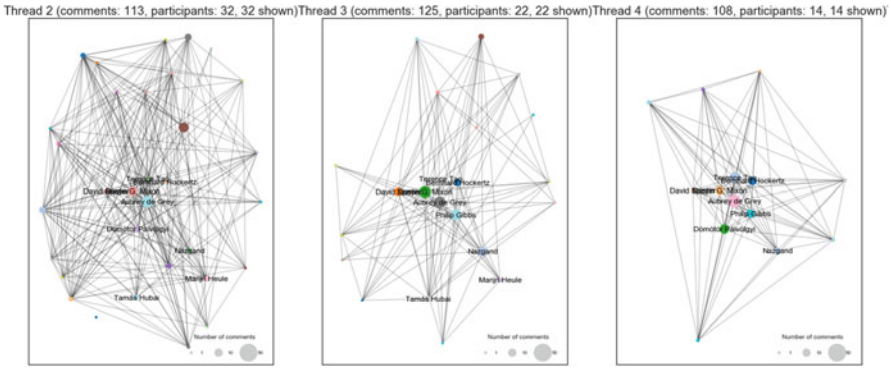


Fig. 9 Active network in Polymath 16 discussion-threads

almost all participants are within a distance of at most 2 of most other participants (We also found that the degree of the most connected participant sticks close to the number of nodes in the network.).

- A smaller project such as Polymath 4 is also based on a smaller network (longest path of 2), with a larger proportion of participants (13 out of 48) within a distance of 1 of most other participants. What is more remarkable is that the Polymath 4 community (presumably due to its small size, with only four research threads) evolved totally differently, and in fact did not visibly grow.
- Projects such as Polymath 5 and Polymath 16 that have a comparable number of participants, but more threads and comments, have again different features. The second has a larger proportion of small distances between participants, whereas the distances between participants are generally larger in the first one.

To understand how differences in the structure of project-networks emerge, it is instructive to consider the active part of the network during each thread. As it turns out, the evolutions we find are extremely different and highly remarkable. The Polymath 8 network, for instance, is best seen as a superposition of minor variations of the same collaborative network. Polymath 4, by contrast, seems to be the result of narrowing down the community to its barest essentials. In Polymath 5, then, the resulting evolution is perhaps best described with Terence Tao's comment that by 2012 Polymath 5 "petered out" (Klarreich 2015). Finally, the development of Polymath 16, reveals an evolution that (at least initially) resembles the development in Polymath 8.

10 Conclusion

For this first overarching quantitative analysis of the 16 Polymath projects, we chose to focus on the scientific community or communities that emerged during these collaborations. In the first sections, we have situated and motivated this endeavor in

general and our approach in particular. We then provided a detailed description of the methodology we developed and carefully laid out the modeling choices we relied on to distil a social network from the traces of the online interactions. For our actual analysis, we subsequently focused on the overarching community that emerged across all projects and the subcommunities associated with specific projects.

Our main findings about the global community concern the distances between participants, and the role of the most active participants. We could conclude that this community is to be situated in the ultrasmall-world regime, where the influence of the hubs is important. We pointed out that this pattern could not be explained by a mechanism of preferential attachment, but that it rather indicated the omnipresence or continuous involvement and availability of the leading participants.

Our main finding about the smaller project communities concerns the evolution of the patterns of interaction throughout the development of a particular project. We observed that there was a substantial variation between projects, with some projects evolving to a very small-scale collaboration and others continually varying around a more permanent core. Further research will have to reveal more detailed typologies for both the structure of the communities and the specific problems they aim to solve.

11 Cross-References

▶ The Social Epistemology of Mathematical Proof

References

Allo P, Van Bendegem JP, Van Kerkhove B (2013) Mathematical arguments and distributed knowledge. In: Abderdein A, Dove I (eds) The argument of mathematics. Springer Netherlands, Dordrecht, pp 339–360

Alstott J, Bullmore E, Plenz D (2014) Powerlaw: a Python package for analysis of heavy-tailed distributions. PLoS One 9(1):e85777

Aron J (2011) How to build the global math brain. New Scientist 210(2811):10–11

Ball P (2014) Strength in numbers. Nature 506(7489):422–423

Barabási A-L (2016) Network science. Cambridge University Press, Cambridge, MA

Barany MJ (2010) But this is blog maths and we're free to make up conventions as we go along. In: Proceedings of the 6th international symposium on wikis and open collaboration. Association for Computing Machinery, New York

Borgatti SP, Halgin DS (2011) Analyzing affiliation networks. In: The sage handbook of social network analysis, vol 1. Sage, Thousand Oaks, pp 417–433

Borgatti SP, Everett MG, Johnson JC (2018) Analyzing social networks. Sage, Thousand Oaks

Castelvecchi D (2010) Problem solved, LOL. Sci Am 302(4):16–18

Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell 24(5):603–619

Cranshaw J, Kittur A (2011) The polymath project: lessons from a successful online collaboration in mathematics. In: Conf Hum Fact Comput Syst Proc Conference on Human Factors in Computing Systems - Proceedings. Association for Computing Machinery, New York, pp 1865–1874

Dunin-Keplicz B, Verbrugge R (2010) *Teamwork in multi-agent systems. A formal approach*. Wiley series in agent technology. Wiley, Chichester

Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the second international conference on knowledge discovery and data mining. AAAI Press, Palo Alto, CA, pp 226–231

Franzoni C, Sauermann H (2014) Crowd science: the organization of scientific research in open collaborative projects. Res Policy 43(1):1–20

Geist C, Löwe B, Van Kerkhove B (2010) Peer review and knowledge by testimony in mathematics. In: PhiMSAMP: philosophy of mathematics: sociological aspects and mathematical practice. College Publications, London, pp 155–178

Gowers WT (2000) The two cultures of mathematics. In: Arnold VI, Atiyah M, Lax PD, Mazur B (eds) Mathematics: frontiers and perspectives. American Mathematical Society, Providence, pp 65–78

Gowers WT (2010) Polymath and the density hales-Jewett theorem. Bolyai Society Mathematical Studies 21:659–687

Gowers T, Nielsen M (2009) Massively collaborative mathematics. Nature 461(7266):879–881

Grossman JW (2002) Patterns of collaboration in mathematical research. SIAM News 35(9):8–9

Klarreich E (2013) Together and alone, closing the prime gap. Quanta Mag, Nov 19

Klarreich E (2015) A magical answer to an 80-year-old puzzle. Quanta Mag, Oct 1

Kloumann IM, Tan C, Kleinberg J, Lee L (2016) Internet collaboration on extremely difficult problems: research versus olympiad questions on the polymath site. In: *Proceedings of the 25th international conference on world wide web*. WWW '16. International World Wide Web Conferences Steering Committee, Geneva, pp 1283–1292

Lane, Lorenzo, Martin, Ursula, Murray-Rust, Dave, Pease, Alison, & Tanswell, Fenner. 2019. Journeys in mathematical landscapes: genius or craft?*:* Hanna, Gila, Reid, David A., & de Villiers, Michael (eds), Proof technology in mathematics research and teaching. Cham: Springer International Publishing. 197–212

Martin U, Pease A (2013) Mathematical practice, crowdsourcing, and social machines. In: *Proceedings of the 2013 international conference on intelligent computer mathematics*. CICM'13. Springer, Berlin, pp 98–119

Meyer ET, Schroeder R (2015) Knowledge machines: digital transformations of the sciences and humanities. MIT Press, Cambridge, MA

Nathanson MB (2010) One, two, many: individuality and collectivity in mathematics. Math Intell 33(1):5–8

Nielsen MA (2010) Introduction to the polymath project and "density hales-Jewett and Moser numbers". In: Bárány I, Solymosi J (eds) An irregular mind. Springer Berlin Heidelberg, Heidelberg, pp 651–657

Nielsen MA (2012) Reinventing discovery: the new era of networked science. Princeton University Press, Princeton

Parsley J, Rusinko J (2017) CRP: collaborative research project (a mathematical research experience for undergraduates). Primus 27(4–5):442–450

Pease A, Martin U (2012) Seventy four minutes of mathematics: an analysis of the third Mini-Polymath project. In: Proceedings of AISB symposium on mathematical practice and cognition II. Society for the Study of Artificial Intelligence and the Simulation of Behaviour, Bath, pp 19–29

Pease A, Martin U, Tanswell FS, Aberdein A (2020) Using crowdsourced mathematics to understand mathematical practice. ZDM 52(6):1087–1098

Polymath DHJ (2014) The "bounded gaps between primes" Polymath project - a retrospective. Newsletter of the European Mathematical Society 94:13–23

Sarvate D, Wetzel S, Patterson W (2011) Analyzing massively collaborative mathematics projects. Math Intell 33(1):9–18

Sauermann H, Franzoni C (2015) Crowd science user contribution patterns and their implications. Proc Natl Acad Sci 112(3):679–684

Van Bendegem JP (2011) Mathematics and the new technologies, part III: the cloud and the web of proofs. In: Schroeder-Heister P, Heinzmann G, Hodges W, Bour PE (eds) Logic, methodology

and philosophy of science. Proceedings of the 14th international congress (Nancy). Logic and science facing the new technologies. College Publications, London, pp 427–439
Van Bendegem JP, Van Kerkhove B (2009) Mathematical arguments in context. Found Sci 14 (1):45–57
Varshney LR (2012) Toward a comparative cognitive history: Archimedes and D. H. J. Polymath. Collective Intelligence 2012: Proceedings (arXiv:1204.3534). MIT, Cambridge, MA
Wagenknecht S (2014) Opaque and translucent epistemic dependence in collaborative scientific practice. Episteme 11(4):475–492
Watson D, Floridi L (2018) Crowdsourced science: sociotechnical epistemology in the e-research paradigm. Synthese 195(2):741–764
Whitley R (2000) The intellectual and social organization of the sciences. Oxford University Press, Oxford
Zadrozny W, de Paiva V, Moss, LS (2015) Explaining Watson: polymath style. AAAI 2015, Palo Alto, CA, pp 4078–4082

Is Mathematics Connected to Religion?

Stanisław Krajewski

Contents

1 On the Approach to the Problem 2758
1.1 What This Chapter Is Not 2759
2 Mathematicians Using Religious Metaphors 2760
3 Historical Interrelations 2762
3.1 Pythagoreans and Their Legacy 2762
3.2 Medieval Developments 2764
3.3 (Early) Modernity 2766
3.4 Cantor and Infinity 2768
4 Religious Sources of Mathematical Concepts 2770
5 Mathematical Models in Theology 2774
6 Conclusion 2776
References 2777

Abstract

Modern mathematics has been shaped by the process of secularization of science. Yet even some present-day mathematicians use religious terms behind the (mathematical) scenes. How essential this is remains debatable. Before modernity, everything, including mathematics, was perceived from a religious perspective. Deeper connections existed for Pythagoreans, and nowadays there is a revival of Pythagoreanism. Mathematics was used by medieval theologians, and even the founders of modern science – for instance, Newton, Leibniz, Boole, Cantor – referred to religion. From the present perspective, those theological aspects seem mostly irrelevant. The concept of actual infinity had a fundamentally religious meaning before the nineteenth century. Now, only a few see mathematical infinity as a notion common to mathematics and theology.

Religious origins of some mathematical concepts can be detected, for example, Indian origins of zero, possible Christian sources of the initiation of the

S. Krajewski (✉)
Faculty of Philosophy, University of Warsaw, Warsaw, Poland
e-mail: stankrajewski@uw.edu.pl

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_77

algebraic approach, the creative function of naming applied to infinite entities. On the other hand, mathematical illustrations of theological concepts are possible, even though genuine mathematical models do not seem to be essential or testable. Generally, religious needs and motivations played a role in the development of mathematics, and mathematical metaphors helped theologians. To perceive this connection as relevant now requires sensitivity to the realm of the religious.

Keywords

Philosophy of mathematics · Theological mathematics · God's mind · Religious metaphors · Religious inspirations · Mathematical models in theology · Infinite sets · God's point of view · Pythagoreans · Cantor

ἀεὶ ὁ θεὸς γεωμετρεῖ
God always geometrizes. (Attributed to Plato)

Mathematica est inimicissima omnino theologiae.
Mathematics is theology's greatest enemy of all. (Martin Luther)

Das Leben der Götter ist Mathematik. . . . Reine Mathematik ist Religion.
The life of the Gods is mathematics. . . . Pure mathematics is religion. (Novalis)

To consider relations of mathematics and religion we must assume we know what these two entities are. It is not easy to define mathematics but in the present Handbook familiarity with this field is presumed. What about religion, or rather religions, and theology, a faith-based rational study of religious issues? It seems hopeless to define religion. The best way to illustrate this difficulty is to quote Wilfred Cantwell Smith: "the question 'Is Confucianism a religion?' is one that the West has never been able to answer, and China never able to ask" (Smith 1964, p. 66). We have no choice but to assume familiarity with religious functioning and thinking about divine matters. The main reference points are the Abrahamic religions: Judaism, Christianity, and Islam; the religions that developed in India; and the legacy of Greek theological philosophy.

1 On the Approach to the Problem

While it may be true that everything is connected to everything else, mathematics–especially the practice of mathematics as we see it nowadays – seems unrelated, indeed totally opposite, to religion and theology. To modern eyes there exist no other two fields as unrelated as numbers and formal structures on the one hand, and God and divine matters on the other; or as calculating and proving on the one hand and praying and prostrating on the other. Yet to premodern eyes the connection was obviously there. The first task in the attempt to answer the question of whether mathematics is connected to religious matters should be to indicate the old

understanding of the relationship and, as a part of it, to indicate those who were both: mathematicians and theologians or passionately religious men. This task can be fulfilled only partially, by giving a few examples (in Sect. 3), since a full presentation would be close to a comprehensive history of the entirety of human thought.

Modern mathematics is the result of a process of secularization of science in general and mathematics in particular. The process took place mainly in the nineteenth century. It would be a mistake, however, to assume that the establishment of modern branches of mathematics was always done in the secular spirit.

It is of special interest to note that even present-day mathematicians use religious terms. To say that mathematics is for us "a divine language" constitutes a standard, if certainly metaphoric, idiom. The presence of such links is illustrated in Sect. 2. The most important (allegedly) common concept is, of course, infinity. It is treated at the end of Sect. 3, but it is also present at almost every point since some reference to infinity appears in most cases where we can observe connections of the mathematical to the religious.

Among more specific connections that can link mathematics and religion are the allegedly religious origins of some principal mathematical concepts (Sect. 4) and attempts to apply mathematics in theology (Sect. 5).

Comprehensive treatment of the subject seems hardly possible at this stage of research. According to Ivor Grattan-Guinness, the "secularization of mathematics has led historians to ignore religious factors in the development of mathematics, and even more of mathematical elements in religion" (Grattan-Guinness 2009, p. 278). To be sure, there are many publications researching aspects of the problem. Noteworthy recent books include the comprehensive volume edited by Teun Koetsier and Luc Bergmans 2005, the collection on infinity edited by Michał Heller and Hugh Woodin 2011, special issue on mathematics of the journal *Theology and Science* vol. 9, No 1, 2011, edited by James Bradley, and the issue of *Studies in Logic, Grammar and Rhetoric* 44 (57), 2016, subtitled "Theology in Mathematics?", edited by Stanisław Krajewski and Kazimierz Trzęsicki. A more popular collection was edited by Lawrence and McCartney (2015). Book length studies include Albertson 2014 dealing with medieval history, Cohen 2007 treating Victorian Britain, and Graham and Kantor (2009), describing the role of Orthodox Christian heresy "Imiaslavie" (see Sect. 4).

1.1 What This Chapter Is Not

The present chapter attempts to avoid prejudgments. The approach is cautious: some connections of mathematics and religion are indicated but no claim is made that they are very deep or absolutely essential unless this can be demonstrated. It is important to distinguish this relatively inconclusive approach from the essentially faith-based position adopted by some authors, especially in the USA, who advocate the religious impact of mathematics or a religious dimension of it. While publications under the rubric "Christian mathematics" can include valuable works, they are all written with the presumption that mathematics is religiously relevant. And this assumption is a

corollary to a more basic belief: everything is God's creation so we can perceive him everywhere, including in mathematics. This approach easily leads to doubtful claims. One can state that "the truth $2 + 2 = 4$ has the attributes of God, such as omnipresence, eternity, omnipotence, and truthfulness. God spoke the world into existence" (Poythress 2015, p. 42). This characterization is not without interest, but it brings no insight when the conclusion is that "when we analyze what $2 + 2 = 4$ really is, we find that arithmetic constantly confronts us with God himself, the Trinitarian God" (p. 25). In addition, "in thinking about arithmetic, we are thinking God's thoughts after him" (p. 25). The image behind all those statements is a theological assumption that used to be standard and normative two or more centuries ago but is totally unconvincing to most scientists today.

To mention a much more serious achievement, Bradley and Howell 2011 present a competent popularization of several mathematical topics. They argue that mathematics is fascinating and beautiful, the points which will be approved by every mathematician, and end each chapter with a statement like "Isaiah 1:18 asks us to reason with God, and the logical structure of mathematics gives us a way to reason effectively" (Bradley and Howell 2011, p. 138). The authors offer a general conclusion saying that "the Christian students would realize that the philosophy of mathematics and Christian faith share a number of interesting touch points. Indeed all things exist and have their being through Christ, including mathematical objects" (p. 217). This does not seem illuminating, for it is just presumed. That is why such books, even when they contain good presentations of mathematics, do not answer the question about the relation of mathematics to religion asked by unprejudiced inquirers.

Rather than trying "to redeem" mathematics, the approach adopted in this chapter is critical, assuming no religious starting point, but at the same time, it is open to the realm of the religious. It is neither assumed that everything must be according to religious convictions nor that nothing religious can be essentially tied to mathematical practice.

2 Mathematicians Using Religious Metaphors

In 1883, Georg Cantor expressed a strikingly modern opinion: the essence of pure mathematics is its freedom (Cantor 1932, p. 182). He wanted to defend the meaningfulness and significance of his pioneering work; he felt that his theory of actually infinite sets was consistent and compatible with the existing mathematics. This opinion is today absolutely dominant among mathematicians. Cantor and David Hilbert were among its first spokesmen. It was Hilbert, who in 1925, made the famous remark on behalf of mathematicians: "No one shall drive us out of the paradise which Cantor has created for us" (Hilbert 1926; Benacerraf and Putnam 1983, p. 191). Is the religious terminology he used only a convenient and universally understandable rhetorical device or does it express something substantial? Even if he just wanted to use a nice metaphor one can still ask why religious metaphors are particularly satisfactory when we talk about mathematics (Krajewski 2016). This

question is naïve, perhaps too naïve, but the ease with which mathematicians utilize theological language must not be ignored. To be sure, this language does not appear inside proofs or definitions, but rather on the side, in informal speech when one wants to say why some topic is important or interesting, and what were one's motivations or associations. On such occasions it can be perfectly natural to talk about God, "God's standpoint," "God's mind." Occasionally traces of this way of talking can be retained in the "official" text. Thus, we can talk about performing infinitely many acts (for instance, acts of choice of an element from a set) as if we had an infinite, "divine" mind; we can refer to a complete knowledge (for instance, taking the set of all sentences true in a given interpretation) as if we were actually omniscient. We can also refer to paradise in Hilbert's sense.

One might say that all such figurative utterances using, directly or indirectly, theological terms are irrelevant and should be ignored in reflection about the nature of mathematics; they are mere chatting, present around mathematics, but not part of it. This loose conversation is, however, part of mathematics, says Reuben Hersh, who asks us to consider seriously the fact that mathematics, like any other area of human activity (as analyzed by Erving Goffman), has a front and a back, the chamber, and the kitchen. The back is of no small importance since the product is made there. The guests or customers enter the front door but the professionals use the back door. Cooks do not show the patrons of their restaurant how the meals are prepared. The same can be said about mathematicians, and for this reason, its mythology reigns supreme. It includes, says Hersh, such "myths" as the unity of mathematics, its objectivity, universality, and certainty (due to mathematical proofs). Hersh is not claiming that those features are false. He reminds us, however, that each has been questioned by someone who knows mathematics from the perspective of its kitchen. Real mathematics is fragmented, it relies on esthetic criteria, which are subjective, proofs can be highly incomplete, and some of them have been understood in their entirety by nobody. Seen from the back mathematics is "fragmentary, informal, intuitive, tentative. We try this or that, we say 'maybe' or 'it looks like'" (Hersh 1991, p. 128; also in 2006). And we do borrow from the religious language. Perhaps the most telling example is the joking contention of Paul Erdös, the famous author of some 1500 mathematical papers (more than anyone else), according to which there exists "The Book" in which God (who else?) has gathered the most elegant proofs of mathematical theorems. Erdös was far from standard religiosity, but he reportedly said, "You don't have to believe in God, but you should believe in The Book" (Aigner and Ziegler 2001, Preface). If in the kitchen of mathematics people use religious language, does it mean the two fields are related?

Probably the most famous example of direct use of the term "theology" in relation to mathematics can be found in the reaction of Paul Gordan, in 1888, to Hilbert's non-constructive proof of the theorem on the existence of finite bases in some spaces. Gordan said, „*Das ist nicht Mathematik. Das ist Theologie.*" Later, having witnessed further accomplishments of Hilbert, he would admit that even "theology" can be useful (Reid 1996, p. 34, 37).

One should not miss the expressions that seem to point to an overlap of the contents of mathematics and theology or at least of the presence of common features.

This possibility is mentioned in some official statements, made for the "front" public. When Hilbert mentioned "the paradise" he could have meant only a rhetorical effect, but in another classic exposition of "Hilbert's Program," in 1930, Rudolf Carnap, while talking about logicism, used the phrase "theological mathematics." According to him, Ramsey's assumption of the existence of the totality of all properties should be called theological mathematics in contradistinction to "anthropological mathematics" of intuitionists; in the latter, all operations, definitions, and demonstrations must be finite. Ramsey "elevates himself above the actually knowable and definable and in certain respects reasons from the standpoint of an infinite mind which is not bound by the wretched necessity of building every structure step by step" (Benacerraf and Putnam 1983, p. 50). Clearly, when infinity is considered, theology is most easily invoked. Or, at any rate, this was happening in the late nineteenth and the early twentieth centuries. Back then it was still far from obvious that we could act as if we could fully control infinite structures, perform infinite sequences of operations, and consider as given the results of an infinite number of acts. By now, such capacities are commonly seen as unproblematic. On a mathematical level, the opposition to infinitistic mathematics seems virtually non-understandable for a contemporary university student. For example, the rejection of the Axiom of Choice by French "semi-intuitionists" was explained, in 1908, by Emile Borel: "the legitimacy of a non-denumerable infinity of choices (successive or simultaneous) . . . appears to me . . . entirely meaningless" (after Moore 1982, p. 102). Nowadays, only the rare species of committed constructivists can express such objections.

The present situation resulted from an exceptional, total victory of Cantor's ideas. Infinitistic mathematics is now hardly seen as connected to theology. Almost always, the theological connection is completely ignored. It is only deep at "the back" that we could say that only God knows the entire decimal representation of the number π. If we were to say that "at the front," we would stress it is just a joke, as did Erdös.

3 Historical Interrelations

Through the eighteenth century nearly everyone was religious, so no wonder mathematicians also were; and later there existed professional mathematicians who were deeply religious; this is still the case. It seems, however, that for many of them the two dimensions of their lives are not genuinely connected but rather just present side by side. The point is whether for some of them the two components could be in harmony. What this can possibly mean in general is hard to say. It is possible, however, to present telling examples.

3.1 Pythagoreans and Their Legacy

The best known and most highly consequential ancient case is that of the Pythagoreans whose founder was both a mathematician and a religious leader. Around 530 B.C.E., Pythagoras began to attract disciples in southern Italy. The Pythagoreans

saw (natural and rational) numbers as the essence of everything – things, relations, musical harmonies, and the celestial harmony, produced by the spheres containing heavenly bodies; at the same time they discovered – to their alarm – incommensurables, or realized that there existed irrational numbers. (For details see Fritz 1945) The most directly religious aspect of their views can be seen in their veneration of the tetractys, or the number 10 represented by the points arranged in an equilateral triangle formed by rows of 1, 2, 3, and 4 points. The upper point represented the essence of all numbers, the two below symbolized the line, the next three symbolized the two-dimensional triangle, and the last four stood for a pyramid. The tetractys was seen as revealing the structure of reality. It was considered holy, and the Pythagoreans swore by it. (Fritz 1981, p. 221.) One of the prayers was, "Bless us, divine number, thou who generates gods and men! O holy, holy tetraktys, thou that containest the root and source of eternally flowing creation!" (Dantzig 1954, p. 42).

The Pythagorean view of the role of numbers sounds far too mystical for present-day scientists and philosophers, yet not only did it remain extremely influential throughout centuries but it has reemerged in a new format in our own time. A century ago, Bertrand Russell wrote that "perhaps the oddest thing about modern science is its return to Pythagoreanism" (Russell 1924, p. 252). He meant, for example, quantum physics allowing only a finite number of states, which brings back the idea that everything is composed of definite units. This idea is further developed in modern physics. For example, Max Tegmark (2014) claims that reality is a mathematical structure. Russell also referred to modern genetics and to the arithmetization of the calculus, leaving only natural numbers as the basis of the continuous. For us, the arithmetization of syntax and especially the ubiquitous presence within our world of the digital realm, which is fundamentally a configuration of numbers, has opened up a new reality that deserves the name of "modern Neopythagoreanism" (Krajewski 2011). Its climax can be seen in the thesis presented by Stephen Wolfram, the author of the widely used program *Mathematica*. In his 2002 book *A New Kind of Science*, he claims that linear automata can code the most complex calculations including the universal Turing machine. (For a critique, for instance of the issue of the complexity of de-coding, see Baldwin 2004.) This claim makes possible the vision of the universe being a computer or even a computer program. Thus, as had been claimed by Pythagoras, literally everything would consist of numbers. To be sure, most of us see no directly religious aspect in such pancomputationalism. Yet for the ancients the interplay of the religious and the mathematical was obvious.

Plato extended Pythagorean ideas, ascribing to mathematical concepts a special role (see, for example, Baum 1973). Plato's philosophy has been at the center of Western thought and mathematics plays an important role in his thought as an example of certain knowledge and also as an educational tool required in his Academy, established at the beginning of the fifth century B.C.E., the first institution of higher learning. According to (later) tradition, the famous inscription at the entrance to the Academy read: "Let no one ignorant of geometry enter." Mathematical examples are presented in *Meno* to prove both the existence of immaterial Forms, which is a central point in Plato's philosophy, and the preexistence of the soul, a religious claim. In the late dialog *Timaeus* the divine "demiurge" is presented

as using mathematics to form the material world. And it is mathematical entities that are in between the ideal Forms in the "divine" realm and the visible objects in our material world. Mueller says, "Plato did have a comprehensive picture of the cosmos divided into a higher divine, intelligible world and a lower human, sensible world . . . A, and I suspect the, crucial link between these worlds is mathematics" (Mueller 2005, p. 117).

Theological presentation of the Platonic vision became common. In pre-Christian Rome, Seneca wrote that numbers existed in the divine mind: "God has these models of all things within himself, and has embraced the numbers and measure of all things which are to be accomplished in his mind. He is filled with those shapes which Plato calls 'Ideas': immortal, immutable, indefatigable" (after Albertson 2014, p. 56).

Around 100 C.E. Nicomachus of Gerasa wrote *Introduction to Arithmetic*, soon translated into Latin. "Nicomachus's works represents the fruition of Pythagoreanizing Platonism in Greek antiquity and the realization of what we can call a mathematical theology" (Albertson 2014, p. 51). It became highly influential: "the continued use throughout Europe of the arithmetic of Nicomachus, as translated by Boethius, are typical of the period from the tenth to the sixteenth century," state Robbins and Karpinski (Nicomachus 1926, p. 145). According to them, "the philosophical arithmetic of the Greeks, of which the arithmetic of Nicomachus is a specimen" includes number theory as we understand it now and at the same time "we have a mystical development, ascribing even magical powers and life properties to numbers" (Nicomachus 1926, p. 4). Indeed, such numerology was present in many cultures as was "sacred geometry," especially insights used in the design and building of religious constructions. In the ancient world also gematria was used, ascribing numbers to words, which is possible due to the fixed numerical values of letters. It was applied primarily to religious texts and is still mentioned in traditional Judaism (Cf. Grattan-Guinness 2009, Chap. 14).

3.2 Medieval Developments

Another Pythagorean achievement, not appreciated today, reigned supreme for centuries. It is the classification of mathematics into the discrete and the continuous with the former involving arithmetic and music, and the latter geometry and astronomy. This was the basis for the famous quadrivium of knowledge, which structured education, religious and nonreligious, in Europe until the early modern era. Its religious significance is stressed by Albertson: "the quadrivium as originally conceived for Latin Christianity by Boethius named the universal principles of first philosophy, or fundamental theology" (Albertson 2014, p. 10). In the fifth century C. E., in the prologue to his commentary on Euclid's *Elements*, Proclus explained that "mathematicals" are "projections of higher principles within soul, just as soul itself is a projection of the divine Mind." Hence mathematics is secondary to theology, and "the highest bond of the quadrivial arts is the divine Mind itself" (Albertson 2014, p. 66). Proclus also believed that "mathematics is the best conceptual language for

expressing the ineffable" (Albertson 2014, p. 66). The connection of theology to mathematics was very strong.

And it remained so for many centuries, although not for everyone. In the fourteenth century, Nicole Oresme, who was a scientist, philosopher, and theologian, presented the first proof of the divergence of the harmonic series. His mathematical work was separate from theology, unlike the scholarly work of Nicolas of Cusa in the fifteenth century. "Cusanus viewed his theological and mathematical explorations as belonging to the same integrated intellectual enterprise." He was praised in surprisingly strong words: "how is even Aristotle comparable to this Cusanus?" wrote Giordano Bruno, and added, "I would readily acknowledge his mind to be not just the equal of Pythagoras, but far superior" (Albertson 2014, p. 1).

Nicolas continued the earlier idea that the proofs of arithmetic, music, geometry, and astronomy lead to knowledge of the Creator. In particular, "mathematicals provide the surest pathway to contemplating God" (Albertson 2014, p. 171). For example, his book *De docta ignorantia*, or "On Learned Ignorance," of 1440, contains a presentation of geometrical figures as tools helping in contemplation, based on the idea that "mathematical knowledge is a *via negativa* to God" (Albertson 2014, p. 181). Thus, given a sequence of circles tangent to a given line in the same point, when their radii grow indefinitely the limit is a line, the common asymptote. And similarly with other figures, for instance indefinitely growing triangles. Therefore, "if there were an infinite line, it would be a straight line, a circle, and a sphere" (Nicholas of Cusa 1985, p. 20). In another work, he describes the game of a spinning top. "If the top spins with 'infinite velocity,' then any two points on the circumference of the top's surface would be co-present with any given point on the fixed circle" (Albertson 2014, p. 259). All the examples show that infinity is not present in our world, and cannot be reached from the domain of the finite. Even potential infinity is impossible. Yet it can be grasped conceptually. The reason for this is the presence of infinity in God. It involves *coincidentia oppositorum*. And since infinity is never reached via approximations by finite objects, it is infinity that is primary: "the one and only Absolute Eternity is Infinity itself, which is without beginning. Consequently, everything finite is originated from the Infinite Beginning" (Nicholas of Cusa 1996, p. 538). And mathematics can reflect God's mind: "we give our name 'number' to number from the Divine Mind" (Nicholas of Cusa 1996, p. 552).

All such considerations strike us now as rather naïve mathematically. Perhaps the most appealing image, also discussed by Cusanus but present since at least the twelfth century, and perhaps much earlier, is offered by the concept of the infinite sphere, whose center is everywhere and whose circumference nowhere. This Latin statement "*Deus est sphaera infinita cuius centrum est ubique, circumferentia nusquam*" (Liber XXIV philosophorum II 1989) can evoke the concept of the circle at infinity in the projective plane. (See Sect. 5 below for a contemporary insight.) The picture can have far-reaching consequences. According to Counet, "Nicholas's universe, which is actually neither infinite nor finite, whose centre is at once everywhere and nowhere, which is not exactly describable through mathematical notions, although mathematics can give an ever more precise account of it . . . the revolution initiated by Nicholas of Cusa is far more radical than that of Copernicus" (Counet 2005, p. 284–5).

This infinite circle was meant as one of the definitions of God. Cusanus also tried to present various mathematically conceived names of God, for example, "infinite Angle" and "infinite Number." He claimed that "an infinite angle would enfold all opposing angles, maxima and minima" (Albertson 2014, p. 247).

3.3 (Early) Modernity

Galileo's famous dictum that the Book of Nature is "written in the language of mathematics" did show the scientific way independent of religion. It is common to treat Galileo as the initiator of modern secular science who was silenced for his anti-Church position. Yet he remained a believer and did not think his views were heretical. In his letter of 1615, he wrote that "the true meaning of the sacred texts . . . will undoubtedly agree with those physical conclusions of which we are already certain and sure enough through clear observations or necessary demonstrations" (after Remmert 2005, p. 353). What is more, Galileo wanted to convince everyone that the mathematical approach is necessary and, according to Remmert, he referred to "the metaphysical legitimation he constructed for the mathematical sciences by eventually taking recourse to God and the divine" (Remmert 2005, p. 349). Similarly, Johannes Kepler "even considered the mathematical way as a kind of divine service" (Remmert 2005, p. 356). And the celebrated founder of mathematical physics, Isaac Newton, was not only deeply involved in theological research, but he should be seen as someone working not in modern science but still in the field of natural philosophy that "considers God and His creation. This orientation towards God was taken for granted" (Pater 2005, p. 463).

Another giant of that era, Leibniz, seems to indicate a much more specific connection of mathematics and theology. His dictum, "As God calculates and executes thought, the world comes into being" (after Breger 2005, p. 489), presents God as a mathematician. Again, he employs theological language to indicate the role of mathematics in the scientific account of the world. Leibniz believed that the same mathematical approach can be applied in theology. By that, he meant the need to use precise logical arguments rather than mathematics as such. One can still ask whether Leibniz applies mathematics in his theological considerations, that is, if it is possible to connect Leibniz's mathematical achievements, such as the creation of contemporary infinitesimal calculus, with his theodicy. An attempt to establish a link has been made. Namely, the idea of our world being the best of possible worlds is an expression of the belief in the applicability of an optimization principle. According to Emil du Bois-Reymond, Leibniz "conceives God in the creation of the world like a mathematician who is solving a minimum problem, or rather, in our modern phraseology, a problem in the calculus of variations – the question being to determine among an infinite number of possible worlds, that for which the sum of necessary evil is a minimum" (du Bois-Reymond 1870, p. 36). Max Planck liked this interpretation, and actually, this approach can be generalized since it is known that many physical theories can be seen as applications of the principle of the least action. Leibniz can be perceived as a forerunner of this approach not just with regard

to physical theories but also in the realm of theology. Yet, one can ask, how essential is mathematics here? The mathematical terminology does not seem to add anything essential to the general idea that our world is the best of all possible worlds. To express this idea no mathematics is needed. Thus, while it is possible that Leibniz was motivated by ideas similar to his mathematical concepts, the connection appears to be primarily psychological. Possibly important in the context of discovery, the connection plays no role in the context of justification.

Similarly, Blaise Pascal who made great contributions to mathematics devoted a lot of attention to religious matters. He emphasized the opposition and the lack of continuity between these areas: the reasons of the mind are opposed to the reasons of the heart, *esprit de géometrie* is opposed to *esprit de finesse*, and, even more important, he knew since his mystical experience of 1654, that the God of Abraham, Isaac, and Jacob is to be distinguished from the God of philosophers. A mathematical mind and the experience of a man who created the calculus of probabilities can be seen in his famous theological argument, the wager, *le paris de Pascal*. While some mathematical inspiration might have been important for Pascal in his theological considerations, this was a general, nonspecific encouragement. The formulation of it does not require mathematics, even if strict mathematical formulation can be helpful. Actually, "probability theory does not seem to have been invoked to elaborate" the high probability of the existence of a Designer, notes Grattan-Guinness (2009, p. 305).

George Boole was a founder of modern symbolic logic that became the foundation for the epitome of our era, computers and the digital realm. Yet, again, writes Daniel Cohen, "it is wrong to assume that the purpose of nineteenth-century pure mathematics and the symbolic logic that arose out of it was to construct a completely scientific, secular realm of philosophy" (Cohen 2007, p. 11). According to his wife, for Boole mathematics and logic was of secondary interest; primary was "the cause of what he deemed pure religion" (Cohen 2007, p. 77). He "seemed never tired of reading what are called 'religious books.' . . . The pursuit of scientific truth was also a religious act" (Cohen 2007, p. 92–93).

Another key logician of that time Augustus De Morgan shared much of Boole's religious approach, yet in the late nineteenth century he "provided a secular approach to those issues that would become increasingly alluring to the growing ranks of professional mathematicians. In a charged cultural environment, De Morgan realized it would be best for everyone if religion and mathematics did not mix" (Cohen 2007, p. 136). In England, Russell completed the divorce, and in France, it was the influence of the Enlightenment that had a strong anti-religious impact. Yet for the creators of the *Encyclopédie* both theology and mathematics were seen as marginal branches of the tree of knowledge. Perhaps the reason was that "both mathematics and theology moved beyond the reach of experience into realms of the perfect and the infinite" (Richards 1992, p. 54).

Some mathematicians are also theologians. Some theologians use mathematical concepts or expressions with mathematical connotations. How deep is the connection of theological discourse with mathematics? Often it seems superficial and inessential, but sporadically the suspicion can arise that a serious relation of contents or method can be detected.

Pascal, Newton, Leibniz, and Boole were leading mathematicians and scientists and at the same time men of faith who, in addition, were devoting a lot of efforts to theology. They all thought that the search for mathematical truths makes it possible to uncover God's thoughts. However, one can doubt if they revealed a genuinely essential connection between the two fields. Can we find a stronger connection between mathematics and religion? In the next Section some examples of possible religious inspirations are presented, and in Sect. 5 mathematical models in theology. Before this, the most important case of an interplay of religion and modern mathematics is presented, that of Cantor.

3.4 Cantor and Infinity

Infinity, understood as a completed entity, was for a long time seen as representing the nature of God. It was basically a theological concept. Mathematicians used potentially infinite processes: extending lines, adding numbers, dividing geometrical objects. The proposals that actual infinities could be studied by mathematicians were not accepted. In the early nineteenth century Gauss, the prince of mathematicians, rejected infinite collections. And then came Georg Cantor. At first, he was a regular mathematician. Studying the sets of convergence of trigonometric series, he introduced the notion of a derivative of a set – that is, the collection of the limit points of sequences with members belonging to the given set – and he noticed that we can consider the intersection of the infinite sequence of subsequent derivatives. What is more, one can take the derivative of the intersection, then the derivative of the derivative, etc. This led him to the notion of a transfinite sequence and transfinite ordinal number. Soon, a whole theory of infinite sets was developed by him and later by others. For several decades, however, mathematicians were against the new theory. The strongest criticism was voiced by the highly influential Leopold Kronecker who was defending the traditional finitism, that is, the view that only finite objects really exist. Cantor claimed that his infinite numbers are of the same nature as the irrational numbers, since both kinds of numbers are the limits of infinite sequences. And both can be accepted because they satisfy certain conditions, which, according to Joseph Dauben, should be interpreted as consistency (Dauben 1977, p. 88). In addition, Cantor had deeper reasons to believe in the reality of his infinite numbers. Deeper means theological. In an 1895 letter to Hermite, Cantor referred to St. Augustine stating that natural numbers "exist at the highest level of reality as eternal ideas in the Divine Intellect" (Dauben 1977, p. 94; also Dauben 1990). Having adopted this view, it is easy to accept the existence of arbitrary infinite objects; after all, there is no reason to impose any limitations on the "divine intellect."

Cantor faced a wall: his approach was rejected not only by Kronecker or by Charles Dodgson, better known as Lewis Caroll, the author of *Alice in Wonderland*, who derided the "bewildering region of Infinities and Infinitesimals" (Cohen 2007, p. 174); even one of Cantor's allies, who had popularized his mathematical results, Magnus Mittag-Leffler advised him, in 1885, not to publish a paper on order types.

As a result, Cantor was pondering withdrawal, abandonment of mathematics. What he did instead was to start an intensive correspondence with theologians. He wrote to Mittag-Leffler that his ideas about infinity were inspired by God. He also stated that by his work he wanted to serve the Catholic Church. His approach was met with interest by some neo-Thomist theologians. For example, Constantin Gutberlet used Cantor's set to argue that actual infinity exists. In correspondence with him, Cantor argued that his infinite numbers make God greater, not smaller. Cantor was able to gain the acceptance of Cardinal Johann Franzelin, a leading theologian of the First Vatican Council, who was satisfied by the Cantorian distinction between absolute infinity and actual infinity. According to Cantor, some infinite structures are so infinite that they are not sets, that is, they cannot be subjected to set-theoretic operations. Due to this distinction, antinomies could be avoided. For instance, the ordinal number ω and the set of natural numbers are actually infinite, but there is no set of all ordinal numbers, because as a "multiplicity" it is "absolutely" infinite. Franzelin accepted the idea that the Absolute Infinite is reserved for God and the (mere) actual infinity is present in the creation. He wrote to Cantor: "there is no danger to religious truths in your concept of the Transfinitum" (Cantor 1932, p. 386).

Cantor firmly believed in the existence of infinite sets. At the same time, he categorically rejected infinitesimals, the infinitely small quantities. He alleged that "the atoms of our physical world form an actually infinite set with cardinality $\aleph_0$ and he believed that the set of the atoms of the universal ether is essentially bigger and possesses the power $\aleph 1$ of the continuum" (Thiele 2005, p. 535). (He assumed the continuum hypothesis.) He was not only deeply interested in theology – for example, he wrote a pamphlet rejecting the immaculate conception – but also was convinced that he was "merely the tool of a higher power" (Thiele 2005, p. 546).

It is apparent that Cantor not only used theological arguments to support his theory, but he also found consolation among Catholic theologians. Granting this, one can still ask whether theologians and mathematicians speak about the same realities. This is unclear (see Oppy 2011; Tapp 2011). Did Cantor use the theological notions in his mathematical considerations? It is doubtful that his religious considerations were necessary or even essential.

The Cantorian distinction between the "regular" and the "absolute" infinity was mathematically right. It gave rise to the mathematical distinction between sets and proper classes, or collections that may not be elements of other classes. However, Cantor's theological-philosophical justification of the distinction did not stay. For him, any set, finite or infinite, could be treated as unity (*Einheit*), or one completed whole, whereas those absolutely infinite collections could not be treated as completed unities. Yet von Neumann, Bernays, Gödel formulated axiomatic theories of classes in which both sets and proper classes, like the totality of all sets, are treated as objects in the same way; they all can equally be treated as unities.

The story does not end here. One can imagine the Cantorian-style procedure continuing. Christopher Menzel invokes this possibility as a comment upon one proof of the existence of God proposed by Alvin Plantinga 2007. The proof takes for granted a later development of the Cantorian vision, that is, the generally accepted cumulative hierarchy of sets: they are formed iteratively by collecting sets, and

possibly other objects, formed at previous stages; the sequence of stages is transfinite. Incidentally, this hierarchy uncovers the role of theological categories: the void and infinite power. Namely, from the empty set one can create "everything," or rather the universe of "pure" sets sufficient for the foundations of mathematics. Nothing, emptiness, is combined with infinite power and a kind of unrestricted will to continue the construction ad infinitum. Terms like "infinite power," "all-powerful will" are unmistakably theological. If Leibniz had known modern set theory, he would have rejoiced, both as a theologian and as a mathematician. He claimed that "all creatures derive from God and nothing" (Breger 2005, p. 491). When he introduced the binary notation, he gave theological significance to zero and one: "It is true that as the empty voids and the dismal wilderness belong to zero, so the spirit of God and His light belong to the all-powerful One" (after Barrow 2000, p. 33).

Now, if idealized "collecting" is treated as executed by an infinite mind, the process can be extended: all sets form the universal class and beginning with it one could consider classes of classes, and form a superhierarchy, and then to continue it further into the transfinite. As Menzel notes, for the realist who believes in the existence of sets (as most mathematicians seem to do) it is unclear "why the cumulative hierarchy is necessarily only as "high" as it is in fact and could not be extended. Call this the realist's impasse" (Menzel 2018, p. 44). His solution would surprise mathematicians: "a satisfying explanation for the existence of the sets in any given world: they are the ones whose members God in fact chose to collect. . . . given that God knows all of the large cardinal properties – certainly at least those we are in principle capable of formulating – it is reasonable to suppose that God does enough collecting to satisfy all of them" (Menzel 2018, p. 49–50). This discussion shows a lasting religious potential in the Cantorian vision.

4 Religious Sources of Mathematical Concepts

There are examples of mathematical concepts that according to some authors have been developed due to religion. None of these instances gives an undoubted exemplification of fundamentally religious origins. Yet each of the cases mentioned below – Indian origins of mathematics and zero, Christian sources of the initiation of the algebraic approach, Maxwell's stress on relations, Grassman's theological inspirations, the creative function of naming – does provide a vista on something religious.

Basic religious behavior involves ritual. Often such rituals are ancient and their sources are of mythic character. Abraham Seidenberg in (1961, 1962), and later papers (e.g., 1978) presented the theory that religious rituals were the source of mathematics. The need to perform ceremonies, for instance, those presenting myths of the origin, led to both the art of calculation and basic geometric notions. Independently of this idea, Seidenberg advocated the diffusion theory: fundamental mathematical concepts were introduced once and gradually migrated to other places to eventually reach the whole globe. In this way, he opposed those historians who believe that basic notions were discovered independently in several places. The

diffusion claim was approved by Bartel van der Waerden: "if we find quite similar ideas about the ritual importance of geometrical constructions in Greece and India, . . . the conclusion that these religious and mathematical ideas have a common Indo-European origin is highly probable" (Van der Waerden 1983, p. 35).

In addition, Seidenberg claims that it was ancient India that was the hypothetical source of basic arithmetic and geometric concepts. This is claimed with reference to interpretations of some old Vedic writings. To give one illustration, "most mathematical problems considered in the Śulba Sūtras spring from a single ritual requirement; namely that of constructing altars that have different shapes but occupy the same area" (Rogers 2015, p. 272). From India, the mathematical knowledge was brought to Babylonia. Seidenberg, a distinguished algebraist, was a major scholar so his proposals should be treated with attention, even though they met with serious criticism. Yet, more recently, further research of ancient mathematics has shown that "some of Seidenberg's opinions may have been justified," says Leo Rogers (2015, p. 269).

While ritual applications could stimulate research, it is less than clear that they were the source of mathematical ideas. A great example of the religious source of concepts, and an Indian one at that, is provided by the invention of zero. Zero as a full-fledged number – a more substantial idea than just an indication of absence in notation systems – emerged in India, not later than fifth century CE. In 628 Brahmagupta presented the rules of using zero in addition and multiplication. And from there it migrated to the Islamic centers and to Europe. The term "zero" was derived from the Arabic *assifr*, meaning void, the translation of Hindu *sunya* (see Barrow 2000).

It is impossible to overstate the importance of zero. It is necessary for the full treatment of arithmetic, for the mature functioning of the decimal notation of numbers, for the operations like binary arithmetic that made computers conceivable. According to John Barrow, "the Indian system of counting is probably the most successful intellectual innovation ever devised by human beings" (Barrow 2000, p. 69). It is so essential and seems to us so simple. Why did the introduction of the number zero take place in India and the brilliant Greek mathematicians not have the idea? One might say that it just so happened, for no special reason. It is much more illuminating to detect cultural circumstances that enabled the innovation, more specifically, the features of Indian culture that were absent in Greece, and helped develop the notion of zero. The relevant difference can be seen in the treatment of nothingness. Zero symbolizes nothing. For Greek philosophers, nothing can result from nothingness, one cannot reasonably state about non-being anything more than its nonexistence. Against that logic, the Indian "religions accepted the concept of non-being on an equal footing with that of being. Like many other Eastern religions, the Indian culture regarded Nothing as a state from which one might have come and to which one might return – indeed these transitions might occur many times, without beginning and without end" (Barrow 2000, p. 65). To have the concept of zero one needs to treat nothing as something, absence and presence at the same time. The Hindu idea of a "positive nothing" indicates a complex notion involving various phenomena, from an empty container to a mystical void. This concept is expressed by the term *sunya*. It is a consistent idea, and despite its contradictory description as non-being and being at the same time, it can be represented by a symbol, zero.

Barrow's account is echoed by William Byers (2007) who treats it as an instance of a more general rule. He argues that some contradictions play a constructive role in mathematics. In particular, new concepts are introduced to overcome incompatible frames of reference. He states, "our culture has tamed the concept of zero by removing its contradictory aspects and incorporating it into a logically coherent mathematical system." Of course, he does not deny the need to have consistent theories. Yet he adds, "reality in general and mathematics in particular are greater than logical consistency" (Byers 2007, p. 109.)

The above account of the history of zero seems to be the best available. The role of religion, Hinduism, has turned out to be essential and non-trivial. A conscious effort of discovering vital Christian sources was undertaken by Gene Chase. He proposed a few examples "wherein Christian theology potentially provided specific insights and motivations for the development of specific new branches of mathematics" (Chase 1996, p. 193). One of them is Cantor, another Boole, both treated as unproblematic cases. Two other stories are of special interest since specifically Christian motives are invoked.

First, William of Ockham (fourteenth century) "wanted to be faithful to the doctrine of transubstantiation but to be consistent with observation. He argued, therefore, for a philosophy which allows the substance to change from bread to body, while the attributes (accidental properties) do not change" (Chase 1996, p. 198). And "his views on the Eucharist were a theological catalyst for his algebra" (Chase 1996, p. 197). Ockhamm "freed mathematics from being about specific concrete objects, or even about specific objects in Plato's heaven" (Chase 1996, p. 198). This is the initial source, claims Chase, of our contemporary mathematics dealing with abstractions, independent of substance.

The second example is provided by the bold move made by James Clerk Maxwell who introduced the concept of field. In the 1860s, he wrote "that an electromagnetic field acts through all space and all time, not just where matter is present. Lord Kelvin called Maxwell's view 'mysticism,' for in it even time and space were contingent upon God's continual sustaining" (Chase 1996, p. 206). Today we think of physical fields as material. Maxwell's view was understood as being non-materialistic, because it was against the mechanist reading of materialism. It was possible for him to make that move, claims Chase, because "in the Calvinistic tradition, God is the author not only of the material universe but also of the laws governing that universe. . . . [Maxwell's approach] has completely reversed the roles of things and their relations" (Chase 1996, p. 206).

In mid-nineteenth century, another development took place, important for later mathematics and related to theology, as has been recently pointed out (Lewis 2011; Achtner 2016). Hermann Grassman in *Ausdehnungslehre* of 1844 introduced revolutionary concepts that, albeit not immediately, initiated vector and tensor calculus and modern linear algebra. "Ernst Cassirer contended, . . . that it was Grassmann who changed the understanding of modern mathematics as a science of magnitudes to a science of structures, functions or relations" (Achtner 2016, p. 112). The point is that Grassmann was primarily a theologian. What is more, he credited the possibility of introducing his new concepts to his teacher Friedrich Schleiermacher, a leading

German theologian of that time, who taught the right approach, and provided his students with the appropriate method. This sounds like a confirmation of the theological source of multidimensional structures, but unfortunately, Grassmann's specific explanations are very general. According to Achtner's summary, "New knowledge is found by constructive abstraction from diversity to the complexity of unity going through different levels of abstraction leaving behind perturbations at a particular level" (Achtner 2016, p. 117). The theological aspects of such statements are unclear, and a scholarly debate continues with regard to the claim that Schleiermacher had a genuine influence on Grassmann.

Two more examples come from the beginning of the twentieth century. Intuitionistic mathematics, created by Luitzen E. J. Brouwer, denies non-constructive entities and even the logical law of excluded middle. At the origin of this vision, there were mystical considerations, important for Brouwer himself. His doctoral dissertation was purged of those thoughts since his advisor did not count them as mathematics. Almost all mathematicians would agree. Yet, the most authoritative study of Brouwer, van Dalen (1999), is entitled "Mystic, Geometer, and Intuitionist," in this order. The story is even more ambiguous. Brouwer was pressed by his mysticism to abandon mathematics as "evil," but his attraction was too strong. In the words of his biographer, "the genesis of the dissertation is a story of temptation, as experienced by all hermits and saints. No matter how much Brouwer fought the evil influence of the world, the fascination of mathematics proved stronger than his Spartan views" (van Dalen 1999, p. 84). It can be said that his research was conducted despite his religious views rather than because of them. At the same time, his mystical predilections could play a role in devising mathematical constructions: "Brouwer's view of the continuum shows a suggestive similarity to the mystic experience of the initial chaotic state of the individual. In this state, also, there are no sharp bounds, everything is flowing and amorphous" (van Dalen 1999, p. 114).

Sources of inspiration should not be ignored, but if they are religious, it is indeed unclear how to account for them, be it in the case of Grassmann or Brouwer. An especially clear case, although much less familiar, is that of Russian mathematicians who about one hundred years ago were developing the so-called descriptive set theory. Some of the scholars of that group, in particular its founder, Nicolas Luzin, were inspired by a heretical (for the Christian Orthodox Church) theology *Imiaslavie* (name worshipping). The story is now known due to the book Graham and Kantor 2009. The theology of name worshipping was heretical because it involved the readiness to identify God with God's name. Luzin, later the head of the Moscow school of mathematics, was influenced by theologian Pavel Florensky who noticed that making God present by an appropriate uttering of God's name is similar to Cantor's claim that a proper naming of an infinite set is sufficient to ensure the set's existence. At the beginning of the twentieth century, it was highly problematic to claim that the naming of an infinite ordinal number is enough, no other proof of its existence is needed. Now, this claim is widely accepted. One of the ways to its acceptance was Luzin's conviction that university stress on logical and empirical sources of knowledge should be supplemented with "intuitive-mystical understanding" (Graham 2011, p. 161). This religious approach made possible the development

of a major mathematical school. While also in this case one can claim that the same research could have developed without that mystical background, the fact is that a hundred years ago some religious ideas and practices encouraged new mathematical constructions.

A more recent parallel has been proposed: Alexandre Grothendieck, the creator of a massive framework that has deeply influenced modern mathematics, was also using the method of creative naming and, at the same time, spoke of a mystical vision of the mathematical universe. "Grothendieck, like Luzin, placed a heavy emphasis on 'naming,' seeing it as a way to grasp objects even before they have been understood" (Graham and Kantor 2009, 200).

Finally, a famous case must be mentioned, that of Indian genius Srinivasa Ramanujan who, a hundred years ago, produced unbelievably complicated theorems, claiming that they were revealed to him by goddess Namagiri. She "would write the equations on his tongue. Namagiri would bestow mathematical insights in his dreams" (Kanigel 1992, p. 36). Ramanujan's story is so exceptional that it is impossible to generalize it.

5 Mathematical Models in Theology

Mathematical illustrations of theological concepts are possible. Those of Cusanus or Leibniz (see Sect. 3) look rather naïve from the present-day perspective. The same can be said about explaining the Trinity by $\aleph_0 + \aleph_0 + \aleph_0 = \aleph_0$, where the symbol "aleph-zero" stands for the first infinite cardinality. (After all, $\aleph_0 + \aleph_0 + \aleph_0 + \aleph_0 = \aleph_0$ also holds, which means that the concept of a "Quadrity" would be equally well justified.) Moreover, it is doubtful that it makes sense to refer to mathematical addition of numbers, finite or infinite, as corresponding to the "addition" of the persons of the Trinity. Perhaps it would be better to refer to the operation of the Boolean union: in Boolean algebras adding the (Boolean) one to itself results in the same one. In addition, it is the only element (of the algebra) that is greater than each of the remaining ones. It has, therefore, a "divine" quality, and since it is unique, it can be seen as modeling monotheism. Yet, obviously, this model is as shallow as it is arbitrary. It is no better than Cusanus's reverence for "the equation $1 \times 1 = 1$ as a sublime name of God" (Albertson 2014, p. 1). Nowadays, mathematical models are so rare that in the monumental volume Diller and Kasher 2013, presenting dozens of models of divinity, they do not appear at all. Yet, mathematical models are sometimes invoked by modern theologians. A notable example appeared in the Victorian era: *Flatland* (Abbott 1952) is a tale about geometric creatures living in a flat world. A three-dimensional being has divine properties from the flatlanders' perspective: it can perceive the whole flat world and when it interferes with the plane this is perceived by the flat figures as a miraculous happening. In this way, the recently discovered multidimensional geometry is presented as not a threat but rather an aid to religious faith. The words of Boethius (sixth century) become understandable: "God abides for ever in an eternal present, . . . and, embracing the whole infinite sweep of the past and

of the future, contemplates all that falls within its simple cognition as if it were now taking place" (Boethius 1897, Book V, Song VI).

To give a contemporary example, James Miller 1998 presented a model, based on the normal (Gaussian) distribution, in order to explain the supposedly paradoxical belief that submission to God's will provides more freedom than does disobedience. While this is difficult to explain in everyday language, it becomes understandable when the Gaussian curve is considered in the space of all options available at a given moment of one's life. On the y-axis the number of options is indicated and on the x-axis their quality. Miller assumes that the most agreeable to God are the options belonging to the "middle way." The acceptable options are not much different from those in the middle, distant no more than one standard deviation. As a result, we model free will, since there are many possible options to choose from. At the same time, since the columns near the middle position are relatively larger, the way that pleases God offers more opportunities. The submission provides more freedom. Furthermore, sin can be illustrated as a deformation and dislocation of the Gaussian curve: then what the sinner takes to be the middle way is in reality distant from the genuine "divine" middle way.

The value of the model is debatable. It is commendable if it helps someone to grasp some theological ideas.

Another proposal has been made, in a way opposite to the well-known idea that the set of true sentences is part of the "divine" knowledge. Making use of logical and set-theoretical notions one can try to express "divinity" as ... the set of all true sentences. John Post (1974) does it in the spirit of Quine's minimalism in order to express the foundations of theism, avoiding "metaphorical, analogical, symbolic, mythical" (Post 1974, p. 736–7) talk about God. He tries to paraphrase the statements about God in a possibly restricted language so that one can get sentences that are "literally true." Thus "Godhead" is defined as the smallest set G, whose transitive closure (that is, its elements, elements' elements, etc.) includes the set of all true sentences about sets, the set of true sentences about justice, virtues, religious experience (whatever it may mean for us), etc. The set G is mysterious, it "combines transcendence with immanence" (Post 1974, p. 740), it contains complete knowledge, and the "mythic" personal God is its "intended object." Even this short summary suggests that this construction can be easily criticized as contributing nothing essential to theology. The model can be easily undermined by asking, e.g., whether devout people really pray to a set.

Another model is presented in Krajewski (2019) (and mentioned in 2011). It illustrates Martin Buber's vision of God who is revealed in interhuman relations. In 1923, he wrote, "The extended lines of relations meet in the eternal Thou" (Buber 1937, p. 75). To anyone acquainted with projective geometry, this picture would seem familiar. In the system of plane geometry, the direction of each straight line constitutes a point of a new type, and this point at infinity is common to all parallel lines. Thus, every pair of straight lines (in the plane) intersects in exactly one point, with no exceptions. For parallel lines this point of intersection is at infinity, but in projective geometry, these infinitely far located points have the same status as do the usual points of the plane. What is more, the totality of the points at infinity constitutes one "straight line at infinity." This line can be seen to correspond to

God in Buber's vision: the line of the relation between two arbitrary individuals always reaches the same God.

The use of this picture can be extended. The well-known model of the projective plane in the Euclidean (three-dimensional) geometry is given by a hemisphere. If the Northern hemisphere together with the equator is imagined all the points above the equator constitute the proper (finite) points of our model, and the points on the equator constitute the points at infinity. Straight lines are the semicircles with the center coinciding with the center of the ball (the Earth). (They form the shortest way between two points, and are called great circles.) One more move must be made. In order to have only one point corresponding to the direction of a line, the opposite points of the equator must be identified. The resulting objects cannot be easily imagined, but it is also possible to assume that the whole line segment joining two opposite points constitutes, *ex definitione*, one "improper point." The totality of these improper points forms a plane figure, or a disc determined by the equator. In our model, this object is the basis of the hemisphere. Thus, God is here the fundament of our world, unattainable directly, but, being the direction of every line joining two points, it is present in every relation between the beings of this world. In addition, the model can be used as a visualization assisting prayer (Krajewski 2019, p. 1014–1015).

The emerging picture is attractive, but mathematical research on the projective plane includes many more deep theorems than what was used in the model. They seem to have no theological parallels. If the geometrical model is seen as an adequate description of the philosophico-theological idea, philosophy and theology look meager compared to mathematics. If, on the other hand, one believes that it is philosophy and theology that deals with important problems of humankind and the world, then the mathematical model becomes no more than an illustration – a suggestive one, but meager and fundamentally irrelevant.

None of the presented examples of applications of mathematics to theology seems truly fundamental. If they help pondering some theological ideas, they are useful. However, they are only metaphorical and lack the features that make mathematical models so powerful in science, especially in physics, where they are creative; in particular, they suggest new concepts, and have predictive power, that is, enable us to foresee events or experimental results, and thereby check the model's adequacy. In theology, nothing so essential seems to be added by mathematical modeling.

6 Conclusion

Relations between mathematics and religion are seen by some as friendly, by others as antithetical – see the mottos to the present chapter. While modern mathematicians usually disregard the connection, some philosophers try to detect it. This can be done by affirming faith, as does Roy Clouser (1991), who argues that the right Biblical approach takes everything as dependent on God, therefore to attribute absolutely independent existence to numbers constitutes paganism. On the other hand, the criticism of Platonic assumptions can be done from an anti-religious position: "the actual Platonic infinity seems to invoke some version of the Greek–Hebraic divinity," states Brian Rotman (1993, p. 49). Any belief in the existence of mathematical

objects leads to the "unstated theism – implicit and unacknowledged – of twentieth-century mathematical infinitism" (Rotman 1993, p. 157). This, one could add, seems to be a major reason for the attractiveness of mathematics. According to Vladislav Shaposhnikov, at the turn of the twentieth century "some key mathematicians of the period made an attempt (perhaps partially unconscious) to substitute pure mathematics for departed theology" (Shaposhnikov 2016, p. 48).

Up to modern times, the relation of mathematics to religion was close. According to Ivor Grattan-Guinness, "the main contexts lay in numerology, sacred geometry, and mechanics" (Grattan-Guinness 2011, p. 144). They are now of historical interest only. With the advent of modernity, the interplay of mathematics and theology continued. Ultimately, modernizing did mean secularization. Yet the process of getting to the present fundamentally irreligious science, including mathematics, was long and slow.

The examples mentioned in this chapter do not provide an indisputable proof that religion was indispensable for mathematical developments or that mathematical modeling was essential for theology. There is no doubt, however, that religious needs and motivations played a role in the development of mathematics, and mathematical metaphors helped theologians. Perhaps to have genuine mathematical models of divine matters another mathematics, still to be developed, is needed?

Is mathematics still connected to religion(s)? The present-day answer depends on how open is one's attitude to the realm of the religious. For those sensitive enough to religion the connection can be perceived.

In addition to possible mutual influences, there exists another method of relating mathematics to theology. It seems that in both disciplines a central role is played by mystery: the more we know the more we realize the limitations of our knowledge. While such limitations are present in all sciences they are particularly prominent in apophatic theology and in mathematical incompleteness phenomena. Some of the main assumptions must be accepted for no other reason than faith. Again, this may be true about knowledge in every field because we often function within a paradigm, chosen without sufficient reasons – which does not mean irrationally. Yet mathematics and theology seem to be more arbitrary than other fields, which is due to the fact that both are more distant from the empirical world than are the other disciplines. Mathematics and theology are controlled by internal consistency requirements, fruitfulness, esthetic values, intuitive appeal. Even though intuition is shaped by empirical and cultural factors, they do not fully determine mathematical or theological assumptions. Positions in the philosophy of mathematics differ as widely as they do in theology.

Acknowledgment The author is grateful to Andrew Aberdein for several suggestions that have helped correct and improve this chapter.

References

Abbott EA (1952) Flatland, Dover, NY. Original ed. 1884

Achtner W (2016) From religion to dialectics and mathematics: Schleiermacher's philosophical-theological contribution to the development of modern tensor calculus in Graßmann's Ausdehnungslehre. Stud Log Gramm Rhetor 44(57):111–131

Aigner M, Ziegler G (2001) Proofs from THE BOOK, 2nd edn. Springer-Verlag, Berlin

Albertson D (2014) Mathematical theologies: Nicholas of Cusa and the legacy of Thierry of Chartres. Oxford University Press, New York
Baldwin J (2004) Review of Wolfram 2002. Bull Symb Log 10(2004):112–114
Barrow J (2000) The book of nothing: vacuums, voids, and the latest ideas about the origin of the universe. Vintage Books, London
Baum RJ (ed) (1973) Philosophy and mathematics: from Plato to the present. Freeman, Cooper Co., San Francisco, CA
Benacerraf P, Putnam H (eds) (1983) Philosophy of mathematics: selected readings, 2nd edn. Cambridge University Press, Cambridge
Boethius (1897) The consolation of philosophy (Translated into English prose and verse by H.R. James). Elliot Stock, London. www.gutenberg.org/files/14328/14328-h/14328-h.htm
Bradley J (ed) (2011) Theology and mathematics. Special Issue of Theology and Science 9(1):1
Bradley J, Howell R (2011) Mathematics through the eyes of faith. HarperOne, New York
Breger H (2005) God and mathematics in Leibniz's thought. In: Koetsier T, Bergmans L (eds) Mathematics and the divine. Elsevier, Amsterdam, pp 485–498
Buber M (1937) I and Thou (Trans: Ronald Gregor Smith of 1923 *Ich und Du*). Clark, Edinburgh
Byers W (2007) How mathematicians think: using ambiguity, contradiction, and paradox to create mathematics. Princeton University Press, Princeton, NJ
Cantor G (1932) In: Zermelo E (ed) Gesammelte Abhandlugen mathematischen und philosophischen Inhalts. Verlag Julius Springer, Berlin. Reprinted Springer, Berlin 1980
Carnap R (1930) The logicist foundations of mathematics. In: Benacerraf P, Putnam H (eds) Philosophy of mathematics: selected readings, 2nd edn. Cambridge University Press, Cambridge, pp 41–52
Chase GB (1996) How has Christian theology furthered mathematics? (Chapter 8). In: Van der Mee J (ed) The role of beliefs in mathematics and the natural sciences (Facets of faith & science, Vol II). University Press of America, Lanham, MD
Clouser RA (1991) The myth of religious neutrality. University of Notre Dame Press, Notre Dame, IN
Cohen DJ (2007) Equations from god: pure mathematics and Victorian faith. The John Hopkins University Press, Baltimore
Counet J-M (2005) Mathematics and the divine in Nicholas of Cusa. In: Koetsier T, Bergmans L (eds) Mathematics and the divine: a historical study. Elsevier, Amsterdam, pp 273–290
Dantzig T (1954) Number: the language of science. Doubleday-Anchor, Garden City
Dauben J (1977) Georg Cantor and pope Leo XIII: mathematics, theology, and the infinite. J Hist Ideas 38(1):85–108
Dauben JW (1990) Georg Cantor: his mathematics and philosophy of the infinite. Princeton University Press, Princeton, NJ
de Pater C (2005) An ocean of truth. In: Koetsier T, Bergmans L (eds) Mathematics and the divine: a historical study. Elsevier, Amsterdam, pp 459–484
Diller J, Kasher A (eds) (2013) Models of god and alternative ultimate realities. Springer, Dodrecht
du Bois-Reymond E (1870) Leibnizische Gedanken in der neueren Naturwissenschaft. In vol. 1 of du Bois-Reymond's collected works, Google Books. https://www.google.com/books/edition/Reden_von_Emil_Du_Bois_Reymond_v_1/SSkJAAAAIAAJ?hl=en&gbpv=0
Graham L (2011) The power of names. Theol Sci 9(1):157–164
Graham L, Kantor J-M (2009) Naming infinity: a true story of religious mysticism and mathematical creativity. Harvard University Press, London
Grattan-Guinness I (2009) Roots of learning: highways, pathways and byways in the history of mathematics. John Hopkins University Press, Baltimore
Heller M, Woodin H (eds) (2011) Infinity: new research frontiers. Cambridge University Press, Cambridge
Hersh R (1991) Mathematics has a front and a back. Synthese 88:127–133
Hersh R (ed) (2006) 18 unconventional essays on the nature of mathematics. Springer, New York

Hilbert D (1926) Über das Unendliche, *Mathematische Annalen* 95, 161-190 (trans: On the infinite). In: van Heijenoort J (ed) From Frege to Gödel: a source book in mathematical logic, 1879–1931. Harvard University Press, Harvard. 1967; and in Benacerraf and Putnam (1983), 183–201

Kanigel R (1992) The man who knew infinity: a life of the genius Ramanujan. Washington Square Press, Little Brown

Koetsier T, Bergmans L (eds) (2005) Mathematics and the divine: a historical study. Elsevier, Amsterdam

Krajewski S (2011) Czy matematyka jest nauką humanistyczną? [Is Mathematics a Humane Science? In Polish]. Copernicus Center Press, Cracow

Krajewski S (2016) Theological metaphors in mathematics. Stud Log Gramm Rhetor 44(57):13–30

Krajewski S (2019) Mathematical models in theology. A Buber-inspired model of God and its application to "Shema Israel". J Appl Log 6(6):1007–1020

Krajewski S, Trzęsicki K (eds) (2016) Issue Theology in mathematics? Stud Log Gramm Rhetor 44(57)

Lawrence S, McCartney M (eds) (2015) Mathematicians and their gods: interactions between mathematics and religious beliefs. Oxford University Press, New York

Lewis A (2011) The divine truth of mathematics and the origins of linear algebra. Theol Sci 9(1): 109–120

Liber XXIV Philosophorum (1989). www.artandpopularculture.com/Liber_XXIV_philosophorum. French ed. by F. Hudry, Millon, Grenoble

Menzel C (2018) The argument from collections. In: Walls L, Dougherty T (eds) Two dozen (or so) arguments for god: the Plantinga project. Oxford University Press, Oxford, pp 29–58

Miller JB (1998) From the garden to Gauss: mathematics as theological metaphor. AAAS dialogue on science, ethics, and religion. https://web.archive.org/web/19990221051548/; https://www.aaas.org/spp/dspp/dbsr/RESOURCE/miller/GAUSS4.htm

Moore GH (1982) Zermelo's axiom of choice: its origins, development, and influence. Springer-Verlag, New York

Mueller I (2005) Mathematics and the divine in Plato. In: Koetsier T, Bergmans L (eds) Mathematics and the divine. Elsevier, Amsterdam, pp 99–121

Nicholas of Cusa (1985) On learned ignorance (trans: Jasper Hopkins). The Arthur J. Banning Press, Minneapolis, MN

Nicholas of Cusa (1996) Idiota De Mente (the layman on mind) (trans: Jasper Hopkins). The Arthur J. Banning Press, Minneapolis, MN

Nicomachus (1926) Introduction to arithmetic (trans: M. L. Dodge, with studies by Frank Egelston Robbins and Louis Charles Karpinski). The Macmillan Company, New York. https://babel.hathitrust.org/cgi/pt?id=mdp.39015005675411&view=1up&seq=9

Oppy G (2011) God and infinity: directions for future research. In: Heller M, Woodin WH (eds) Infinity*:* new research frontiers. Cambridge University Press, Cambridge, pp 233–254

Plantinga A (2007) Two dozen (or so) theistic arguments. In: Baker D-P (ed) Alvin Plantinga. Cambridge University Press, Cambridge, pp 203–227

Post JF (1974) New foundations for philosophical theology, Quine with God. J Philos 7:736–748

Poythress VS (2015) Redeeming mathematics: a God-centered approach. Crossway, Wheaton, IL

Reid C (1996) Hilbert. Springer Verlag, Berlin

Remmert VR (2005) Galileo, god and mathematics. In: Koetsier T, Bergmans L (eds) Mathematics and the divine: a historical study. Elsevier, New York, pp 347–360

Richards JL (1992) God, truth, and mathematics in nineteenth century England. In: Nye MJ et al (eds) The invention of physical science. Kluwer Academic Publishers, Dordrecht, pp 51–78

Rogers L (2015) Historical epistemology: professional knowledge and proto-mathematics in early Civilisations. In: Barbin E, Jankvist UT, Kjeldsen TH (eds) History and epistemology in mathematics education. Proceedings of the seventh European Summer University ESU 7. Danish School of Education, Aarhus

Rotman B (1993) The ghost in Turing's machine: taking god out of mathematics and putting the body back in. Stanford University Press, Stanford
Russell B (1924) Atoms in modern physics. In: Essays on language mind and matter 1919-26. Allen and Unwin, London. 1988. https://archive.org/stream/in.ernet.dli.2015.273552/2015.273552.Bertrand-Russell_djvu.txt
Seidenberg A (1961) The ritual origin of geometry. Arch Hist Exact Sci 1:488–527
Seidenberg A (1962) The ritual origin of counting. Arch Hist Exact Sci 2:1–40
Seidenberg A (1978) The ritual origin of mathematics. Arch Hist Exact Sci 18:301–342
Shaposhnikov V (2016) Theological underpinnings of the modern philosophy of mathematics. Part I: mathematics absolutized, Part II: the quest for autonomous foundations. Stud Log Gramm Rhetor 44:31–54 and 147–168
Smith WC (1964) The meaning and end of religion. Mentor Books, New York
Tapp C (2011) Infinity in mathematics and theology. Theol Sci 9(1):91–100
Tegmark M (2014) Our mathematical universe: my quest for the ultimate nature of reality. Knopf, New York
Thiele R (2005) Georg Cantor (1845–1918). In: Koetsier T, Bergmans L (eds) Mathematics and the divine: a historical study. Elsevier, New York, pp 523–548
van Dalen D (1999) Mystic, geometer, and intuitionist: the life of L. E. J. Brouwer. Oxford University Press, Oxford, p 1999
van der Waerden BL (1983) Geometry and algebra in ancient civilizations. Springer-Verlag, Berlin
von Fritz K (1945) The discovery of incommensurability by Hippasus of Metapontum. Ann Math 46(2):242–264
von Fritz K (1981) Pythagoras of Samos. In: Pitcairn A, Rusch B (eds) (gen. ed. C. S. Gillispie) Dictionary of scientific biography, vol 11. Ch. Scribner's Sons, New York, pp 219–225
Wolfram S (2002) A new kind of science. Wolfram Media, Champaign, IL

Recommended Readings on Mathematics and Religion/Theology (Selected from the Above-mentioned References)

Abbott EA (1952) Flatland, Dover, NY Original ed. 1884
Albertson D (2014) Mathematical theologies: Nicholas of Cusa and the legacy of Thierry of Chartres. Oxford University Press, New York
Baum RJ (ed) (1973) Philosophy and mathematics: from Plato to the present. Freeman, Cooper Co., San Francisco, CA
Bradley J (ed) (2011) Theology and mathematics. Special Issue of Theology and Science 9(1):1.
Cohen DJ (2007) Equations from God: pure mathematics and Victorian faith. The John Hopkins University Press, Baltimore
Chase GB (1996) How has Christian theology furthered mathematics? (Chapter 8). In: Van der Mee J (ed) The role of beliefs in mathematics and the natural sciences (Facets of faith & science, Vol. II). University Press of America, Lanham, MD
Dauben JW (1990) Georg Cantor: his mathematics and philosophy of the infinite. Princeton University Press, Princeton, NJ
Graham L, Kantor J-M (2009) Naming infinity: a true story of religious mysticism and mathematical creativity. Harvard University Press, London
Grattan-Guinness I (2009) Roots of learning: highways, pathways and byways in the history of mathematics. John Hopkins University Press, Baltimore
Heller M, Woodin H (eds) (2011) Infinity: new research frontiers. Cambridge University Press, Cambridge

Hersh R (ed) (2006) 18 unconventional essays on the nature of mathematics. Springer, New York
Koetsier T, Bergmans L (eds) (2005) Mathematics and the divine: a historical study. Elsevier, Amsterdam
Krajewski S, Silvestre R (eds) (2019) Special issue "Concept of God". J Appl Log 6(6)
Krajewski S, Trzęsicki K (eds) (2016) Issue "Theology in Mathematics?". Stud Log Gramm Rhetor 44(57)
Lawrence S, McCartney M (eds) (2015) Mathematicians and their gods: interactions between mathematics and religious beliefs. Oxford University Press, Oxford
Seidenberg A (1978) The ritual origin of mathematics. Arch Hist Exact Sci 18:301–342

Section XIII

New Perspectives

Valeria Giardino

New Perspectives: An Introduction

Valeria Giardino

Contents

1 New Perspectives: Where Do They Come From? 2786
2 Summary of the Chapters in the Section 2787
3 Cross-References 2792
References 2792

Abstract

In this short Introduction to the present section, I will first briefly point out the reasons why the chapters collected here present original research in the context of the philosophy of the practice of mathematics, and open even newer perspectives. It is important to note that one crucial issue for future research will be to explore the connections within these chapters and with other chapters included in other sections – in particular, but not exclusively, the sections on Proof, on "Experimental" mathematics, and on the Semiology of Mathematical Practice. Second, I will present a short – and not exhaustive – summary of the themes that are tackled in the chapters. There will be overlaps and common threads: These chapters are indeed the final product of several discussions and exchanges within the community of the philosophers of the mathematical practice, which is extending toward other regions of philosophy thus blurring standard boundaries.

Keywords

New perspectives · Interplay of practice · Experiments in mathematics · Social ontology · (Social) Epistemology · Philosophy of action · Mathematics education · Corpus linguistics

V. Giardino (✉)
Chargée de Recherche/Research Fellow (CRCN) Centre National de la Recherche Scientifique, Institut Jean Nicod, Paris, France
e-mail: valeria.giardino@ens.psl.eu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_126

1 New Perspectives: Where Do They Come From?

As it is well known – and the number of pages of this Handbook proves it – the attention to the practice of mathematics has extensively grown in the past years, thus becoming a trend that has gained more and more momentum in philosophy as well as in other neighboring disciplines.

Now that the philosophical interest in considering the more dynamic aspects of mathematics has been settled, it is time to ask for more details and in particular to clarify the implications for the study of mathematics, possibly opening new research perspectives. It should be noted that the interdisciplinary nature of the attention to the practice of mathematics has been emphasized at many occasions, from the beginning; however, there is a sense in which, at least in the case of philosophy, such an approach can also be seen as "intradisciplinary": As it is evident from the chapters collected in this section, the notions that have emerged in the context of the philosophy of the practice of mathematics can be connected to and enriched by the consideration of lively debates in other territories of philosophy, such as philosophy of science, social ontology, epistemology, and philosophy of action. Indeed, the aim of this section is to see whether examining the research in other fields of philosophical inquiry can reply to questions emerged in the philosophy of the practice of mathematics and open even newer perspectives, which might also imply the application of new methods.

All chapters present very recent and original work, as testified in the lists of suggestions for further research that are found in the conclusions of each of them. Moreover, despite the fact that the chapters seem to go in different directions, they are not necessarily incompatible and many of the topics that are touched upon can be related to one another. Before summing up the contents of the chapters, I will highlight three themes that seem to me to be recurrent in the characterization of new perspectives for the philosophical approach to the practice of mathematics. I will leave to the reader the challenge of finishing "connecting the dots" and finding these and other relevant links among the chapters of the section.

The first theme is the relation between a form of knowledge that is recognized as *knowing-how* and the traditional definition of knowledge as *knowing-that*. If the philosophy of mathematical practice has been stressing the "procedural" aspect of mathematical expertise, some authors – for example, Ferreiros and De Toffoli in this section – propose a framework where mathematical knowledge is still intended as a form of knowing-that, where beliefs do not lose their central place; however, this is done in a renovated fashion, so as to emphasize as well the processes – the knowing-how – that have brought to it.

A second theme is the focus on the social aspects of mathematics: What does it mean precisely that mathematics is a collective enterprise? This social element comes to the fore in all chapters, for example, in connection with social ontology or social epistemology. In particular, Cantù discusses at length the growth and the development of mathematical knowledge in a social context, with its repercussions on traditional notions such as objectivity, while Hamami stresses the role for mathematical knowledge of the actions that are learnt and performed – collectively – by

the mathematical subjects. An approach focused on the social aspects of the practice of mathematics can also make a difference in considering mathematics education, as shown in Tanswell and Inglis' chapter.

A third theme that appears in the background of the chapters is the tension between formal and informal methods in mathematics, in particular in relation to the notion of proofs.

2 Summary of the Chapters in the Section

The first two chapters in the section, Ferreiros and van Bendegem's, present innovative and more general frameworks to account for the practice of mathematics. It is important to note that both Ferreiros and van Bendegem are among the founding fathers of the *Association for the Philosophy of Mathematical Practice,*[1] which was created in 2009 – the first conference took place in Brussels in 2010.

In his chapter, Ferreiros comes back to his very influential book from 2016, and to his picture of mathematics as based on an interplay of practices. At the time, Ferreiros' book had surely the merit of putting forward an exhaustive account of mathematical knowledge as requiring a "web of practices." In this chapter, he outlines some of the ideas presented in the book and makes them even more explicit. First, he clarifies that his framework presupposes that mathematical knowledge is knowing-that: Mathematics has for him a specific theoretical orientation besides mathematical research. However, it is clear that an approach directed to the practice of mathematics will attend to the process of mathematics and look for epistemic values and goals. At the forefront, there are indeed the dynamics of knowledge; for example, we might look at the prehistory of concepts and therefore go also beyond mathematics. However, mathematics is not a natural product of the brain but involves the reference to external representations: writing, notations, and geometrical figures; this symbolic part of mathematics is important not only for reasoning but also for communication. It should be pointed out that mathematics is not a symbolic practice as any other, but a specific one where the contents involved are the notions of number and form; in general, it is the study of patterns – an idea that we will find also in van Bendegem's chapter. Ferreiros also points out that his interplay of practice does not entail a form of social constructivism. Such interplay is in fact characterized by many constraints at different levels, our biology for a start but also the several symbols that are introduced in the practice; moreover, mathematics has cultural preconditions and requires social institutions. We will come back to this point when presenting Cantù's paper.

Another interesting element in Ferreiros' approach is that it allows for the study of other proper practices that are interested in knowing-that beside Western mathematics: This is the case of Chinese or Indian mathematics, even if they do not have a notion of "theorem proving." For this reason, Ferreiros distinguishes his view from

[1] For more information about the Association, see http://www.philmathpractice.org

ethnomathematics. Interestingly, according to him, cognitive science and ethnomatematics share the inclination to identify something as innate or panhuman in mathematics; in contrast to this idea, Ferreiros reminds that there is a danger in retroprojection, that is, in looking with modern eyes at practices that are in fact not the same as ours. This may happen, for example, in the case of the concept of "function," which he discusses in detail, that cannot be thought as a simple extension of previous number concepts. In presenting the case of function, Ferreiros introduces Cavaillès' *thematization* and Piaget's *reflecting abstraction* as having an important role in accounting for the historical emergence of this notion, but also for the way students grasp this concept. It looks like there is some kind of "meta" element that is very important for Ferreiros in considering the development and the stabilization of mathematical concepts, and their emancipation from other previous more technical notions. This opens a new perspective in the history of mathematics.

van Bendegem's chapter considers mathematics as an experimental science and refers to some of the studies presented in the section of the Handbook on "experimental" mathematics. According to van Bendegem, experiments in mathematics are far from being a fiction, because they are a fact, and a fact that will keep being a part of the practice of mathematics for long; for this reason, they deserve philosophical attention. He invites the reader to think "within" his framework, which he calls a "scheme," and possibly to individuate new questions and solutions, hoping for what he calls a "*Mendelejev-effect.*" The framework presents a list of elements that characterize what can be called an experiment and is a sort of working grid that is discussed and modified throughout the chapter, following the analysis of concrete examples and the comparison of them to what is found in the empirical sciences. van Bendegem's strategy is in fact to start from a scheme that can be proposed for scientific experiments, and proceed by analogy to formulate the corresponding one to talk about experiments in mathematics. Of course, the notion of experiment is a matter of debate also in philosophy of science; however, for the sake of the argument, van Bendegem does not go into details of these discussions and simplifies them in order to focus on the case of mathematics. He specifies that his schema is just the beginning of the story, as it is plainly first order; new research might be able to provide a second-order version of it.

The connection with philosophy of science is thus made and presents frictions. In fact, if in science to make a statement and to make an observation seem to imply different things, in mathematics to make a conjecture and to acknowledge some specific instance of it look like they are of the same kind, that is, they give two mathematical statements as a result. A second problem is that while an experiment in science needs to be performed, how can we perform a mathematical experiment? In fact, it seems that in mathematics we simply prove (or disprove) a statement. Another difference is that there seem to be far more surprises in the universe that science is concerned with than in the world of mathematics. For van Bendegem, the only way of resolving these tensions is to open a new perspective, which will not focus on confirmation in mathematics but on insights and on the generation of ideas and concepts, something similar to what happens when we explore new territories. One possibility is to come back to the old notion of "probing" instead of proving,

where the mathematician looks for patterns to show up, and to consider tacit knowledge – knowing-how – as well.

van Bendegem offers several examples of experiments and introduces possible extensions starting from his schema. Among others, one possible new direction of research would be to ask whether given a mathematical theory and a scientific theory, there exist experiments that are related to the scientific theory but are also relevant for the mathematical theory. If this were the case, then such an experiment should be considered at least in part a mathematical experiment. Further research concerns the role of thought experiments in mathematics, and their relation to mathematical proofs.[2] In van Bendegem's view, a "grander" scheme should be envisaged so as to account for empiricism in the philosophy of mathematics – getting rid of the Lakatosian "quasi." He concludes his chapter by claiming that mathematical experiments are an essential ingredient of the activity of mathematicians to the same extent that understanding the path that brings to a proof is as important as obtaining the proof.

The next three chapters in the section discuss more specifically debates in three particular areas of philosophy – social ontology, epistemology, and philosophy of action – in relation to notions that have emerged in the philosophy of the practice of mathematics. In all three cases, the aim is to capitalize on these discussions to clarify notions that are still understudied in the philosophy of the mathematical practice, and conversely to evaluate whether the research about the practice of mathematics can shed a new light on them.

Cantù's very rich chapter opens up the new perspective of considering concepts from social ontology as relevant to explain the social constitution of mathematics. In her view, a philosophical-normative framework is necessary to guide the discussion on the objectivity of mathematics and on the kinds of semantics that are introduced when the attention is focused on the practice. It is only once such a framework is specified that we can look back at episodes in the history and philosophy of mathematics and see what an approach based on the practice of mathematics can teach us that we did not know already – starting from more standard pictures of mathematics – about them. As an example, Cantù takes the case of the axiomatic practice; if mathematics is assumed as having crucial informal elements, what does it remain to say about its more formal aspects, such as axiomatics or formal derivations? Is it the case that a sharp distinction between formal and informal methods in mathematics cannot be drawn? In a sense, Cantù's chapter resonates with Ferreiros', as she insists on the need to specify more what is intended by practice or practices and why they should still be subject of philosophy.

The chapter starts with a brief introduction to social ontology, which is undoubtedly very useful for the reader of this Handbook who is not necessarily acquainted with it. She discusses at length the philosophy of John Searle, who was not interested in mathematical phenomena as social facts but in social phenomena such as money,

[2] See Starikova's article in this volume and Giardino (2022), where I distinguish between thought experiments, real experiments, and quasi-experiments in mathematics.

property, or marriage; however, a lot of social ontology developed in reaction to his views. Francesco Guala criticizes Searle and claims that the focus should shift from theories about the mind to the sociohistorical study of institutions: Social ontology can be naturalized, and the dependence of social phenomena on representation can be explained in purely causal terms, without the need to introduce collective intentionality.

The point is then to compare social facts and mathematical objects, and to do that Cantù considers the views of Feferman, Ferreiros, and Cole. Feferman considered mathematical objectivity as a special case of intersubjective social objectivity, and his view, according to Cantù, can be integrated by looking at Lorenzen's research. Ferreiros – as we know it from his book and from his chapter in this section – bases his notion of practices on the presence of agents and of symbolic frameworks; however, he does not agree with Cole, who claims that mathematical objects are institutional facts, since he believes that mathematical objects are more stable than institutions. Carter, Giardino, and Pantsar's proposals on how to intend mathematical objectivity are then introduced; Cantù highlights that three different guiding principles emerge from the literature: naturalization, historicization, and normatization of mathematical practices and mathematical agents. For what regards the forms of semantics that are introduced when considering mathematical objects in this new perspective, according to Cantù's reconstruction, the semantics is representational when the focus is on mental states, and nonrepresentational when the focus is on the practice. Finally, as mentioned above, Cantù takes the example of axiomatics conceived as an institution along the lines of Guala's social theory: Axiomatics can be considered as a coordination tool and not necessarily as having a justificatory role.

Some of the elements already discussed in Cantù's chapter will be taken up again in the next two chapters, opening other complementary new perspectives. De Toffoli's objective is to reconcile analytic epistemology and the philosophy of the mathematical practice, and to this aim she introduces a fallible conception of mathematical justification.[3] In her view, beliefs, which are the main topic of epistemology, are still central to the consideration of the practice of mathematics, no matter if mathematicians are recognized as embodied, fallible and embedded in a specific context. In line with Ferreiros' insistence on mathematics as theoretical knowledge, she argues in favor of a *human* epistemology of mathematics, where subjects should not be considered in isolation. She also offers some points of contact between the mathematical subject and the characterization of subjects in social and feminist epistemologies: Knowledge can in fact be indexed to groups rather than to individuals.

Furthermore, De Toffoli shows how traditional topics in epistemology can be applied in mathematics – notably, there are Gettier cases also in mathematics – provided that we substitute the traditional infallible notion of mathematical justification that is standardly associated with proofs with its fallible counterpart; at the

[3] On this point, she draws on previous work. See De Toffoli (2021).

center of our epistemological inquiry, there are fallible human subjects and not anymore highly idealized ones. In her view, justification is fallible and therefore there is a gap between justification and truth: Whether a subject is justified is not only a matter of logic but depends also on social considerations, which does not mean that we have to endorse a form of relativism. De Toffoli's knowing subject is obviously related to Ferreiros and Hamami's agents; however, she defends the idea that it is crucial to keep separate the subjects of epistemology from the agents of ethics and opts for the use of "subject" as the appropriate term.

Hamami's chapter aims at making sense of the claim that mathematics is an activity, which has been one of the "mottos" of the philosophy of the mathematical practice from the start. It is evident that such a claim goes against standard philosophy of mathematics, which is agentless; however, what does it mean really? What is agency in mathematics? According to Hamami, it is not enough to acknowledge that mathematics has an active dimension, but we should figure out how to study it. Ferreiros is always on the background, since also for Hamami there is no practice without practitioners; closely to De Toffoli's project, he wants to account for what a subject needs to do to acquire knowledge.

Hamami surveys several themes that have emerged in the literature concerning mathematical activities and their relation to artifacts, proofs, and texts. Mathematical knowledge has been considered as rooted or grounded into elementary human activities: Kitcher has followed this line of thought as based on an ideal agent, while Ferreiros has chosen to talk about the interplay of practices. Several new perspectives are now open: the consideration of what it means to do things with artifacts such as diagrams, symbols, or graphs – referring to the notion of epistemic action, or of mathematical cognition. An important point of the chapter is the interaction of an agent with mathematical informal proofs. In his reconstruction based on previous work (see Hamami and Morris 2021), proof activity refers to the sequence of deductive inferences that correspond to a mathematical proof, which is in turn the result of the execution of a plan that is produced by a mathematical agent – what Poincaré described as the proof's *architecture* or *unity*. Moreover, proof plans can be related to proof understanding.

Hamami discusses also the question of how mathematical agents can realize mathematical actions based on mathematical texts, which directly brings to the subject of the last chapter in the section. Finally, he stresses how the most obvious areas for future research are social agency, computer agency, and extended agency. This new perspective raises questions that are both theoretical and methodological in an interdisciplinary fashion, in line with the characterization of the philosophy of the mathematical practice.

In the last chapter in the section, Tanswell and Inglis apply methods from corpus linguistics to look for empirical evidence for the prevalence of imperatives and instructions in mathematical preprints. Their focus, as highlighted by Hamami, is on the relation between agency and texts. In previous publications, Tanswell has argued for a *recipe model* of proofs (Tanswell forthcoming): In his view, proofs are like recipes, that is, they give instructions for mathematical actions to be carried out. In this chapter, together with Inglis, he offers a systematic, large-scale,

computational analysis of mathematical texts, so as to focus on the relationship between the language used in written, textual mathematics, and the way in which we should understand what proofs are. Another merit of the chapter is that it surveys interesting studies in mathematics education on the use of instructions. Methodologically, the chapter is very careful in making the hypotheses explicit before presenting the analysis in detail, and in pointing out the limits of the particular study, thus paving the way for new quantitative analyses.

As evoked above, all the chapters in this section end with a list of suggestions for further research, so as to prepare the ground for even newer perspectives. From the discussions in the chapters, we see that the interdisciplinary nature of the philosophy of the mathematical practice is confirmed today, in line with the pioneering works published at the beginning of the 2000s; furthermore, the chapters collected here also show that assuming an "intradisciplinary" approach within the regions of philosophy can enrich the debates and raise new questions. I cannot but encourage the reader to dive into them.

3 Cross-References

► Thought Experiments in Mathematics: From Fiction to Facts

Acknowledgments I wholeheartedly thank Bharath Siraman for his work and patience in the course of these long years. My thanks also go to the authors, who have taken the time and the energy to respond to the challenge of the theme of this section and present new perspectives in the philosophy of the mathematical practice. I am sure that their chapters will be a reference for future research.

References

De Toffoli S (2021) Groundwork for a fallibilist account of mathematics. Philos Q 71(4):823–844

Ferreirós J (2016) Mathematical knowledge and the interplay of practices. Princeton: Princeton University Press

Giardino V (2022) Experimenting with triangles. Axiomathes 32(Suppl 1):55–77

Hamami, Morris (2021) Plans and planning in mathematical proofs. Rev Symb Log 14(4): 1030–1065

Tanswell F (forthcoming) Go forth and multiply: on actions, instructions and imperatives in mathematical proofs. In: Bueno O, Brown J (eds) Essays on the work of Jody Azzouni. Cham: Springer

What Are Mathematical Practices? The Web-of-Practices Approach

José Ferreirós

Contents

1 Defining Mathematical Practices: Cognitive Aspects 2796
2 Defining Mathematical Practices: Sociocultural Aspects 2801
3 The Limits of Mathematical Practices 2802
4 The Web of Mathematical Practices: The Case of Functions 2806
4.1 The Web-of-Practices Approach 2806
4.2 The Function Concept 2808
5 1600 Onward: A Mixture of Mathematical and Scientific Practices 2810
6 Basic, Low-Level Practices in the Case of Functions 2814
7 Cross-References 2817
References 2817

Abstract

This chapter can be considered as made up of two parts, a general discussion of the notion of mathematical practice and the limits of its use, comprised by the first three sections, and a particular case study that is presented in order to exemplify the idea of the web of practices, which occupies the remaining three. The presentation of my approach to the notion of mathematical practice is brief and synthetic but more articulated theoretically than in a previous book (Ferreirós 2016). Considerations in Sect. 3 about mathematical cognition and ethnomathematics have to do with the "limits" of mathematical practice, i.e., with practices that do not fall under that rubric.

The case study concerns the evolution of the function concept, which offers a complex example of the web-of-practices approach. We start with early Modern developments that led to the eighteenth-century idea of function (as "analytical expression") and the nineteenth-century "logical" idea of function, showing how complex is the spiderweb of interrelated mathematical and scientific practices

J. Ferreirós (✉)
Universidad de Sevilla, Sevilla, Spain
e-mail: josef@us.es

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_47

knotted around the function concept. In the last section, I consider the difficult question of the most basic practices and cognitive experiences that form the basis of the idea of function. I propose that two conceptions – a dynamic and an algebraic or operational one – coalesced to give rise to the early Modern idea of function.

Keywords

Mathematical cognition · Ethnomathematics · History of mathematics · Function concept · Pragmatism about mathematics · Agent-based approach

The line of studies called *philosophy of mathematical practice* has been very active during the last 20 years. It is a lively community of research, including not only philosophers but also other relevant disciplines. Here is how it is described in the webpage of the APMP:

> Over the last few years, approaches to the philosophy of mathematics that focus on mathematical practice have been thriving. Such approaches include the study of a wide variety of issues concerned with the way mathematics is done, evaluated, and applied, and in addition, or in connection therewith, with historical episodes or traditions, applications, educational problems, cognitive questions, etc. We suggest using the label "philosophy of mathematical practice" as a general term for this gamut of approaches, open to interdisciplinary work.[1]

The aim of this Association was to stimulate research in philosophy of mathematics "related to mathematical activity, past and present." From this statement of goals, it should be clear already that historical studies, cognitive studies, and educational research, among others, together with philosophical and foundational studies, are relevant lines of research and analysis for people in the field.

Traditional philosophy of mathematics has been described as focused on problems of ontology and epistemology of mathematics, with a tendency to see mathematics as a fixed body of theory – usually either arithmetic (PA) or set theory (ZFC and extensions). Ontological and epistemological problems keep occupying philosophers of mathematical practice seriously, but they argue for the need to expand "the epistemology of mathematics . . . well beyond its present confines" (Mancosu 2008, p. 1) and, often, also for the need to relax ontological concerns.[2] The idea is that the more traditional questions and analytical tools leave out some highly relevant epistemological issues. As a result, careful attention has been paid to questions of evidence, fruitfulness, visualization, explanation, purity, etc., sometimes employing similar tools to those that were being exploited already in the analytical literature (see Mancosu's collection *The Philosophy of Mathematical Practice*, 2008) but sometimes using also quite different tools.

[1] See http://www.philmathpractice.org/

[2] However, this is a rather prominent trend nowadays, arguing for minimalist or "thin" ontology. See, e.g., Linnebo (2018).

As should be expected, the field is not free of disagreements. In an overview paper, Jessica Carter (2019) identifies three different (albeit in some cases overlapping) "strands" in the philosophy of mathematical practice: (1) an *agent-based* strand, with two different orientations, (2) the *historical* strand, and (3) the *epistemological* strand (so-called for lack of a better term) perceived as closer to traditional philosophy of mathematics. Here is an example of tensions: some proponents of the agent-based orientation have emphasized sociological questions, to be studied with the empirical methods of the social sciences, while some proponents of the "epistemological" strand object and argue against giving such prominence to sociological factors. In many other cases, of course, there is cooperation and/or complementarity, rather than disagreement.

One obvious, but deep, problem facing this thriving field – which is actually underscored by the wide variety of perspectives and research methodologies – is the lack of a clear definition of its intended object of study. What is meant by mathematical practice? And, by the way, singular or plural, practice or practi*ces*? Some critics argue (Salanskis 2014, p. 44) that, if the program of the practice turn is to be significant, it must recommend looking at science or mathematics in a specific way, which can be contrasted with other ways: "This absolutely requires that we have at our disposal a non-universally encompassing notion of practice." Indeed, if anything and everything, from perceiving corners in a room and counting coconuts to doing statistical analysis of big data, is "mathematical practice," then this notion says almost nothing and will be of no use.

The three strands identified by Carter, with their various sub-strands, may actually have different implicit conceptions of mathematical practice. At the very least, they have very different methodological orientations, regardless of whether they study past case studies or present-day material. In another overview, Giardino (2017, pp. 24–25) has called attention to four different "replies" to the question about mathematical practice. One suggests that maths is a *historically situated* human activity, another focuses on its being a *semiotic* practice, the third calls for an *expanded epistemology*, and the fourth understands itself as a *pragmatist* approach.

In fact, the main example of a pragmatist approach offered by Giardino is my own proposal in *Mathematical Knowledge and the Interplay of Practices* (2016), which also happens to be agent-based and historical.[3] That book constitutes an attempt to articulate an *epistemology* of mathematics, in a sense of the word that is perhaps broader than usual in the analytical tradition. Some of the claims are that understanding mathematical knowledge requires careful attention not only to its cognitive roots but also to its semiotic and cultural roots, in ways that can be illuminated by the ideas of pragmatism, with the result that one must follow or accompany (epistemologically) the historical dynamics of this form of knowledge.

Here I would like to revisit some general aspects of the proposal I made for the analysis of practices, in particular the characterization of mathematical practices and the idea of an intricate spiderweb that extends from "technical"

[3] Carter (2019) distinguishes two kinds of agent-based approach: the sociological and the pragmatist; my orientation is along the latter line.

(non-mathematical) practices to high-level mathematical ones, also including links to scientific practices of modeling, data control, etc. We shall (i) reconsider and refine the working definition of what a mathematical practice is; (ii) analyze the relations and differences between mathematical practices and other kinds of cultural practices, e.g., those studied in ethnomathematics; and (iii) consider a particular case study, the web of practices and theory related with the central concept of function, which articulated a large portion of the mathematics developed in the last 300 years.

The reader can profitably consider the chapter as made up of two parts, a general discussion of the notion of mathematical practice (hereafter MP) and the limits of its use, comprised by Sects. 1, 2, and 3, and a particular case study that is presented schematically in order to exemplify the idea of the web of practices, which occupies Sects. 4, 5 and 6 – this concerns the evolution of the concept of *function*, which, to be sure, is very complex. The presentation of my approach to the notion of mathematical practice intends to be brief and synthetic but more articulated theoretically than in the book (Ferreirós 2016). The considerations in Sect. 3 about mathematical cognition and ethnomathematics have to do with the "limits" of MPs, i.e., with what does not fall under that rubric.

1 Defining Mathematical Practices: Cognitive Aspects

Chapter 2 of Ferreirós (2016, p. 33) proposed a "working definition" of mathematical practice, restricted to the limited purposes of an epistemological analysis. This was meant to acknowledge that studies of mathematical practice may have goals different from the philosophical ones – not the case here, – and that for those purposes it may be useful to characterize mathematical practice differently. (It also meant to emphasize that I preferred to keep the "definition" simple, in order to employ it as a tool for analysis; it would have been possible, but probably of little use, to present things in a much more sophisticated way.) Here we aim to revise and update that definition, making it a bit more precise and self-contained. In any event, I should warn the reader that this first part of the chapter is a kind of précis and clarification of aspects of the book, but some important elements cannot be elaborated upon in detail. Thus I'll be referring you to the book itself for elaboration and details.

By a "practice" in general we mean a recognizable type of activity that is *done*, and can be *taught* and *learned*, by human agents. We always associate mastery of a practice with one or more particular ways of solving some problems, and we only acknowledge that somebody knows or masters a practice when some *success conditions* are met – they know *how to* do it. In many cases there will be explicit norms and rules associated with a practice, but in *all* cases, there are implicit norms.[4]

[4] Think of the practice of riding a bike or sailing a boat: we have explicit rules in the form of traffic regulations and also rules for competitions, but, even at the most basic level, there are success conditions and associated norms (the ability to maintain control and balance, to brake at or negotiate a curve, and so on).

In relation to such implicit norms, and how a given agent may learn them, other agents typically act as regulators – the community is most relevant to an individual's attainment of proficiency. Any reasonable analysis of the factors that regulate mathematical practice, making mathematical knowledge possible, must include the resources and abilities of the single mathematician and her or his interaction with others in the community. So,

First, reference to the community of mathematicians is essential, and one can state briefly the main reason: there is no practice without practitioners.

But next,

Second, each mathematical practice is linked with a recognizable type of symbolic framework—formulas or diagrams or signs of a certain kind—employed by the agents to solve problems in the pursuit of particular goals and to share information.

The symbols we are considering here are external representations, and it is important to keep in mind that diagrams are counted among them.[5] Mathematics is public and shared; it is "common knowledge" (see the Liu Hui quotation below).

At this point, it's relevant to introduce an aside, a note of warning that will accompany our reflection in what follows. Notice that, in retrospect, it will always be possible to link such mathematical practices with particular mathematical theories and thus to *reconceive* them from a theoretical standpoint, given that we, today, possess a large body of mathematics. But in fact, for our purposes, the adequate thing is to keep separate, at the level of analysis, the practice of mathematics and its products (see Sects. 4 and 6). This implies also to be careful distinguishing a given practice from any possible (and, in the chronological order, future) theoretical reconstruction. We come back to this below with remarks on retroprojection (Sect. 4), but let's continue with the exposition.

Since our notion of practice aims to help understand mathematical knowledge, i.e., developing an epistemology of mathematics, it is important to underscore that mathematics, as we understand it here, is not just practical knowledge but has at its core theoretical knowledge. Thus,

Third, we understand by mathematical practice one that is oriented to the production of knowledge, with a bias toward *know-that* rather than know-how at the level of goals.[6]

Mathematical practices incorporate a theoretical orientation, which is characteristically linked with attention to some epistemic *values*, such as accuracy or precision, consistency, and simplicity, and with the *goals* of the practice, such as generality, understanding, and certainty. As an example, consider the gou-gu rule (this is the ancient Chinese name for what we call the Pythagorean theorem – when three lengths form a right-angled triangle, they stand in the relation $a^2 = b^2 + c^2$): for practical purposes of building and the like, *particular* instances are most useful, e.g.,

[5] Diagrams are not just figures; there ought to be rules of construction and/or rules of operation on the elements of the diagram (see De Toffoli and Giardino 2014). On the symbolic component of mathematical practices, see Ferreirós (2016, pp. 46–50).

[6] Jessica Carter comments here that know-how and know-that are intertwined. Granted. They are interconnected, but what we call mathematics gives pride of place to some theoretical goals, to know-that and associated values (examples are given immediately in the text above).

the fact that triangles of sides 3-4-5 or 8-15-17, which comply with the rule, are right-angled. From a theoretical standpoint, however, the theorem as proved by Euclid or by Liu Hui is most valuable due to the great *generality* of the result. This is akin to the difference between approximative procedures, aimed at practical solutions, and perfectly exact or precise results.[7]

Attention to *processes*, and not merely to more or less perfect results, is characteristic of studies of mathematical practice. From the standpoint of practice, theories are one and *only one* of the products created by the mathematical community (besides, any given theory is in a dynamic process of reconception and changes in presentation). Other highly relevant products put forward by mathematicians are problems, conjectures, and methods, and so:

Fourth, we shall consider *problems* and *proofs* (including proof-*methods*) and *theories* as central elements of mathematics, or if you wish as the central products of the processes that conform mathematical practices.

Problem-solving, theorem proving, and theory shaping are some of the most important aspects of mathematical activity. Conjectures and methods can be made sense of as revolving around those key goals, e.g., in the process of solving a problem, some conjecture may be proposed and subsequently refined (see Lakatos 1976; on conjectures, Mazur 1997). Or take what we just called *theory shaping*, a very important activity: it aims to select those methods that are judged most appropriate and relevant for a given subject matter and to rework systematically a theoretical corpus so as to develop it in accordance with that approach. The theory will be reconceived and the proofs of theorems will be fine-tuned accordingly; often the process will lead to decisions about the principles and methods that are to be employed preferentially, about the assumptions that are decisive for establishing what one regards as the central results in the field or about the axiomatic basis on which the theory is to be erected. Value judgments are of great importance in such cases – judgments about the most adequate conceptual framework; about the "purity" of the methods employed with respect to the subject matter; about which methods or ways of proving theorems are more fruitful and more explanatory; and so on.[8]

Summing up: we have proposed that it is adequate for our purposes to concentrate on theoretical approaches to methods and results, clearly setting the mark above the so-called practical mathematics, which allows us to make a difference between mathematical practices and their prerequisites, among which we prominently find (i) basic cognitive skills and (ii) technical practices.[9] What I call "technical practices" are *techniques* such as counting, or practical measuring: practices in the general sense, which may involve the use of instruments (a rod, a knotted cord of rope, a vessel) but which do not depend on external representations or especial

[7] More on this topic in Ferreirós (2016, pp. 115–117, 119–122). On the question of generality, see Chemla, Chorlay and Rabouin (2016).

[8] There's a lot of literature on these issues, see, e.g., Mancosu (2008) and other chapters in this Handbook.

[9] Meant are cognitive skills having to do with number (or rather, the "feeling" for numerosity) and with visuospatial abilities of orientation in the environment, and of identification of shapes, see, e.g., Hohol (2020), Van der Ham et al. (2017) and Dehaene (2011).

symbolic components – oral language is involved, of course, but writing is not needed. They have to do with practical problems such as measuring fields or creating a building with right-angled corners.[10]

There are also *constraints* associated with the role of agents, with the cultural level at which we shall locate practices (in connection with symbolic frameworks, problems, results, methods, and their justification), with our interest in know-that (not just know-how), and with the main goals of practice that will be our concern. For our limited epistemological purposes, then:

(I) *Mathematical practice is what a community of agents do when, on the basis of their cognitive abilities and cultural techniques, they employ the resources of symbolic frameworks and methods* (perhaps together with other instruments) *to solve problems, justify results and procedures, shape theories, and eventually elaborate new frameworks*.

Let me insist that problem-solving, theorem proving (or perhaps better, to avoid the Western notion of *theorem*: the justification of results and procedures),[11] and theory shaping are some of the most important goals of mathematical activity.

One possible objection is that my definition does not single out *mathematical* practices. There are other kinds of symbolic practice – especially writing, but also the employment of musical notation – that comply with that definition. By writing we solve certain problems (e.g., communication between people or the storage of relevant information), sometimes we justify certain results and ideas (e.g., by logical argument and critique), and we may even articulate and develop theories (say, philosophical theories). I accept this objection. In fact, I don't think it is possible to characterize mathematics without reference to some of the *contents* involved. Mathematical practices have to do with number and form (i.e., spatial figures) or at least with results and patterns that can be linked with numbers and forms or instantiated in numerical and figural instances (e.g., geometric configurations).

Notice that concepts of number and concepts of geometric form are articulated prior to the introduction of mathematical practices in the proper sense: one can say that they belong to proto-mathematics.[12] So we add to the previous clause (I) the following:

(II) *those practices are recognized as* mathematical *due to their links with the notions of number and form* (i.e., spatial figures)*, which may be explicit or implicit;*

[10] This notion of a "technical" practice, applied to measuring and to drawing (geometric) designs, is discussed in Ferreirós (2016, pp. 38–42, 113–114). See also Sect. 6. Counting may involve instruments (like a stick and a tool to mark tallies) or may just involve fingers or other body parts.

[11] This ought to make room for including ancient Chinese maths, and Indian maths, as proper practices; Liu Hui (see Chemla and Shuchun 2005) or Bhaskara (see Plofker 2009) were clearly interested in know-that, far beyond the realm of know-how.

[12] For this idea of proto-mathematics elaborated in the case of geometry, see Ferreirós and García-Pérez (2020); interestingly, we agree on many points – including the use of this expression, proto-geometry – with Schemmel (2016), a work that only came to my attention recently.

generally speaking, mathematics has to do with the study of patterns*, numerical and geometric patterns being prominent*.

Please take note, then, that the definition of mathematical practice should include both clauses, (I) and (II).

The ancient Chinese classic of mathematics, *Nine Chapters of mathematical procedures* (*Jiŭzhāng Suànshù,* 九章算术), may serve as a paradigm example of a mathematical practice defined as per (I) and (II). We find there the symbolic ingredients; procedures for the solution of problems, which are formulaic and explicitly recorded; mathematical instruments like the try square and compass;[13] the search for patterns and interconnections between those procedures; and the systematic organization of knowledge into nine "branches" of the tree of mathematical knowledge.

This is how the great Chinese mathematician Liu Hui (in the third century) writes about mathematical training, as it existed around 200 BCE, in the preface he wrote to the *Nine Chapters*:

> Mathematics (*suan*) is part of the six arts: the ancients employed them to select people of talent, to instruct the children of high dignitaries. . . . As regards the transmission of methods, one can certainly make common knowledge, as with the trysquare (*ju*), the compass (*gui*), the numbers, and measurement; there is nothing there particularly difficult. (My translation, following Chemla and Shuchun 2005, p. 127)[14]

In the Chinese tradition, as in the Greek one of Euclid, problems of drawing figures and measurement problems were considered in one sweep, and they were approached using traditional methods based on compass and ruler, a most noteworthy coincidence.[15] This underscores commonalities linked with the above condition (II). You may think that the theoretical orientation is not sufficiently explicit in the *Nine Chapters*; nevertheless the work of Liu Hui (in his commentaries) has an undeniable theoretical bias – an explicit interest in justifications, in knowing what is *precise* and what is not, an orientation toward know-that.

Notice too that, in the definition given before, we have not required a focus on mathematical innovation or research, the elaboration of new problems, or new results. In Ferreirós (2016) I intentionally did *not* adopt the position that only

[13] The trysquare is an L-shaped instrument, a *gnomon* in Greek terminology, used to draw straight segments and right angles and to find distances. For further details about the *Nine Chapters*, see Chemla and Shuchun (2005).

[14] The above translation does not coincide with Shen et al. (1999, p. 53), where the last sentence reads: "The course of [the arithmetical arts] is not particularly difficult using the methods handed down from generation to generation, just like the compass and trysquare in measurement, with which we draw figures."

[15] This is not so different from early Modern views: D'Alembert in Vol. 7 (1757), p. 629 of the *Encyclopèdie* says: "Geometry is the study of the properties of area, inasmuch as one considers this simply as extent and form." And Legendre started his classic *Éléments de Géométrie* saying that "La géométrie est une science qui a pour objet la mesure de l'étendue," the measurement of extension.

research mathematicians do mathematics, properly speaking, since we should be interested in how individuals attain knowledge of mathematics. Doing research and developing new mathematics is certainly of enormous interest, but the question is also how aspiring members of the community learn to do mathematics and mathematical research and how they obtain their professional know-how.

2 Defining Mathematical Practices: Sociocultural Aspects

We emphasized that mathematics is the collective work of a community of agents. It should be a truism that mathematical knowledge is a collective product: no single individual could have developed the cumulative knowledge that was mastered by Liu Hui (or by Archimedes), which can only be amassed due to the existence of social institutions making it possible. To emphasize this point is by no means to argue for any strong form of social constructivism – that would imply to forget the role of biological conditions, cognitive abilities, semiotic components, logical aspects, and so on. However, it has to be made clear that mathematics has cultural preconditions and requires social institutions.

Mathematicians are a group of learned practitioners; they are erudites, *savants*, philosophers in the ancient sense of the word; and one can always see that their collective work is possible, thanks to the existence of some institutions. That is, their practices are made possible by the stratification of societies and the emergence of division of labor: Babylonian mathematics was made possible by the institution of the scribes, just like Greek mathematics was made possible by the institution of schools like the Academy in Athens and the Museion in Alexandria. The reinvention of mathematics in the nineteenth century was intimately connected with the emergence of a strictly professional class of mathematicians within industrialized societies. Fermat was an amateur, and Gauss was still (professionally) an astronomer, but these kinds of situations increasingly disappeared as the nineteenth century advanced, leaving the scene to professionals.[16]

As regards the cultural preconditions, we have assumed that mathematical practices are, at core, semiotic practices – this is why symbolic frameworks have been emphasized. Thus they depend *inherently* on forms of *external representation*, such as writing, numerical notations, and figural designs such as those usually called geometric.[17] Of course, the cultural existence, development, and preservation of such symbol systems have further preconditions in material culture, such as the development of writing media: cuneiform tablets made of clay and worked with a stylus c. 3200 BCE at Uruk in Sumer; the Egyptian papyri and the scrolls preserved in ancient libraries like the famous one of Alexandria; modern printed books, together

[16] See, e.g., C. Goldstein's chapter in *Eléments d'histoire des sciences*, ed. Michel Serres (Paris: Bordas, 1989; English version Oxford: Blackwell, 1995).

[17] On this matter, see the papers by Giardino in the Proof section and by Waszek in the Semiology of Mathematical Practice section of this Handbook. On the last topic, see Schemmel (2016).

with the book commerce and the organized libraries that accompanied them; or systems of digital archives in our computerized internet age.

Semiotic components, what I've called *symbolic frameworks*, belong among the cognitive preconditions of mathematical thinking, which by no means is merely a "natural" product of the brain. But there is more. The semiotic component is central to the emergence of communication, to the sharing of information and the interconnection of cognitive processes among agents, and thus to the intersubjective character of mathematics that turns it into true knowledge. For those reasons, mathematical cognition depends on sociocultural niches like the educational ("paidological") niches that I talked about in Ferreirós (2016, pp. 19–22); see Menary (2015).

Babylonian mathematics was made possible by the institution of the scribes, which in turn depended on the invention of writing and numbers. Notice also that the "appearance of texts clearly representative of theoretical thinking is a markedly uncommon phenomenon in the ancient world" (Boltz & Schemmel, in Schemmel ed. 2016a, p. 142), and the presence of external representations of knowledge does not lead by itself to theoretical knowledge. For instance, we find "no sources from ancient Egypt or Mesopotamia that document a theoretical reflection on spatial language" (*op. cit*., 123), which however we find in ancient Greece and ancient China, during the Warring States period.

It would be of little use to enter more deeply here into such components of cultural and mathematical history, which are well studied in a number of works.[18] Let us now turn to another question, about the lower reaches of MPs.

3 The Limits of Mathematical Practices

The previous working definition of MP may be perceived by some as too restrictive. Wouldn't it be advisable to work with a more inclusive approach that opens the way to acknowledge the mathematical aspects present in *all* human cultures? Many contributions of the last decades, to the history of non-Western mathematics, the study of mathematical cognition, and not least ethnomathematics, have adopted such a stance. Let us briefly review their perspectives.

A. *Cognitive science of mathematics* (*mathematical cognition*): according to some authors, the number concept has been wired into our brains by evolution (and into the brains of rats, pigeons, dogs too); the same is true of the space concept, Euclidean space. This means that the basic notions that we have underscored in (II) as definitory of the contents of mathematics are, one may say, "natural" and "universal" (see, e.g., Izard et al. 2011). At most, then, there could be cultural novelties having to do with the methods for dealing with such notions and

[18] See, among others, Netz (2000), Hoyrup (2002), Chemla and Shuchun (2005), Robson and Stedall (2009), Schemmel (2016), and Overmann (2019).

developing them. But mathematics is in essence, according to these authors, pan-human and innate. (Cultural elements are often acknowledged as contributing to the emergence of elaborate concepts, e.g., by Carey (2009), although the role of the cultural in cognition is mostly ignored; see Beller et al. (2018).) Such claims are in conflict with traditional philosophy of mathematics.

The crucial point for many experts in mathematical cognition is that there exist specific numerical (and spatial) capabilities in our basic cognitive setup and that they are independent from other cognitive capabilities, especially linguistic ones.[19] Yet we should make clear that speaking of an innate "number sense" is ungranted by the available data or at the very least misleading: the number concept, in a minimally precise rendering of this notion, goes beyond the basic cognitive abilities of subitizing and of judging size (what is sometimes called the AMS or "analogue magnitude system").[20] So, what should the philosophy of mathematics answer to the claims of cognitive scientists?

B. *Ethnomathematics*: experts in this field have found "mathematics" in basically all cultures: the famous *quipu* of the Incas, string games of Papua New Guinea, the patterns of Arabic geometric decoration, polyrhythmic music from the pygmies, and so on (see Ascher 1991; Eglash 2000; Gerdes 2004). Often these experts suspect ethnocentrism whenever more restrictive conceptions of "mathematics" are defended:

> Ethnomathematics holds that mathematical ideas are pan-human and are developed within cultures. ... The imposition of order on space, for example, is universal, but its particular expression varies with and within culture and may change over time. (Ascher and Ascher 1994, p. 1545)

Notice also that an important part of ethnomathematics deals in effect with applications of modern mathematics (formal languages, graph theory, knot theory, combinatorics) to the study of non-Western cultural products. The mathematical nature of string games (Vandendriessche 2015) or polyrhythmic music (Arom 1985; Chemillier 2002) is gauged by applying modern mathematical tools. There is here an intriguing combination of the opposition to ethnocentrism and the imposition of math-centrism.

C. *History of non-Western mathematics*: one finds very significant bodies of mathematical information in cultures that did not develop the famous "Greek" or Euclidean notion of proof and systematic deductive organization. Cultures and civilizations that have carefully elaborated mathematical notions and methods, to solve problems in ways that can be compared with so-called "Western" mathematical practices: Babylonian or Egyptian mathematics many centuries before the common era, Chinese mathematics from before C.E., Indian mathematics long

[19] See, among others, Dehaene et al. (2006) and Izard et al. (2011). For discussion and criticism of such views, see Núñez (2009), Overmann (2011, 2019), and Ferreirós and García-Pérez (2020).

[20] These concepts are abundantly discussed in the literature on mathematical cognition. More on the argument outlined above, in Ferreirós (2016, sections 3.4 and 3.5, and Chap. 7).

before 1300, and so on. In fact, this body of work can be analyzed with the approach to PM that I outlined in the previous two sections – for it does not presuppose a given way of proving, neither the Euclidean nor what one might call the "Hilbertian" one.

In all of these cases, we find results or ideas *comparable* to results in the mathematics of the West, but not always commensurate with them. In all of them, the old wisdom of hermeneutics applies: we can only interpret other languages, other cultures by analogy with our own; even historians, in the midst of all their sophisticated attempts to break the hermeneutic circle, cannot avoid to relate the ideas and techniques of other times or spaces with our own. But also, hermeneutics has always emphasized the need to make a conscious effort to question one's own cultural assumptions while developing an acquaintance with the foreign. If you study a non-urban, unstratified society with oral culture, it would not be wise to assume that they will display marked differences between, say, the areas of religion and sports, like the ones we may be accustomed to. I assume it wouldn't be polemical to say that some cultures do not have "sports,"[21] and so one may ask why would it be polemical to say that some cultures do not have "mathematics." A matter of cultural pride, there's little doubt.

In the historical and ethnographical cases, we find ideas or practices that can be profitably compared to mathematical results, or studied with mathematical methods, but of course the possibility to interpret something with mathematics does not imply that such a thing is mathematics. The classical example of this is music. Django Reinhardt was not like Bach. Is the musician who plays by rote learning, with no erudite knowledge of musical notation, counterpoint, and the like, still unknowingly a mathematician?

If we wanted to embrace the approach of ethnomathematics, we would have to revise the previous notion of mathematical practice. Notice, by the way, how the ethnomathematical perspective may also be seen as a natural by-product of the nativist ideas of some cognitive scientists. Both trends are inclined to emphasize that there is something innate or pan-human in mathematics. But to make room for this, one would have to modify the basic definition of mathematical practice. Not only should the emphasis on knowledge – on know-that, the theoretical – be avoided, but also we'd have to cease requiring that there exist semiotic components, formulas or diagrams, or some kind of external means of representation, as a basic element of the practice.

However, one can offer rather robust arguments for the view that the abovementioned experts in mathematical cognition are wrong: there are no concepts

[21] *Sport* describes a category of activities that only coalesced in the West in the nineteenth century and was then carried around the globe by Western colonialism and imperialism and later globalization (Guttmann 1994). One may find it wise to broaden the notion of sport so as to include Greek Olympic games, Roman gladiator games and chariot races, and Mesoamerican ball court games, even though this must be done rather carefully. But it would be uninformative to adopt such a broad notion of sport that one would *ipso facto* find them in all cultures.

of number, nor concepts of spatial configurations, without cultural practices involving external representations. The case has been made by experts such as Overmann and Núñez as regards number or Schemmel for space (see also Ferreirós and García-Pérez 2020).

Adopting the views of these experts, coming from fields as diverse as cognitive psychology, archeology, and historical epistemology – so that their views are based on careful consideration of cognitive science, anthropology, and history – it's natural to argue that one should differentiate three levels of cognition:[22]

- **Basic cognition**, of the kind that is developed from innate cognitive abilities, "naturally," by exposal to typical phenomenal experiences[23]
- **Proto-mathematics** based on basic cognition amplified by external representations, which requires some sociocultural conditions, including a certain degree of organizational and material complexity[24]
- **Mathematics** properly speaking, which is the product of theoretical reflection on certain forms of proto-mathematics, made possible by the creation of specialized external representations.

Let me give an example. The "sense of numerosity" (to use Dehaene's phrase) is pan-human, available at the level of basic cognition, and it *must be* distinguished from the possession of a concept of number; the latter emerges at the second level,[25] not in all cultures but in many of them, often in relation to measuring practices (measuring with concrete tools requires counting and thus presupposes number). Mathematics properly speaking begins when there is not only methods for the resolution of concrete problems, practical problems like counting or creating a building with right-angled corners, but also a next-higher interest in reflection about such methods and the patterns that emerge in them – questions such as whether a given method can be extended further, the exact limits of their validity, their degree of precision, etc.[26] These seem to have always required systems of writing, the elaboration of technical vocabulary, and in many cases special symbol systems for the practice of mathematics.

I claim that the distinction of those three levels is highly informative and even necessary. The labels may be judged more or less felicitous, and perhaps the expert

[22] See also Giardino (2016).

[23] Here we mean experiences characteristic of human life on earth, such as walking or running, regardless of cultural environment.

[24] On those preconditions, see, e.g., Overmann (2011, 2019).

[25] The so-called number sense is precise only at the range of subitizing but is fuzzy in general; external representations (counting with sticks, words, or body parts) make it possible to obtain a crisp general concept of number. Instead of the "number sense," one should perhaps speak of a *quantic sense*, making a contrast of crisp concept (number) with imprecise intuition (of quantity).

[26] For instance, the *gou-gu* rule is valid for certain triangles, like the 3-4-5 or the 5-12-13 triangle, but is it valid more generally for any right-angled triangle? The ratio of circumference to radius is close to 157/50, is this result fully precise?

in ethnomathematics will prefer to label 2. as "mathematics" and 3. as "theoretical mathematics." I have nothing to object to such a terminological option, provided the relevant preconditions and presuppositions are made clear. Of course, a change in the definition implies a change in the content and the conditions of applicability.

4 The Web of Mathematical Practices: The Case of Functions

In Ferreirós (2016), I presented and discussed the thesis that mathematical knowledge typically requires a web of practices, interconnected by systematic links. The meaning and function of mathematical statements can be linked with a plurality of practices, starting from what I called "technical practices" (see Sect. 1) and going all the way to sophisticated scientific practices. This can be easily seen for any of the core concepts of mathematical knowledge, such as the notions of natural number, of real number, of geometric form (in the plane), or of function. Intentionally left aside in Ferreirós (2016) were more complex notions such as function, but it may be useful to elucidate the thesis of the web of practices with this case in view.

4.1 The Web-of-Practices Approach

Let me first present a few of the basic tenets of this approach. There is a multiplicity of coexisting practices, i.e., different levels of practice, and associated knowledge, that coexist and are systematically linked in an intersubjective way (as shown by the fact that all this is taught and learned, collectively shared). To summarize the main theses, they are:

Coexistence: several different levels of practice and knowledge coexist; this means coexistence both historically in a given period and "cognitively" so to speak.[27]

Interplay: the interconnections between practices are central, i.e., their *systematic* links;[28] analysis of one practice requires to consider its relations to other MPs.

Symbolic nature: mathematical practices, properly speaking, elaborate on symbolic frameworks (semiotic components) and conceptual frameworks.

Roots in techniques: at the roots of the web of practices lie "technical" practices; also, in due course, there arise links of MPs with scientific practices.

Notice also that the web-of-practices approach *requires* an agent-based perspective: since the links between different strata (or levels of practice) are not abstractly given,

[27] A single agent usually masters many different practices; consider the examples given in the text.

[28] I will not give details about the idea of systematic links here, but it is central to my analysis; see Ferreirós (2016, pp. 83–88, 34–43) and *passim*. For details about frameworks, see Ferreirós (2016, Chap. 3).

but concretely and "cognitively" in a broad sense of this word (see Ferreirós 2016, p. 59), the whole web of interconnected practices can only be sustained by agents.[29]

I presented elementary examples of three different levels of practice interconnected, in cases such as the natural numbers, geometry, and proportion theory (see Ferreirós 2016, pp. 38–39). But of course there usually are more levels involved, as happens with core mathematical ideas such as the real numbers.

The concept of real number, even in its formative stages, links back to measuring practices (an example of a "technical" practice) and to the calculus of fractions and plane geometry (properly MPs); in its subsequent evolution, it connects with many more levels of mathematics and science. The web-of-practices approach underscores the idea that an adequate epistemological analysis of our knowledge of the real numbers should consider those practices separately but must also analyze their systematic interconnections.

Notice that this goes against some traditional tendencies in the philosophy of mathematics: from foundational studies we inherited a tendency to adopt a single mathematical system (e.g., axiomatic set theory) as a reductive basis for all of mathematics. I claim that such a reductionist perspective makes it *impossible* to properly analyze the cognitive bases of mathematical knowledge and its historical development, i.e., the dynamics of knowledge.

With a view to the configuration of mathematical knowledge prior to the seventeenth century, one can mention schematically the following three practices (Ferreirós 2016, p. 39):

> Measuring practices / fraction arithmetic / proportion theory.[30]

This interconnected web became a basis for introducing a new layer of practice and knowledge, the real numbers. Some mathematicians, notably Stevin and Newton, proposed to *define* number as "that by which the quantity of any thing is explained" (Stevin in 1585), which linked the idea of number with proportion theory; with Newton (in 1720) this became explicit, since number is for him "the abstracted ratio [proportion] of any quantity, to another quantity of the same kind" (see Ferreirós 2016, Chap. 8). Notice that this represents an enormous shift from some ancient conceptions of number, in particular the Greek conception. The arithmetic of fractions is subsumed here, the system has the classical algebraic properties, and the underlying assumption that some domains of quantities – to which the theory of ratios is applied – are perfectly continuous implies the continuity of the real number system.

[29] This applies to typical cases up to twentieth-century mathematics, but this does not preclude the possibility of higher-level interrelations (e.g., between structures) being explicitly thematized (see Ferreirós 2023).

[30] The reader can probably reconstruct on her own the systematic links involved here: to give only an example, measuring practices lead to a system of fractions as soon as we consider the idea of using submultiples of the basic measuring unit, in order to refine measurements (an example being aliquot parts of the unit). For proportion theory and the reals, see, e.g., Stein (1990).

The real numbers **R** are a veritable knot linked with all kinds of MPs. When we examine the real numbers, a complex spiderweb should have to be considered, linking number and geometry, the discrete and the continuous, algebra and analysis, and topology and more. But this is not all: one can (and should?) always consider the "applications" of those ideas in scientific or engineering contexts,[31] with astronomical models usually leading the way (in the historical order of development). There are always connections with one or more "technical" practices but also with the so-called "applications," i.e., with scientific practices (Ferreirós 2016, pp. 39–43).

When dealing with advanced mathematical concepts, the emerging picture can be truly complicated and difficult to explore. The situation with functions is correspondingly complex, but let's try to offer a survey.

4.2 The Function Concept

It should be obvious that the notion of function is central to modern mathematics as developed from the eighteenth century onward, until the beginning of the twentieth century. A very large amount of advanced mathematics can be explained and justified in terms of this central notion, which has also been at the core of innovative ways of understanding physical reality. If a great many mathematical theories can be considered as "extensions of the general theory of sets" (Bourbaki 1949, p. 7), exactly the same can be said about the theory of functions.[32] However, complex cases like that of the emergence and development of the function concept are truly defying for historical-philosophical analysis.

In a highly commendable paper, offering a general overview, Israel Kleiner wrote:

> Mathematical concepts usually develop gradually, in response to mathematical needs. While the function concept dates back about 300 years, the "instinct for functionality" may be said to be about 4000 years old. (Kleiner 1993, p. 184)

I wholeheartedly agree that mathematical concepts develop historically, and the function concept is an excellent example. But the last part of Kleiner's statement is far from being clear; indeed it is far from a formulation of "fact" – at most it should be regarded as the formulation of a problem. How far back can we follow the idea of function? Maybe the "instinct for correlation" that appears to be extremely ancient could be regarded already as an "instinct for functionality"? Or perhaps not. How do you know? How *can* you know?

[31] I use scare quotes around the word applications, because at the time those were not *external* to the mathematics, but rather perfectly natural and internal. Only with the advent of "pure mathematics" in the nineteenth and twentieth century did it make sense to regard the applications as something else, a novel addition, so to say.

[32] Set and function are interdefinable concepts: to see the point fully clear at a technical level, consider von Neumann's original version of the NBG axiomatic system (von Neumann 1925) or else the system ETCS (see Leinster 2014). After all, axiomatic set theory is, to a large extent, just a basis for developing the number systems and the theory of functions.

Fig. 1 Tablet #92698 in the British Museum (Old Babylonian period), featuring squares of the numbers up to 59. Here we find a record of calculations, but no functional thinking. . . (Neugebauer, *Mathematische Keilschrifttexte* I, pp. 68–73 (especially p. 71 n. 2). Budge, *Guide to the Babylonian and Assyrian antiquities* (London, 1922), p. 161)

Notice that in such cases there is a great danger involved in what I have called retroprojection. It's easy to see "functional thinking" when we look with modern eyes into documents that come from practices of which (very likely) that thinking was *not* part. Say, a Mesopotamian table of squares (Fig. 1; we come back to this example in the last section). The phenomenon of retroprojection is familiar from studies of mathematical cognition, where some authors go so far as to speak of the "number sense" as a system for "set detection."[33] The phenomenon is also familiar from studies of ethnomathematics, many of which adopt the methodology of analyzing cultural practices with modern mathematical tools (here retroprojection is an explicit part of the methodology, which makes it admissible, but this also invalidates some possible conclusions or extrapolations from the case studies). In our case, since we aim at analyzing the historical dynamics of mathematical concepts, we must be especially aware of retroprojection and regard it as damaging to our aims.

[33] Spelke et al. (2010, pp. 874–875) Unlike the notion of class, which in its early stages was vague and imprecise, the idea of set is very modern; it didn't exist before 1860, and of course there are no sets in the environment with which we might interact (for details, see Ferreirós 2007, especially the Epilogue).

In sharp contrast with the sentiment expressed by Kleiner, math educator W. L. Schaaf wrote in 1930:

> The keynote of Western culture is the function concept, a notion not even remotely hinted at by any earlier culture. And the function concept is anything but an extension or elaboration of previous number concepts—it is rather a complete emancipation from such notions. (Schaaf 1930, p. 500)

The idea expressed here may also be exaggerated, but as far as I know, there is no evidence that the concept of function can be found early and in particular not in the so-called origins of Western culture – ancient Greece or the Hellenistic period.

In the next two sections, I will develop some ideas about the web of practices that is knotted around the function concept. We start with Early Modern developments that led to the eighteenth-century idea of function (as "analytical expression") and the nineteenth-century "logical" idea of function. Then, in the last section, I'll enter into the difficult question of the most basic practices and cognitive experiences that form the basis of the idea of function. In the end I propose that two old conceptions – a dynamic and an algebraic or operational one – coalesced to give rise to the early Modern idea of function, but this was only the start of a longer process.

5 1600 Onward: A Mixture of Mathematical and Scientific Practices

As concerns the historical evolution of the function concept, from my perspective one should differentiate four phases: (1) the *prehistory* of the concept, up to about 1650 (I leave open the question when does it begin, but perhaps in the Middle Ages); (2) the *early history* in analytical geometry and infinitesimal analysis, leading up to the explicit formulations of the concept in the eighteenth century;[34] (3) the *classical period* that can be located in the nineteenth century, marked by the debate around Dirichlet's proposal of "arbitrary functions" and characterized by progressive expansion of the field investigated by mathematicians (e.g., discontinuous functions); and (4) the "modern" or *contemporary period*, including full extensionalization of the function concept but also featuring a series of innovative structural conceptions (e.g., function spaces).[35] However, it's not my purpose here to write a history of functions, but to investigate the web of practices that surrounds this crucial mathematical idea.

To speak about the prehistory, I surmise, there has to be a more or less explicit conceptual understanding of notions that can be recognized as leading to the idea of function. But let us for the moment ignore the difficult question of the most

[34] By Bernoulli (1718), quoted below. It is well known that, at this point, functions were mostly conceived to be (or defined as) analytical expressions; for details, see the expositions by Luzin (1936), Youschkevitch (1976), and Kleiner (1989).

[35] The evolution of the function concept has not stopped, but it falls outside the scope of this chapter to try to analyze its development during the twentieth century.

elementary practices that are linked with our conceptions of function (I come back to this in Sect. 6). We will assume that the prehistory of the function concept saw the emergence of a *dynamic* conception of the interrelation between variable magnitudes,[36] under the impulse of the study of models in astronomy and mechanics.

Such models had to do with the understanding of physical phenomena in terms of functional dependence, where *time* figures prominently and rises to the status of independent variable par excellence. If this is right, the emergence of an idea of function didn't just occur in connection with any calculation or correspondence of values, but was crucially linked to a conceptual understanding of *dynamic interrelations*. And thus the rise of functional thinking would have been inextricably linked with scientific topics in astronomy, optics, mechanics, and perhaps other fields of study.

An obvious example can be Galileo's study of accelerated motion, "the intimate relationship between time and motion" that he discovered (Galileio 1638, p. 197). Notice that we find here a clear conceptual understanding of dynamic interrelations between (physically significant) variable magnitudes, but without any explicit technical language for the concept of function. Let me quote from the Third Day of *Two New Sciences*:

> Some superficial observations have been made, as, for instance, that the free motion [*naturalem motum*] of a heavy falling body is continuously accelerated; but to just what extent this acceleration occurs has not yet been announced; for so far as I know, no one has yet pointed out that the distances traversed, during equal intervals of time, by a body falling from rest, stand to one another in the same ratio as the odd numbers beginning with unity. (Galileo 1638, p. 153)

Galileo manages to produce a precise mathematical correlation between times and spaces, expressed in the classical language of proportionality: if a mobile falls from rest, under gravity, in equal times the spaces are in the same proportion as the odd numbers 1, 3, 5… Thus he complements the idea of uniformity of motion (equal distances traversed during equal time intervals) with the picture of uniformly accelerated motion (during any equal intervals of time, equal increments of speed; see Galileo 1638, pp. 197–198).

His studies allowed him also to associate clear geometrical patterns with other physically significant phenomena: projectiles describe a curved path, which in fact is a parabola. The Galilean method is controlled by experiment but depends also on idealization; he relies on the hypothetical construction of analogical models (see Crombie 1994).

In the next phase of development, analytical geometry is formed by linking geometric practices with the general (still imprecise) idea of real numbers, and with algebraic equations, via the Cartesian way of representing and studying geometric curves by means of algebra. This adds a new layer to previous mathematical practices, incorporating innovative methods, and it also adds to the methodological

[36] Functional thinking was moving *beyond* causality, understood in the simple classical sense of basic causal schemes linked to "manipulative" causation.

tools available for the construction and study of mathematical models of astronomical systems, mechanical systems, optics, etc.[37]

The second generation of Cartesians will go beyond the finitary restrictions of the master, also deploying infinite series – a great conceptual and methodological innovation, in a context that continues to be a mixture of mathematical practices, mechanics, optics, and astronomical practices. Notice that the context is *non-mathematical* in the twentieth-century sense (i.e., not pure mathematics); this is where the crucial step of explicit introduction of the function concept will be prepared. It's an "impure" background of *mixed* mathematics: astronomy being the context for the introduction of trigonometrical operations, mechanical or astronomical systems and their evolution being the frame for emergence of the dynamic conception of how variable magnitudes are interrelated, and the dynamics of such systems being a key goal for those who introduced infinite series.

The dominant conception of functions, from their inception in the early eighteenth century, was as "analytical expressions." Bernoulli defined in 1718: "We call function of a variable magnitude, a quantity composed in any manner of this variable magnitude and of constants"; essentially similar definitions can be found in Euler and others, even in (Cauchy 1821).[38] The quantity so composed is again a "variable magnitude," whose variation can be computed from the variation of the independent magnitude. Thus, conceptually a function is just an interrelation of variable magnitudes, with one independent and one dependent variable. But the mathematical, operational aspect of things dominates the picture, leading authors to insist on "analytical" representability by means of some basic operations (still, this left room for broadening the conception by introducing new basic operations, e.g., sine and cosine).

Nevertheless, as is very well known, there was tension with alternative ideas inspired in physical images: the idea of free graph (to speak in later terms) is defended by Euler, against the ideas of D'Alembert and Bernoulli, in the celebrated controversy on the vibrating string.[39] The debate centered on admissible initial forms of the vibrating string, whether it had to be given by a single analytical expression; Euler argued that more than one law should be admitted, and even more, freely drawn curves should be admitted. This prepared the way for an extension of the function concept, a liberation from the viewpoint of explicit operations or expressions. Euler's intervention was saluted by Truesdell as "the greatest advance in scientific methodology in the entire [18th] century" (1957, p. 248).

[37] Perhaps we shouldn't forget engineering, artillery, and so on, which were branches of seventeenth- or eighteenth-century mathematics.

[38] See Euler (1748), Klügel (1803), and Cauchy (1821).

[39] See Truesdell (1957) and Bos (1980). The vibrating string controversy had to do with the problem of solving the wave equation, a partial differential equation employed to describe mathematically the vibrations of an elastic string (think of a violin). This is described in any good exposition of the evolution of the function concept, e.g., Luzin (1936), Youschkevitch (1976), and Kleiner (1989). Apparently the first time that this episode was emphasized, thanks to Dirichlet's advice, was in Riemann's exposition written in 1854 for his *Habilitation* thesis.

The mixed or impure context of physical and mathematical practices intermingled continues to be decisive in the subsequent evolution of the function concept. The next crucial steps came with the triumph of trigonometric series in the hands of Fourier and his followers. P. G. Lejeune Dirichlet, who studied in Paris and was a young member of Fourier's circle, made a decisive contribution with his rigorous study of the representation of functions by trigonometric series and his clear but challenging conception of "arbitrary functions." This was introduced in an 1829 paper published in Crelle's *Journal* but also (and the detail is worth notice) in a chapter of Dove's *Repertorium der Physik* (1837).[40] A function was now to be just any (many-to-one) correspondence of real values, which is exactly the idea expressed by the modern notation f: $\mathbf{R} \rightarrow \mathbf{R}$:

> it is not at all necessary that y depend on x in the whole interval according to the same law, indeed one does not need to think of a dependence that is expressible by mathematical operations. . . . This definition does not prescribe a common law to the different parts of the curve; one can think of it as composed of the most heterogeneous parts or even totally lawless [*gesetzlos*]. ... As long as one has determined a function for only a part of the interval, the manner of its continuation over the remaining interval will be totally left to arbitrariness [*ganz der Willkür überlassen*]. (Dirichlet 1837, pp. 152–153)

The idea of arbitrary function is meant seriously here.

One can say that this is a purely *logical* conception of functions. In effect, Dirichlet's 1829 result established that any arbitrary function that is piecewise continuous (and complies also with another condition) will be representable by a trigonometric series. Thus, a conceptual determination of the functions, by very general properties, leads to the existence of an analytical representation. That throws light on some of the deep reasons for the sharp *conceptual turn* that Dirichlet promoted, reasons behind his famous remark on "the ever more prominent tendency in modern analysis to replace calculations by thoughts."[41]

The history of the function concept continued with many ups and downs (Riemann, Weierstrass, and so on), but this is not the place to consider the details. To summarize, we have reviewed some of the main mathematical and scientific practices that were involved in the first explicit emergence of the function concept. They include on the "mathematical" side Euclidean geometry, proportion theory, real numerical magnitudes, and algebraic equations and on the side of "physics"[42]

[40] The only paper in the section Mathematical Physics: "On the representation of completely arbitrary functions by series of sinus and cosinus" (Dirichlet 1837).

[41] See Ferreirós (2007, pp. 28–31). Similar remarks were offered by Eisenstein in 1843, Riemann in 1851, Dedekind in 1871, Hilbert in 1897, and others. Riemann in particular (1851, Sect. 1) offered a very clear reflection on the import of Dirichlet's contribution. Eisenstein remarked on the strong contrast between the older school and the "fundamental principle of the new school, founded by Gauss, Jacobi, and Dirichlet" (full quote in Wussing 1984, p. 270).

[42] I place those words in quotes because they represent a categorization that is anachronistic; more or less they reflect our present conception, but this only emerged in the nineteenth century.

astronomical and mechanical models (e.g., the vibrating string) and the theory of heat (Fourier). More precisely, let us mention:

- Geometric models like those of Galileo for uniformly accelerated motion, which were interpreted in terms of proportion theory
- Other geometric models in astronomy (Kepler), in the study of the path of projectiles (Galileo), and many more
- Geometric models linked with the notion of real magnitudes and with algebraic equations and manipulations, in the "analytical geometry" of Descartes. (This was elaborated both at the level of pure geometry and at that of physical studies, optics, etc.)
- Further developments of those models (mixing geometry and algebra) with the introduction of new methods and ideas, such as infinite series, and in due course trigonometric series, and more.

Again, let us emphasize that all this is not separable from their uses in the study of astronomical phenomena, mechanical systems like the vibrating string, the theory of heat, etc.

All of these elements allowed for the articulation of a basic general conception of function, in the sense of a dynamic interconnection between variable magnitudes, that had been ripening in contexts such as astronomy and mechanics – e.g., in Galileo's work, where geometrically represented physical magnitudes appear as functions of time.[43] This underlays the earliest explicit definitions of "function" as analytical expressions (Bernoulli in 1718, Euler in 1748), which however emphasize algebraic features and which were already incorporating new ingredients from the emerging infinitesimal calculus (infinite series).

We have seen how the process starts with a "dynamic" conception, couched first in somewhat rigid algebraic language, which leads to an "analytical" notion of function. And also that this early conception evolved into the modern "logical" one of Dirichlet ("arbitrary" functions understood as relations between real numbers), again in connection with models of physical situations and with innovative mathematical methods such as the trigonometric series of Fourier.

6 Basic, Low-Level Practices in the Case of Functions

Leaving aside the further development of the function concept, I'd like to say a few more words on its early stages of development and its connections with very basic practices. In my view, the prehistory of this concept (from geometric models in astronomy, until about 1650) would be particularly interesting and promising as a field of future studies. But this prehistory is particularly understudied because it

[43] Even if lacking the conceptual and methodological resources that would be introduced along the following century, this is still expressed in the language of proportion theory.

requires us to go beyond "mathematical" history of mathematics: to analyze it in detail, we have to enter into transdisciplinary considerations. We will need to carefully explore cases in which the mathematical ideas are emerging in the context of astronomy, geometry of intensions (Oresme), the rise of mechanics (interrelations between inhomogeneous variable magnitudes, e.g., time and distance and speed), and so on.

Beyond that, there's still the question about the "technical" practices – in my sense of the word – and the most basic cognitive experiences at the basis of the idea of function. This is a fascinating topic, still awaiting careful exploration, although much material on it has been amassed by educationalists. Here I'd like to present only two hypotheses, which I will introduce with some remarks on Cavaillès and Piaget. A phenomenon that is often found in the development of mathematics is what Cavaillès called *thematization* (Benis Sinaceur 1994, p. 77), namely, progress by making explicit a notion that was implicit in a previous practice. This move takes place by reflection on that practice, making an aspect of it the thematic focus of a new symbolic framework and new methods (new concepts, a new kind of object). This is typically found in the "duality operation-object" (also "method-result") that Cavaillès emphasizes in the development of mathematics (Benis Sinaceur 1994, p. 81).

Let me offer two examples from nineteenth-century mathematics: at one stage we find groups of operations (e.g., permutations) as a method for the study of a certain body of mathematics; later we encounter a new generation of mathematicians studying groups themselves as their object, via higher-level methods. What was previously a tool in a particular theoretical setting has become a new thematic object. Second example: at one stage we find bijective correlations employed as a tool, be it in projective geometry or in Cantor's work; subsequently the general concept of mapping offers a thematic focus for a new theory that subsumes previous ideas of function, bijection, and so on. This idea of thematization seems to be intimately linked with that which Piaget called "abstraction réfléchissante," *reflecting abstraction*: some aspect of, say, action is projected onto a higher level of representation, which involves its reconstruction and reorganization on this higher plane.[44]

Thematization or reflecting abstraction seems to play an important role in the historical emergence of the notion of function but also in the way students of mathematics grasp the function concept. Paz and Leron (2009) argue that the formal function concept is intimately linked with "intuitions of actions on familiar objects." More particularly, they identify an "algebraic" conception of functions, which they present as based on the image of an input-output machine, and the "dynamic" conception that we discussed, based on the image of covariation of two magnitudes.[45] My first hypothesis is that the basis of this "algebraic" conception emerges from thematization of some aspects of basic practices of calculation or reckoning.

[44] See selected quotes in Montangero and Maurice-Naville (2013), citing the 1977 work: *Recherches sur l'abstraction réfléchissante*.

[45] There is a terminological problem, since Paz and Leron call this last the "analytical" conception, but as we've seen historically, the idea of function as an "analytical expression" is closer to their "algebraic" conception. I make my own terminological choice with a view to integrating history and cognitive analysis.

Children who learn mathematics are drilled into the practice of obtaining an output from a given input (e.g., to compute the square of a given number), in a kind of symbolic or semiotic action that, naturally, we regard as intimately connected with mathematics. It is a significant part of "the mathematical experience." Such input-output aspects of our conscious behavior can also be found in non-mathematical practices, to be sure – say, when a child is given a sentence and is assigned the task to parse it grammatically into its components. But usually, such situations are based on deployment of an explicit symbolic framework (be it numbers and calculation with them or letters and writing), and thus the basis for thematization is not just action on familiar objects, but actions on familiar symbolic systems – semiotic action. Another example could be the execution of the steps in the algorithm for solving a second-degree equation, with output the value of x.[46]

My second hypothesis would be that the historical emergence of the function concept was a product of the combination of the two conceptions of function just described. On the one hand, based on scientific practices of modeling (astronomy, optics), there emerged a conceptual understanding of dynamic interrelations, the "image of covariation of two magnitudes." On the other hand, the rise of algebra implied reflecting abstraction on reckoning operations, or their thematization, to become new focal concepts of algebraic practice. At some point, *cross-hybridization* of these two lines of development led to the idea of *representing dynamical interrelations by means of algebraic operations*, with geometric models as a background and with time as the independent variable. This was the birth of the function concept, and it is less interesting to know the "who" of this invention, than the "what" was involved. Perhaps several generations of mathematicians (from Galileo to Bernoulli, say) were needed to distil the idea and fully elaborate the new conception.

Notice that the situation is mixed or impure, as I have stressed, and also that both conceptions had ties with geometric thinking. The birth of the function concept *was not purely mathematical*, unlike, say, the concept of set in the late nineteenth century. The result was a decisive new idea of the early Modern "New science," which would be articulated around the technical terms: *function* and *variable magnitude*. (In the nineteenth century, linked with the rise of the Dirichlet conception of functions, the idea of a variable magnitude receded and was replaced by the conception of the system of real numbers. But again, let's not enter into the historical details too much.)

The subsequent development, guided by refined methods of expression and study of functional interrelations, such as infinite series and trigonometric series, led in due course to the rise of a third key conception – the "logical" conception of functions. This is a more abstract, static conception that came to be known as the "Dirichlet concept" of "arbitrary" function. Part of the complexity of the evolution of the concept of function, and the obstacles encountered by those who learn about functions (see, among many other works, Paz and Leron 2009), have to do with the conflicts between those different conceptions, which give rise to conflicting "intuitions" about functions, variables, and values.

[46] Also relevant may be intuitions of "technical" actions, such as a sequence of measuring actions (repetitive) and its output, a measured length or volume.

7 Cross-References

► Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research
► Signs as a Theme in the Philosophy of Mathematical Practice

Acknowledgments I would like to thank Valeria Giardino, John Mumma, Silvia de Toffoli, María de Paz, and Jessica Carter for their valuable comments on a previous version of this chapter. Research funded by the Spanish Government (FFI2017-84524-P, PID2021-123297NB-I00) and Junta de Andalucía (P12-HUM-1216).

References

Arom S (1985) *Polyphonies et Polyrhythmies instrumentales d'Afrique Centrale*. Paris: S.E.L.A.F.
Ascher M (1991) *Ethnomathematics: a multicultural view of mathematical ideas*. Brooks/Cole Publishing Co, Pacific Grove
Ascher M, Ascher R (1994) Ethnomathematics, Chapter 12.1. In: Grattan-Guinness I (ed) *Companion encyclopedia of the history and philosophy of the mathematical sciences*. Routledge, London, pp 1545–1554
Beller S, Bender A, Chrisomalis S, Jordan FM, Overmann KA, Saxe GB, Schlimm D (2018) The cultural challenge in mathematical cognition. *J Numer Cogn* 4(2). https://doi.org/10.5964/jnc.v4i2.137
Benis Sinaceur H (1994) *Jean Cavaillès: philosophie mathématique*. P.U.F, Paris
Bos H (1980) Mathematics and rational mechanics. In: Rousseau GS, Porter R (eds) *The ferment of knowledge: studies in the historiography of the eighteenth century*. Cambridge University Press, Cambridge, pp 327–355
Bourbaki, N (1949) Foundations of mathematics for the working mathematician. *The Journal of Symbolic Logic* 14(1):1–8.
Carey S (2009) Where our number concepts come from. *J Philos* 106(4):220–254
Carter J (2019) Philosophy of mathematical practice – motivations, themes and prospects. *Philos Math* 27(1):1–32
Cauchy AL (1821) *Cours d'Analyse de l'École Polytechnique*. Imprimerie Royale, Paris. English translation, Springer, 2009
Chemillier M (2002) Ethnomusicology, ethnomathematics. The logic underlying orally transmitted artistic practices. In: Assayag G, Feichtinger HG, Rodrigues JF (eds) *Mathematics and music*. Springer, Berlin, pp 161–183
Chemla K, Shuchun G (2005) *Les neuf chapitres: le classique mathématique de la Chine ancienne et ses commentaires*. Dunod, Paris; D.L. 2005
Chemla K, Chorlay R, Rabouin D (eds) (2016) *The Oxford handbook of generality in mathematics and the sciences*. Oxford University Press, Oxford
Crombie AC (1994) *Styles of scientific thinking in the European tradition*. Gerald Duckworth & Co., London
De Toffoli S, Giardino V (2014) Forms and roles of diagrams in knot theory. *Erkenntnis* 79(4): 829–842
Dehaene S (2011) *The Number Sense*. Oxford University Press, 2nd edn.
Dehaene S, Izard V, Pica P, Spelke E (2006) Core knowledge of geometry in an Amazonian indigene group. *Science* 311(5759):381–384
Dirichlet PGL (1837) Über die Darstellung ganz willkürlicher Functionen durch Sinus- und Cosinus-Reihen. In: Dove HW, Moser L (eds) *Repertorium der Physik*, sec. II: Mathematische Physik, vol 1. Veit & Comp, Leipzig, pp 133–160
Eglash R (2000) Anthropological perspectives on ethnomathematics. In: Selin H (ed) *Mathematics across cultures: the history of non-western mathematics*. Kluwer, Dordrecht/Boston

Euler L (1748) *Introductio in analysin infinitorum*. In: *Opera Omnia*, ser. 1, vol 8. English translation: *Introduction to analysis of the infinite*. Springer, 1988
Ferreirós J (2007) *Labyrinth of thought: a history of set theory.* Basel, Birkhäuser
Ferreirós J (2016) *Mathematical knowledge and the interplay of practices*. Princeton University Press, Princeton
Ferreirós J, García-Pérez M (2020) Beyond natural geometry: on the nature of proto-geometry. *Philosophical Psychology* 33(2):181–205
Ferreirós J, (2023) Conceptual structuralism. *J Gen Philos Sci* 54:125–148 https://doi.org/10.1007/s10838-021-09598-8
Galileo G (1638) *Dialogues concerning two new sciences*. Dover Publications, New York, 1954
Gerdes P (2004) Ethnomathematics as a new research field, illustrated by studies of mathematical ideas in African history. In: Kjeldsen TH, Pedersen SA, Sonne-Hansen LM (eds) *New trends in the history and philosophy of mathematics*. University Press of Southern Denmark, Odense
Giardino V (2016) ¿Donde situar los fundamentos cognitivos de las matemáticas? In: Ferreirós J, Casanave AL (eds) *El árbol de los números: cognición, lógica, prácticas*. Editorial Universidad de Sevilla, Sevilla, 2015
Giardino V (2017) The practical turn in philosophy of mathematics. *Phenomenol Mind* 12:18–28
Guttmann A (1994) *Games and empires: modern sports and cultural imperialism*. Columbia University Press, New York
Hohol M (2020) *Foundations of geometric cognition*. Routledge, New York
Høyrup J (2002) *Lengths, widths, surfaces: a portrait of old Babylonian algebra and its kin*. Springer, New York
Izard V, Pica P, Dehaene S, Hinchey D, Spelke E (2011) Geometry as a universal mental construction. In: Brannon E, Dehaene S (eds) *Space, time and number in the brain*. Elsevier Academic, San Diego, pp 319–332
Kleiner I (1989) Evolution of the function concept: a brief survey. *Coll Math J* 20(4):282–300
Kleiner I (2012) *Excursions in the history of mathematics*. Birkhäuser, Basel, Chaps. 5 and 6
Kleiner I (1993) Functions: historical and pedagogical aspects. *Sci Educ* 2:183–209
Klügel GS (1803) *Mathematisches Wörterbuch*, vol I–III, 1803–08. Schwickert, Leipzig
Lakatos I (1976) *Proofs and refutations*. Cambridge University Press, Cambridge
Leinster T (2014) Rethinking set theory. *Am Math Mon* 121(5):403–415
Linnebo Ø (2018) *Thin objects*. Oxford University Press, Oxford
Luzin NN (1936) Function. Part I and II. *Am Math Mon* 105(1):59–67; 105(3) (1998), 263–270
Mazur B (1997) Conjecture. *Synthese* 111(2):197–210
Mancosu P (ed) (2008) *The Philosophy of Mathematical Practice*. Oxford University Press.
Menary R (2015) Mathematical cognition – a case of enculturation. In: Metzinger T, Windt JM (eds) *Open MIND*: 25(T). MIND Group, Frankfurt am Main. https://doi.org/10.15502/9783958570818
Montangero J, Maurice-Naville D (2013) *Piaget or the advance of knowledge: an overview and glossary.* Psychology Press, New York
Netz R (2000) *The shaping of deduction in Greek mathematics: a study in cognitive history.* Cambridge University Press, Cambridge
Núñez R (2009) Numbers and arithmetic: neither hardwired nor out there. *Biol Theory* 4:68–83
Overmann KA (2011) The prehistory of number concept. *Behav Brain Sci* 34(3):142–144
Overmann KA (2019) Materiality and the prehistory of number. In: Overmann KA, Coolidge FL (eds) *Squeezing minds from stones: cognitive archaeology & the evolution of the human mind*. Oxford University Press
Paz T, Leron U (2009) The slippery road from actions on objects to functions and variables. *J Res Math Educ* 40(1):18–39
Plofker K (2009) *Mathematics in India*. Princeton University Press.
Riemann B (1851) Grundlagen einer allgemeinen Theorie der Funktionen complexer Grössen. In: Riemann B (ed) *Gesammelte mathematische Werke*, vol 1892. Teubner, Leipzig, pp 3–45
Robson E, Stedall JA (2009) *The Oxford handbook of the history of mathematics*. Oxford University Press, Oxford

Salanskis JM (2014) Some notions of action. In: Soler L, Zwart S, Lynch M, Israel-Jost V (eds) *Science after the practice turn in the philosophy, history, and social studies of science*. Routledge, New York/London, pp 44–57

Schaaf WL (1930) Mathematics and world history. *Math Teach 23*(8):496–503

Schemmel M (2016) *Historical epistemology of space*. Springer, Cham

Schemmel M (ed) (2016a) *Spatial thinking and external representation*. Max Planck Research Library, Berlin

Shen et al. (1999) *The Nine Chapters on the Mathematical Art*. Oxford University Press.

Spelke E, Lee SA, Izard V (2010) Beyond core knowledge: natural geometry. *Cogn Sci* 34(5): 863–884

Stein H (1990) Eudoxos and Dedekind: on the ancient Greek theory of ratios and its relation to modern mathematics. *Synthese* 84(2):163–211

Truesdell C (1957) The rational mechanics of flexible or elastic bodies, 1638–1788, in Euler. *Opera Omnia*, Series 2 11(2):240–262

Van der Ham IJM, Hamami Y, Mumma J (2017) Universal intuitions of spatial relations in elementary geometry. *J Cogn Psychol* 29(3):269–278

Vandendriessche E (2015) *String figures as mathematics? An anthropological approach to string figure-making in oral tradition societies*. Springer, Dordrecht

Von Neumann J (1925) An axiomatization of set theory. In: van Heijenoort J (ed) *From Frege to Gödel: a source book in mathematical logic*. Harvard University Press, Cambridge, pp 393–413

Wussing H (1984) *The genesis of the abstract group concept*. MIT Press, Cambridge, MA

Youschkevitch AP (1976) The concept of function up to the middle of the 19th century. *Arch Hist Exact Sci* 16:37–85

Experiments in Mathematics: Fact, Fiction, or the Future?

Jean Paul Van Bendegem

Contents

1 Introduction 2822
2 A First-Order Scheme for Discussing Experiments 2823
3 To Compute = To Experiment? 2826
4 A New Research Field: Experimental Mathematics 2829
5 Can There Be "Real-World" Mathematical Experiments? 2832
6 Thought Experiments 2837
7 Opening up the Scheme 2841
8 General Conclusion 2844
9 Cross-References 2844
References 2844

Abstract

In this chapter, the possibility of experiments in mathematics is examined. A general scheme is proposed as a tool to handle the different forms of experiments that are being used in mathematical practices: computations, "experimental mathematics" as a new research domain in mathematics and computer science, real-world experiments, and thought experiments. In a final section, extensions of the scheme are proposed that further support the conclusion that mathematical experiments are indeed facts and the future.

Keywords

Mathematical experiment · Computation · Pattern searching · PSLQ · Plateau's problem · Thought experiment · Lakatos

J. P. Van Bendegem (✉)
Center for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium
e-mail: jpvbende@vub.ac.be

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_32

1 Introduction

At first sight, the idea of experiments in mathematics, from the philosophical point of view though not exclusively, does not seem to make a lot of sense. After all, so the argument goes, mathematicians do not have laboratories where experiments are performed and mathematical theories are not rejected because of real-world data that contradict the theory. In fact, there are still plenty of philosophers to be found who defend such a thesis, see, e.g., Brown (2022), Buzzoni (2011, 2022) and, as it happens,[1] Van Bendegem (2003). This lack of a bodily interaction with the real world is put forward as the crucial distinction between scientific experiments and mathematical "experiments." The latter can only claim a metaphorical status at best. So, if a choice has to be made between the three "f"'s in the title, surely, it should be a fiction with little or no future. However, at second sight, simply looking at the table of contents of this handbook, a different and complementary story must be told. There is in fact a separate section, edited by Jessica Carter and Henrik Kragh Sørensen, entitled "Experimental" Mathematics,"[2] containing chapters such as "Numerical calculations as experiments in 19th century mathematics" (C. Goldstein), "On computer simulations as experiments in the natural sciences" (A. Barberousse), and "Thought Experiments in Mathematics" (I. Starikova). So surely should we rather not conclude that it is clearly a fact and very likely to stay for the future. Or, in a weaker form, that there are a number of philosophers who do believe that genuine experimentation in mathematics is indeed a possibility. In fact, as I will try to show in this survey of what can count as an experiment in mathematics, the wealth and richness of available material needs some framework to understand the connections and relations between the different forms of such experiments.

To achieve this goal, the underlying strategy of this chapter is built on the following idea. From the point of view of mathematical practices, there definitely are practices, as the chapters just mentioned demonstrate. These practices are clearly not "purely" mathematical and they do seem to share at least some characteristics of what could be considered to be an experiment of some kind, commonly understood.[3] So, if we have a closer look at these practices, what can we say about their status as experiments? Are they in some sense "genuine" experiments, similar to scientific experiments, or are they "experiments" in some metaphorical or analogical sense, or are they a new brand of experiment in its own right? To be able to answer these questions, I propose in the next section a scheme that will structure the rest of the chapter and that will help to come to some kind of first-order classification of what mathematical experiments could be. Although I had hoped for a *Mendelejev effect* –

[1] To understand this turn of phrase, see note 30.

[2] Note that the editors themselves used quotation marks to indicate that the interpretation of "experimental" is not necessarily a standard one.

[3] At this introductory stage, this description should suffice. I am referring here to the standard view of experiments being performed by physicists, chemists, usually in laboratories or specialized spaces such as observatories. Or, to put it differently, the way experiments are standardly presented in scientific documentaries for lay audiences.

i.e., a classification usually has some blank spots that no one really expected but that may happen to be filled in later on[4] – this will not happen in this chapter, though there is no reason to assume that the classification is in some sense complete, since it is after all first-order. That being said, it does allow to accommodate many of the known examples of mathematical experiments and that is precisely what will be presented in the sections following the presentation of the scheme. At least the richness and diversity of the classification are very strong arguments to indeed drop the second "f" of the title of the chapter and to insist both on experiments in mathematics being a fact with great promises for the future. Many of these developments are quite recent, as they involve computing machinery, but it seems rather clear that in years to come these experimental practices will "solidify" within the more global structure of all mathematical practices.

2 A First-Order Scheme for Discussing Experiments

The scheme that is presented below should be seen as a first-order characterization of what an experiment could be in a quite general sense. There is of course in the philosophy of science an immense literature on the topic but in order to avoid a survey of this vast domain, I will restrict myself to a "skeleton" version or, if you like, a list of elements that need to be present whenever we talk about an experiment. I do realize that this way of presenting carries a rather strong "flavor" of standard logical empiricist jargon, preferably to be expressed in the language of some logical calculus, but that is not the aim here. The aim is not to arrive at a clear-cut definition but rather to look at what the idea of an experiment presupposes (or, if you prefer, to list conditions, necessary to qualify as an experiment). The skeleton suggestion, throughout this chapter, is that an experiment E presupposes at least the following elements:

- A theory T that is related to the experiment E,
- A "realm" outside of T, usually "nature" or "the world" W,
- (A) procedure(s) P on how to execute or perform E in W, possibly involving instruments M,
- A set of relations R between the outcomes of E and T (confirmation, falsification, a Bayesian evaluation method, or some kind of support in more general terms, . . .).

First one may consider this skeleton to be somewhat naïve. In fact, I believe that the opposite is the case. The reader must have noticed that ample use of quotation marks. The reason is simple. I do not want to enter here and now into discussions about how we should regard W, whether it is a world existing independent from us, whether the person or persons performing E are themselves part of the world W and

[4] I do realize that historians and philosophers of science will consider this description rather naive but it is the idea rather than the actual historical processes that I am considering here. Gaps occur in classifications all the time and that generates the interesting question what to do with them.

what the consequences are of that parthood[5] . . . I also do not want to be too specific about the relations R.[6] What, on the other hand, is important is that this scheme covers just as well elementary experiments – drop an object and measure the distance covered in function of time covered[7] – as experiments performed with the *Large Hadron Collider* as experiments in behavioral economics or experimental psychology.

Given this elementary and general description of an experiment E, the strategy I am proposing here is to see what happens when we perform a replacement, whereby the scientific theory T is replaced by a mathematical theory MT. Again, I want to emphasize that what distinguishes T from MT is itself a tricky business[8] (although in Sect. 5 of this Chap. I will have to address the matter in some more detail). In schematic form, the crux of the matter is this: given the scheme below, what can we fill in for the question marks?

- Scientific theory T → Mathematical theory MT
- The world W → ?
- Procedures P → ?
- Instruments M → ?
- Relations R → ?

At first sight, this does not look very promising. It is not helpful at all to simply add "mathematical" everywhere, although the possibility should be considered seriously that instead of the "real" world W, we consider a mathematical world MW, which somehow leads an existence, "independent" from any mathematician. If, instead of such a mathematical world MW, we use the expression "mathematical universe U," this should sound quite familiar to many mathematicians and philosophers. It corresponds to what might be called, "folk platonism." I use this term to

[5] So, at this point, the definition should be acceptable to a naive scientific realist, to a social constructivist, to a classical view or a one or other quantum-mechanical model, where the observer is implicated in the experiment itself, co-determining its outcome.

[6] Frankly, I would not know where to start to provide an overview of the possibilities for the relations R. The few examples listed such as confirmation and Bayesian updating . . . are mere illustrations, nothing more. For the same reason I do not provide here references but limit myself to direct the reader to the Stanford Encyclopedia of Philosophy, where lemmata such as "Experiments in physics," "Computer Simulations in Science," and "Theory and Observation in Science" can be consulted.

[7] Although we might discuss whether this is indeed an elementary experiment. It requires devices for measuring distances and, even more importantly, devices for measuring time (as Galileo knew very well).

[8] For one thing, it raises the issue what the relations are between "pure" and "applied" mathematics (and, from the historical perspective, one cannot ignore the importance of "mixed" mathematics), assuming that the labels "pure," "applied," and "mixed" are clearly understood, which they are not. Sufficient to note here that the very idea of "pure" mathematics has undergone some dramatic changes in time. Or, as Ian Hacking calls it in his [Hacking 2014]: "A very wobbly distinction" (164 ff.).

make the distinction with "sophisticated platonism," where all the philosophical and mathematical considerations have their part to play, leading to subtle distinctions between (forms of) platonisms and (forms of) present-day-platonisms, connected to just as subtle considerations related to the underlying logic(s).[9] "Folk platonism" is the way that most mathematicians and many philosophers understand the idea that the mathematical universe U is a kind of independent entity, existing somewhere, where not only all mathematical objects (though not always clear whether that includes mathematical structures as well) are to be found but also the famous *Book*, containing all the godly proofs.[10] Seen thus, there is a neat way to complete the scheme, namely as follows:

- Scientific theory T → Mathematical theory T
- The world W → Mathematical universe U
- Procedures P → Mathematical intuition
- Instruments M → Mental imagery
- Relations R → Establishing mathematical "facts"

The accessibility question for U is of course a well-known and deep problem, see Linnebo (2018) for an overview, but the reference to "mathematical intuition" is surely one of the preferred candidates for an experimental procedure. And in so far as instruments are involved, it seems reasonable to assume that "mental imagery" should do. Again, a wealth of issues are coming together here – can this (in)famous intuition be the result of educational training or not? Is it innate in some sense? To what extent are we talking here about "images"? Is it necessarily visual? . . . – but these need not be answered here. This also holds for the last item, namely the establishment of mathematical "facts." In what sense can we indeed talk about facts (keeping in mind that in the philosophy of science too, facts are not straightforward entities that scientists are supposed to deal with)? Granted that in a second-order version, this should all be taken into account, here it suffices to note that plenty of quotations can be offered from mathematicians, logicians, and philosophers, who support the scheme as it is filled in above. No doubt the most famous one being the Gödel statement:

> "But, despite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, as is seen from the fact that the axioms force themselves

[9] As a special case, I mention Balaguer [1998] where so-called "full-blooded platonism" is defended, suggesting that, similar to quantum mechanics, we should better talk of mathematical multiverse U instead of single, unique universe. "Folk platonism" has little or nothing to do with such philosophically challenging theories.

[10] I am referring here, of course, to Aigner and Ziegler [2018]. The first edition appeared in 1998 and the number of pages has increased by almost 50%. This seems to suggest that the book is not all that difficult to read. Of course, the most interesting worry is whether the book is finite in size or rather something similar to the book Alpha in the Borges' story, counting a denumerably infinite number of pages.

> upon us as true. I don't see any reason why we should have less confidence in this kind of perception, i.e., in mathematical intuition, than in sense perception . . .." (Feferman et al. 1990, 268)

This analysis leads to a rather curious conclusion. If the question is asked whether experiments are possible in mathematics, then a "folk platonist" can, actually should answer with a clear yes. But that means that in that sense we have been experimenting in mathematics for centuries, ever since we had the idea that mathematical objects and structures have an independent existence in a separate universe. In that sense, mathematical experiments have been around all of the time. The question is therefore not whether experiments are possible, but rather whether the scheme can be completed in other ways than the one sketched above. As one might expect the answer, as mentioned in the introduction, is doubly positive. Firstly, yes, there exist other ways to fill in the scheme and, secondly, these ways are diverse and plentiful. To be a bit more specific, the two sections that follow suppose that we keep (something like) the mathematical universe U and explore what other choices there are for the procedures P, the instruments M, and the relations R. In Sect. 5 and, to a certain extent, also in Sect. 6. I will look at the scenario where mathematical experiments are performed in the real world W itself and investigate what the implications are of such a view.

3 To Compute = To Experiment?

Let me start with a well-known example, namely Goldbach's conjecture: for $n \geq 2$, there exist primes p_1 and p_2 such that $2n = p_1 + p_2$. As it is a statement that starts with a universal quantifier, we can always look at particular instantiations and check whether the conjecture holds for that particular instance. Taking, e.g., $n = 100$, then it is straightforward to check that $100 = 3 + 97$. Actually, if all possibilities are considered, then there are several solutions: $100 = 11 + 89 = 17 + 83 = 29 + 71 = 41 + 59 = 47 + 53$, six in total. In fact, at present the conjecture has been checked for values of n up to 4.10^{18}, this result dating from 2012. Here is the question to answer: can the computation that leads to 100 being equal to the sum of 3 and 97 be considered to be an experiment, testing Goldbach's conjecture.

At first sight, the answer seems to be yes. Take an example in the natural sciences, say biology. We have a hypothesis H and we perform an experiment E to check whether H is supported or not. If H states that all ravens are black and E is the observation of a black raven then we have support for H.[11] At second sight, however, a case can be made against this positive answer. First note that we are not in W but in U and all we seem to be dealing with are mathematical statements.[12] In the scientific

[11] Here I am actually using the analytical way of writing, though here too it carries no specific philosophical commitments. Terms such as "observation" and "support" are used here in their ordinary language (complex) meaning.

[12] That is, for all of us who are not endowed with such marvelous mathematical intuition as, e.g., Kurt Gödel.

case, we can make a distinction between H and E, the former one being a statement, the latter one an observation. But in U this is a different matter. The conjecture and the specific instance are of the same kind, they are both mathematical statements. The case is then further strengthened by an additional difficulty, namely we need not "perform" a mathematical experiment, in whatever sense of performance, we can actually *prove* that $100 = 3 + 97$. This means that the underlying strategy that is being followed here, is not that we perform an experiment for the special case $n = 100$ in support of the Goldbach conjecture/hypothesis. Rather we reason as follows: if there is proof for Goldbach's conjecture then there should also be a proof of the special case $n = 100$. Now, there is a proof for the latter, hence this supports the conjecture. So what we are really talking about here, are relations between proofs and that has little to do with experiments.[13] In summary, the answer to the question should be negative. Calculating special cases of a general statement cannot be considered as performing an experiment.[14]

This, however, need not be the end of the story. There is another way of looking at computations as experiments that, actually seems far more interesting. When we look at actual mathematical practices, as already indicated above, we note that mathematicians do not perform isolated computations, rather they check a range of cases. Of course, it seems plausible to assume that from the standard confirmation-falsification point this makes sense. A larger set of positive instances increases the likelihood of the conjecture, one negative instance and it can be rejected. Granted that a negative instance is indeed a counterexample, it does not seem to be the case that mathematicians think in terms of confirmation. The strongest argument why this is so, is the fact that there are too many examples around of universal statements that are "confirmed" up to "large" n, yet counterexamples do exist for higher values. A well-known example is Pólya's conjecture: more than half of the numbers less than n have an odd number of primes. The smallest counterexample is in the neighborhood of 1845×10^{361}. In terms of the worlds W and U, it seems that U behaves quite differently from W and holds far more surprises and therefore the confirmational approach is not really favored. Now, given the fact that mathematicians do perform these massive computations, this raises the unavoidable question: why then do they do these computations nevertheless?

The answer is that they are looking for patterns or structures to show up. A fine example is precisely Goldbach.[15] A derived function is the G-function, where

[13] I will come back to this matter near the end of this chapter.

[14] In a more elaborated version of this chapter, a comparison should be included with the ideas of Ludwig Wittgenstein on computation as an experiment, to be found in the *Remarks on the Foundations of Mathematics* (1956).

[15] There are plenty of other examples to be found in the mathematical literature. An example probably just as famous as Goldbach's conjecture is the Collatz conjecture, also known as the hailstone problem (see Van Bendegem 2005). Given a natural number n, if n is even, replace n by $n/2$, if n is odd, replace by $3n + 1$. The conjecture is that repeated applications of this transformation must produce 1 (and then go in a cycle, $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$). Interesting graphs show, e.g., the number of steps needed to get to 1, or, the highest value reached during the process. These graphs are quite informative. Another famous example is Ulam's spiral representation of distribution of primes.

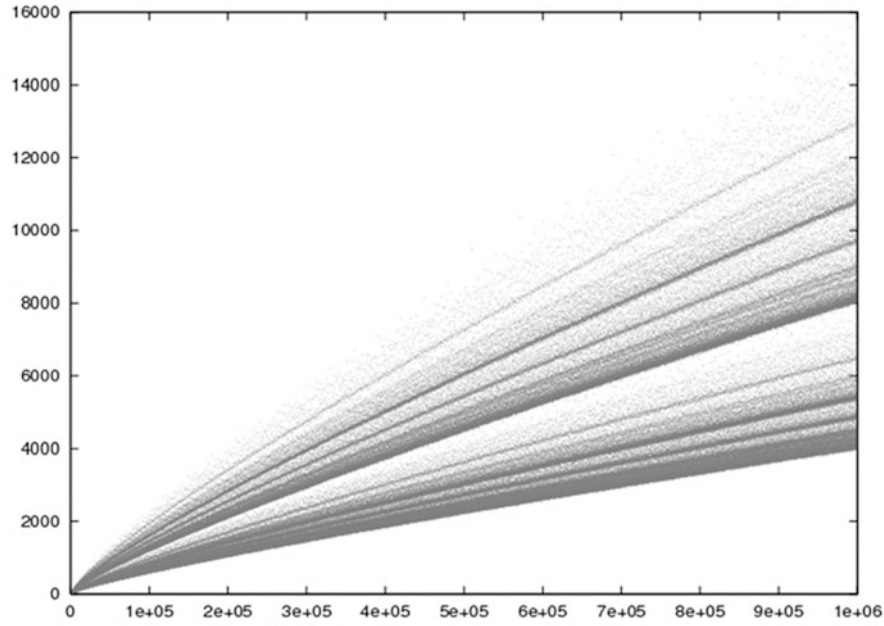


Fig. 1 Graph of the G-function

G($2n$) = number of ways of writing $2n$ as a sum of 2 primes. As mentioned above, G (10) = 6. Given this function, it is sufficient to show that G($2n$) $\geq$ 1, for all n, as this statement is equivalent to the original problem. One way of trying to prove that statement, it is sufficient to prove that, for all n, G($2(n + 1)$) $\geq$ G($2n$). If we now look at the graph of G for an initial range of values, this is the result (see Fig. 1). Clearly this is an increasing function but, in addition, we get some information about the size of the increase. We also learn something about the density that, for a given n, clearly does not appear to be uniformly distributed but rather comes in clusters. We are not talking here about confirmation but rather about insights and generation of ideas and concepts, a topic that will be revisited in this chapter. It also implies that we are not interested in the proofs of all the individual statements, such as G(10) = 6, as that would not be helpful. We are, in a real sense, performing an exploration of unknown territory and, in that sense, it definitely qualifies as an experiment. And it compares well with scientific experiments as often they too are explorations. The outcomes of the *Large Hadron Collider* almost literally produce maps that are the result of a huge amount of data. For sure, they are used to confirm theoretical ideas, such as the existence of the Higgs boson, but, at the same time, it helps researchers in their exploration, hoping that things will crop up that they did not expect.

Coming back to the scheme, it leads to the following completion:

- Scientific theory T → Mathematical theory T
- The world W → Mathematical universe U

- Procedures P → "Number crunching"
- Instruments M → Calculating power.
- Relations R → Getting insights in patterns or structures.

We can however do more. The completed scheme refers to "number crunching" and "calculating power," and it seems clear that this can only cover a small part of possible procedures and instruments. Is there a way to generalize these items? The answer is, I believe, a firm yes, especially if for the instruments computer power is used. Of course they are perfectly suited for number crunching procedures but they can do a lot more. As far as the procedures is concerned, what number crunching does is to deliver interesting graphs but the insights that may result from an inspection of and a reflection on those graphs, seem to be governed more by chance than by a more systematic approach. One happens to stumble onto something, as the expression goes. So, the question becomes whether there exist procedures to arrive at those insights in a more systematic way? The answer, once again, seems to be positive, as there are strategies and heuristics that can help and/or guide us. This leads us to the following scheme:

- Scientific theory T → Mathematical theory T.
- The world W → Mathematical universe U.
- Procedures P → Strategies and heuristics.
- Instruments M → Computer programs.
- Relations R → Getting insights in patterns or structures.

In fact, this description fits a new research domain in mathematics, namely experimental (or "experimental") mathematics. As the label suggests, it is considered to be experimental but the question remains of course whether that label is indeed justified. In the next section we will have a closer look and defend a positive answer.

4 A New Research Field: Experimental Mathematics

Let me start with a quote from two researchers in the field:

> The English word "prove" – as its Old French and Latin ancestors – has two basic meanings: to try or test, and to establish beyond doubt. The first meaning is largely archaic, though it survives in technical expressions (printer's proofs) and adages (the exception proves the rule,[16] the proof of the pudding). That these two meanings could have coexisted for so long

[16] This is not of real importance, but I cannot resist to point out that the adage "The exception proves the rule" means something different from what is usually assumed (as in this quote). In Latin the verb "probare' is indeed used so "to test' is certainly a correct translation but what was meant is that, if in court a party claims that something is an exception, then this presupposes that there must be a rule to which it is an exception. So the exception tests the idea that there must be a rule (and, hence, in legal practice it is justified for the other party to now ask what that rule is).

> may seem strange to us mathematicians today, accustomed as we are to thinking of "proof" as an unambiguous term. But it is in fact quite natural, because the most common way to establish something in everyday life is to examine it, test it, probe it, experiment with it. (Epstein and Levy 1995, 670).

By making a reference to an older use of the verb "to prove," the authors want to make clear that, because the older version is sufficiently different from the other (and recent) use, it merits to be called something else and, because of the idea of "trying" and "testing," the label "experimental" seems justified. What I have done here, is to follow another route that leads to the same conclusion: if "number crunching" can be considered to be an experimental procedure, then an extension of it should also count as such. However, let us look first at an example to get some clearer idea of what we are talking about.

The example concerns, perhaps not really unsurprisingly, the number π. There is a multitude of ways of expressing π in terms of sums of series and the like. A recent addition is this formula:

$$\pi = \sum_{i=0}^{\infty} \frac{1}{16^i}\left(\frac{4}{8i+1} - \frac{2}{8i+4} - \frac{1}{8i+5} - \frac{1}{8i+6}\right).$$

It is worthwhile to say a few words about how the formula was found. First, it is important to distinguish two phases in the "discovery" process,[17] namely (a) how was the formula found, and (b) is there a proof?

As to (a), the computer program, called PSLQ,[18] was used, based on the following strategy. Starting from the observation that numbers such as *log* 2 can be expressed as the sum of a series, namely

$$log\ 2 = \sum_{k=0}^{\infty} \frac{1}{k2^k}$$

It might be interesting to look at a generalized form of this expression, namely

$$\sum_{k=0}^{\infty} \frac{p(k)}{q(k)2^k}$$

where $p(k)$ and $q(k)$ are integer polynomials. The next step is to consider equations of the form $a_1.\alpha_1 + a_2.\alpha_2 + \ldots + a_n.\alpha_n = 0$, where the α_i are real numbers, the a_i are integer constants such that $\alpha_1 = \pi$ and the other α_i are constants as mentioned above. What one does is, given an initial choice, to apply a form of matrix reduction-

[17] Who says discovery often says justification, especially in philosophy of science, where the distinction between the context of discovery and the context of justification seems to be still very much alive. I will return to this issue near the end of this chapter.

[18] PSLQ stands for **P**olynomial-time algorithm, involving the **S**um-of-square and **LQ** decomposition schemes, where LQ refers to one of the methods to decompose matrices.

decomposition in consecutive steps, to see whether or not the equation can be brought to 0, preferably as close as possible. In this particular case, an interesting result come up, namely

$$\pi = 4F(1/4, 5/4; 1; -1/4) + 2\,arctan(1/2) - \log 5,$$

where F is a hypergeometric function. "Cleaning up"[19] this result leads precisely to the formula above. It is important to emphasize here that, at this stage of the process, no proof has been delivered (yet), as the following quote makes clear:

> The ratio between the smallest and the largest y entry when a relation is detected can be taken as a "confidence level" that the relation is *a true relation* and not an *artifact* of insufficient numeric precision. Very small ratios at detection, such as 10^{-100}, almost certainly denote a true relation (although, of course, such results are experimental only, and do not constitute rigorous proof). (Borwein and Bailey 2004, 231–232, my emphasis)

The distinction, made by the authors, between a "true relation" and an "artifact" is an interesting piece of support for considering such explorations as experiments. If we make, once more, the comparison with scientific experiments, then it is clear that the phenomenon of instruments producing artifacts is a well-known feature of scientific investigation. It is therefore an important characteristic that mathematical and scientific experiments are sharing, thus further supporting the case for genuine mathematical experiments.

As to (b), there is in fact a proof of the formula. Quite amazingly, it is no more than half a page and involves a clever use of connections between series and integrals. But, unless one has the formula as a guidance, it seems very implausible that one would think of such a connection. At the same time, it is not the case that the strategies and heuristics used to find the formula were a direct help for the proof itself. Thus, to use the terms of Epstein and Levy, the first part is indeed about probing and the second part is about proving. The two stages can be clearly distinguished from one another and the first stage has a sufficient number of features to merit the label "experimental."

Although the label "experimental" might seem more than justified, there is a counterargument that needs to be addressed. In its simplest formulation it states that, even if the search for the formula, using the program PSLQ, can be described as experimental, once the formula has been proven, it unavoidably ceases to be experimental, as it has now become a mathematical theorem. In contrast, so the argument continues, in the sciences, an experiment in a sense remains an experiment, as it can be repeated over and over again. But it can be argued that this is a misleading

[19] I will not explore this matter here but there are good arguments to claim that the development of notations and diagrams also involves experiments to see what notations do and do not "work." This is however a separate theme as this involves semiotics, specifically of notations and diagrams. See, e.g., Pietarinen et al. (2020) to have some idea of the richness of the domain. That being said, I cannot exclude that fitting this kind of experiment into the scheme of this chapter might lead to a revision (but hopefully an extension) of the scheme.

representation of how experiments function in the sciences. None of us is presently in the situation of, say, Johannes Kepler, who was trying to find out what the trajectory of a planet around the Sun looks like, based on observational data. In any decent present-day astronomy course, the movement of such a planet is now derived from basic equations, such as the law of gravitation and no observations are needed, which corresponds quite neatly to an experimentally obtained equation being proved. In addition, there seems to be little or no difference between, on the one hand, familiarizing a student with the PSLQ program and, on the other, repeating experiments in a physics class that have been done a billion of times all over the planet. This process has more to do with a form of "re-enactment": understanding the historical process by recreating parts of it. It also goes together with the acquisition of "tacit" knowledge, as Polyani in his (Polanyi 1966) already argued for years ago, a form of knowledge that requires practices rather than theories. Seen thus, it seems that mathematical and scientific experiments are perhaps more similar than we imagined. And that raises a very important question. If indeed the similarities are that strong, could it be that mathematical experiments are possible not only in the (abstract) mathematical universe U but also in the real world W itself? The exploration of that question will be the theme of the next section.

5 Can There Be "Real-World" Mathematical Experiments?

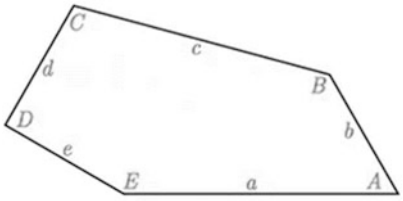
To get started, consider this simple example. Imagine that you have a pile of little squares at your disposal and "playing" around with them, you notice that you can decompose a 5×5 square into two smaller squares of sizes 4×4 and 3×3. You have in fact "discovered" an instance of the Pythagorean theorem, $5^2 = 4^2 + 3^2$. Could this count as a "real-world" *mathematical* experiment? On the one hand, the answer seems to be yes, if we take into account that a "negative" experiment is possible. So, imagine that you have 25 squares and, no matter what you do, you do not succeed in decomposing this square into two smaller squares. If that were to happen, this would constitute a serious problem for mathematics. This means that the outcomes of the experiment can make a difference. On the other hand, the answer seems to be no. To see why, imagine that we are in the close vicinity of a very dense black hole. The surrounding space-time will be strongly warped and distorted and surely very strange things will happen. Whatever comes out of Pythagorean experiments in those circumstances will not have any effect on the Pythagorean theorem itself. Actually, for one thing, the notion of a square will become problematic itself. An additional argument in favor of the negative answer is that in the description given above I made no reference whatsoever to specific properties of the little squares. Certainly not physical properties, apart from being a square and it is precisely that property that is crucial in the black hole scenario. The material, the color, the density, the solidity[20] . . ., it all makes no difference. Can we then not claim

[20] Perhaps solidity is a borderline property. One could imagine that the squares are made of ice, straight out of the freezer but that start to melt at room temperature. On the other hand, a square that is half melted is no longer a square and so the solidity is dependent on the property of being a square.

that we are actually dealing with a mathematical object rather than a physical object, even though we will be manipulating the physical object? There is another example that helps to understand what is said here. There is a famous puzzle that asks, given an amount of (real) oranges, whether it is possible to build both a square-based and a triangle-based pyramid with that given amount. It is straightforward to calculate the number of oranges needed in both cases and that leads to the equation $n(n+1)(n+2) = m(m+1)(2m+1)$. We now know that this Diophantine equation has no solutions,[21] except for the trivial one, $n = m = 1$. Imagine that there was a solution involving very large values for n and m. Then it does not make sense to claim that that solution is not acceptable as the bottom layer of a pyramid of that size will be squashed to orange juice. Quite similar to the squares, it means that the oranges have ceased to be oranges and hence we have another problem. In short, we need to look further.

A potential candidate could be tessellations. Similar to the previous discussion on computations, calculations, and the search for patterns and structures, the Pythagorean example is a "brute-force" case, simply trying out a finite number of cases, whereas tessellations involve a certain amount of not necessarily highly sophisticated mathematics. In that sense it is much closer to experimental mathematics, as discussed in the previous section.

The example illustrated in Fig. 2 is the result of such a search to find a pentagonal tile that tiles the plane.[22] However, looking more closely, the problem does boil



$$A = 60° \qquad a = 1$$
$$B = 135° \qquad b = 1/2$$
$$C = 105° \qquad c = \frac{1}{\sqrt{2}(\sqrt{3}-1)}$$
$$D = 90° \qquad d = 1/2$$
$$E = 150° \qquad e = 1/2$$

Fig. 2 Plane-filling pentagonal tile

[21] As a matter of fact, the history of this particular problem is quite interesting in itself. It started as a question asked in the advanced problems section, problem no. 3951, of the *American Mathematical Monthly* in 1940. One has to wait until 1966 for the answer (in the same journal) that the trivial solution is the only solution, see Finkelstein [1966]. The proof does involve some sophisticated mathematics (in itself not that surprising since we are dealing here with Diophantine equations). Later on connections are made with elliptic curves and measures of the irrationality of certain algebraic numbers. I am mentioning all of this because it shows that, in contrast with the Pythagorean example, it is excluded to consider a set of physical actions that one could undertake to transform one pyramid into another.

[22] Actually this problem is itself a quite curious and interesting case study. It starts a bit more than a century ago, with the construction by Karl Reinhardt of five types of pentagons that tile the plane. The classification was thought to be complete for about half a century. The example found by Casey Mann, Jennifer McLoud and David Von Derauad was solution number fifteen(!), apparently interesting enough to get mentioned in the newspapers. So far no proof has been found (after peer review) that this list of fifteen cases is a complete list.

down to finding solutions of a set of equations, relating angles and lengths of the sides of the pentagon, and thus to a mathematical problem, where the search can be done by computer:

> We discovered the tile using a computer to exhaustively search through a large but finite set of possibilities, said Casey. We were of course very excited and a bit surprised to find the new type of pentagon.[23]

There seems to be no need whatsoever to work with concrete tiles made from particular materials and trying out different patterns. Even if we are thinking of a counterexample. Imagine that a mathematician proposes a particular candidate for tiling the plane and it turns out if one actually tries it out, that a non-eliminable gap appears. Do we need really need the concrete tile to find this out? The answer is negative as a calculation of angles and sides can produce the same result. "Playing around" with concrete tiles can be inspirational but so is "messing around" with equations and trying out specific values. That is not enough to count as an experiment (though I will come back to this issue in Sect. 7).

The reasoning above does lead to an important observation: if indeed specific properties of the objects, involved in the experiment, should play a role, such as being a square, an orange, or a tile, we need arguments and justifications related to these properties to show that their physicality is somehow crucial, if not necessary. To provide arguments and justifications requires a theoretical framework, in short we need theories. What the examples above show, it cannot be a *mathematical* theory but rather it should be a scientific, say, a physical theory. If so, this leads us the following idea: given a mathematical theory MT and a scientific theory ST, can there exist experiments, related to ST, that have a relevance for MT? If so, such an experiment can be interpreted as a (perhaps partially) mathematical experiment. In fact, the answer to that question turns out to be quite deep. But, first, an example to see what we are talking about.

The example is a "great classic" and involves the existence of minimal surfaces. Figure 3 shows a wire framework and a soapy solution, attached to it, that forms a surface of minimal tension. The mathematical partial differential equations that describe this surface are well known. In fact, how otherwise could we know that they are indeed minimal? At first sight, we seem to be no better off than in the Pythagorean and tessellation examples. But let us have a closer look. What can happen when we perform experiments? I see at least three possible scenarios:

- *Case 1*: a minimal surface appears that is known. It is clear that the outcome can count as a confirmation for the physical theory PT (making abstractions of all the subtleties that surround the idea of confirmation). But for MT, this makes little or no difference. It is nothing but a known solution to a given equation. To claim that it

[23] See https://www.theguardian.com/science/alexs-adventures-in-numberland/2015/aug/10/attack-on-the-pentagon-results-in-discovery-of-new-mathematical-tile, consulted on 30 April 2021.

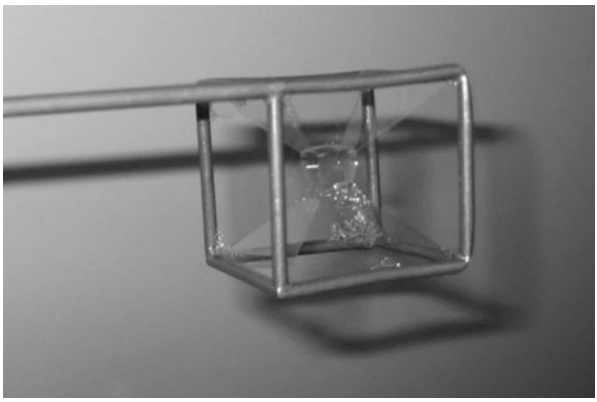


Fig. 3 A surface of minimal tension

also confirms the mathematical theory begs the question how the physical realization of a known mathematical object can add anything extra? Phrased differently, nothing comes from this case that would generate new mathematical ideas, concepts, insights, problems, proofs . . . This case therefore can be dismissed.

- *Case 2*: a minimal surface appears that is not known. Here something interesting may happen. Typical for the equations involved is that, taken together with the initial and boundary conditions, firstly, it is not certain at all that there is a solution, and, secondly, not certain at all that the solution is unique (or not). In both cases, the experiment is mathematically speaking relevant. In the former case, because it shows us a solution, this can strongly simplify the search for the mathematical solution since one might expect that mathematical characteristics can be derived from the physical form we are observing.[24] And, in the latter case, if we are in a situation where a solution is known but unicity has not been established, and where we witness a second solution[25] then the search for a unicity proof can be abandoned and that surely is a mathematically relevant result. This case at least holds a promise for mathematical experiments via the real world W into the mathematical universe U. But there is a third case to be considered.
- *Case 3*: a surface appears that is not minimal. And, let us assume, that we know that the surface is not minimal because we can mathematically prove that it is not minimal. What conclusions can be drawn from this situation? Probably the most

[24] Think, e.g., of continuous and discontinuous features of the surface. Is it smooth all over or does the surface shows the presence of sharp ridges, showing a form of discontinuity? Such considerations might lead to an enlargement of the set of solutions (and even to a better understanding of the equations themselves).

[25] What constitutes a second solution is itself a delicate matter. The surface has to be "sufficiently" different from the first one but that means some features have to be selected such that a difference can be established. A possibility could be whether or not the solution contains holes, a typical topological characteristic. This of course implies that the difference is itself co-determined by mathematical considerations. Such forms of circularity need not be harmful, only when they are vicious care should be taken.

straightforward conclusion to draw is that there is something wrong with the physical theory PT. Such an outcome of the experiment constitutes a simple refutation of PT (again, leaving aside all the intricacies surrounding the notion of falsification). Hence the mathematics remains "untouched" and it is PT that needs mending.[26] But that need not be the end of the story. As it happens, the standard mathematical treatment of such surfaces will introduce some idealizations, such as the soap film being infinitely thin, that is, as a two-dimensional structure. A strong motivation for such idealizations is of course mathematical tractability. If however (variable) thickness is taken into account as an additional parameter, the minimality of the surface can lead to new solutions and explain why the surface, that appeared to be non-minimal, turns out to be minimal after all. Such a result is both physically and mathematically relevant. From the physical point of view, it shows that the equations have to be changed to include additional parameters (that can be left out in special circumstances but not necessarily always) and that can be seen as a response to the falsification that occurred. From the mathematical point of view, it is relevant in the study of geometrical structures in relation to dimension (which in fact has happened in the study of this problem). This is perhaps less clear-cut than the second case but at least the weak conclusion seems justified that *both* PT and MT benefit from this real-world experiment.[27]

It is less evident how we should complete the scheme, considering the different cases. Trying to keep things as general as possible, I suggest the following as a possibility:

- Scientific theory T → Mathematical theory MT.
- The world W → The world W
- Procedures P → Performance of a scientific experiment
- Instruments M → Instruments as used in scientific experiments
- Relations R → Different forms of impact both on the underlying scientific theory and on the underlying mathematics used

[26] One might argue that this way of reasoning has a Quinean flavor. In a one-line summary: if revisions are to be made, start with the scientific theories, then the mathematics, and, if there is really no other option, the underlying logic (see, e.g., Quine and Ullian 1978). Seen from the perspective of this chapter, this order of revision seems to lead to a denial of mathematical experiments, if PTs are concerned. Whatever happens, will first of all be attributed to the PT. To deal with the full intricacies of this issue requires a separate paper.

[27] As one might expect, what is presented here is only an extremely rough outline of a case study worth pursuing. The interaction between mathematics and physics, when studied in more detail, becomes more and more intricate, up to the point where it becomes clear that there is a two-way traffic between the two domains. Especially the generalizations into surfaces of dimension k in a space of dimension n, greater than k, mainly done by mathematicians, has led to physically interesting consequences, e.g., concerning the occurrence of singularities. A nice survey is presented in Almgren (2001).

Note that the difference resides almost entirely in the relations R. In case 2, it is the mathematics that benefits most, whereas in case 3 it is a mutual relationship. That being said, case 3 in a more detailed study will most likely lead to special subcases. One such subcase could be the situation where there are very good and accepted arguments to not doubt the PT, that is, the equations are deemed to be a correct description of the processes occurring in the actual world, so the outcome of the experiment is not considered to be a falsification. In such a case, there would be a genuine problem for the mathematics itself. A possible answer could be that we have made mistakes in deriving solutions from the equations. Such an answer could lead to a reconsideration of the mathematics itself and count as some form of falsification as the experimental outcome forces us to change the mathematics.

Although I will not explore this idea any further – other chapters in this handbook deal with this issue in a more profound way – there seems to be a link with the discussion surrounding explanation of physical events, involving mathematics, and where the topic is the question whether *genuine* mathematical explanations are possible in such contexts.[28] By extension, this section raises the whole complex of problems and issues of the applicability of mathematics in the sciences (see Sect. 7 for some comments on this issue). The same goes for another phenomenon, involving the notion of experiment, and that occurs both in the sciences and in mathematics. And that is thought experiments, a deep and complex issue, the topic of the next section.

6 Thought Experiments

In the philosophy of science one finds a rather extensive and still growing literature on the topic.[29] A host of issues surrounds the topic, ranging from whether thought experiments (TEs) should be seen in first instance as experiments or as something else, e.g., rather as an argument (as I myself defended back in 2003, see Van Bendegem (2003)[30]). This of course seriously complicates the issue if within philosophy of science itself it is not clear whether a TE is an experiment or not. At the same time this generates a great deal of freedom vis-à-vis the question whether *mathematical* thought experiments (MTEs) exist or not. Let us note first of all that the factual question is easy to answer: indeed, MTEs exist in the sense that certain "items" that are being discussed carry that label. But that in itself is of course not sufficient, the question here is whether they can actually claim to be experiments (in some sense).

Although in this section I will be presenting evidence for the idea that there are very good reasons to consider the possibility of MTEs, it is important to realize that

[28] An excellent overview is to be found in Mancosu (2018).

[29] A standard work of reference is Stuart et al. (2018).

[30] To which must be added that in an earlier paper, Van Bendegem (1998), I defended the idea that there are no such things as experiments! The most positive reading of this multiple change of position is that it is indeed a hard and difficult topic to deal with.

they are distinct from the previous forms of mathematical experiment, in the sense that they are more closely connected to mathematical proofs. In that sense, it has many points in common with "experimental" mathematics, as discussed in the previous sections, but it does stand apart. (At the same time, it should be noted that MTEs will also fit within the scheme that is the backbone of this chapter.) As it turns out, here too we notice that the role of insight and understanding is one of the more important elements, as is also the case in classical TEs. Similar to the PSLQ algorithm, presented in Sect. 4, they do not lead necessarily in a straightforward way to a mathematical proof itself but it provides at least some elements to obtain such a proof. This somewhat vague description cries out for a specific example.

In Starikova and Giaquinto (2018), such an example is given:

> It is not news that we often make discoveries or find reasons for a mathematical proposition by thinking alone. But does any of this thinking count as conducting a thought experiment? The answer to that question is "yes," but without refinement the question is uninteresting. Suppose you want to know whether the eq. [$8x + 12y = 6$] has a solution in the integers. You might mentally substitute some integer values for the variables and calculate. In that case you would be mentally trying something out, experimenting with particular integer values, in order to test the hypothesis that the equation has no solution in the integers. Not getting a solution first time, you might repeat the thought experiment with different integer inputs. The fact that there are such mundane thought experiments is no surprise and does not answer the question we are really interested in. (257)

Note that they do consider this example to be a MTE, though perhaps not an interesting one but that is a different matter. Of course, in the quote they refer to simple calculation, i.e., the case we dealt with in Sect. 3 and then we could be skeptical about MTEs being experiments at all. But in line with the position taken in Sect. 4, other strategies can be constructed. An easy example would be to generalize and look at the eq. $8x + 12y = n$, for n an arbitrary integer. An elementary technique that is familiar to us all is that the eq. $8x + 12y$ can be equivalently rewritten as $4.(2x + 3y)$, hence that n has to be divisible by 4 for a solution to exist (or, if one is familiar with modulo arithmetic, that $n = 0 \pmod 4$). But 6 is not $-6 = 2 \pmod 4$ – hence there are no solutions. In this case, the "extended" thought experiment does produce the proof but that has to do with the basic techniques involved. Formulated in this way, it looks that MTEs of this type are close to the examples in Sect. 3 and certainly in that sense we can talk about the genuine possibility of MTEs. It is worthwhile to reiterate one more time that, once the proof has been found, the experiment as such can be "forgotten," although for a MTE a case could be made that the experiment remains as it contributes in the best of cases to the understanding[31] of the proof.

[31] I just mention here that there is a discussion going on in both the philosophy of science and of mathematics, what understanding a mathematical proof (or a mathematical concept for that matter) can mean and how it relates to scientific and mathematical explanation. They are definitely not the same but apparently it is not an easy matter to characterize mathematical understanding. A nice analysis is to be found in Folina (2018).

If we now look at the more interesting examples, one striking feature is that the MTEs tend to be visual, up to the manipulation of specific objects in the real world W (which brings us back to the previous section). Examples, discussed in the already mentioned Starikova-Giaquinto paper, are:

- *Knots*, as described in knot theory. If a set of moves are specified then one of the interesting questions is to decide whether a given knot can be transformed into the unknot. In their own words: "So one may really be experimenting, in the sense of performing a series of actions in order to test the hypothesis that the original diagram is a diagram of the unknot" (261).
- *Cayley graphs*: "Why bother with the visible diagrams of Cayley graphs? Because they can help us grasp the nature of the Cayley graphs they represent; they can help us reason about them; they can suggest to us hypotheses about them; they can help us to discover or explain facts about them" (268).
- *Cycle graphs*: "The results of the experiment are the forms we find; we use them as input to further thinking leading to our mathematical conclusion" (266).

When we referred to the actual manipulation of objects, the last example can provide a perfect illustration. A cycle graph is basically a closed chain so there is an "easy" visual representation. If we define the distance between two vertices as the length of the shortest path between them and the length as the number of edges, then we can define the diameter of a graph as the maximum of all distances. Mathematically one can prove that, if the number of vertices $= n$, then the diameter $= \lfloor n/2 \rfloor$. There is a straightforward MTE to "see" this. Imagine the graph is an actual necklace, that is, the beads have a weight and we are on the surface of the Earth so gravitation is present. If we now take the necklace by one of the beads and lift it up, the following figure (s) should result (Fig. 4). What this example suggests, is that we are now moving more in the direction of experiments as described in Sect. 5. We could in fact even perform the experiment in the real world. And, in addition, we are using here properties of the physical objects themselves (and not just their mathematical properties). This can be made clear simply by pointing out that the MTE would probably not work in the neighborhood of the already mentioned strongly warped black hole, hence the specific structure of real space enters into the MTE to be effective.

As said above, these examples emphasize the visual nature of MTEs,[32] but it would not be correct to say that these few examples are to be considered as prototypes or "exemplars" of MTEs. Two quick examples to justify this claim:

- *Negating the continuum hypothesis*: in this MTE, devised by Chris Freiling (1986), it is stated in the abstract that: "We will give a simple philosophical

[32] In that sense visual MTEs are related to so-called "proofs by looking" (or "proofs without words"), i.e., where, instead of a written-out text with standard mathematical symbols a picture is used and a mental exploration of that picture should produce the same result as the written version. See, e.g., Nelsen (1993, 2000, 2015) for an elaborate set of examples.

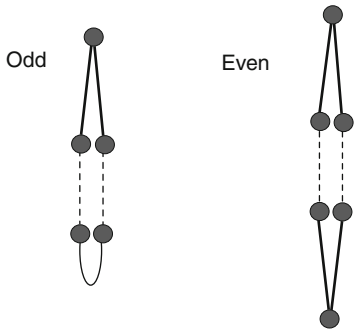


Fig. 4 A thought experiment about cycle graphs

"proof" of the negation of Cantor's continuum hypothesis (CH)." This in itself is quite interesting: the author himself prefers to speak of a "proof" (and, in addition, philosophical!) rather than a thought experiment. However its very description points in a straightforward way to a MTE: "Suppose we were to throw a random dart at the real number line (i.e., the interval [0, 1]) and ask whether the dart landed on a rational number. The outcome is, of course, predictable. We could say in advance that the dart will (with probability one) land on an irrational number." (191) Even if we can have a visual representation of the real number line, it is hard, if not impossible, to understand what it can mean "to hit a rational number." If this is a MTE then it is almost the counterpart of the necklace example above. Note also that it does not directly provide elements for a possible proof. It is not immediately clear what kind of insights are presented here. They do seem to illuminate something but what exactly? Apart from these questions, it is justified to conclude that, if we count this as a MTE, then the scope of MTEs will be wide.[33]

- *The sum of inverse squares*: this example occurs almost everywhere in the literature on mathematical practices. It concerns the famous "proof" – the quotation marks are justified because whatever it is that Euler has done, it is not to have written a proof – of the sum of the inverse squares being equal to $\pi^2/6$. The core of the matter is that Euler extends a procedure that is perfectly justified for the finite case to the infinite case where such justification is lacking.[34] Could this count as an experiment? It differs clearly from all preceding examples but one could argue that the object of what Euler did was to obtain the number $\pi^2/6$ without an acceptable proof and in that sense it can be seen as an experiment as it uncovers a particular number in the very same way that a scientific experiment delivers us data that need to be interpreted and made sense of in the framework of a theory.

[33] It must be noted that Freiling's "experiment" has received highly mixed responses but that should be a topic for a separate paper.

[34] An excellent historical presentation is to be found in Calinger (1996) and in other chapters of this handbook.

The final outcome can be counted as an insight as it guides the search for the proof as we now have a specific number to look for.

All of this shows that the field of MTEs is a complex and heterogeneous one, and that observation leaves us with one final challenge to be dealt with: how shall we complete the scheme? Here is a proposal:

- Scientific theory T → Mathematical theory T
- The world W → Mathematical universe U or the World W
- Procedures P → (Visual) arguments, (symbolic) experiments,
- Instruments M → Mental imagery
- Relations R → Getting insights in and understanding of proofs and, possibly, gathering elements to construct an actual proof

Do note that for the procedures P, I have deliberately used both the terms "arguments" and "experiments" as this particular handbook chapter is not the right place to settle the discussion whether one of the two terms is reducible to the other. We refer to the already mentioned chapter of Irina Starikova in this handbook. There is however something that the reader must have noticed: the completion of the scheme tends to become more and more general. It seems therefore legitimate to ask the question what the limits of this scheme are and to wonder whether a "grander" scheme should be envisaged. This will be the topic of the next section.

7 Opening up the Scheme

Let me start with a warning: I will not in fact present here such a "grander" scheme but rather sketch a number of roads that depart from the original scheme I have used in this chapter. I expect these roads to lead toward other chapters in this handbook for more detailed and in-depth analyses. It is worth remembering that all of the previous sections were meant to demonstrate the existence and importance of mathematical experiments. To the extent that that task has been successfully completed, we should ask ourselves what the next steps could be. What follows is nothing but a non-exhaustive list of possibilities that should include at least the following items:

- Up to now the starting point has been a mathematical theory MT. There is however no particular reason to limit oneself to MTs. I list here three examples. Firstly, mathematicians not only use theories but, as indicated in the scheme also methods: geometrical constructions, algebraic algorithms, particular proof methods such as induction, reduction, and proof by cases[35] …. Finding out what method is most suitable, effective, simple (to implement), justifiable … all of this can be done in an experimental way in the sense of an exploration. Secondly,

[35] A fuller treatment is to be found in Van Bendegem (2004).

what goes for methods, certainly goes for (symbolic) notations. The history of mathematics is also a history of the development of notations – a famous example[36] is of course the introduction of variables x, y, z . . . – involving experiments in the sense of trying out and see what works and what not. Thirdly, going back to MTs, we can look at pairs of MTs and translations between them. 2 in ordinary arithmetic becomes s(s(0)) in Peano Arithmetic, {{Ø}} or {Ø,{Ø}} in set theory, or (ιn)(Pr(n) & (∀m)(Pr(m) → m ≥ n))[37] in first-order predicate logic. And, of course, we are all familiar with the standard translation(s) between the geometry of the Euclidean plane and real and complex number theory. Note, in addition, that the three routes sketched here are not independent from one another. Translations between methods and/or notations are also possibilities to explore.

- In previous sections it has been mentioned that one of the purposes of mathematical experiments is to guide us toward proofs. In the case of the search for patterns and structures, computing power becomes important and hence the hardware, implementing the software, is an element that has to be taken into account. Two examples that come to mind are quantum computing and DNA computing.[38] It should be noted that in both cases the understanding of the computing processes requires an elaborate theoretical background, respectively, in physics and in biology. This implies that we are also dealing here with experiments in the sense of real-world experiments. What if, e.g., the quantum computer tells us that a mathematical pattern is present in the data that seems to conflict with mathematical results already known? Is mathematics in trouble, physics or technology or . . .? At the same time we should not forget that human search processes can challenge us as to how insights are generated. Let me simply mention Ramanujan. It still remains a deep problem to understand the mental processes that generated all the beautiful and deep results he produced.[39]
- Another road that certainly needs to be explored involves the proofs themselves. The well-known Euler example mentioned in the previous section is a fine example of an "almost" proof, so to speak, but it would be too narrow a view to assume that this can be the only form of "experiment." A fine example are the

[36] See, e.g., Heeffer (2009).

[37] The logical formula is perhaps somewhat unusual but it shows that we do not need a particular name for 2 in order to be able to define it. The only thing that we have to assume is that a primitive predicate Pr exists that states that a number n is a prime number. Once we have that predicate and the iota-operator "ι', the statement reads as follows: the number n, if it exists, that is a prime number and such that for any other number m, if that number is prime, then it must be at least as great as n. The second half is a logical translation of the property that 2 is the smallest prime number.

[38] For an overview of these and other forms of so-called "natural' computing, see Nunes de Castro (2006).

[39] See Kanigel (1991) for a biography of Ramanujan that shows the quite extraordinary mathematical mind he possessed. It is usually assumed that it is as good as impossible to understand his reasoning processes but recently in Raayoni et al. (2021) the so-called Ramanujan machine is introduced that seems to mimic to some extent some of his mathematical powers. The method outlined there relies on polynomial continued fractions and is related to the PSLQ algorithm mentioned in this chapter.

so-called probabilistic proofs. At play is an intriguing interaction between a rigorous proof on the one hand and a theorem that states that an object has a certain property with a probabilistic measure. A specific case concerns tests for primality that lead to conclusion that a given natural number n is prime or composite with a certain probability.[40] A related area that needs to be explored is proof verification: the computerized control of the correctness of given proofs. In ideal circumstances this control is itself of "proof quality," i.e., there is full guarantee of the proof being correct. In real circumstances it is often a probabilistic claim that is being made.[41] This leads in a rather straightforward way into the idea of tests in a statistical setting and statistical thinking in general as the relations between "pure" mathematics and statistics are complex and intricate. This opens up not so much a road as an avenue.

- This avenue is itself part of a larger network that looks at the relations between "pure" mathematics and "applied" mathematics. As the applications of mathematics are mostly in the sciences,[42] this necessitates looking at the philosophy of science. Via this route relations can be investigated between the cluster concept of experiment in the sciences versus the cluster concept of experiment in mathematics, as presented in the previous sections of this chapter. This investigation will imply a second look at mathematical experiments, now from the perspective of what scientific experiments are. This in its turn might lead to a re-evaluation of the classification presented here and eventually lead to a "reflective equilibrium,"[43] thereby firmly establishing the very notion of a mathematical experiment.

As said, this is an incomplete list for, among many other things, nothing has been said about mathematics education and the role of experiments (in any of the sense discussed above) within the educational process. And, of course, there is still that other domain to deal with, namely philosophy. Apart from the reference to "folk platonism," I have tried to keep the presentation philosophically "neutral," in the sense that, whether one is a Platonist, formalist, intuitionist, constructivist, nominalist …, a notion of mathematical experiment will need a place somewhere in the chosen philosophical framework. This thought leads straightforward to the conclusion of this chapter.

[40] The basic format is, given a number n that is to be tested, take any number $a < n$. Calculate some equation involving a. If the result is positive then a is a witness and the number is composite. If not, then chose another a and repeat. As more and more numbers do not lead to a positive result, the probability increases that the number n is prime (and the probability can be effectively calculated).

[41] The core problem is the complexity and length of some of the present-day proofs such as the four colour theorem (that has now been fully checked) and the classification of finite, simple groups. See Alama and Kahle (2013) for a nice philosophically motivated discussion of proof checking.

[42] It is a separate study to look at the interactions between mathematics and the arts.

[43] Reflective equilibrium is a method first proposed by John Rawls and refers to a back-and-forth movement to check and countercheck one's beliefs and arguments, make necessary changes and adaptations, go for another round until a form of equilibrium is reached. See Daniels (2020).

8 General Conclusion

Instead of a conclusion in the form of a summary, I think it is worthwhile to return to that philosopher who is considered as (at least one of the strong favorites to be) one of the founding figures of the philosophy of mathematical practices, namely Imre Lakatos. We are all familiar of course with *Proofs and Refutations* but one should not forget that he is also responsible for the introduction of the intriguing term "quasi-empiricism" in the philosophy of mathematics (see Lakatos 1978). It is, of course, all too easy to say that Lakatos was too strongly focused on Popperean falsification and on mathematical proofs and accompanying proof activities but the first introduction dates back to 1967, more than 50 years ago. What has happened since then is not merely an elaboration of Lakatos' approach to mathematics but also a "strengthening" of his views. If what has been presented in this chapter supports the idea of genuine experiments in mathematics then it is time that we can drop the "quasi" and instead speak about full-fledged "empiricism" in the philosophy of mathematics. In terms of the title of this chapter, experiments are indeed a fact and, on the assumption that understanding how to get to a proof is just as important and essential as actually having the proof, it is clear that mathematical experiments are an essential ingredient of what is what mathematicians do when they do mathematics and thus the final answer is that indeed they are the future.

9 Cross-References

- ▶ Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research
- ▶ Experimental Mathematics in Mathematical Practice
- ▶ Heuristics and Mathematical Practice
- ▶ Mathematical Experiments on Paper and Computer
- ▶ Non-deductive Justification in Mathematics
- ▶ Proofs, Arbitrary Exemplifications, and Inductive Generalizations in Euler's Mathematical Practice
- ▶ The Role of Experiments in *Experimental Mathematics*
- ▶ Thought Experiments in Mathematics: From Fiction to Facts
- ▶ The Intimate Interplay Between Experimentation and Deduction: Some Classroom Implications

References

Aigner M, Ziegler GM (2018) Proofs from THE BOOK. Springer, New York. (6th edition, 1st edition 1998)

Alama J, Kahle R (2013) Checking proofs. In: Aberdein A, Dove IJ (eds) The argument of mathematics. Springer, New York, pp 147–170

Almgren FJ Jr. (2001) Plateau's Problem. An Invitation to Varifold Geometry. AMS, Providence, RI. (revised edition of 1966)

Balaguer M (1998) Platonism and anti-Platonism in mathematics. Oxford University Press, Oxford

Borwein J, Bailey D (2004) Mathematics by experiment. Plausible reasoning in the 21st century. Natick, A.K. Peters
Brown JR (2022) Rigour and thought experiments: Burgess and Norton. Axiomathes 32(1):7–28
Buzzoni M (2011) On mathematical thought experiments. Epistemologia XXXIV:61–88
Buzzoni M (2022) Are there mathematical thought experiments? Axiomathes 32(1):79–94
Calinger R (1996) Leonhard Euler: the first St. Petersburg years (1727–1741). Hist Math 23: 121–166
Daniels N (2020) Reflective equilibrium. In: The Stanford encyclopedia of philosophy (summer 2020 edition), Edward N. Zalta (ed), URL = https://plato.stanford.edu/archives/sum2020/entries/reflective-equilibrium/
De Castro LN (2006) Fundamentals of natural computing. Basic concepts, algorithms, and applications. Chapman and Hall/CRC, New York
Epstein D, Levy S (1995) Experimentation and proof in mathematics. Notices of the AMS 42(6): 670–674
Feferman S et al (eds) (1990) *Kurt Gödel. Collected works volume II: publications 1938–1974*. Oxford University Press, Oxford
Finkelstein R (1966) On a Diophantine equation with no nontrivial integral solution. Am Math Mon 73(5):471–477
Folina J (2018) Towards a better understanding of mathematical understanding. In: Piazza M, Pulcini G (eds) Truth, existence and explanation. Springer, New York, pp 121–146. (*Boston Studies in the Philosophy and History of Science*, vol 334)
Freiling C (1986) Axioms of symmetry: throwing darts at the real number line. J Symb Log 51(1): 190–200
Hacking I (2014) Why is there philosophy of mathematics at all? Cambridge University Press, Cambridge
Heeffer A (2009) On the nature and origin of algebraic symbolism. In: Van Kerkhove B (ed) *New perspectives on mathematical practices. Essays in philosophy and history of mathematics*. World Scientific, Singapore, pp 1–27
Kanigel R (1991) The man who knew infinity. A life of the genius Ramanujan. Scribner's, New York
Lakatos I (1978) A renaissance of empiricism in the recent philosophy of mathematics? In: Worrall J, Currie G, Lakatos I (eds) *Mathematics, science and epistemology. Philosophical Papers volume 2*. Cambridge University Press, Cambridge, pp 24–42
Linnebo Ø (2018) Platonism in the philosophy of mathematics. *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed), URL = https://plato.stanford.edu/archives/spr2018/entries/platonism-mathematics/
Mancosu P (2018) Explanation in mathematics. *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed), URL = https://plato.stanford.edu/archives/sum2018/entries/mathematics-explanation/
Nelsen RB (1993) *Proofs without words. Exercises in visual thinking*. MAA Press, Washington, D.C.
Nelsen RB (2000) *Proofs without Words II. More exercises in visual thinking*. MAA Press, Washington, D.C.
Nelsen RB (2015) *Proofs without Words III. Further exercises in visual thinking*. MAA Press, Washington, D.C.
Pietarinen A-V, Chapman P, Bosveld-de Smet L, Giardino V, Corter J, Linker S (eds) (2020) 11th International Conference, Diagrams 2020 Tallinn, Estonia, August 24–28, 2020. Proceedings diagrammatic representation and inference. New York: Springer
Polanyi M (1966) The tacit dimension. Routledge & Kegan Paul, London
Quine WV, Ullian JS (1978) The web of belief. Mcgraw-Hill, New York
Raayoni G, Gottlieb S, Manor Y, Pisha G, Harris Y, Mendlovic U, Haviv D, Hadad Y, Kaminer I (2021) Generating conjectures on fundamental constants with the Ramanujan machine. Nature 590:67–73
Starikova I, Giaquinto M (2018) Thoughts experiments in mathematics. In: Michael T. Stuart et al., op.cit.

Stuart MT, Fehige Y, Brown JR (eds) (2018) The Routledge companion to thought experiments. London: Routledge

Van Bendegem JP (1998) What, if anything, is an experiment in mathematics? In: Anapolitanos D, Baltas A, Tsinorema S (eds) *Philosophy and the many faces of science*, (CPS publications in the philosophy of science). Rowman & Littlefield, London, pp 172–182

Van Bendegem JP (2003) Thought experiments in mathematics: anything but proof. Philosophica 72:9–33. (date of publication: 2005)

Van Bendegem JP (2004) The creative growth of mathematics. In: Gabbay D, Rahman S, Symons J, Van Bendegem JP (eds) Logic, epistemology and the Unity of science (LEUS), vol 1. Kluwer Academic, Dordrecht, pp 229–255

Van Bendegem JP (2005) The Collatz conjecture: a case study in mathematical problem solving. Logic and Logical Philosophy 14:7–23

Wittgenstein L (1956) Bemerkungen über die Grundlagen der Mathematik. Anscombe GEM, Rhees R, von Wright GH (eds) Frankfurt: Suhrkamp, 1984. (Originally published in 1956. English translation: Ludwig Wittgenstein: *remarks on the foundations of mathematics*. Edited by von Wright G. H., Rhees R. & Anscombe G. E.M., translated by Anscombe G. E.M. Oxford: Basil Blackwell)

The Social Constitution of Mathematical Knowledge: Objectivity, Semantics, and Axiomatics

Paola Cantù

Contents

1 Introduction 2848
2 A Brief Introduction to Key Issues in Social Ontology 2852
3 The Social Constitution of Mathematical Objectivity: A Survey 2856
3.1 Solomon Feferman's Conceptual Structuralism 2856
3.2 José Ferreirós Interplay of Practices 2858
3.3 Julian Cole's Institutionalization of Mathematics 2861
4 The Social Constitution of Mathematical Knowledge: Key Issues 2863
4.1 Objectivity With or Without Objects? 2863
4.2 A Representational or Non-representational Semantics? 2867
5 Axiomatics as a Social Mathematical Practice 2869
6 Conclusion 2872
7 Cross-References 2874
References 2874

Abstract

The philosophy of mathematical practice sometimes investigates the social constitution of mathematics but does not always make explicit the philosophical-normative framework that guides the discussion. This chapter investigates some recent proposals in the philosophy of mathematical practice that compare social facts and mathematical objects, discussing similarities and differences. An attempt will be made to identify, through a comparison with three different perspectives in social ontology, the kind of objectivity attributed to mathematical knowledge, the type of representational or non-representational semantics adopted, and the justificatory or coordinative role entrusted to axiomatics. After a brief introduction to key issues in social ontology, Sect. 3 of the chapter offers a survey of contributions by Feferman, Ferreirós, and Cole, highlighting a difference between approaches based on rules, practices, and intentional states,

P. Cantù (✉)
Centre Gilles Gaston Granger, Aix-Marseille University, CNRS, Aix-en-Provence, France
e-mail: paola.cantu@univ-amu.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_57

respectively. Section 4.1 also discusses results by Carter, Collin, Giardino, and Pantsar, focusing on differences between the objectivity of knowledge and objectivity of objects and showing that many authors combine a realist, structuralist, and pragmatist perspective, as well as the idea that objectivity comes in degrees. Section 4.2 focuses on the kind of semantics that is adopted in socially oriented approaches. A representational semantics is preferred by authors grounding their views on mental states, whereas a non-representational semantics best fits with views based on practices. Section 5 considers how axiomatics can be understood in a social perspective. Axiomatics generally plays a justificatory role also in theories that aim to explain the social constitution of mathematical knowledge. Yet, if one shifts from approaches based on mental states to approaches anchored on rules or habits, axiomatics can be viewed as an institution based on obligations, functions, coordination problems, and agent's actions and roles, thus playing an organizational and coordination role.

Keywords

Mathematical practice · Social ontology · Objectivity · Semantics · Axiomatics

1 Introduction

The philosophy of mathematical practice is a constellation of researches having various aims and interests. Many authors share some form of anti-foundationalism, attention to history, and interest for ontological and epistemological questions that arise from mathematical case studies, partly in opposition to classical trends in analytic philosophy. Recent surveys have detailed the origins and the history of the study of mathematical practice, recalling its interdisciplinary nature, its multiform approaches based on agency, epistemological concerns or historical studies.[1] Most authors agree that mathematical knowledge is not obtained through the analysis of abstract properties of entities that can be analyzed independently of the practices in which they emerge. They claim that it should be understood through the study of human activities. However, it is not always clear to what extent the study of such activities can still be the prerogative of philosophy and not precisely of other disciplines such as history, sociology, psychology, cognitive sciences, or educational theory. Indeed, one may wonder whether the pragmatic turn does not end up undermining the normative viewpoint in favor of a merely descriptive approach.[2] The question we will ask here is whether mathematical knowledge understood as a product of intersubjective social practices exhibits radically different characteristics (e.g., in terms of objectivity) or raises new logical problems (e.g., the development of

[1] For a survey of the different notions and traditions included in the philosophy of mathematical practice, see Mancosu (2008), van Bendegem (2016), Giardino (2017), Carter (2019), Hamami and Morris (2020b), and of course other chapters of this handbook.

[2] See, for example, Gandon's (2013) criticism.

non-representational semantics) or allows for a better understanding of the axiomatic practice.

The aim of this article is to shed light on some more recent philosophical trends that analyze how mathematical objects can be socially constituted and at the same time exhibit a kind of objectivity that is, if not identical to the objectivity granted by platonistic approaches, at least sufficient to explain the stability and interculturality of parts of mathematical knowledge. In a sense, now that the presence of a whole range of prelinguistic, pre-symbolic, and informal practices is deemed legitimate, it becomes possible but also urgent to question how these practices were constituted and whether they can give rise to the mathematical objects that are typically investigated by philosophy. For example, the development of the philosophy of mathematical practice goes hand in hand with the development of the idea that mathematics is also an informal discipline. But what then of formal demonstrations and axiomatics? Do the objects studied in the philosophy of mathematical practice really have the same degree of objectivity as formal demonstrations, axioms, and theorems understood as abstract propositions? If mathematical objects are no more considered as independent entities, but are viewed as a byproduct of interactions between agents, how can one account for their objectivity? This chapter will offer a survey of several recent theories that attempt to solve this problem, explaining how intersubjectivity gives rise to practices and practices give rise to objects, considered either as the outputs of certain processes or as a reification of roles or social habits. This has various consequences, including the discovery that mathematical objectivity is not a matter of all or nothing, as it is in ontology (an object exists or does not exist), but a matter of degrees (an object is more or less robustly constituted by certain practices).[3] That mathematical objectivity is a matter of degree leads to questions about the possibility of drawing a sharp line between formal and informal methods in mathematics. On the one hand, formal approaches cannot disregard the analysis of intentional contents (e.g., the intended model of an axiomatic theory). On the other hand, the notion of interpreted formal language is gaining more and more legitimacy, suggesting that formalized axiomatics need not be based on uninterpreted symbols.[4] In both cases, the study of intentionality and interpretation refers back to epistemic actions and interactions that suggest the importance of studying underlying human practices and activities.

The study of the social constitution of mathematical objects is thus of interest not only to philosophers who believe that mathematics is essentially a human practice and who investigate its informal aspects but also to philosophers who wish to account for the formal aspects of mathematics, because the latter cannot be fully understood except against the backdrop of a more pointed analysis of the practices associated with the

[3] On the effects that this change of viewpoint has on mathematical objectivity, see the TOPOI special issue "Mathematical practice and social ontology," where different degrees of objectivity are compared (Cantù and Testa 2023).

[4] See Sundholm (2002) on formal interpreted languages and expressions without meaning and Cantù (2021) on the distinction between uninterpreted and as-if uninterpreted symbols in the Peano School axiomatics.

construction of formal axiomatic systems. For this very reason, this chapter will not only present a survey of some promising lines of research on the social constitution of mathematical objects, discussing key features of mathematical knowledge such as objectivity and semantics, but also offer new insights on a mathematical practice that is generally regarded as a formal activity and rarely described using the tools of social ontology: axiomatics.[5] Comparisons between social ontology and mathematical practice serve on the one hand to show how investigations into the social constitution of mathematics are not beyond the scope of philosophical inquiry into mathematics; indeed they even bear similarities to metaphysical approaches to the study of social interactions such as those conducted by social ontology. On the other hand, they suggest the need to make explicit the philosophical-normative frameworks underlying much research in philosophy of mathematical practice.

The approach followed in this chapter opens a new perspective on the philosophical investigation of mathematical practice, because attention is paid to authors who question the relationship between social and mathematical facts. Even if tendencies in the direction of a disciplinary crossover between philosophy of mathematics and social ontology can be traced, bridges between the two disciplines still have to be developed and no general survey has been offered yet of existing preliminary explorations. Since it is not usual to analyze mathematical practice from the concepts of social ontology, we will provide in Sect. 2 a brief introduction to the latter, focusing on some of its characterizing aspects, such as social constructivism, collective intentionality, and regulative and constitutive rules, which we hope will be helpful in orienting readers experienced in mathematical practice studies but less familiar with the key concepts of social ontology. We will recall John Searle's ideas, which are necessary for understanding Julian Cole's contributions, and then the criticisms advanced from Guala and other naturalist philosophers.

We will then move in Sect. 3 to the presentation of recent studies that describe mathematics as socially constituted knowledge. It is certainly not possible to offer here an exhaustive and organic survey of all the literature analyzing mathematics as a social activity. Reuben Hersh was among the first authors who explicitly claimed that mathematics is a human activity and a social, cultural, and historical phenomenon. Hersh (1997) claimed that mathematical objects are created by humans, not arbitrarily, but as a byproduct of interactions with existing objects and from the needs of science and daily life. Hersh created a history of what he called (following Aspray and Kitcher 1988) the Maverick tradition including modern (Peirce, Poincaré, Wittgenstein, Lakatos, Pólya, Polányi, White, Wilder) and contemporary authors (Popper, Bunge, Kitcher, Piaget, Vygotsky, Ernest), and ethnomathematics. He opposed mathematical foundationalism and especially Platonism and formalism. We will start where Hersh stops, including twenty-first-century authors who were not already mentioned in Hersh or other existing surveys in the literature, and which

[5] A notable exception to the investigation of hypothetico-deductive axiomatics as a mainly formal activity is Schlimm, who, quoting Bourbaki, claims that formalization is just one aspect and the least interesting of axiomatics (Schlimm 2013, 41).

we consider particularly relevant to a comparison with the processes of social construction, language, intentionality, and rules discussed in social ontology.[6]

The starting point is Solomon Feferman, who explicitly brings to the table many of the issues we are interested in: he asserts the structuralist, intersubjective, and objective nature of mathematical knowledge and its similarity to other social constructions. And indeed, Ferreirós and Cole's research, which we will analyze afterward, can be seen as two relatively independent developments of Feferman's central ideas: the social constitution of mathematics from empirical activities and the study of mathematical concepts as social facts.[7] In their perspective two central issues are the explanation of the possibility for mathematics to be both socially constituted and objective and the development of a non-representational semantics capable of expressing the inferential component of mathematics without presupposing abstract entities existing independently of human activity.

After outlining some key points of the research of Feferman, Ferreirós, and Cole, we will move on to a comparative analysis of their work, considering further research by Carter, Collin, Giardino, Pantsar, Hartimo, and Rytilä. We will see how their views interact with three different approaches in social ontology, centered on intentional states, rules, and habits.

Points of contact and divergences between social ontology and approaches in philosophy of mathematics that defend the social constitution of mathematical objectivity might include (1) different ways of characterizing objectivity – as a property of mathematical knowledge or as a property of mathematical objects; (2) the function assigned to intentionality, language, and rules in the constitution of the intersubjectivity necessary to justify a certain degree of objectivity; and (3) different ways of analyzing mathematical agents: through a naturalization of mathematical practice that requires the analysis of the cognitive and epistemic abilities of agents, through a historical and comparative study of the development of mathematical abilities in different cultures or through a normative approach, which considers agents as dialectical roles. In the second section of this chapter, we will mainly focus on objectivity and semantics, leaving aside the discussion of the other directions of research.[8]

[6] We will not discuss the work of Paul Ernest either, who, like Hersh, analyzed the contributions of Maverick authors, but was more interested in discursive practices and in characterizing mathematical knowledge. Adopting Bloor's theory of objectivity, he distinguishes between explicit and implicit mathematical knowledge, claiming that both produce an objective knowledge, i.e., a knowledge that is shared and intersubjective, deployed and used in public, for persons to witness (Ernest 1998: 144ff). For recent developments of Ernest's views, see chapter ▶ "The Ethics of Mathematical Practice," this volume.

[7] See Ferreirós (2022) for an explicit discussion of Feferman's conceptual structuralism, in which, however, the author explicitly distances himself from the idea that the kind of objectivity proper to mathematics, called robust, is akin to the objectivity of social facts.

[8] On collective agency we refer to Yacin Hamami's article ▶ "Agency in Mathematical Practice," this volume. On language, symbolization, and artifacts, see Valeria Giardino 2018b and 2023. For further directions of research see also M. Kremakova's article "The Sociology of Mathematical Practice," this volume and the paper by M. Inglis and K. Weber "Education and Mathematical Practice," this volume.

To illustrate how this comparative approach between disciplines that rarely interact can suggest new perspectives in the study of mathematical practice, we will provide in the last section some new results suggesting how axiomatics, a practice combining formal and informal elements, can appear in a different light when analyzed as the fruit of a social constitution. Should axiomatics be understood only as a mathematical method based on relations (inference rules) between given sets of sentences (axioms, theorems, definitions) expressed in a formal language or also as an institution based on rules, obligations, functions, and agent's actions and roles? Is its goal merely foundational or includes coordination and organization of knowledge?

2 A Brief Introduction to Key Issues in Social Ontology

This section briefly introduces the notions of collective intentionality, regulative and constitutive rules, and status functions, as introduced by Searle. This choice is based on the fact that many of the later theories in social ontology were developed in response to Searle's views. However, Searle was not specifically interested in mathematical phenomena as social facts, but in social phenomena such as "money, property, marriage, governments, elections, football games, cocktail parties and law courts" (Searle 1995: xii and ff.).

Using linguistic theories and tools developed in earlier work – the theory of speech acts and performatives – he grounded social ontology on intentionality and in particular collective intentionality and rule-governed behavior. He opposes social facts, depending on human activities and language, to brute facts, which are taken to be what makes human utterances true.[9] To ground the opposition between brute and social facts, Searle assumes a distinction between properties that an object inherently has (being made of wood) from properties that depend on an observer (being a screwdriver). Among the latter he identifies properties that are assigned or imposed on a natural object or artifact by a given function, acquiring a normative value (e.g., "That river is good to swim in" or "That computer is user-friendly").[10] Searle rejects any reduction of functions to causal chains and believes that whenever we talk about the function of an object, we do so because we insert the causal chain into a normative system defined by purposes, goals, and values. This holds both for non-agentive functions (causal processes to which the observer attributes a function)

[9] According to Searle's correspondentist theory of truth, an utterance is true if and only if it corresponds to the fact that makes it true. However, with Strawson, Searle agrees that facts cannot be considered as groups or agglomerations of things and that a picture theory of meaning, according to which sentences have the meanings they do because they are conventionalized pictures of facts, does not work. Facts are conditions in the world that satisfy the truth conditions of the statements.

[10] For example, to discover the function of the heart is to discover causal processes but also to place them in a teleological framework, which distinguishes hearts that function well from hearts that function poorly, depending on how they contribute to the maintenance of what we assume to be the highest value: life (Searle 1995: 14–15).

and for agentive functions, which "are matters of the use to which agents put entities," and to which belong semantic functions, that symbolize, represent, stand for, or signify something (Searle 1995: 23). The terms denoting social facts are nodes in a network of practices (owning, buying, selling, earning, paying for services, paying off debts, etc.), and in a sense the word "money" is a placeholder for the linguistic articulation of these practices (Searle 1995: 33–51).

Language is used to distinguish cooperative animal behavior from collective intentionality: it is only when a collectively recognized status is continuously attributed to a function (only when there is the imposition of a *status function*) that an institutional fact is created, and this requires language, because the linguistic label *Y* in the formula "*X* counts as *Y*" is a constitutive part of social facts (Searle 1995: 37). The role of language is crucial, and indeed it is considered as the primary institution that conditions the existence of all others. The key point is that when asserting and believing that a physical object *X* counts as *Y*, agents attribute a symbolic role to *Y*, because *Y* stands for a pattern of activities. What is required is thus at least a primitive form of symbolization. We will see in Sect. 3 how the notion of status function is further developed by Cole into that of surrogacy function.

Even if Searle does not specifically discuss the status of mathematical facts, he mentions a mathematical example suggesting, according to my reading, that mathematics' dependence on symbolism does not appear to have the same necessity as language dependence proper to social facts. Mathematical facts, e.g., the addition $371 + 248 = 619$, seem to be a thought that cannot be realized without the aid of symbols, but for Searle this is an empirical and not a logical impossibility: it depends on how we are made and is something that could change in the course of evolution. On the contrary, social facts are language dependent because symbolic mental representations are constitutive of them. Can symbols be entirely eliminated in mathematics? One could think the result of the addition operation even without expressing it linguistically (at least, if one were constituted otherwise), but there certainly are other mathematical operations where there is no prelinguistic way to represent the Y element and a symbol is needed. Many of the studies in cognitive science evoked by philosophers of mathematical practice can be taken as confirming Searle's claim. Since a part of arithmetic can be achieved by prelinguistic operations of subitizing, etc., at least a part of human arithmetic abilities are shared with animals, while for larger numbers and more complex number systems, the introduction of a symbolization becomes necessary. In both cases the role of language in the differentiation between animals and humans plays a key role.

Searle argues that neither an internalist phenomenological approach nor an external behaviorist approach can explain the functioning of institutional facts, but only an intentional approach, whether expressed in the first person or through the vocabulary of capacities, abilities, tendencies, and dispositions (Searle 1995: 2–5). Intentionality is the capacity of the mind to represent objects and states of affairs in the world other than itself. Intentionality is a capacity of complex systems of physical particles that are organized to constitute living organisms endowed with consciousness: intentional acts are either conscious or accessible at least in principle to consciousness and are guided by intentional states (beliefs, desires, etc.).

Collective intentionality is animal capacity to engage in cooperative behavior sharing intentional states. It is a primitive phenomenon, i.e., one that is not traceable to a form of individual intentionality to which shared beliefs are added.

Another relevant feature of Searle's theory is the distinction between regulative rules, which serve precisely to regulate or coordinate a phenomenon that exists prior to the creation of the rule (driving on the right) and constitutive rules, which create the possibility of a certain activity (the rules of chess) and are, unlike the former, unconventional and therefore non-arbitrary (Searle 1995: 22–28).[11]

In social ontology there are alternative theories to Searle's that do not attribute a similar role to language, to collective intentionality, to the distinction between regulative and constitutive rules, nor to the opposition between brute and social facts in the characterization of social facts and institutions. Francesco Guala, e.g., develops a perspective starting from the presumption of the importance of social institutions as instruments of economic growth and human flourishing and analyzing the mechanisms by which they arise and are maintained. Structures originate for Guala as rules that constitute a tool for coordinating human activities and are maintained in place through a system of incentives that serve to ensure that social agents follow the rules. In the book *Understanding institutions*, a key role is played by game theory and by coordination games with multiple equilibria and dilemmas of cooperation. The coordination promoted by institutions is based on the understanding of deontic powers as costs that transform individual incentives in a strategic game. Institutions are "behavioral patterns that tend to persist because individuals have no incentive to deviate from the pattern unilaterally (unless everyone else does the same)" (Guala 2016a: xxiii). Two general claims interest us here.

There is an essential difference between tokens of institutions, which are solutions to coordination problems, and types of institutions, which are characterized by the general function of facilitating coordination and cooperation (see chapter Reform). The emphasis is not on the cognitive abilities of agents, or their intentional mental states or the representational role of language, but rather on the organization of knowledge and the reasons that might have made one institution prevail over another. Institutions depend on the mind, but in a causal way, and language is one of many possible devices that can be used to send signals that people use to converge on new equilibria, creating expectations about what each agent will do in the interaction. The focus thus naturally shifts from the investigation of theories of mind to the

[11] To compare social facts and mathematical facts, it is useful to have an idea of several phenomena peculiar to social facts that Searle intends to explain with his theory (although they do not apply indiscriminately to all social facts): (1) self-referentiality (for something to fall under the concept of money, it must be believed to be used or be regarded as money, which is also expressed by saying that being a type of money is a constitutive property of money); (2) the creation by explicit performative utterances as in the example "I appoint you chairman"; (3) the existence of social facts on top of brute facts, which are logically prior; (4) the existence only in a set of systematic relations to other facts (e.g., money cannot be conceived independently of some exchange system and property ownership); (5) social objects are just the continuous possibility of a social activity, because the process is prior to the product; and (6) language is a constitutive property of social facts (Searle 1995: 33–37).

sociohistorical study of institutions, to understand the motivations and incentives proper to token institutions, which are characterized by a precise geographical and historical location. The macroscopic investigation of institutions proposed by Searle (2005) is complemented by the possibility of a microscopic investigation of institutions, which can include traffic regulation as well as marriage.[12]

Another key criticism concerns rules: Guala and Hindricks claim that all constitutive rules can be obtained from regulative rules via the introduction of theoretical terms (Guala 2016a, Chap. 5). Constitutive rules are naming devices for regulatory rules and are not constitutive of institutions: they have the function of "term-introducing principles that state the conditions of application of the theoretical terms that we use to label institutions. They are, first and foremost, naming devices for regulative rules" (Guala 2016a: xxvii). Symbolization becomes crucial when regulative rules are considered as constitutive rules, but the key elements to explain social mechanisms are only regulative rules and therefore intentionality is not presupposed.[13]

One of the ideas that Guala recognizes as primary motives for the development of his game-theoretic approach is a kind of naturalization of social ontology, aimed at showing that the dependence of social phenomena on representation can be explained in purely causal terms, without the need to introduce collective intentionality (Guala 2016b). Other theories of social ontology criticize Searle's anti-naturalism, not only to reduce the role of language as presupposition of all social interactions but also to criticize the relationship of logical primacy between individual and group. If Guala explains intentionality through the motives and interests of individuals in a strategic game, Italo Testa, reconstructing Dewey's social ontology, explains collective intentionality as a second nature generated by habits and thus as a point of arrival and not a presupposition of social interaction (Testa 2017). Indeed, the habits of individuals are explained through the habits of a group, the interactions with the interests and habits of other groups, and with the environment. Applying this perspective to the study of mathematical practices could help to better understand both the interactions between mathematical groups (schools, research traditions, scientific journals, etc.) in the constitution of mathematical facts and the role of the applicability of mathematics, whose centrality would be explained by the interactions with the environment. In this framework, a pragmatist approach prevails, but also a tendency to bridge social ontology and cognitive sciences (Testa 2020), as is also increasingly common in the philosophy of mathematics after Lakoff and Núñez (2000).

[12] The token-type distinction could be applied to mathematics, asking whether and how the number type differs from tokens related to specific number systems. Applying this theory to mathematical knowledge, it would then be a matter of explaining that some extension of the number system is non-arbitrary, not so much because it is the result of constraints imposed by status functions but because it results from a strategic interaction between agents guided by motivations, incentives, and expectations.

[13] An interesting research direction would be to compare abstraction procedures in mathematics with the way in which constitutive rules are obtained from regulatory rules through the introduction of theoretical terms.

3 The Social Constitution of Mathematical Objectivity: A Survey

3.1 Solomon Feferman's Conceptual Structuralism

In a paper on the continuum hypothesis, Feferman characterizes mathematical objectivity as a special case of intersubjective social objectivity. According to Feferman, research in mathematical practice has been dominated by structuralist views at least from Dedekind onward,[14] since all questions about zero, negative quantities, irrational magnitudes, imaginary numbers, and infinitesimal quantities are questions about the coherence of these ideas with other meaningful ideas of mathematics (Feferman 2011: 8). This view together with the idea that "the source of all mathematical thought [is] in human conceptions," forms the core of the version of structuralism proposed by Feferman and that differs from those proposed by Hellman, Resnik, Shapiro, Chihara, and Isaacson in the insistence that mathematical entities are human conceptual constructions (Feferman 2011: 10). Without summarizing all ten theses that Feferman introduces to characterize conceptual structuralism, let us bring attention to a few key points for understanding the social constitution of mathematical concepts.

All mathematical knowledge proceeds from mental concepts, but these concepts have their origin in the empirical activities of counting, ordering, matching, combining, separating, and locating in space and time, which give rise to pre-mathematical ideas of order, succession, collection, relation, rule, and operation (Feferman 2011: 10–11). From this pre-mathematical knowledge, theoretical or pure mathematics emerges through an awareness of the repeatability over time and on different objects of these activities, their interconnection into structures, and the subsequent conceptual expansion of these structures. What guarantees the objectivity of mathematics is its relative stability over time and the independent activity of different agents but also the fact that what does not withstand intersubjective critical scrutiny is abandoned (Feferman 2011: 10–11).

With social ontology Feferman shares the basic question, "How can there be an objective reality that exists in part by human agreement?" For Feferman, in fact, the objectivity of mathematics does not fail once its social constitution is explained. On the contrary, quoting Searle 1995, Feferman believes that mathematical objectivity (intended as some kind of non-arbitrariness of human conventions) is explained by the constraints that social facts and institutions impose on our choices (Feferman

[14] On the relationship between structuralism and mathematical practice, see Reck and Schiemer (2020), presenting in the first part mathematical practices related to structuralism from Grassmann to McLane. See also Cantù and Patras (2023) and Carter (2008). Carter distinguishes different ways in which structures operate in mathematics: structures are not only objects of study but also ways of studying sets initially conceived as devoid of structure, of which one investigates whether they have or can be given one or more structures. Cantù and Patras claim that structures often play a dual role in axiomatic practice: e.g., in Bourbaki (1948) the notion of structure plays the role of an epistemic object (formalized in the theory through an axiomatic definition) as well as the role of an epistemic concept (used in the architectural construction of the axiomatic framework).

2011: 12–13). While we cannot go into detail here about the construction of the structure of natural numbers and its subsequent expansions, it seems necessary to point out that in Feferman there appears, albeit in an implicit form, the idea of a scaffolding of mathematical knowledge and of different degrees of objectivity characterizing it. In the case of the sequence of integers, there is no doubt about the definite meaning of the statements describing it, while in the case of the continuum, there are distinct concepts with different degrees of structuring and axiomatization, which make the question of the definiteness of the problem much more complex. In other words, there are more degrees of arbitrariness than in the case of natural numbers, whose structure is modeled on the activity of counting (Feferman: 13–22).

Feferman explicitly acknowledges, citing Cole 2013, a similarity between mathematical concepts and other social institutions: both allow substantial intersubjective agreement and have concepts that are understood without assuming reification. But he adds that there are also differences: "What makes mathematics unique compared to institutions and games is its endless fecundity and remarkable elaboration of some basic numerical and geometrical structural concepts" (Feferman 2014). We will see how Ferreirós develops this idea of the uniqueness of mathematical concepts compared to other social institutions.

Feferman distinguishes not only between pre-mathematical and mathematical activities but also between basic mathematical activities and axiomatized mathematical activities or subject to systematic logical development (Feferman: 11).[15] The analysis of a prelogical and a basic level of mathematical knowledge was one of the central themes of Paul Lorenzen's research, although its primary goal was not to explain the objectivity of socially constituted mathematical knowledge. His two main aims were to highlight a mathematical core common to formalists and intuitionists in the dispute over foundations and to characterize a part of mathematics as operational, picking up on Dingler's views (Lorenzen 1955: 4). As we have pre-mathematical activities in Feferman, and basic skills that cannot be further analyzed in Dingler (1931),[16] we have in Lorenzen "protological" activities underlying logic and arithmetic, that is, practical faculties of operating with figures, recognizing them as equal or different, without thereby already possessing the assertion (*Behauptung*) of their equality (Lorenzen 1955: 7). A part of mathematics is in fact regarded as a theory of calculus (an idea Lorenzen derives from Curry (1951) but which he restricts to a part of mathematics only), where a calculation is not necessarily formulated in a language. The axiomatic method intervenes only at a later stage not so much to justify but to organize operational mathematics and enable a better understanding of the relationships between the various parts of "concrete" mathematics (Lorenzen 1955: 8).[17]

[15] We will return to this point in the next section, but it is important to note that axiomatization is conceived by Feferman as a formal activity that presupposes the choice of one language and one logic and whose goal is basically to transform undefined into defined statements.

[16] On Dingler see also Gödel (1932).

[17] See also Heinzmann and Wolters (2021).

A further aspect of Lorenzen's research is of interest to us here, namely, the development of propaedeutic logic as a dialogical theory of rational argumentation, in which logic emerges from *Lebenspraxis*. In particular, the truth-verification principle of simple and compound statements is intersubjective and dialogical, and the meaning of logical constants is their use in a dialogue. A proposition *Q(b)* is true when a speaker "rightly" attributes the predicate *Q* to the object denoted by *b*. Attributing "rightly" means that any other speaker who is knowledgeable (open to the ideas of the partner in the dialogue) and rational (capable of formulating discourse that is not determined only by emotions, traditions, habits) would, after appropriate verification, attribute that predicate to that object. To determine whether a compound statement is true, it is necessary to imagine a dialogue between two ideal speakers playing different and asymmetrical roles (a proposer and an opponent), set the logical rules (establishing the meaning of the logical constants) and the general dialogue rules (e.g., who starts speaking first, that the interlocutors have the floor in turn, that the proposer has to defend only the assertion made in the immediately preceding turn by the opponent, while the opponent has to defend each of his own assertions, etc.), and finally see who wins: if the proposer wins, the compound statement is true; if the opponent wins, it is false.[18]

The main reason why we discussed Lorenzen's view in the analysis of Feferman's perspective is because we think that it might well make explicit what remains implicit in Feferman's brief survey. According to Lorenzen, (1) mathematics is constructed in stages from some practical activity (and this is true not only for geometry, of which a material, empirical, semi-empirical, and abstract component had long been distinguished in relation to educational needs,[19] but also for arithmetic); (2) the main reason for isolating an operational part of mathematics is to show that although conventional, mathematical knowledge is not arbitrary; (3) the formal axiomatic method, which requires the choice of a language and a logical system, is not part of operational mathematics but serves to structure its parts; (4) formal logic itself, presented in dialogical form, reflects the intersubjective constitution of truth and shows how agreement on the truth of an utterance is the result of a dialogical and dialectical practice centered on ideal agents, who embody roles and follow rules.

3.2 José Ferreirós Interplay of Practices

José Ferreirós (2016) defines mathematical practice as an activity supported by individual and social agents and characterized by stability, reliability, and intersubjectivity. The two essential elements of the definition of practice are agents and a framework. There are two kinds of frameworks: theoretical and symbolic (Ferreirós

[18] See Kamlah and Lorenzen (1967) and Cantù and Testa (2006).

[19] See, for example, Veronese (1909) and Veronese and Gazzaniga (1895–97), discussed in Cantù (2000). Another example suggested to me by Dirk Schlimm, whom I thank for several precious comments to a preliminary version of this chapter, would be Moritz Pasch, who claimed that both arithmetic and geometry can be constructed in stages from some practical activity. See in particular Pasch 1923.

2016: 26–28).[20] Agents are individuals, but their practices are communal, so it is necessary to consider groups, schools, or entire research traditions. This notion of communities of agents shows that Ferreirós's approach is far from the idea of reducing the study of agents to the investigation of individual mental states, values and goals, but considers the historical development of schools and research traditions as a key element in defining what a practice is. Ferreirós's notion of practice differs from Kitcher's (1984) in at least two respects: it explicitly introduces the notion of agent and renounces defining practices as frameworks. Practices also include ways in which agents use frameworks and thus also exemplars (in Kuhn's sense), questions of style, and other cultural differences in the configuration of knowledge.[21] Practices can be recognized by the fact that they can be taught and learned and that there are success conditions that should be met for one to know them, which allows Ferreirós to include the notion of norms (implicit more often than explicit) and norm transmission among the characteristics that identify a practice (Ferreirós 2016: 28).[22] The approach is agent-based and cognitive, pragmatist or practice-oriented, historical, and hypothetical (Ferreirós 2016: 3). Agents are investigated through their biological and cognitive abilities, but the latter are mediated by culture, so that no established model from cognitive science is applied as such to mathematics.[23] The practices are of various kinds: practices of everyday life, technical practices, specifically mathematical practices and scientific practices. The approach is historical because it is based on the idea that historical changes in practices and their interconnections affect mathematical concepts themselves.

Mathematics progresses through conceptual diversification, and the latter is due to the transformation of networks of practices and knowledge. Whenever a practice is modified or new practices are developed, new connections are made between these practices. Since concepts are socially constituted by practices, then, whenever practices change, concepts also change. It follows that "agents who command

[20] The distinction between the two is taken to be fuzzy, but symbolic frameworks generally contain ideograms, technical expressions, and symbolic methods (symbolic frameworks need not be only formal nor only linguistic), while theoretical frameworks include statements or propositions, proofs, theoretic or proof methods, and open questions (Ferreirós 2016: 55).

[21] One of Ferreirós's goals is precisely to avoid the identification between frameworks and paradigms or disciplinary matrices à la Kuhn: that is why he claims that different frameworks can coexist in the same historical period and even in the same agent (Ferreirós 2016: 37).

[22] Among the examples cited by Ferreirós, one finds traffic regulation (Ferreirós 2016: 29) – a paradigmatic example in social ontology studies on institutions as instruments of social coordination (see Guala 2016).

[23] Indeed, the biological and cognitive component of human abilities is included – it was recalled as one way to explain the cultural invariance of mathematics in Ferreirós 2005 (p. 64) – but the historical dimension clearly prevails. For example, Ferreirós claims that the number concept is a cultural product, even if it is based on some specific cognitive abilities of the human species (Ferreirós 2016: 68). Besides, the second part of the volume on the interplay of practices is entirely dedicated to historical case studies ranging from ancient China to contemporary mathematics. Furthermore, the term cognitive is used in such a broad sense (basic abilities such as perception, memory, and language) that cultural components also enter into such abilities (Ferreirós 2016: 77).

different webs of practice and knowledge may differ in their conceptions" (Ferreirós 2022b: 101). This could lead, from a discontinuous historical perspective, to question the commensurability of concepts constituted in different networks of practices, but Ferreirós rather focuses on the possible cooperation between mathematicians with different expertise.

The hypothetical component of the approach relies on the fact that mathematics is built up from different layers, i.e., from the interplay of different kinds of practices, going from basic skills involving perception to more abstract practices of hypothetical reasoning. Some of these layers (advanced mathematics) are grounded in hypothetical elements, i.e., on a non-representational use of language.[24] This allows for different degrees of objectivity of different parts of mathematics (Ferreirós 2016: xiv). If only elementary parts of mathematics such as the theory of natural numbers can be known with certainty, advanced and hypothetical mathematics also undergo constraints, because new hypotheses are linked with previous strata of mathematical knowledge and force certain results upon agents: "objectivity can still be found in a form of knowledge that falls short of perfect certainty" (Ferreirós 2016: 161).

Despite his insistence on the origin of mathematics from the interplay of practices that are not only specifically mathematical, Ferreirós aims to ensure some kind of specificity to mathematical knowledge, rejecting a holistic approach à la Quine. Mathematical knowledge is distinguished both from scientific knowledge in general and from other forms of knowledge peculiar to the human sciences. Science is based in a different way on experiential practices (or at least on different experiments, dealing not only with symbols and diagrams but also with activities on physical objects) and makes essential use of mathematical results (particularly in the processes of conjecture formation, model formulation, and theory).[25] On this same basis, Ferreirós contrasts the objectivity of mathematics, which is robust, with the objectivity of other socially constituted forms of knowledge. Against those who identify mathematical objects with institutions, Ferreirós claims that the stability of mathematics is much stronger than that of marriage or private property: the latter are highly variable among different cultures.[26]

[24] As in the pragmatist tradition, meaning is elucidated in terms of use, and mathematical understanding in terms of elementary practices and cognitive roots, but Ferreirós tries at all costs to avoid reductionism, suggesting that it is a form of substantiation (Ferreirós 2016: 5).

[25] See Ferreirós (2016, p. 14). In other words, Ferreirós assumes that there is a difference between mathematical and physical experiments and retains a version of the indispensability thesis, even if the emphasis on the interplay between mathematical and non-mathematical practices apparently dissolves the problem of applicability.

[26] Ferreirós, unlike the enculturation tradition (see Pantsar 2023 for a review of enculturation studies), does not provide an explanation of this difference based on cognitive science. He does not attempt to explain but merely notes this difference, suggesting that historical evidence supports a cross-cultural universality of mathematics (Ferreirós 2023).

3.3 Julian Cole's Institutionalization of Mathematics

If Ferreirós develops Feferman's idea of an interplay between mathematical and pre-mathematical practices, Julian C. Cole develops another suggestion by Feferman, i.e., the effort to explain the non-arbitrariness of mathematics by analogy with the case of social facts, which impose constraints on us even though they are social constructions. Cole explicitly introduces mathematical objects as institutional facts, referring to Searle's theory of collective intentionality.

The initial impetus for investigating the social constitution of mathematical objects comes to Cole from structuralism (see Cole 2010). His goal is not so much to show that mathematics is a social activity, a result he believes has already been sufficiently proved by Hersh (1997) and Ernest (1998), but to discuss whether and how it is possible to show that mathematical domains (a general term under which structures fall) are socially constructed. Cole reprises Haslanger's (1995: 98) distinction between causal social construction (when causal factors bring into existence a social fact) and constitutive social construction (when social factors enter into the definition of a social fact) and further distinguishes the class of pure constitutive social constructs, i.e., constructs that exist wholly in virtue of the undertaking of certain acts, decisions, or practices of social significance, but do not require any kind of associated physical object. His central thesis then is that "mathematical domains (and the items of which they are composed) are pure constitutive social constructs constituted by mathematical practices" (Cole 2008: 115). The way these practices constitute mathematical domains is through a consistent characterization, culminating, as Cole notes in the case of Euclid's two-dimensional geometry, with axiomatization (Cole 2008: 117). Indeed, Cole does not deny that the characterization of a domain is the product of an individual, but he believes that it is detached from its originator by becoming an object of study for a larger community and that it is objective precisely because of the logical nature of its axiomatic characterization. Cole, however, also discusses the objectivity of mathematical entities and proposes a version of realism, called practice-dependent, which assumes that "many mathematical domains genuinely exist and have objective features, but their existence is dependent on the existence of mathematical practices that constitute them" (Cole 2008: 121). Cole even goes so far as to say that mathematical objects share many properties with abstract objects, such as acausality, non-spatiotemporality, eternality, and changelessness. Why then do we prefer practice-dependent realism to Platonism? Cole claims that the latter is explanatorily and justificationally superfluous since it is neither necessary to account for the truth value of mathematical utterances nor to construct an adequate semantics (Cole 2008: 126–127).

Cole's adherence to a form of structuralism becomes even stronger in later papers, in which he introduces a distinction between "pure" mathematical facts "that wholly concern mathematical domains" (Cole 2013: 10) and obtain in virtue of collective agreement, and other mathematical facts Cole focuses on the imposition of a particular kind of functions – representational functions – that facilitate our abilities to represent, analyze, reason about, and discover truths. To explain how they constitute social objects, Cole draws on Searle's theory of language acts (via

declarations). To explain how it is possible to account for the fact that there are, apparently, two kinds of objects, physical and symbolic, Cole takes up Amie Thomasson's solution, introducing a hybrid theory of reference and a related meta-ontological perspective. Mathematical domains are thus "freestanding institutional entities that, at least typically, are introduced to serve representation functions" (Cole 2013: 27).

The difference from nominalism or fictionalism is that an individual's commitment to a mathematical domain is sufficient to guarantee its existence and that it is maintained in existence by the collective recognition of the deontic powers attributed to that mathematical domain. Mathematics results from a successive series of layers: in the first layer mathematical entities serve to represent non-mathematical aspects of reality, in the second layer to represent elements of the first, and so on. In particular, the representation function must apply to all (external) possibilities of some kind and thus be free of modal and spatiotemporal restrictions, which would allow for the derivation of the characteristics usually attributed to mathematical entities: being abstract and atemporal. Not only that, but mathematics would also be robustly objective because representational functions are subject to greater constraints than status functions and in the case of the first layer are constrained through abstraction principles from non-mathematical aspects of reality (Cole 2013: 32).

Having thus claimed that mathematics can be objective, abstract, timeless, and even appear to be necessary, Cole believes he has refuted the main arguments against the thesis that mathematics is a social construction (Cole 2015). He also tries to account for the plurality of different theories that can perform representational functions (e.g., classical mathematics or constructive mathematics) by insisting on various possible operations of generalization and extension of a system of objects compatible with their application. Cole introduces the idea of surrogate object: "we find it significantly easier to engage in various representational activities when their subject matter is treated as an object in our representations" (Cole 2015: 1109). For example, the natural numbers are surrogates that are useful when we engage in representations of the cardinalities of finite pluralities: the representational function is thus considered as a property of an object, but this depends on collective intentionality.

Cole's problem then becomes that of answering Benacerraf's (1965) dilemma. If different collections of ordered sets serve the representational function performed by the natural numbers, how can the choice of the latter be not arbitrary? Cole answers that non-arbitrariness concerns only the core surrogacy function, which is to form an omega-sequence (Cole 2015: 1118). In other words, mathematical structures are surrogates, characterized only by positions and relations, and thus metaphysically incomplete (Cole 2017: 11). Structuralism thus becomes, I would say, a necessary correlate of social constructivism, and non-arbitrariness can be characterized only up to structural isomorphism.

In a more recent paper, the scaffolding of mathematics is considered not only through different core surrogate functions but also through different degrees of non-arbitrariness: a social construction is constrained by collective agreements, but the strength of the constraint exerted by those collective agreements depends on their

degrees of codification and explicitness and on different levels of convergence on what follows from a given semantic content. The strongly constrained semantic objectivity of mathematics is explained by a greater symbolic codification of the collective agreements on which institutional mathematical entities are founded, on a greater convergence on the interpretation of the consequences of a semantic content, and on the applications considered as relevant in a particular context (Cole 2023: 10–11).

4 The Social Constitution of Mathematical Knowledge: Key Issues

Social ontology is a branch of metaphysics that investigates various objects of study – including social groups, institutions, language, artifacts, and artworks – the components of social things, and the way they are socially constituted. It investigates not only the grounding of social facts, i.e., their constituents – psychological states of individuals, actions, and practices – but also the anchoring of social facts, i.e., the sources of social constitution: attitudes, language, and causal patterns (Epstein 2018). When analyzing a social fact or institution, one can ask what it is made of, as in mathematics we ask what the features of the concept of number are, or one can ask what makes something a social fact, as in mathematics when we ask how humans came to determine or create numbers. This distinction can be fruitfully applied to mathematical knowledge in two ways. On the one hand, one might claim that social constitution of mathematics is aimed at anchoring and thus is a complementary approach to Platonism, which, being primarily concerned with ontological issues, is instead aimed at grounding. On the other hand, one could suggest a classification of the various theories of the social constitution of mathematical knowledge as based on different groundings: cognitive states, practices, or coordination rules.[27] Indeed, Cole's and Ferreirós's proposals are respectively based on intentionality and practice, as we will claim by analyzing the role of objectivity and semantics in Sects. 4.1 and 4.2, respectively. We will then suggest how axiomatics can be seen as a social institution anchored in coordination rules in Sect. 5.

4.1 Objectivity With or Without Objects?

Ferreirós analyses the objectivity of mathematical knowledge without reference to mathematical objects or rather as an explicit inquiry on objectivity without objects, a tradition he refers to Kreisel and Putnam (Ferreirós 2016: 9). Mathematical

[27] Epstein distinguishes three main kinds of grounding: (a) cognitive states, as in the theories of Searle and Levi-Strauss; (b) practices, considered as embodied routines in the social world combined with physical actions and symbolic systems (Bourdieu) or in the interaction with social environment (Giddens and Sewell); and (c) functions, as in Millikan and Boyd's theory (Epstein 2016: 245–246).

objectivity is granted by the cognitive intersubjectivity and the applicability of mathematical knowledge. This is achieved by abandoning the referential semantics and the correspondentist theory of truth that is usually assumed for natural language. Consequently, the semantics of mathematics differs from the semantics of natural language. It is not representational, in the sense that mathematical terms do not refer to objects existing outside our minds. Mathematical hypotheses have the peculiar characteristic of performing a constitutive and non-representational function, in the sense that they constitute or lay out a mathematical domain (Ferreirós 2016: 156). And yet for Ferreirós, Platonism remains an obligatory reference, insofar as he ends up explaining (especially in the article on Feferman) how a moderate Platonic theory, e.g., the theory of thin objects (Linnebo 2018), would turn out to be compatible with his own approach.

What constrains mathematical knowledge? First, how we are made, that is, certain biological and cognitive traits proper to the human species; second, certain fundamental technical activities, such as counting, measuring, etc.; and third, inferential relations between mathematical statements and between mathematical statements and those of other sciences. It clearly appears that there are different cores that contribute to objectivity, which therefore is not an all-or-nothing affair but something emerging by degrees and for that very reason modifiable and reformable.[28]

Objectivity does not mean certainty (even if elementary mathematics, as the theory of numbers, is certain): it is compatible with the hypothetical nature of mathematics (this holds at least for advanced mathematics). Certainty is conditional (due to proofs that transmit the certainty of the premises to conclusions) but is constrained from a strong form of intersubjectivity that limits the choice of new hypotheses (depending on related practices) (Ferreirós 2016: 159–160). Historical contingency, hypotheticity, and objectivity go along together, and there is no full autonomy of mathematical practices from other scientific practices: yet holism is sweetened by a pragmatist and cognitive flavor. The focus on agents is often obscured by the primacy accorded to the analysis of texts and thus of byproducts of an activity, rather than to the activity itself and the specific characteristics of agents. Objectivity is akin to what was formerly called progress, a necessary or rational development with respect to the historical change of practices and the involvement of historical agents in them. In this sense, Ferreiros' project could also be seen as a historical epistemology, as his reference to Cavaillès also confirms.

If Ferreirós characterizes his own approach as pragmatist in the practice-oriented sense, other authors take a pragmatist approach in a proper sense, inspired by Peirce's philosophy. Such is the case with Jessica Carter, who combines a structuralist approach with a form of pragmatist realism to explain the non-arbitrariness of mathematical knowledge. Although she assumes mathematical knowledge as an interrelationship of mathematical domains Jessica Carter seems specifically interested in explaining why many mathematicians are realists and take mathematical

[28] "Mathematics is not given a priori, but is the 'science that is sought' (to abuse words of Aristotle), constantly being reshaped and reconstructed." (Ferreirós 2016: 314).

utterances at their face value, when they speak of the existence of certain mathematical objects (Carter 2005). Combining the hypothetical nature of mathematics with a form of realism, Carter argues that mathematical hypotheses do not assert the existence of abstract entities, as in ontological realism, and yet are pragmatically real because the truth of mathematical propositions is pragmatically dependent on the validity of other utterances (and especially utterances about primary substances at some prior level). If fictionalism rightly suggests ways to talk about truth without stating that the discourse is about independently existing abstract objects (we can tell stories about such objects), yet the way mathematical objects are postulated is different, because the choice of new hypotheses is severely limited by previous methods, local and global concerns (the object of one story might interact with the objects of other stories), as well as the activity of reasoning itself (Carter 2023: 211–212). The pragmatist component is strong, and in fact Carter feels no need to justify the timelessness, or abstract nature of mathematical entities nor to guarantee a degree of objectivity to mathematical entities comparable to that offered by Platonism.[29]

Cole is mainly interested in the objectivity of mathematical objects, rather than in the objectivity of our knowledge about them. The starting point is opposed to that of Ferreirós. Cole's interest in objects has a twofold origin. On the one hand, it is driven by the analogy between mathematical objects and institutions. On the other hand, it is driven by the effort to answer a typical criticism against the lack of objectivity of mathematical objects considered as social constructions. Finally, it arises from the social ontology perspective, which aims, as the term "ontology" clearly indicates, to investigate the kind of objects that are socially construed.

Let's recall Searle's insistence on the priority of brute facts over social facts and compare it with the philosophy of mathematical practice. For Searle, every imposition of a status function requires something on which this function can be imposed, and even if it is itself a status function, one must necessarily arrive at a level where the imposition occurs on a brute fact. The question is whether in the philosophy of mathematical practice there is an analogous level of natural facts (objects that are counted, extended surfaces that are measured), which are regarded as logically prior to the concepts of number or measurement or not, and whether this leads, as in Searle's case, to a form of realism.[30] This is a point developed by Carter, who offers two reasons why mathematical objects differ from social constructions in Cole's sense: they are introduced to reason with them (and this holds not only for advanced

[29] On the analysis of the relationship between the practical turn in philosophy of mathematics and developments in pragmatism as a philosophical method, a topic we cannot develop here, see Heinzmann (2015).

[30] Of particular interest for a comparison with mathematics is Searle's claim that is possible to reiterate the imposition of status functions, affirming formulas in which *X* occurs as *Y* in some preceding formula: for example, when I affirm that "only a citizen of the United States can become President." Another point he makes – and that might be relevant in a perspective, like that of Ferreirós or Carter, based on an interplay of practices – is that the persistence of a social fact depends on the possibility of continuously reactivating not only a single social activity but also its interactions with other social activities (Searle 1995: 80–81).

mathematics but also for elementary mathematics) and "can be used to tell us something about the real world of physical phenomena." Carter offers a reformulation of Peirce's principle, claiming that "the reality of mathematical entities thus depends on the truth or validity of statements concerning a more primary substance" (Carter 2023: 212–213). Comparing Cole, Carter, and Searle's approaches, it becomes clearer why realism is considered a plausible option. All three theories seem to assume the existence of physical or brute facts as the starting point for the construction of mathematical facts. Besides, in all three authors, realism is not associated with Platonism but with structuralism, even if different versions of structuralism are in play.

Another element of interest is the link that Searle establishes between the relational nature of social facts and their being placeholders for patterns of activities and thus their being primarily processes and not products (in fact, a paper ticket is money only insofar as there is continual renewal of the activity that assigns to object *X* the role *Y*). I would like to make the point that in philosophy of mathematics, in general, it is not common to identify an explicit relationship between the reasons for a structuralist view and the reasons for holding that mathematical objects are social processes. But this survey may prove otherwise. Comparing Ferreirós's explanation of the mathematical concept of number as structurally connected to a network of interrelated practices, it becomes clear why several theories of mathematical practice focusing on the social constitution of mathematical facts tend toward a form of structuralism: they consider mathematical objects as the result of social processes, and social processes are structurally related through agents' interactions.

Despite this common proximity to structuralism, three different guiding ideas should be distinguished: naturalization, historicization, and normatization of mathematical practices and mathematical agents. Many attempts to explain the objectivity of mathematics without going through the objectivity of mathematical entities are anchored in human practices, which are themselves constrained by biological and cognitive features of agents (Ferreirós, Pantsar, Giardino), by semantic relations between utterances (Carter, Collin), by layered collective agreements in which each level produces constraints on subsequent levels (Cole), by the function of representing certain aspects of reality (Cole, Carter) or certain intersubjective practices (Ferreirós, Giardino), and by cultural and historical factors (Ferreirós, Hartimo, and Rytilä).

Markus Pantsar (2023) opts for an approach aimed at the naturalization of mathematical practice and appeals to empirical evidence. He claims that neither Cole's nor Ferreirós's account explains objectivity, but only biologically determined proto-arithmetical abilities can account for the intercultural stability of mathematics.

Valeria Giardino offers a good example of two different but complementary strategies for providing a normative framework for the philosophy of mathematical practice. The goal is to combine a primarily descriptive theory – Hutchins' distributed cognition – with a normative semantic theory: Brandom's inferentialism. Mathematics can thus be described not only as a symbolic human activity, distributed across a community of agents who interact with each other and with the environment, but also as a specifically symbolic activity based on an interplay of actions that give rise to a dialectic of entitlements and commitments in the game of

giving and asking for reasons (Giardino 2023: 261–264). Giardino takes naturalism as an inescapable starting point, at least insofar as it allows us to explain how human practices are anchored in cognitive attitudes common to other nonhuman animals and not necessarily linguistic. Mathematical reasoning is know-how, but it is not limited to giving and asking for reasons linguistically: it also includes the ability to manipulate material objects that serve as anchors for conceptual models (Giardino 2018a). While Ferreirós proposes a historicization of agents, Giardino simultaneously offers a naturalization and normatization of them.

The idea that Brandom's semantic inferentialism can be developed to account for mathematical facts as social facts is also developed by James Henry Collin (2023), who uses it primarily to explain the objectivity of mathematical knowledge. Collin proceeds very differently, however, because his goal is to show that semantic inferentialism, although it is based on norms and not on mental states, can account not only for the semantic content of utterances through inferential actions, for their subsentential components through substitution and anaphoric actions – as proposed by Brandom – but also for the way semantic content is socially instituted. There is normatization but no naturalization in Collin's view. Social facts and mathematical facts differ, because in the latter case, there is no physical object to which a normative status can be attributed. In mathematics this function is performed by the principles of abstraction, which play the role of rules of introduction and elimination. Collin thus succeeds in accounting for the abstract and objective nature of mathematics, without, however, presupposing the existence of abstract entities but rather accounting for the object-directedness of human discourse through human actions of deontic scorekeeping.

Mirja Hartimo and Jenni Rytilä make a different proposal, which combines a realistic conception of mathematical objects – viewed as independent of individual mathematical practices – with the idea that they are "internal to the sociohistorically generated mathematical practice." They combine a Husserlian phenomenological approach to the investigation of mathematical practice with theories on the social constitution of mathematical objects (Hartimo and Rytilä 2023). The phenomenological approach is aimed at capturing "the mathematicians' view of the purpose of their practices," but the latter, being exclusively internal to mathematical practice itself, is not sufficient to provide a complete picture of the ontology of mathematics, a picture that is then sought in the social construction of mathematical objects. They claim that the social constitution of mathematical practice not only explains the phenomenological appearance of the object as transcendent and separate but also accounts for the pluralistic view of phenomenology by considering historical and social factors that can explain the variety of methods used in mathematics.

4.2 A Representational or Non-representational Semantics?

Another point we would like to make is that three different trends emerge in the discussion of what is the semantics of mathematics. All three have to do not so much with grounding, to use an expression dear to Epstein, but with anchoring. What

explains how a certain term came to take on a certain meaning? For Cole it is clearly a form of collective intentionality; for inferentialism (Collin, Giardino) it is through mechanisms of anaphora, etc., that the meaning of terms is fixed in relation to the meaning of other terms; and for Ferreirós and Carter, intersubjective practices and mathematical domains interact with each other, giving rise to rules that constitute the meaning of mathematical terms themselves. One wonders, then, in which approaches mathematical utterances are really taken at face value and which approaches require a radical modification of semantics. To explain social facts as the result of collective intentional processes of imposing functions, Searle takes three elements as relevant: the imposition of functions, collective intentionality, and constitutive rules. As we will see in the brief survey that follows, these three elements are based on the key role of language and on a correspondentist semantics.

While Searle maintains a representationalist semantics for both linguistic terms denoting natural facts and linguistic terms denoting social facts, many philosophers of mathematics who explain the social constitution of mathematical knowledge rely on a non-correspondentist semantics and do not necessarily speak of social facts (see, e.g., inferentialist semantics above). Cole expands Thomasson's semantics, ensuring that objects falling under a given concept are determined by the coapplication and application criteria related to representational functions. The semantic content is determined by the functions imposed on social entities (or rather collective agreements – e.g., on what is a number and on what land mass is the territory of the United States – are partially responsible for the truth value of the statement). However, thanks to the notion of surrogate introduced in later writings, Cole also accounts for the origin of a representational semantics (based on objects and properties), which can be justified and retained for mathematical language.

Ferreirós aims at granting objectivity to mathematical knowledge without grounding it on mathematical objects and explicitly assumes that mathematical language, unlike natural language, has a non-referential semantics, where terms constitute rather than represent objects: mathematical hypotheses are "constitutive and not representational, in the sense that they constitute or lay out a mathematical domain" (Ferreiros 2016: 156). Yet, it is not strictly necessary to oppose natural language and mathematical language: adopting the Brandomian theory of semantic inferentialism, Collin (2023) adopts the same semantics for natural language and mathematical language, focusing on an activity rather than an object domain. If, however, in Ferreirós the focus is on activities that may also be prelinguistic, such as counting, measuring, etc., in Collin all activities are linguistic. One question that remains open is to what extent naturalization and historicization are necessary to account for all prelinguistic aspects – including not only different varieties of iconicity (Giardino and Greenberg 2015; Schlimm 2021) but also different types of manipulation with artifacts, who seem to be excluded from exclusively normative approaches.

Although the positions of the different authors mentioned above differ in their choice of the most appropriate semantics for mathematics, authors such as Searle or Cole, who base the social constitution of mathematical facts on mental states, tend to maintain a representationalist semantics. In contrast, authors who seek to take into

account cultural and historical factors and interrelated mathematical practices seem more sensitive to the need to formulate a non-representationalist semantics for mathematics.

5 Axiomatics as a Social Mathematical Practice

The goal of Feferman's paper was to explain why and in what formulation the continuum problem can be considered as a definite statement (i.e., liable to be true or false). In answering this question, it seems to us that Feferman also implicitly brought to the table a new challenge to understanding mathematical practice in social perspective: what can be the role of axiomatics in socially constituted mathematical knowledge? If language or at least symbolic activity seems to be essential in all the authors discussed, it would be interesting to see if it is possible to develop a different analysis of axiomatics – often considered the quintessential mathematical formal activity – an analysis in which neither language nor symbolization occupies the stage, but the notions of rule and function come to the forefront.

Axiomatics as the adoption of a method based on the distinction between propositions assumed without proof and propositions in need of a proof has been regarded as one of the leading activities of mathematicians at least from Euclid's *Elements* onward, even if the role and meaning of axiomatics changed with time. Axiomatics is nowadays often considered to be the main aim of formal logic, a way to demarcate formal from social sciences, and especially a means to justify or ground the truth of the former. What we want to understand in this section is whether axiomatic objects (axioms, theorems, rules of inference) are described differently when studied as the result of intersubjective practices.[31] Analyzing axiomatics as a social institution, does it suggest new roles for axiomatics alongside the justificatory one?

Recent literature has shown that axiomatics need not only have a justificatory role but might have a variety of objectives. Sieg distinguished between formal and structural axiomatics in Hilbert's works (Sieg 2014). Schlimm (2013: 39) distinguished a descriptive from a normative function and claimed that the same set of axioms can be employed in different roles (semantic or syntactic). Cantù (2022: 6–12) suggested a variety of objectives of axiomatics: heuristics, conceptual analysis, quest for rigor, modular analysis, coordination tool, and descriptive, genetic-historical, architectural, and pedagogical aim. Ferreirós distinguished between a

[31] First, one might ask why one should use the tools of social ontology to investigate axiomatics in logic and mathematics. The chosen ontology might have consequences on the kind of problems and puzzles one must solve in the development of a given science. Taking axiomatics to be a method or a system is not the same as taking it to be a mode of discourse or symbolic order; an informal institution based on norms, conventions, customs, laws, organizations, groups, and roles; or an entity sustained in existence by collective recognition and serving status functions. The explanation of axiomatics' persistence in time has been partly investigated by an analysis of the function of the axiomatic method, and eventually of its changes in time, but has not been related to an analysis of its specific constituents.

formal-logical side to axiomatics, attentive to the inferential structure of a theory, and a conceptual-mathematical side, aiming to make new connections in mathematics (Ferreirós 2022b). In the survey of the philosophies of mathematical practices centered on the social constitution of mathematical knowledge and objects, formal and informal axiomatics played the role of adding constraints on the arbitrariness of social constructions. For example, axiomatic symbolization added further degrees of codification in Cole, making the objectivity of mathematics more robust. So, in a sense, axiomatics was regarded primarily as a symbolic activity, and its role appeared to be mainly related to increasing mathematical objectivity, and thus, at least in a broad sense, as a justificatory activity. The question is whether considering axiomatics in light of a different theory of social ontology, one that does not presuppose language and intentionality as key elements, might lead to different results.

The question is whether the philosophy of mathematical practice, framed in a normative framework such as that of Guala's rather than Searle's social ontology, (a) explains the intersubjective and social origin of axiomatics and (b) accounts for it as the result of human interactions producing rules, obligations, and coordination functions.

There are some very simple problems of coordination that axiomatics facilitates: (a) it avoids proving the same theorem again; (b) it offers a common framework to compare different proofs of the same theorem; (c) it facilitates the interaction and the coordination between different researchers, research teams, and research projects in the proof of new results; and (d) it helps organizing scattered results in a whole. Axiomatics – as the social institutions studied by Hindriks and Guala (2015) – induces several obligations, which are independent from selfish desires, but probably not from any desire whatsoever. Examples of general obligations are consistency (it is undesirable that a statement and its negation are axioms or theorems of a system) and completeness in Hilbert's geometrical sense (it is desirable that the axiom system has an adequate level of generality). Examples of specific norms are logical rules (the quantifiers introduction and elimination rules, the modus ponens, the argument by transitivity) and some specific rules such as the law of induction in arithmetic. Yet, axiomatics should have other constituents than functions, obligations, and rules in order to be a social institution: social agents and social roles.

Let's briefly consider two examples. One is Lorenzen's dialogical logic and the other is Peano's *Formulario*.

In Lorenzen's logic, and more generally in dialogical logic (Hintikka's game-theoretic approach is another good example)[32], the truth of an elementary sentence is nothing else but the willingness of an expert and rational interlocutor to accept that the predicate holds for a given object. Similarly, the truth of a complex sentence is determined by a dialogue in which each participant has a specific role, that of proponent or that of opponent, and each role is in turn determined and specified

[32] Even if Lorenzen's operative approach is non-axiomatic, it is "compatible with the axiomatic approach in its non-fundamentalist interpretation" (Heinzmann 2021: 15).

by a set of rules, some of them being logical (determining the meaning of logical constants as their possible use in a dialogue) and some of them being rather dialogical (the proponent starts talking, then participants have the word in turn; the proponent must defend only the last assertion made by the opponent, whereas the opponent has to defend all her assertions, because she can assert elementary sentences, whereas the proponent cannot). Logical and dialogical rules are thus abstraction of roles played by people (nowadays also by computer agents). What is said here about logic can be extended to argumentation theory, where the roles of the participants in a dialogue are regulated by less rigid norms, some of which might be conventions, habits, and actions for which people have incentives, because they often obtain general acceptance (argument by analogy, appeal to the expert opinion).[33]

Teaching and education offer other interesting examples of how axiomatics, and the logics on which it is based, are taught and developed in school, thanks to targeted incentives that reward students who better apply the logical and argumentation rules (see the rating grid items, which usually include logical structuring, logical formulation, or logical order as key competences). Further examples can be driven from mathematics education, e.g., the independency of a system of axioms, which is sometimes considered as a rule in formal axiomatic systems (an axiom should not be a theorem that can be inferred from the other axioms), but is certainly not a rule in teaching, where theorems that are too long or difficult to prove with respect to the level of the students are simply assumed as postulates. Further rules and roles emerge from recent investigations of proofs as planning activities (see, e.g., Bratman 2013, Hamami and Morris 2020a) and from inferentialist analyses of language (see, e.g., Brandom 1998 and Collin 2023).

I come now to a second example: Peano's *Formulario*, i.e., an interactive enterprise aiming at constructing a logico-mathematical dictionary that should contain all sentences used in logic and mathematics. The enterprise was at the same time descriptive and prescriptive. Peano published a call for contributions in his journal *Revue des Mathématiques*, asking readers and researchers to contribute to the project (see Peano 1892: 76). Peano held the idea that one should not enumerate a list of derivation rules, but rather first collect all sentences written in mathematical language, order them in axiomatic form, and then investigate the derivation rules thereby used (Cantù 2022).[34] This anatomic investigation underlines how rules come out of mathematical practices and techniques, because they bring some advantage to the coordination, organization, and division of labor. Vailati, for example, considered the normativity of mathematical logic as based in its increased efficiency in the achievement of mathematical results, even if the costs required by the substitution of

[33] A good example is Walton's *New Dialectic* (1998).

[34] Similarly, Otto Hölder (1924) considers that mathematical logic should be a normative logical theory that is derived from the investigation of mathematics itself, as in Lambert's unachieved anatomy of mathematics.

old equipment with new might for some time counterbalance it, just as the change of machinery is a costly way to rationalize production (Arrighi et al. 2010: xlvii).

To summarize, axiomatics can be viewed as based on several abstract roles: proponent and opponent in a dialogue logic, axiom and theorem in an axiomatic system. Instead of considering axioms and theorems, or claims and counterclaims as sentences, they are considered as dialectic roles codified by sets of rules and strategic moves obtained by generalization of behavioral rules of human participants in an argumentative game.[35] Axiomatics can be conceived as an institution along the lines of Guala's social theory: institutions are not the result of some act of collective acceptance that assign them status functions, but rather the result of a coordination device, whose rules need not be constitutive in Searle's sense. Axiomatics need not play a justificatory role but can be viewed as a coordination tool. Indeed, there are several behavioral rules in mathematics that can hardly be defined or explained according to justificatory criteria. For example, simplicity, be it syntactic, semantic, or epistemic parsimony, has been analyzed both as a virtue of axiomatic theories and as a virtue of mathematicians (Inglis and Aberdein 2015; Aberdein et al. 2021) but might profit from an analysis as social strategy.

6 Conclusion

The philosophy of mathematical practice investigates the social constitution of mathematics but does not always make explicit the philosophical-normative framework that guides the discussion. This chapter investigated some recent proposals in the philosophy of mathematical practice that compare social facts and mathematical objects, discussing similarities and differences. An attempt has been made to enucleate, through a comparison with different perspectives in social ontology, the kind of objectivity attributed to mathematical knowledge, the type of representational or non-representational semantics adopted, and the justificatory or coordinative role entrusted to axiomatics. After a brief introduction to key issues in social ontology, Sect. 3 of the chapter offered a survey of contributions by Feferman, Lorenzen, Ferreirós, and Cole, highlighting a difference between approaches based on rules, practices, and intentional states. Section 4.1 also discussed results by Carter, Giardino, Pantsar, and Collin, focusing on differences between the objectivity of knowledge and objectivity of objects and showing that many authors combine a realist, structuralist, and pragmatist perspective, as well as the idea that objectivity comes in degrees. Section 4.2 focused on the kind of semantics that is adopted in socially oriented approaches. A representational semantics is preferred by authors grounding their views on mental states, whereas a non-representational semantics best fits with views based on practices. Section 5 considered how axiomatics can be understood in a social perspective. Axiomatics generally plays

[35] For a historical and philosophical survey of dialogical logic and argumentation theory, see Cantù and Testa (2006). On dialogical deduction see Dutilh Novaes (2020).

a justificatory role also in theories that aim to explain the social constitution of mathematical knowledge. Yet, if one shifts from approaches based on mental states and practices to one anchored on rules, axiomatics can be viewed as an institution based on rules, obligations, functions, coordination problems, and agent's actions and roles, thus playing an organizational and coordination role.

This chapter can in no way be regarded as an exhaustive survey of the new results that bridges between social ontology and mathematical practice could lead to. Interesting research directions that we could not discuss here include (1) the role of intentionality, language, and rules but also of epistemology in the constitution of mathematical social objectivity;[36] (2) an inquiry of collective agency and planification;[37] (3) different ways of analyzing mathematical agents: through a naturalization of mathematical practice that requires the study of cognitive and epistemic abilities of agents, through a historical and comparative study of the development of mathematical abilities in different cultures, or through a normative approach, which considers agents as dialectical roles; (4) a comparison between the symbolic and diagrammatic dimensions of mathematics and social theories studying the origin of human artifacts;[38] (5) a comparison between experimental approaches to mathematics and experimental methodologies in use in social disciplines;[39] (6) whether there is a distinct place assigned to mathematical kinds within the distinction between natural and social kinds when building a bridge between social ontology and mathematical practice;[40] (7) whether intersubjective mathematical objectivity is necessarily connected to a structuralist position or can be compatible with Platonism, logicism, intuitionism, and naturalism; (8) whether axiomatic knowledge can be explained without reference to mathematical structures as abstract objects, given that the objectivity of mathematical knowledge can be explained without reference to mathematical objects; and (9) whether there would be any difference in grounding intersubjective mathematical objectivity on intentions (phenomenological, planned, or shared intentions), on rules, or on habits.

[36] For an epistemological approach to mathematics, see, e.g., De Toffoli (chapter ▶ "The Epistemological Subject(s) of Mathematics," this volume).

[37] See, e.g., Hamami (chapter ▶ "Agency in Mathematical Practice," this volume) and Hamami and Morris (2020a).

[38] See, e.g., Waszek (chapter ▶ "Signs as a Theme in the Philosophy of Mathematical Practice," this volume), Schlimm (2018), Schlimm (2021a, b), and Dutz and Schlimm (2021) on notation. See Giardino (2023), Giardino (2018b), and Toffoli and Giardino (2014) on diagrams and artifacts.

[39] See, e.g., Tanswell and Inglis (chapter ▶ "The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts," this volume), Schlimm and González (chapter ▶ "Mathematical Experiments on Paper and Computer," this volume), van Bendegem JP (chapter ▶ "Experiments in Mathematics: Fact, Fiction, or the Future?," this volume), and more generally the section on experimental mathematics in this volume.

[40] Cole (2023) and Epstein (2014).

7 Cross-References

- ▶ Agency in Mathematical Practice
- ▶ Experiments in Mathematics: Fact, Fiction, or the Future?
- ▶ Inverting Hierarchies: The Sociology of Mathematical Practice
- ▶ Mathematical Experiments on Paper and Computer
- ▶ Signs as a Theme in the Philosophy of Mathematical Practice
- ▶ The Epistemological Subject(s) of Mathematics
- ▶ The Ethics of Mathematical Practice
- ▶ The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts

Acknowledgments I would like to thank Italo Testa, because without our discussions the project of comparing social ontology and mathematical practice would never have come into being, and also the institutions that made it possible to organize two workshops on the topic (the Faculty of Arts at the University of Aix-Marseille and the Department of Humanities at the University of Parma), the participants in these workshops, two referees very generous with valuable advice, and numerous colleagues (especially Dirk Schlimm and Jessica Carter) who carefully read a preliminary version of this chapter.

References

Aberdein A, Rittberg CJ, Tanswell FS (2021) Virtue theory of mathematical practices: an introduction. Synthese 199(3–4):10167–10180

Arrighi C, Cantù P, de Zan M, Suppes P (eds) (2010) Logic and pragmatism: selected essays by Giovanni Vailati. Center for the Study of Language and Information, Stanford

Aspray W, Kitcher P (1988) History and philosophy of modern mathematics. University of Minnesota Press, Minneapolis

Benacerraf P (1965) What numbers could not be. Philos Rev 74:47–73

Bourbaki N (1948) L'architecture des mathématiques. In: Le Lionnais F (ed) Les grands courants de la pensée mathématique. Cahiers du Sud, Paris

Brandom RB (1998) Making it explicit reasoning, representing, and discursive commitment. Harvard University Press, Cambridge, MA

Bratman ME (2013) Shared agency: a planning theory of acting together. Oxford University Press, New York

Cantù P (2000) L'insegnamento della geometria nelle scuole medie inferiori. Una lettera inedita di Giuseppe Veronese a Giovanni Vailati. Il Voltaire 5:109–118

Cantù P (2021) Peano's philosophical views between structuralism and logicism, Chapter 9. In: Boccuni F, Sereni A (eds) Origins and varieties of logicism. A foundational journey in the philosophy of mathematics. Routledge, New York

Cantù P (2022) What is axiomatics? Ann Math Philos 1(1):1–24

Cantù P, Patras F (2023) Les structures bourbakistes: objets ou concepts épistémiques ? Philos Sci 27(2):3–29

Cantù P, Testa I (2006) Teorie dell'argomentazione. Un'introduzione storico- filosofica alle logiche del dialogo. Testi e Pretesti. Bruno Mondadori, Milano

Cantù P, Testa I (2023) Introduction: from social ontology to mathematical practice, and back again. Topoi 42(1):187–198

Carter J (2005) Motivations for realism in the light of mathematical practice. Croat J Philos 5(1):17–29

Carter J (2008) Structuralism as a philosophy of mathematical practice. Synthese 163(2):119–131

Carter J (2019) Philosophy of mathematical practice – motivations. Themes Prospects Philos Math 27(1):1–32
Carter J (2023) Mathematical practice, fictionalism and social ontology. Topoi 42(1):211–220
Cole JC (2008) Mathematical domains: social constructs? In: Gold B, Simons R (eds) Proof and other dilemmas: mathematics and philosophy. Mathematics Association of America, Washington, DC, pp 109–128
Cole JC (2010) Mathematical structuralism today. Philos Compass 5(8):689–699
Cole JC (2013) Towards an institutional account of the objectivity, necessity, and atemporality of mathematics. Philos Math 21(1):9–36
Cole JC (2015) Social construction, mathematics, and the collective imposition of function onto reality. Erkenntnis 80(6):1101–1124
Cole J (2017) Humanism about abstract objects. In: Sriraman B (ed) Humanizing mathematics and its philosophy: essays celebrating the 90th birthday of Reuben Hersh. Birkhäuser, Cham, pp 151–165
Cole JC (2023) Some preliminary notes on the objectivity of mathematics. Topoi 42(1):235–245
Collin JH (2023) Of marriage and mathematics: Inferentialism and social ontology. Topoi 42(1): 247–257
Curry H (1951) Outlines of a formalist philosophy of mathematic. North-Holland Publishing Company, Amsterdam
De Toffoli S, Giardino V (2014) Forms and roles of diagrams in knot theory. Erkenntnis 79(4): 829–842
Dingler H (1931) Philosophie der Logik und Arithmetik. Reinhardt, München
Dutilh Novaes C (2020) The dialogical roots of deduction: historical, cognitive, and philosophical perspectives on reasoning. Cambridge University Press, Cambridge
Dutz J, Schlimm D (2021) Babbage's guidelines for the design of mathematical notations. Stud Hist Phil Sci 88:92–101
Epstein B (2014) How many kinds of glue hold the social world together? In: Galloti M, Michael J (eds) Social ontology and social cognition. Springer, Dordrecht
Epstein B (2016) Social ontology. In: McIntyre AR (ed) The Routledge companion to philosophy of social science. Routledge philosophy companions. Routledge, New York, pp 240–253
Epstein B (2018) Social ontology. In Zalta EN (ed) The stanford encyclopedia of philosophy. Metaphysics Research Lab, Stanford University
Ernest P (1998) Social constructivism as a philosophy of mathematics. State University of New York Press, New York
Feferman S (2011) Is the continuum hypothesis a definite mathematical problem? Manuscript available at http://logic.harvard.edu/EFI_Feferman_IsCHdefinite.pdf
Feferman S (2014) Logic, mathematics, and conceptual structuralism. In: Rush P (ed) The metaphysics of logic. Cambridge University Press, Cambridge, UK, pp 72–92
Ferreirós J (2005) Peano and Dedekind. In Grattan-Guinness I (ed) Landmark writings in western mathematics 1640–1940. Elsevier, Amsterdam, San Diego, Oxford and London, pp 613–626
Ferreirós J (2016) Mathematical knowledge and the interplay of practices. Princeton University Press, Princeton
Ferreirós J (2022) Conceptual structuralism. J Gen Philos Sci 54:125
Ferreirós J (2022b) The two sides of modern axiomatics: Dedekind and Peano, Hilbert and Bourbaki. In: Ferreira F, Kahle R, Sommaruga G (eds) Axiomatic thinking I. Springer, Cham, pp 83–103
Ferreirós J (2023) Degrees of objectivity? Mathemata and social objects. Topoi 42(1):199–209
Gandon S (2013) Quelle philosophie pour quelle mathématique ? Arch Philos 76(2):197–216
Giardino V (2017) The practical turn in philosophy of mathematics: a portrait of a young discipline. Philos Mind 12:18–28
Giardino V (2018a) Tools for thought: the case of mathematics. Endeavour 2(42):172–179
Giardino V (2018b) Manipulative imagination: how to move things around in mathematics. Theoria 33(2):345–360

Giardino V (2023) The practice of mathematics: cognitive resources and conceptual content. Topoi 42(1):259–270
Giardino V, Greenberg G (2015) Introduction: varieties of iconicity. Rev Philos Psychol 6(1):1–25
Gödel K (1932) Review of Dingler 1931. Monatshefte Math Phys 89:3
Guala F (2016a) Understanding institutions. Princeton University Press, Princeton
Guala F (2016b) Naturalism and anti-naturalism in the philosophy of social science. In: Humphreys P (ed) The Oxford handbook of philosophy of science. Oxford University Press, New York, pp 43–61
Hamami Y, Morris R (2020a) Plans and planning in mathematical proofs. Rev Symb Log 14(4): 1030–1065
Hamami Y, Morris RL (2020b) Philosophy of mathematical practice: a primer for mathematics educators. ZDM 52:1113–1112. https://doi.org/10.1007/s11858-020-01159-5
Hartimo M, Rytilä J (2023) No magic: from phenomenology of practice to social ontology of mathematics. Topoi 42(1):283–295
Haslanger S (1995) Ontology and social construction. Philos Top 23(2):95–125
Heinzmann G (2015) Pragmatism and the practical turn in philosophy of mathematics: explanatory proofs. In: Agazzi E, Heinzmann G (eds) The practical turn in philosophy of science. Franco Angeli, Milan, pp 113–129
Heinzmann G (2021) Operation and predicativity: Lorenzen's approach to arithmetic. In Heinzmann G, Wolters G (eds) Paul Lorenzen – mathematician and logician, logic, epistemology, and the unity of science. Springer International Publishing, pp 11–22
Heinzmann G, Wolters G (eds) (2021) Paul Lorenzen – mathematician and logician. Springer International Publishing, Cham
Hersh R (1997) What is mathematics, really? Oxford University Press, New York
Hindriks F, Guala F (2015) Institutions, rules, and equilibria: a unified theory. J Institutional Econ 11(3):459–480
Hölder O (1924) Die mathematische Methode. Springer, Berlin
Inglis M, Aberdein A (2015) Beauty is not simplicity: an analysis of mathematicians' proof appraisals. Philos Math 23(1):87–109
Kamlah W, Lorenzen P (1967) Logische Propädeutik: Vorschule des vernünftigen Redens, 3rd 1996 ed. Springer. Stuttgart-Weimar
Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press, New York
Lakoff G, Núñez RE (2000) Where mathematics comes from how the embodied mind brings mathematics into being. Basic Books, New York
Linnebo Ø (2018) Thin objects: an abstractionist account. Oxford University Press, Oxford
Lorenzen P (1955) Einführung in die operative Logik und Mathematik. Springer, Berlin
Mancosu P (ed) (2008) The philosophy of mathematical practice. Oxford University Press, Oxford
Pantsar M (2023) From maximal intersubjectivity to objectivity: an argument from the development of arithmetical cognition. Topoi 42(1):271–281
Peano G (1892) Sopra la raccolta di formule di matematica. Rivista di Matematica 2:76–77
Reck EH, Schiemer G (eds) (2020) The prehistory of mathematical structuralism. Oxford University Press, New York
Schlimm D (2013) Axioms in mathematical practice. Philos Math 21(1):37–92
Schlimm D (2018) On Frege's Begriffsschrift notation for propositional logic: design principles and trade-offs. Hist Philos Logic 39(1):53–79
Schlimm D (2021a) Peano's dot notation and design principles for notation. In The Peano school: epistemology: logic, epistemology and didactics, Special Issue ed. by P. Cantù and E. Luciano, Philosophiae Scientiae 25(1):95–126
Schlimm D (2021b) How can numerals be iconic? more varieties of iconicity. In Basu A, Stapleton G, Linker S, Legg C, Manalo E, Viana P (eds) Diagrammatic representation and inference. 12th International Conference, Diagrams 2021, Virtual, Proceedings. Springer, pp 520–528
Searle JR (1995) The construction of social reality. Free Press, New York
Searle JR (2005) What is an institution? J Inst Econ 1(1):1–22

Sieg W (2014) The ways of Hilbert's axiomatics: structural and formal. Perspect Sci 22(1):133–157
Sundholm G (2002) What is an expression. In: LOGICA yearbook 2001. Filosofia Publishers, Czech Academy of Science, Prague, pp 181–194
Testa I (2017) Dewey's social ontology. A pragmatist alternative to Searle's approach to social reality. Int J Philos Stud 25(1):40–62
Testa I (2020) A habit ontology for cognitive and social sciences methodological individualism, pragmatist interactionism, and 4e cognition. In: Caruana F, Testa I (eds) Habits: pragmatist approaches from cognitive science, neuroscience, and social theory. Cambridge University Press, Cambridge, pp 269–295
van Bendegem JP (2016) Philosophy of mathematical practice: what is it all about? In: Paul E (ed) The philosophy of mathematics education. Springer, New York, pp 13–18
van Bendegem JP (2023) Experiments in mathematics: fact, fiction or the future? In Sriraman B (ed) Handbook of the History and Philosophy of Mathematical Practice, Springer
Veronese G (1909) Elementi di geometria intuitiva ad uso delle scuole tecniche. Drücker, Padova
Veronese G, Gazzaniga P (1895–97) Elementi di Geometria, ad uso dei licei e degli istituti tecnici (primo biennio). Drucker, Verona/Padova
Walton DN (1998) The new dialectic: conversational contexts of argument. University of Toronto Press, Toronto

The Epistemological Subject(s) of Mathematics

Silvia De Toffoli

Contents

1 Introduction 2880
2 With or Without Subjects? 2884
2.1 The Epistemology Without a Subject 2885
2.2 Ideal Rationality 2886
2.3 Non-Ideal Rationality 2888
2.4 Social Epistemology 2889
2.5 Feminist Epistemology 2890
2.6 Subjects or Agents? 2892
3 The Analysis of Knowledge 2894
3.1 True Propositions 2895
3.2 Justification 2899
4 Conclusion 2901
References 2902

Abstract

Paying attention to the inner workings of mathematicians has led to a proliferation of new themes in the philosophy of mathematics. Several of these have to do with epistemology. Philosophers of mathematical practice, however, have not (yet) systematically engaged with general (analytic) epistemology. To be sure, there are some exceptions, but they are few and far between. In this chapter, I offer an explanation of why this might be the case and show how the situation could be remedied. I contend that it is only by conceiving the knowing subject(s) as embodied, fallible, and embedded in a specific context (along the lines of what has been done within social and feminist epistemology) that we can pursue an epistemology of mathematics sensitive to actual mathematical practice. I further suggest that this reconception of the knowing subject(s) does not force us to abandon the traditional framework of epistemology in which knowledge requires

S. De Toffoli (✉)
Scuola Universitaria Superiore IUSS Pavia, Pavia, Italy
e-mail: silvia.detoffoli@iusspavia.it

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_51

justified true belief. It does, however, lead to a fallible conception of mathematical justification that, among other things, makes Gettier cases possible. This shows that topics considered to be far removed from the interests of philosophers of mathematical practice might reveal to be relevant to them.

Keywords

Mathematical justification · Mathematical practice · Social epistemology · Feminist epistemology

1 Introduction

In a first analysis, it might seem that contemporary analytic epistemology[1] cannot be fruitfully applied to the philosophy of mathemtical practice. In this chapter, I suggest that this first impression is wrong-headed and that a systematic exchange between the two disciplines would be fruitful.[2]

Applying epistemology to mathematical practice might be thought as problematic for two reasons. First, the standard apparatus of general epistemology, in which (propositional) knowledge involves justification, truth, and belief, has been deemed to be overly narrow for capturing interesting features of mathematical practice.[3] Second, the application of epistemology to mathematics has been traditionally guided by the desire to eliminate the knowing subject as much as possible in order to get an objective, logical analysis of mathematical knowledge. Ideally, the resulting analysis would then be independent of actual mathematical practice.[4]

Although they might appear to be unrelated, these two reasons are connected. As I shall explain, it is only by reconceiving the knowing subject(s) of mathematics as embodied, fallible subjects embedded in specific mathematical contexts that the traditional analysis of knowledge (appropriately developed along the lines traced by social and feminist epistemologists) can be fruitfully applied to mathematical practice.

Let me start with some platitudes about general epistemology to bring out why its application to mathematical practice might seem uncalled for. The almost exclusive focus of epistemologists is directed toward *propositional attitudes,* that is, attitudes like *believing* and *knowing* that are related to propositions. Beliefs attach to propositions,

[1] Throughout the chapter, I will use *epistemology* to mean *analytic epistemology*. This clarification is needed in a context like the one of the present volume, which assembles heterogenous contributions belonging to different traditions. As I use the term, analytic epistemology is the theory of knowledge (and justification) in the analytic philosophical tradition of the Anglo-American world – a very different topic compared, for example, with *épistémologie* in the French tradition, which has to do with the critical and historical study of science and scientific knowledge.

[2] To be sure, there are significant exchanges already, albeit not systematic ones – I will turn to some of those in the conclusion. See, for example, Azzouni (2006), Easwaran (2015), and Tanswell and Kidd (2020).

[3] Lakatos (1976), Kitcher (1984), and Ferreirós (2016).

[4] See, for example, the discussion in Kitcher (1992).

and propositional knowledge is *knowledge that* such-and-such is the case, where this such-and-such is a proposition. Typical questions are: When is a belief that p (where p is a proposition) justified? When does it constitute knowledge? This seemingly innocuous starting point already clashes with a widespread commitment of philosophers interested in scientific and mathematical practice. This commitment is voiced explicitly by Reviel Netz in the context of his analysis of ancient Greek geometry:

> [W]hat unites a scientific community need not be a set of beliefs. Shared beliefs are much less common than shared practices. This will tend to be the case in general, because shared beliefs require shared practices, but not vice versa. [. . .] Whatever is an object of belief, whatever is verbalisable, will become visible to the practitioners. What you believe, you will sooner or later discuss; and what you discuss, especially in a cultural setting similar to the Greek, you will sooner or later debate. But the real undebated, and in a sense undebatable, aspect of any scientific enterprise is its non-verbal practices. (Netz 1999, 2)

Being verbalizable, the objects of belief can be thought of as corresponding to sentences, which can be associated with propositions. But non-verbal practices, says Netz, must also be taken into account if we want to paint an accurate picture of scientific inquiries. It is safe to say that Netz's commitment can be associated with a widespread view among philosophers interested in mathematical practice, namely, that mathematical knowledge cannot be reduced to the knowledge of propositions alone.[5] Accordingly, all forms of knowledge should be considered, including *knowing how* to do something. The epistemologists' analysis is then seen as overly narrow and in need of broadening up beyond propositional attitudes.[6]

This point is well taken. Still, accepting that beliefs (and other mental states that take propositions as their content) are insufficient to accurately describe scientific or mathematical practices does not force us to marginalize them. We can still think of them as *central* to scientific practices. Indeed, this aligns with a practice-oriented tradition in the philosophy of science and mathematics – it will be enough to refer to Thomas Kuhn's (1962) description of scientific paradigms and Philip Kitcher's (1984) analysis of mathematical practices. More recently, José Ferreirós (2016) emphasized the importance of going beyond propositional attitudes to understand mathematical practices but admitted that "[m]athematics, as we understand it, is not just practical knowledge but has at its core theoretical knowledge" (Ferreirós 2016, 30). This theoretical knowledge, I submit, can be cashed out in propositional form. And in fact, according to Ferreirós, it is always possible, albeit in some cases just in retrospect, to associate a mathematical practice with a particular mathematical *theory* (*ibid.* 29).

The above considerations show that the fact that epistemology is almost exclusively focused on propositional attitudes does not thereby preclude its application to mathematical practice. What is more, recent efforts in epistemology are bringing into

[5] This view is explicitly endorsed in Ferreirós (2016) and implicitly assumed in a plethora of works, such as the ones discussing (like Netz's) the roles of diagrams and notations in mathematics; see, for example (Chemla 2018; Manders 2008).

[6] Thanks to Gisele Secco whose criticism helped me realize the importance of this issue.

focus the analysis of *knowing how* and assessing the hypothesis that it is not always reducible to propositional knowledge, or *knowing that*.[7] These efforts could contribute to a better understanding of what unites scientific practices besides shared beliefs, to use Netz's expression in the quotation above. That is, philosophers of mathematical practice could take advantage of the tools developed to talk about knowledge how in general and apply them to their preferred domain.

The first obstruction to the application of epistemology to the study of mathematical practice is thus cleared. This much will suffice to convince the reader, I hope, that the traditional focus on propositional attitudes should not be a reason for philosophers of mathematical practice to brush aside epistemology. But this is by no means the sole obstruction.

Philosophers of mathematical practice might also be dissatisfied with how epistemology developed. Here is a cartoon version of the story. Going way back to Plato's *Theaetetus*, (propositional) knowledge has been thought of as something like justified true belief. And then, in the 1960s, Gettier (1963) came around and showed that justification, truth, and belief are not jointly sufficient for knowledge. This is because there are cases in which a subject is justified in believing a true proposition, but the proper connection between justification and truth is severed by some sort of epistemic luck that is incompatible with knowledge.

Suppose that I am justified in believing the proposition <Valeria is in Paris>. I am therefore justified in believing <Valeria is in Paris or p> for any proposition p since this is a simple logical consequence of <Valeria is in Paris>.[8] Therefore, even if I have no justification for believing that Valeria has a car, I am justified to believe proposition q = <Valeria is in Paris or she has a car>. As it turns out, unbeknown to me, Valeria just bought a new car and drove to Barcelona with it. In this case, q turns out to be true, but not for the reason I thought. I believe q, I am justified in believing in q, and q is true. But something has gone awry. I do not know q. Gettier cases like this one are cases in which a subject holds a justified true belief that does not constitute knowledge because some type of luck is at play.

For a while, after Gettier's publication in 1964, epistemologists set out to find an elusive fourth condition for knowledge. This new enterprise, however, did eventually stall. Many epistemologists decided to let go of knowledge to focus exclusively on justification. Others took knowledge as a primitive notion.[9]

The problem is that Gettier cases seem to be utterly irrelevant to mathematics. Even setting aside the vexed issue of how we should conceive the truth of mathematical propositions, Gettier cases seem to be impossible in mathematics because justification in mathematics has been traditionally associated with proofs, and these entail their conclusion, leaving no gap between justification and truth – at least if we

[7] See, for example J.A. Carter and Pritchard (2015), Bengson and Moffett (2011), and Habgood-Coote (2019).

[8] Note that this "or" is inclusive. It is the standard disjunctive connective of classical logic: "A or B" is true unless both A and B are false.

[9] This lead to *knowledge first* approaches in epistemology; see Williamson (2000).

take mathematical propositions to be conditional on their premises (i.e., the axioms). Moreover, mathematical propositions are necessary, and thus it is unclear how epistemic luck could enter the picture.

Furthermore, several of the analyses of knowledge developed by epistemologists after Gettier did explicitly exclude mathematical knowledge. For one, Alvin Goldman's causal theory of knowledge is designed to apply to empirical beliefs exclusively.[10] And this is not considered to be a problem at all:

> My concern will be with knowledge of empirical propositions only, since I think that the traditional analysis is adequate for knowledge of nonempirical truths. (Goldman 1967, 357)

The causal theory was then replaced by *reliabilism* (Goldman 1979). Very roughly, this view says that a belief is justified if and only if it is produced by a reliable belief-producing process. Reliabilism applies to mathematics, but it is not immediately clear how it can help us with the specificities of *mathematical* justification, especially if this is connected with the notion of proof. This is because it delivers an overly idealized account of mathematical beliefs based on proofs. Since the premises of genuine proofs entail their conclusions, it is plausible to think that beliefs justified by genuine proofs will be produced by an ideally reliable process. But this is certainly not what happens in practice. A broadly reliabilist framework could, however, be used to paint a faithful picture of mathematical practice. However, to do that, we would have to move away from an infallible conception of mathematical justification and begin to explain how our belief-forming processes at play in mathematical practice are in fact reliable, even if not ideally reliable – for an effort in that direction, see Avigad (2021).

Other attempts to characterize knowledge invoke modal notions (e.g., sensitivity and safety) to get at some sort of robustness that is intuitively peculiar to knowledge. Modal notions do not, however, apply to necessary truths in non-trivial ways.[11]

This cartoon story makes it clear that the concerns of general epistemologists do not meet the ones of philosophers of mathematics, let alone the ones of philosophers of mathematical practice. What is more, in the background of many epistemological discussions about knowledge lurked the problem of skepticism. How can we revindicate the fact that we know most things we think we know? Like that we have hands, that water is H_2O, or that Trenton is the capital of New Jersey. And here, too, the focus is on ordinary beliefs rather than mathematical beliefs.[12]

One strategy to show that epistemology still has something to offer to philosophers of mathematical practice is to observe that there are other topics besides the mainstream ones. I mentioned already recent efforts toward the analysis of

[10] It is therefore somewhat bizarre that the famous *problem of access to abstract objects* originally formulated in Benacerraf (1973) involved Goldman's causal theory of knowledge (more on this later).

[11] There are ways to apply them to mathematics, however. See, for example, Clarke-Doane (2020).

[12] This has not always been the case. For instance, Descartes's most radical form of skepticism involved mathematical beliefs, but these do not feature in most contemporary discussions.

knowledge how. I will later indicate a few other such themes. My main goal, however, is to suggest something stronger. Namely, that even these traditional topics that seem so far from the concerns of a practice-oriented philosophy of mathematics could be applied to it. Among other things, I will draw on previous work (De Toffoli 2021a, 2022) to show that Gettier cases are indeed possible in mathematics, and that thinking about them can be instructive to philosophers of mathematical practice. Gettier cases, I will indicate, arise when we substitute the traditional infallible notion of mathematical justification associated with proofs with a fallible counterpart. This substitution is natural when we put at the center of our epistemological inquiry fallible human subjects rather than highly idealized ones.

Summing up, it is not hard to explain why it seems that mathematics is not relevant to general epistemology. Moreover, when epistemology has been applied to mathematics, it is not the mathematics that philosophers of mathematical practice seek to investigate – but rather a mathematics conceived as a realm of necessary truths, the knowledge of which can be gained a priori. To my mind, this gives a plausible (if only partial) answer to why the philosophy of mathematical practice and epistemology have been overlooking each other, with just a few (but notable) exceptions. I want to suggest that it is time for a change. My goal here is to prepare the terrain for further work at the intersection of the philosophy of mathematical practice and epistemology by showing that we can fruitfully think about mathematical practice by adopting long-established epistemological frameworks.

Here is the plan for the chapter. Section 2, which is the bulk of the chapter, is devoted to exploring how to characterize the knowing subject(s) if we want to pursue a successful epistemology of mathematical practice. I will indicate some points of contact with the characterization of the subjects by social and feminist epistemologists. In Sect. 3, I turn to the traditional analysis of knowledge and discuss how it can be applied to mathematics. I build on previous work (De Toffoli 2021a) to sketch a fallibilist account of mathematical justification that makes it plausible to think that Gettier cases could arise in this context. I then sum up the discussion and point to further epistemological themes relevant to mathematical practice.

2 With or Without Subjects?

I start my analysis of the knowing subject(s) by discussing how the epistemology of mathematics (and of science more generally) has been conceived by Karl Popper and others as an enterprise aimed at the study of "knowledge in the objective sense" in which subjects do not figure at all. Afterward, I consider how highly idealized subjects entered the picture. I then sketch a different type of subject: a fallible subject whose rationality is bounded. It is only with this type of subject in mind that we can pursue a *human* epistemology of mathematics. But in order for this epistemology to be tethered to actual practice, subjects should not be considered in isolation but connected with other subjects and embedded in a specific mathematical context. I then explore the possibility of indexing knowledge to groups rather than to individuals – as has been done by some social epistemologists. I also consider how this

reconception of the knowing subject(s) could bring philosophers of mathematical practice close to feminist epistemologists. I wind up the section by canvassing the importance of conceiving epistemic subjects as not entirely passive, but as active members of specific mathematical practices.

2.1 The Epistemology Without a Subject

Descartes, Locke, and Hume are illustrious among the so-called *belief-philosophers*. This is a label introduced by Popper to describe those philosophers that are "interested in our subjective beliefs, and their basis or origin" (1968, 334). According to Popper, epistemology is a theory of scientific knowledge in which the knowing subject plays no role:

> Knowledge in the objective sense is knowledge without a knower: it is knowledge without a knowing subject. (Popper 1968, 335)

In Popper's view, there are three worlds. The world of material objects, the world of subjective ideas, and a third world, which is the world of *objective contents of thought*. Driven by a similar motivation of separating subjective ideas from objective non-material items, Frege had already introduced a *third realm*.[13]

General epistemology – to be clear, the one pursued by belief-philosophers such as Descartes and later by analytic epistemologists – focuses on the second world and, according to Popper, should be reclassified as psychology or sociology, being utterly irrelevant to his conception of epistemology, which refers exclusively to the third world. According to him, epistemology relates to items such as scientific problems, conjectures, and theories. These items are created by human animals but, most importantly, are *autonomous*. Crucially, Popper's objective knowledge is not constrained by the cognitive abilities of any subject. Indeed, it even includes pieces of knowledge that no human subject will ever acquire.

Why was Popper so adamant in rejecting the knowing subject? It is because he wanted, like Frege before him, to shun *subjectivism* at all costs.[14] In Frege's case, the threat was *psychologism*,[15] and what had to be saved was mathematics. Frege assumed that *ideas* (i.e., things belonging to the second realm) would inevitably vary between subjects. If mathematical entities were ideas, then we would be forced

[13] Frege's third realm, however, differently from Popper's third world, only accommodates necessary immutable truths, leaving out the laws of science.

[14] As Susan Haack (1979) explains, it is not trivial to pin down the exact locus of the negative feature of the type of subjectivism Popper had in mind. First, the objective/subjective dichotomy can be conceived in several different ways. Second, the notion of *intersubjectivity* can be placed in between, thus leaving the dichotomy to give space to something more nuanced.

[15] For an analysis of the reasons (some of which are institutional in nature) why psychologism started to be conceived as a major sin in the German philosophy departments of the beginning of the twentieth century, see Kusch (1995).

to distinguish, for example, between *my* Pythagorean theorem and *your* Pythagorean theorem. But that is absurd.[16] Popper's worries have a similar flavor.

Popper's attitude toward epistemology is not unique to him. It is endorsed by other players in the twentieth-century philosophical landscape. The "epistemology without a subject," as Ferreirós (2016, 11) labels it, found fertile ground in the realm of mathematics, which, after all, was where Frege's inquiries took place. In France, Jean Cavaillès (1937, 1938) developed a philosophy of mathematics focused on concepts, rather than on subjects.[17]

Focusing on mathematics is congenial to an epistemology without a subject because mathematics is prima facie a realm of necessary truths that admit indefeasible justification.[18] After all, it is reasonable to think that we could not be wrong that $1 + 1 = 2$. This infallibility indicates that there are no relevant individual differences in mathematics. The main assumption of philosophers of mathematics was that mathematical justification (and knowledge) is underwritten by proofs, which were conceived as deductions from axioms. The focus thus was to justify mathematical theories rather than mathematical beliefs. The epistemology without a subject ignored individual beliefs and the inner working of mathematics and focused on very general issues.[19]

If the application of epistemology to mathematics led to an *epistemology without a subject*, it is not surprising that philosophers of mathematical practice felt the need to look elsewhere for philosophical inspiration. But this is not the end of the story. Besides efforts toward the justification of mathematical theories, in the second half of the twentieth century, philosophers of mathematics started tackling different epistemological issues. And this time, subjects were needed. The problem, however, is that these subjects, as we are about to see, were too idealized to be taken as the subjects of inquiry for a philosophy of mathematical practice.

2.2 Ideal Rationality

In 1973, Benacerraf published what would become one of the most influential papers in the philosophy of mathematics of the twentieth century, "Mathematical Truth" (1973). The paper poses a dilemma. If mathematics is, as it seems to be, about mathematical objects, and if mathematical objects are abstract, then how is knowledge of such objects possible at all? A quandary arises from combining a view of mathematical objects as not spatiotemporally located (and thus causally inert) with

[16] To be sure, this is a simplification. Frege critique of psychologism is sophisticated and developed across multiple works.

[17] See Sinaceur (2019).

[18] And in fact, the most virulent form of psychologism, against which Frege's criticism were levelled, arises from the application of psychology to logic and mathematics.

[19] Think, for example, about conventionalism and Quine's holistic empiricism. For an elaboration on these ideas, see De Toffoli (2021a).

the causal theory of knowledge. *Benacerraf's dilemma*, as it is now commonly referred to, has generated a massive amount of scholarship. The causal theory of knowledge is surpassed by subtler alternatives, but Benacerraf's dilemma has proven to be robust with respect to different epistemological theories of knowledge.

Benacerraf's dilemma has to do with individual knowledge, and thus it introduces individual subjects into the epistemology of mathematics. But the problem it raises is very general and has to do with metaphysical considerations that practicing mathematicians are generally oblivious to. From such a general vantage point, the knowledge that 2 is a prime number poses no more and no fewer problems than the knowledge of Fermat's Last Theorem. This is because Benacerraf's dilemma focuses exclusively on our knowledge of mathematical objects *in general*. In other words, the type of discussion of mathematical knowledge it generates does not require considering any sophisticated mathematics at all.

What is more, mathematical reasoning is not relevant at all to Benacerraf's type of questions. The only propositions that matter for such an epistemological inquiry are the starting points. Given the axioms, the rest follows unproblematically by means of deductions, or so the story goes. That is, given the axioms, we can infallibly know all the rest of mathematics by proofs. If there is a fallibilist component at all in our justification of mathematical facts, it is relegated to our beliefs about the axioms. Any other type of fallibility would have to do with human (in)competence. It should be borne in mind that nobody denies that we are fallible and at times make mistakes – what is rejected is that such fallibility is of any interest in the epistemology of mathematics.

These considerations do not square with a philosophy of mathematics interested in mathematical practice. Mathematicians seldom concern themselves with the axioms and are generally indifferent to the choice of a particular formal system.[20] More crucially still, the subjects for which only the starting points matter are logical omniscient subjects.

Note that although it is ill-advised to study mathematical practice by conceiving subjects as logically omniscient, this is not the case for other domains. In the context of formal epistemology, it is perfectly suitable:

> [I]n many standard applications of the Bayesian machinery, the assumption [of logical omniscence] is natural. Suppose, for instance, we're trying to use it to decide between two hypotheses about the chance with which a coin will land heads when it is tossed. [. . .] In this case, logical omniscience tells us that we must be certain that, if the coin didn't land heads on the first toss, it landed tails; it tells us to be at least as confident that it landed heads on the first toss as we are that it landed heads on the first two tosses; and so on. In these cases, logical omniscience seems a reasonable requirement. And it was such cases where Bayesianism first found application. (Pettigrew 2021, 9992)

In cases such as the one described in the quote above, it is plausible to think that we are forced to assign the same probability to all logically equivalent propositions, on pain of irrationality. And then, Richard Pettigrew explains, it is natural to extend

[20] See Burgess (2015, 145) illuminating discussion of different types of *indifference* in mathematical practice.

further such a requirement. Logical omniscience is a requirement that arises from the Dutch Book Arguments, which govern the norms that regulate rational credences and were developed by Frank P. Ramsey and Bruno de Finetti.

But even in the context of Bayesian epistemology, logical omniscience can become a *problem* since it clashes with our ordinary judgments of rationality. And in fact, there are several attempts to *solve* the problem of logical omniscience and to apply formal epistemology to logical learning and other cases of bounded rationality.[21] But these efforts have not considered in detail sophisticated mathematics, which arguably is the main target of study for philosophers of mathematical practice.

2.3 Non-Ideal Rationality

If logical omniscience is a requirement of rationality, it is not a very useful one to evaluate the epistemic achievements and falls of *human* subjects. Nobody would be able to pass the bar. As David Christensen (2004) explains:

> When we call someone "irrational," we are saying that he is deficient *relative to a contextually appropriate standard*, which need not be—and typically is not—the standard of absolute rational perfection. (152, *emphasis added*)

Developing *contextually appropriate standards* allows us to distinguish between different types of failures. One thing is a careful mathematician who makes a subtle mistake. An entirely different case is the one of an absent-minded mathematician who comes up with a bogus proof for some result or another. As Daniel Greco and Brian Hedden (2016) explain:

> [I]t is important to distinguish being irrational in the sense of falling short of the rational ideal from being irrational in the sense of being less rational than (most of) the rest of us. (3677)

The type of rationality we are concerned with when analyzing mathematical practice is *non-ideal,* and it is going to be indexed to a particular mathematical context. The challenge is then to focus on specific mathematical practices and spell out the context-dependent norms that regulate such practices. And this is perfectly in line with a philosophy of mathematics interested in mathematical practice.

As a matter of fact, several works produced by philosophers of mathematical practice can be understood as a reply to such an epistemological challenge. As a way of example, consider the multiple studies on diagrams in mathematics. Albeit focusing on different practices and endorsing different methodologies, most of these studies aim at unveiling the (often implicit) norms governing the use of diagrams in different contexts. For example, there is ample literature discussing the epistemic role of diagrams in Euclidean geometry and the way in which their use was regulated (Netz 1999; Manders 2008; Macbeth 2010; Panza 2012). Diagrams from other historical practices have been considered as well – in this regard, the work

[21] For a recent endeavor in that direction, see Pettigrew (2021).

of Karine Chemla (2018) on diagrams in ancient Chinese mathematics is a prime example. Examples of studies considering diagrams in contemporary mathematics are Carter (2019), De Toffoli and Giardino (2014), and De Toffoli (2023).

These studies put at the center of their inquiry fallible, limited subjects with a specific cognitive makeup and inhabiting particular contexts. It is precisely because of their cognitive shortcomings that representational choices (such as the ones to use diagrams) are relevant in practice.

2.4 Social Epistemology

Putting at the center of epistemological inquiries non-ideal subjects embedded in a specific mathematical practice obliges us to look at the social dimension of mathematics. In point of fact, contrary to the romantic image of the mathematician working in isolation, most members of a mathematical practice interact systematically with each other and with the broader mathematical community. It is for this reason that the epistemic norms at play in mathematics present a social dimension.

For instance, what counts as a proof is not fixed once and for all: what is acceptable in one context might not be acceptable in another. This does not mean that we must renounce to a context-independent notion of logical validity, but that arguments put forward as proofs cannot be evaluated independently from the audience to which they are addressed. By way of example, in topology but not in other branches of mathematics, arguments involving sequences of visualizations might be acceptable as proofs.[22]

Moreover, it has been pointed out that some of the norms governing justification in mathematics guarantee the possibility of the mathematical community to perform a self-monitoring activity on the results it produces, thus aiding single mathematicians to overcome their individual limitations.[23] These considerations point to the fact that to paint a faithful picture of mathematical practice, the social dimension of mathematics should be considered. To do so, it is possible to import into the philosophy of mathematics some of the tools that have been developed by social epistemologists.

There are already some efforts in this direction. For example, Kenny Easwaran (2009) has argued that proofs should be *transferable* – that is, roughly, that they should convince the relevant experts of their correctness without any appeal to testimony.[24] He then explains:

> Philosophers of mathematics have long focused on individual epistemic norms for mathematics, and deductive logic has been very useful in helping to understand these norms.

[22] See De Toffoli (2021b).

[23] See Easwaran (2009) and De Toffoli (2021a).

[24] In my view, transferability can also be understood in terms of the a priori – given a minimal, fallible conception of a priori justification. In this volume, Danielle Macbeth (2021, 13) also connects transferability to the a priori: "The proof is, as it is sometimes put, "transferable" (Easwaran 2009). Indeed, it is in just this sense that mathematical knowledge is a priori: It is a priori not because it is infallible but because and insofar as it does not rely on empirical evidence, whether that of one's own senses or that given on the testimony of another."

> However, transferability is a social norm—it can help the community develop a better grasp on the knowledge of its members, even though it may not have any advantages for the individual considered in isolation. (Easwaran 2009, 356–357)

Although general epistemology has traditionally focused on individual subjects in isolation, social epistemology is growing very rapidly. This is a branch of epistemology that puts particular emphasis on the fact that the subjects under inquiry are embedded in specific social contexts.

Sanford Goldberg (2018) ventured so far as to trace the very source of epistemic normativity to our social nature. In his view, interacting with others has many benefits, but it also puts us in a vulnerable position. That is why we are entitled to require specific standards of the doxastic lives of others – and these entitlements give rise to the normativity of epistemological notions such as justification and knowledge. His proposal applies to beliefs in general, but it would be worth considering how it could be modulated in specific for the mathematical case.

One respect in which social epistemology could be applied to the study of mathematical practice is the following. Social epistemologists are not only interested in how social norms and social contexts do influence epistemic notions such as individual justification and knowledge, but also in how such notions *themselves* can be indexed to a plurality of subjects. This suggestion could be evaluated and tested by considering specific cases from the practice of mathematics. For example, several proposals have been put forward on how to characterize group beliefs (Gilbert 2004; Lackey 2016, 2020). The literature can be broadly divided into two camps: the one formed by those who endorse an *inflationary* approach and the one formed by those defending a *deflationary* approach. Briefly, according to the latter but not to the former, group beliefs are reduced to the beliefs of the members of the group.

Looking at the social structure of mathematics could offer important data in support of one or the other position. More specifically, contemporary mathematics offers us examples in which it is helpful to think of groups of practitioners themselves, rather than individual mathematicians, as epistemic subjects. *Big proofs*, that is, proofs that are not within reach of a single mathematician, are an ever more common phenomenon in contemporary mathematics – besides computer-assisted proofs, prime examples of big proofs are proofs that involve large collaborations. A famous example of the latter is the proof (or the multiple proofs) of the theorem of the classification of finite simple groups.[25] Focusing on such a case or similar ones could, for instance, be used to evaluate and support the inflationary approach to justified group belief.

2.5 Feminist Epistemology

I started this section by surveying the *epistemology without a subject* and, passing through an epistemology considering logical omniscient subjects, I ended with the consideration of fallible, limited subjects inhabiting specific contexts. I then suggested that subjects are better conceived as entertaining systematic relations with each other

[25] See Habgood-Coote and Tanswell (2021).

rather than as operating in isolation. Finally, I indicated that, in some cases, it might even be useful to think of multiple subjects rather than individuals as the relevant subject to which to index belief, justification, and knowledge.

Moving from a very abstract and disembodied way of thinking about the epistemology of mathematics to a more concrete and embedded way is in line with the methodological guidelines of feminist epistemology. And, in fact, I want to suggest that the philosophy of mathematical practice and the epistemology approaches that are most suitable for it share some joint tenets with feminist epistemology.

To be sure, feminist epistemology is a cluster of heterogenous views rather than a unified position.[26] Notwithstanding the great variety among the particular theories, they all share some core commitments. As Helen Longino (1998) has emphasized, one of such core commitments is to put at the center of investigation non-ideal subjects situated in a specific context:

> The shift from a transcendent and disembodied subject to empirical, embodied, and differentiated subjects is often represented as a loss, encouraging a representation of the aspects of situatedness as interfering with knowledge or cognitive access. [. . .] however, we might better think of them as focusing devices, or cognitive resources, directing our attention to features of that which we seek to know that we would otherwise overlook. (335)

Indeed, as I have suggested, without this de-idealization of the knowing subject(s), it would be impossible to appreciate the epistemic norms governing actual mathematical practice.

Longino also individuates another aspect that unites feminist approaches to epistemology: the consideration of social factors. Subjects are often considered to be interdependent. In her words,

> One further consequence of acknowledging the embodied character of knowers that feminists have explored is the dependence or interdependence of knowers. (*ibid.,* 335)

The methodology of feminist epistemologists is thus seen to converge with the one of social epistemologists.

It is therefore clear that an epistemology suitable to the study of mathematical practice conceives of the knowing subject(s) in a way that is in line with how these subject(s) are conceived within feminist epistemology. This is not surprising. After all, some of the feminist approaches, like Longino's own approach,[27] emerge precisely from what can be called the *practice turn in the philosophy of science*[28] – a trend that started well before[29] the *practice turn in the philosophy of mathematics*.[30]

[26] See Garavaso (2018).

[27] See Longino (2002).

[28] See, for example, Soler et al. (2014).

[29] Its beginning could be pinned to Kuhn's (1962) influential work on scientific revolutions; later on, emphasis on practice was prominent among historians and sociologists of science.

[30] Although there are important precursor, like (Lakatos 1976), the philosophy of mathematical practice took off much later. See, for example, Mancosu (2008).

2.6 Subjects or Agents?

Before moving on to the analysis of knowledge, I want to address a last issue on the nature of epistemic subjects. Ferreirós favors talking of *agents* rather than *subjects*. Although I share some of Ferreirós's motivations, I will take sides against his proposal.

Ferreirós strives to emphasize the active nature of mathematicians producing and sharing knowledge in a specific mathematical practice. The problem is that etymologically, "subject" has a passive connotation ("sub-jacere": to throw under). In his words:

> I dare say that the *oblivion of action* was the worst defect of traditional epistemology, its main shortcoming that, in my view, has made impossible a well-grounded account of human knowledge. Even perception, which modern philosophers from Descartes and Locke onward understood as a passive reception of impressions, is in the light of neurobiology and cognitive science a complex system that results from the interplay of sensory input and motor output (not anything like a primitive faculty of the mind, as philosophers tended to think). Input and output, sensation and action–without this feedback, one cannot even make sense of perception, let alone the further complexities of knowledge production. (11)

I agree with Ferreirós that traditionally epistemic subjects have been conceived of as entirely passive – and that this is ill-suited to study mathematical practice. However, naturalized epistemologists do nowadays consider the lessons of cognitive science and are aware of the important feedback between sensory input and motor action. Moreover, epistemologists that are sympathetic to the embodied cognition movements and its cognates are certainly happy to consider the subjects of epistemology as embodied. As we saw, this is also one of the main tenets of feminist epistemology:

> One of the consistent themes running through the feminist rethinking of the subject of knowledge is the insistence on its embodiment. (Longino 1998)

Still, there is an important sense in which the *agents of ethics* and the *subjects of epistemology* are better kept separate. What we believe is not *up to us* in the same way as what we do. This is clear for perception: most of my perceptual beliefs are formed without conscious deliberations. Trivially, I might influence what my perceptual beliefs are by controlling what sensory stimuli I expose myself to. For instance, if I turn my head right and keep my eyes open, which I can control, I will inevitably see out of the window and form the belief that the sun is shining. There might also be other forms of indirect control. However, even for reflective beliefs such as mathematical beliefs, it is plausible to think that they often result automatically from exposure to the relevant evidence.[31] There are various positions in the doxastic voluntarism debate that indicate several ways in which we might be

[31] For instance, this would be in line with Kornblith (2012).

able to have some kind of control over our beliefs,[32] but they generally take for granted the difference between belief and action.

When I go through a proof, and I understand it, or when I read a compelling argument, I cannot help but form a belief.[33] There is no direct choice there. But action is different. If I see a child drowning in a pond, I can choose between saving him and ruining my beautiful Italian boots or watching him drowning and preserving the boots.[34]

Crucially, saying that belief and action are different in this respect does not entail that we cannot talk about normative notions and responsibility in epistemology. To be sure, we can still talk about justified and unjustified beliefs. Moreover, there is a substantive body of literature on how we can be considered responsible for our beliefs – and the *ethics of beliefs* is a sub-field in epistemology that has grown exponentially in the last few years.[35]

There is yet another sense in which the subjects of epistemology can be conceived as active. Namely, the status of *being justified* may be connected to the *activity of justifying*. As Ferreirós notes, this can be traced to the tradition of the American pragmatists and goes against what Adam Leite (2004) calls the "Spectatorial Conception," which he characterizes as follows:[36]

> The fundamental idea shared by these views is that being justified is something which happens to you. According to these theories, the justificatory status of a person's belief is determined by certain facts which obtain prior to and independently of the activity of justifying. The activity itself plays no role in determining justificatory status; it is simply a secondary and optional matter of attempting to determine and report, as far as is conversationally necessary, the prior and independent facts which determine the justificatory status of one's belief. Consequently, even if things go badly wrong in the course of the activity, that will not determine one's actual justificatory status. On this conception, one stands in a primarily theoretical or epistemic relation to the justificatory status of one's beliefs: positive justificatory status is something which one finds out about, not something which one brings about. I therefore call this view the Spectatorial Conception. (222)

Leite proposes an alternative account in which the *basing relation*, that is, the relationship that exists between a belief based on a reason and the reason on which it is based, is something that can be described in terms of the activity of justifying. This is another example of how a traditional epistemological debate, the one about the nature of the basing relation, can be applied to a philosophy of mathematical practice – and how in turn considering mathematical practice could bring the debate forward. For instance, in De Toffoli (2021a), I proposed an account of the basing relation for

[32] See, for example, Hieronymi (2006).

[33] Note that such a belief might not be an unqualified belief in the conclusion of the proof but can be the conditional belief that the conclusion is implied by the (explicit or implicit) premises.

[34] The example is adapted from the classic Singer (1972).

[35] See Chignell (2018).

[36] This is also related to John Dewey's critique to traditional epistemology according to which the subject of knowledge, like the classical subject of perception, is passive.

mathematics that is in line with Leite's and that puts the emphasis on our ability to provide reasons for our beliefs rather than on their causal origin.

Let's take stock. I agree with Ferreirós that appropriate subjects of an epistemology of mathematics taking the practice seriously should be considered as embodied and situated in a particular context. I further agree that they should be conceived as active members of their communities. Still, I prefer to keep in line with the tradition and call them epistemic *subjects* merely to emphasize that the focus of our inquiry is belief rather than action – which, to be sure, are interrelated.

Having discussed how epistemic subjects can be characterized, we can turn to the traditional analysis of knowledge and evaluate whether it can be fruitfully applied to mathematical practice.

3 The Analysis of Knowledge

In general epistemology, it is common to deploy formulas beginning with "S knows that p. . ." We now have a sense of how to conceive of the subject S. Rather than a disembodied, highly idealized subject in a vacuum, our subject will either be a fallible and embodied subject embedded in a social context or be a group of such subjects.

And then? After "S knows that p. . ." we can satisfy ourselves with an "only if" rather than an "if and only if." Gettier's lesson was precisely that justified true belief is necessary but not sufficient for knowledge. We have: "S knows that p only if p is true, S believes p, and S's belief that p is justified." More schematically:

$$\text{Knowledge} => \text{Justification} + \text{Truth} + \text{Belief}$$

If a subject knows that p, then she believes p. The *belief condition* is the least controversial. Moreover, it is plausible to think that if it presents problems at all, these are not peculiar to mathematics. For this reason, I will set it aside.

The other components require some discussion. The fact that propositional knowledge is knowledge of a true proposition is taken for granted by epistemologists. However, there are reasons to think that this is a non-starter for a philosophy of mathematical practice. I will address this point before turning to an account of justification that allows for Gettier cases in mathematics.

For reasons of space, it is not possible for me to go beyond these necessary conditions and discuss a full analysis of knowledge. I leave the topic for further work. However, let me just briefly list three promising strategies that can be deployed to develop an account of knowledge that fits mathematical practice. One is to tweak modal notions (such as safety or sensitivity) so that they can be applied to necessary truths.[37] A second option is to deploy the virtue epistemology machinery developed by Ernest Sosa (2007).[38] A third approach (which is compatible with the

[37] Some steps in this directions have been taken in (Clarke-Doane 2020).

[38] See Tanswell (2016).

second) is to go knowledge-first. Following Timothy Williamson (2000), the idea here is to take knowledge as a primitive notion and explain justification in terms of it rather than the other way around – note that although Williamson and other proponents of the knowledge-fist approach are infallibilist about justification (and thus in my view their theories can hardly be applied to analyze mathematical practice), fallibilist knowledge-first theories of justification have been developed as well.[39]

Let us then turn to the truth and justification conditions for knowledge.

3.1 True Propositions

If knowledge requires justified true belief, it must be knowledge of true propositions. I have already discussed in the introduction that to analyze a scientific practice, it might not be enough to consider beliefs exclusively. Beliefs play, however, a central role in any practice that has a theoretical orientation.

In the tradition of Robert Stalnaker (1976), propositions have been coarsely identified using the semantics of possible worlds. A proposition can be conceived as the set of possible worlds where it is true. But this implies that there is just one necessary proposition. David Lewis (1996) is explicit about this point:

> What I choose to call 'propositions' are individuated coarsely, by necessary equivalence. For instance, there is only one necessary proposition. (551)

But plausibly mathematical propositions are necessary propositions. And if there is just one mathematical true proposition, then propositional knowledge in mathematics will be quite uninteresting since it won't even be possible to differentiate our knowledge of different mathematical propositions. Lewis continues:

> So the necessary proposition is known always and everywhere. Yet this known proposition may go unrecognised when presented in impenetrable linguistic disguise, say as the proposition that every even number is the sum of two primes.[40] Likewise, the known proposition that I have two hands may go unrecognised when presented as the proposition that the number of my hands is the least number n such that every even number is the sum of n primes. [. . .] These problems of disguise shall not concern us here. Our topic is modal, not hyperintensional, epistemology. (*ibid.*, 552)

It is clear, however, that the problem of disguise must concern us *here*. And in fact, we need to individuate propositional content in a finer way to talk about mathematics at all! This problem should not be, however, a roadblock but a challenge to overcome. Several methods are available. One option is to index beliefs

[39] See, for example, Silva (2022).

[40] Actually, the Goldbach conjecture says that every even number *larger than 2* is the sum of two primes.

to sentences rather than to propositions.[41] Another is to adopt one of the strategies that have been developed under the label of *hyperintensionality*.[42]

Besides problems arising from the very conception of propositions, another issue concerns talking about *truth* in mathematics. Here is Ferreirós (2016) on the matter:

> From the time of the ancient Greeks, with their definition of knowledge as "justified true belief," to the twentieth century under the influence of denotational semantics, it has been common to consider mathematics as a body of truths (otherwise, it is felt, there couldn't be knowledge of them). Math has been regarded as the discipline that provides justifications of the strongest kind, namely deductive proofs. But what are mathematical truths true of? (6–7)

Let us bracket the claim about justification (to which I will return shortly) and focus on truth. To dodge the question "what are mathematical truths true of?" and avoid talking about truth altogether, Ferreirós completely abandons the classical analysis of knowledge. He prefers to focus on objectivity instead of truth.[43] I agree with him that concentrating on objectivity rather than truth is better suited to studying mathematical practice. "The problem," as Kreisel famously put it, "is not the existence of mathematical objects, but the objectivity of mathematical statements."[44]

However, I want to indicate here that we are not forced to abandon the traditional analysis of knowledge (that has at its center the notion of true propositions) altogether. The issue is massive; in this context, I just want to suggest that talking about true mathematical propositions could be applied even to Ferreirós's own account of mathematics, or something in line with it, anyway.

Ferreirós asks, "what are mathematical truths true of?" It is plausible to think that mathematical truths are truths about mathematical objects. This reflects Benacerraf's (1973) desideratum (shared by many philosophers of mathematics) of having a homogeneous semantics between mathematical and non-mathematical discourse. How we conceive of these objects is another matter, however. Moreover, there are alternatives to this simple answer. As a way of example, consider Geoffrey Hellman's (1989) modal structuralism. In his view, mathematical truths are about possible systems of objects that exemplify structures. In his words:

> [O]rdinary mathematical statements are construed as elliptical for hypothetical statements as to what would hold in any structure of the appropriate type [. . .] Absolute reference to mathematical objects is eliminated entirely. (53)

Hellman's position is inspired by Hilary Putnam's (1967) seminal (and controversial) paper "Mathematics Without Foundations," in which the author claims that

[41] This is done in Pettigrew (2021).

[42] See, for example, Berto and Nolan (2021).

[43] A similar approach can be (and has been) adopted with respect to science. See for example the discussion in Bird (2010).

[44] Reported in Dummett (1978, xxviii).

it is possible to give two different interpretations of the same mathematical fact: platonistic or modal.

We can endorse our favorite position in the ontology of mathematics – we can be platonists or nominalists, or neither of them.[45] We can adopt a structuralist strategy and think of mathematical objects as positions in abstract structures. This should not, however, force upon us a specific epistemology of mathematics. However, it is true that some metaphysical views may rule out certain epistemologies. For example, some versions of fictionalism take mathematical statements like "2 is a prime number" at face value to be false because they are about objects that do not exist.

Another relevant consideration is that, looking at the practice, it is often natural to focus on *conditional* mathematical propositions. Although traditional inquiries in the epistemology of mathematics have been focused on our knowledge of axioms, from the perspective of the working mathematician, the truth of axioms is seldom addressed.[46] What matters the most from a practice-based perspective is whether certain premises entail certain conclusions.[47]

When we consider mathematical knowledge in practice, we think about mathematicians producing correct proofs for some mathematical proposition. We have a pretty good logical story to understand what *correct* means[48] – although this story might not be readily applied to the practice.[49] It is then plausible to think that mathematical knowledge does not require a commitment to the absolute truth of the axioms we use to prove our theorems, but only a commitment to the correctness of our proofs.

In contemporary mathematics, it is often accepted that mathematical proofs are supposed to establish that their conclusions are true *if* their premises are true. Moreover, when investigating certain domains, it is common practice to discharge certain premises. In Euclidean geometry, when we talk about a theorem T, such as the Angle Sum Theorem, our real conclusion is not the stated conclusion of T, but rather: *in any Euclidean plane*, T.[50] Obviously, the same stated conclusion would not be a truth of hyperbolic geometry. Nowadays, everybody accepts that both Euclidean and non-Euclidean geometry are legitimate mathematical domains, and that we can know geometrical truths.

Thinking of mathematical propositions (or at least some of them) as conditional on the axioms is indeed Ferreirós' preferred approach. He offers a *hypothetical*

[45] See, for example, Burgess (1983).

[46] There are exceptions, however. Sometimes the truth of certain axioms (such as the axiom of choice) is debated by practicing mathematicians.

[47] This is in line with Peirce's view of mathematics. According to him, in deductive reasoning we start from an "hypothetical state of things and are led to conclude that, however it may be with the universe in other respects, wherever and whenever the hypothesis may be realized, something else not explicitly supposed in that hypothesis will be true invariably" (1998, 212).

[48] Even though this story is not universally accepted. For example, it is criticized by constructivist mathematicians. This unveils other instances (in this case linked to intuitionism) in which metaphysical positions can indeed lead one to endorse particular epistemological views.

[49] See Avigad (2021).

[50] To be sure, this holds nowadays, not of the historical practice.

conception of advanced mathematics[51] in which the axioms are only conditionally accepted. In so doing, he indexes truth to specific theories:

> My use of the word 'truth' at this point must be relativized by implicit or explicit reference to a mathematical theory. (Ferreirós 2016, 8)

It is because of this relativization that Ferreirós claims that it is advisable to talk of objectivity rather than truth – but this is no roadblock to understanding even mathematical knowledge in terms of true propositions.[52,53] And indeed Ferreirós himself does not renounce to such talk:

> I shall try to provide a strong account of how we arrive at–and share–mathematical truths, without invoking mathematical objects in any strong ontological sense. (*ibid.*)

Crucially, endorsing Ferreirós's hypothetical view, we do not have to postulate the existence of any mathematical object in a strong ontological sense. This is because rather than *describing* a given mathematical domain, axioms would *constitute* it. A final note on the hypothetical view proposed by Ferreirós is in order. It should not be conflated with a simple form of *if-thenism* in mathematics because, in his view, not all hypotheses are on a par – in short, this is explained by Ferreirós appealing to the internal relationships of different mathematical practices as well as to the relationships of mathematical practices with scientific and technical ones. It is by observing the close connection of mathematics with science and even technical practices (from the elementary ones of counting and measuring) that we can explain the non-arbitrariness of mathematics.

Let me also note *en passant* that talking about truth in mathematics does not need to parallel talking about truth in other contexts. Alethic pluralists contend that there are different ways of being true, which are appropriate to different domains of discourse.[54] For our concerns, it suffices to note that what truth amounts to in mathematics might not be the same compared to what truth amounts to when we talk about ordinary objects such as chairs and tables. For instance, while a correspondence theory of truth might be attractive for the latter, for the former, a coherence theory of truth might be more appropriate.[55] Moreover, there could be ways of modulating the same theory in different domains. For example, in her

[51] It is worth noting that he proposes a different account for elementary mathematics.

[52] Interesting similarities could perhaps be drawn from Putnam's *internal realism*, but Ferreirós does not discuss them.

[53] Ironically, Justin Clarke-Done (2020) uses examples similar to the ones by Ferreirós to argue for mathematical realism and *against* mathematical objectivity. In his view, an objective question admits a single answer. He uses the case of non-Euclidean geometry precisely to highlight the non-objective character of mathematics. Although there is much more to be said about this, for reasons of space I have to postpone the discussion to another time.

[54] To be sure, this claim can be interpreted in many ways – see Pedersen and Wright (2018).

[55] A coherence notion of truth for mathematics is developed by Crispin Wright (1992, 1999). For a detailed discussion on how to apply Wright's ideas to mathematics, see (S. Shapiro 2007).

investigations on logic, Gila Sher (2016) has argued in favor of the correspondence theory of truth. In her view, however, what this view amounts to changes in different contexts – she explicitly considers the case of mathematics.

How we should understand truth in mathematics raises monumental questions that I cannot even start to discuss. But I do not need to do so in this context. My aim was just to indicate that Ferrerirós's own proposal contains the resources needed to apply the traditional analysis of knowledge to mathematics. More generally, the suggestion here is that the epistemology of mathematics – especially an epistemology that aims at understanding mathematics as it is actually practiced – can be (mostly) separated from the metaphysics of mathematics.

Let us then turn to the notion of justification.

3.2 Justification

"Math has been regarded as the discipline that provides justifications of the strongest kind, namely deductive proofs," says Ferreirós. It is for this reason that Gettier cases are generally thought to be impossible in mathematics:

> [I]n general epistemology Gettier cases typically involve instances in which S's justification for p is understood to either *proceed through false assumptions* (as in Gettier's original examples) or be such that it is only "by luck" that it delivers a true belief (as in Goldman's Barn County example[56]). *The former issue is ruled out in the mathematical setting in virtue of the fact that a correct mathematical proof cannot proceed through false lemmas* (and that the presence of such fallacious steps can be decided by examining the structure of PROOF itself – a feature which is not true in the case, e.g., of perceptual evidence). And the latter is ruled out in the mathematical setting because the counterfactual dimension of the relevant notion of "luck" (e.g. that S happened to be looking at the one genuine barn in the field as opposed to a fake one in a nearby world) has no clear mathematical analog.[57] (Dean and Kurokawa 2014, 207, *emphasis added*)

In previous work (2021a), focusing on mathematical propositions conditional on axioms, I argued that there is a plausible account of mathematical justification according to which correct mathematical proofs are not the only arguments that provide mathematical justification – even excluding non-deductive methods. In my view, flawed arguments that look like proofs to the relevant practitioners may confer justification as well. Endorsing a fallibilist account of mathematical justification opens the possibility of Gettier cases arising in mathematics, at least those cases involving false assumptions.

To be sure, in the literature, different forms of mathematical fallibilism have been proposed – many authors have focused on how we justify axioms and urged us to leave behind the idea that they are self-evident. For one, Kurt Gödel (1964) himself

[56] This example is actually due to Carl Ginet.

[57] I changed the notation for the subject and the proposition to make it coherent with the one adopted in the rest of the chapter.

was explicit in claiming that some axioms could be justified only *extrinsically* in terms of their consequences. But as I mentioned in the previous section, it is natural for the philosophers of mathematical practice to focus on mathematical propositions other than axioms.

Imre Lakatos (1976) has linked his fallibilism to the quasi-empirical nature of mathematics and to the fluidity of mathematical concepts. However, the type of fallibilism proposed here is different than Lakatos' because it applies straightforwardly to contemporary mathematics as well, not only to the history of mathematics. Rather than deriving from the open-ended nature of mathematical concepts, it originates from our human inability to get things right 100% of the time.

If the subjects at the center of our inquiry are fallible and have limited resources, then it is reasonable to think that justification in general is not factive, that is, that a subject can be justified in holding a false belief or in holding a true belief for wrong reasons. Mathematical justification should be no exception.

Roughly, in my account, whether a subject is justified is not just a matter of logic but hangs on social considerations as well: arguments that are in line with a community's good inferential practices but that contain subtle mistakes may confer justification. In my terminology, these are *simil-proofs*:

> (SIMIL-PROOF) An argument is a Simil-Proof (SP) when it is shareable, and some agents who have judged all its parts to be correct as a result of checking accept it as a proof. Moreover, the argument broadly satisfies the standards of acceptability of the mathematical community to which it is addressed. (De Toffoli 2021a, 835)

Of course, this is a general proposal that needs to be specified in a case-by-case manner. For instance, we should be able to determine what the standards of acceptability of a mathematical community are. However, it is a start toward a fallibilist account of mathematical justification.

It is worth stressing that claiming that mathematical justification is not merely a matter of logic but depends on social and cognitive factors does not force us into some sort of relativism. There is a fact of the matter whether a simil-proof is a genuine proof or not, but we might not be able to establish this fact at a given time – and a notion of mathematical justification that is faithful to actual mathematical practice should be sensitive to it.

That justification is fallible means that there is a gap between justification and truth. Linda Zagzebiski (1994) showed that when this is the case, we can always construct Gettier cases. Analyzing such cases can give us special insight into mathematical practice because it leads us to the consideration of putative proofs that were accepted by the mathematical community for a time and were later deemed to be inadequate. One such example is Kempe's original proof of the 4-color conjecture – it was first published in 1879 and only found wanting 11 years later.[58] Another is Dehn's Lemma:

[58] See the discussion in De Toffoli (2022).

> (DEHN) In 1910 Max Dehn published an argument for what is now called Dehn's Lemma. In 1929 Hellmuth Kneser discovered that Dehn's argument contained a significant gap and thus that the lemma has not been proved (although it had been used as a premise to prove many other results). It was only in 1957 that Christos Papakyriakopoulos published a proof of it. (De Toffoli 2021a, 837)

It is my contention that Dehn was initially justified in believing his result. He had a justified true belief that did not amount to knowledge. Moreover, in the years in which Dehn's simil-proof was accepted by the mathematical community, it would have been perfectly fine in some contexts to use his lemma to establish other results. And this could lead to additional Gettier cases in which a subject is justified in believing a result, the result is true, but the justification is somewhat disconnected from its truth. In these cases, the subject's justification would proceed through false assumptions.

Building on these ideas, Neil Barton (manuscript) has argued that there are different kinds of Gettier cases in mathematics. Besides cases analogous to the one mentioned, Barton also discusses how Gettier cases arise with respect to our beliefs in axioms. In so doing, he builds on a more general conception of mathematical justification, one that does not apply only to conditional mathematical propositions but to axioms as well.

In light of these considerations, it is fair to say that the possibility of Gettier cases in mathematics was dismissed too quickly. And this in turn shows that even traditional concerns of epistemology, which seem so remote from mathematical practice, can be relevant to it.

4 Conclusion

Although the interests of epistemologists and of philosophers of mathematical practice might seem to be disjoint, they are not. I discussed the issue at a very general level.

To sum up, in order to fruitfully apply epistemology to the philosophy of mathematical practice, we have to reconceive the epistemic subjects along lines that have already been traced by social and feminist epistemologists.[59] The subjects at the center of our epistemological theorizing should be embodied, fallible, and embedded in a specific mathematical context. Bringing down the level of idealization at play for the knowing subjects leads us naturally to a fallibilist view of mathematical justification. This, in turn, opens the door to Gettier cases, which, contrary to what one might expect, can be of interest to philosophers of mathematical practice. This is a signal that other unexpected points of contact between the two disciplines could be found.

Moreover, I also stressed that epistemological concerns are better kept separate from metaphysical ones. It is my contention that we can pursue an analysis of

[59] Other relevant approaches in epistemology are ones inspired by the embodied cognition research program, broadly understood to include embedded, enactive, and extended cognition (Shapiro and Spaulding 2021).

mathematical knowledge, even within the traditional epistemological framework, without having to endorse any specific position in the metaphysics of mathematics.

Let me end by listing some specific themes in epistemology that are directly relevant to the philosophy of mathematical practice – themes that I could not discuss for reasons of space. I mentioned that the traditional focus on individual propositional knowledge could be broadened to include different forms of social knowledge as well as an analysis of knowledge how. Moreover, I alluded to the fact that the case of mathematics could be used to inquire about the nature of the basing relation – the relationship that holds between a belief and the reasons on which it is based. There are many other themes in epistemology that have already been applied, or that could be applied to the study of mathematical practice. Here are a few examples: epistemic defeaters (Easwaran 2015), themes in virtue epistemology (Aberdein et al. 2021), epistemic injustice (Rittberg et al. 2020), and disagreement (Aberdein 2023; De Toffoli and Fontanari 2023).

My goal was to lay the ground for more such specific applications and therefore to indirectly encourage them. More generally, I hope that clarifying how we should reconceive the epistemological subject(s) if we want to pursue an inquiry about justification and knowledge in mathematics that is faithful to the practice will dispel some skepticism with respect to the application of epistemology to the philosophy of mathematical practice.

Acknowledgments I am indebted to the following people for helpful feedback: José Ferrirós, Pieranna Garavaso, Valeria Giardino, Andrea Sereni, and Guido Tana.

This research was financed by MUR – Ministero dell'Università e della Ricerca tramite PRIN PNRR Missione 4 "Istruzione e Ricerca" – Componente C2 Investimento 1.1, "Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN)" Finanziato dall'Unione Europea Next GenerationEU (Progetto: "Understanding Scientific Disagreement and its Impact on Society," n. P2022A8F82) – CUP I53D23006880001.

References

Aberdein A (2023) Deep disagreement in mathematics. Glob Philos 33(17):1–27

Aberdein A, Rittberg CJ, Tanswell FS (2021) Virtue theory of mathematical practices: an introduction. Synthese 199(3–4):10167–10180

Avigad J (2021) Reliability of mathematical inference. Synthese 198:7377–7399

Azzouni J (2006) How and why mathematics is unique as a social practice. In: Hersh R (ed) 18 unconventional essays on the nature of mathematics. Springer, pp 201–219

Barton N (manuscript) Mathematical Gettier cases and their implications

Benacerraf P (1973) Mathematical truth. J Philos 70(19):661–679

Bengson J, Moffett MA (2011) Knowing how: essays on knowledge, mind, and action. Oxford University Press

Berto F, Nolan D (2021) Hyperintensionality. In: Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2021/entries/hyperintensionality/

Bird A (2010) The epistemology of science – a bird's-eye view. Synthese 175(S1):5–16

Burgess J (1983) Why I am not a nominalist. Notre Dame J Formal Logic 24(1):93–105

Burgess J (2015) Rigor and Structure. Oxford University Press

Carter J (2019) Exploring the fruitfulness of diagrams in mathematics. Synthese 196:4011–4032

Carter JA, Pritchard D (2015) Knowledge-how and cognitive achievement. Philos Phenomenol Res 91(1):181–199
Cavaillès J (1937) Réflexions Sur le Fondement des Mathématiques. Travaux du IXe Congrès International de Philosophie 6:136–139
Cavaillès J (1938) Methode axiomatique et formalisme. Hermann
Chemla K (2018) The proof is in the diagram: Liu Yi and the graphical writing of algebraic equations in eleventh-century China. Endeavour 42(2):60–77
Chignell A (2018) The ethics of belief. In: Stanford encyclopedia of philosophy. https://plato.stanford.edu/archives/spr2018/entries/ethics-belief/
Christensen D (2004) Putting logic in its place: formal constraints on rational believing. Oxford University Press
Clarke-Doane J (2020) Morality and mathematics. Oxford University Press
De Toffoli S (2021a) Groundwork for a fallibilist account of mathematics. Philos Q 71(4):823–844
De Toffoli S (2021b) Reconciling rigor and intuition. Erkenntnis 86:1783–1802
De Toffoli S (2022) Intersubjective propositional justification. In: Silva P Jr, Oliveira L (eds) Propositional and doxastic justification: new essays on their nature and significance. Routledge, pp 241–262
De Toffoli S (2023) Who's Afraid of Mathematical Diagrams? Philosophers' Imprint 23(1):1–20
De Toffoli S, Fontanari C (2023) Recalcitrant disagreement in mathematics: an endless and depressing controversy in the history of Italian algebraic geometry. Glob Philos 33:1–29
De Toffoli S, Giardino V (2014) Roles and forms of diagrams in knot theory. Erkenntnis 79(3): 829–842
Dean W, Kurokawa H (2014) The paradox of the knower revisited. Ann Pure Appl Logic 165: 199–224
Dummett M (1978) Truth and other enigmas, vol 31. Harvard University Press
Easwaran K (2009) Probabilistic proofs and transferability. Philos Math 17(3):341–362
Easwaran K (2015) Rebutting and undercutting in mathematics. Philos Perspect 29(1):146–162
Ferreirós J (2016) Mathematical knowledge and the interplay of practices. Princeton University Press
Garavaso P (2018) Introduction to epistemology. In: Garavaso P (ed) The Bloomsbury companion to analytic feminism. Bloomsbury, pp 3–16
Gettier EL (1963) Is justified true belief knowledge? Analysis 23(6):121–123
Gilbert M (2004) Collective epistemology. Episteme 1(2):95–107
Gödel K (1964) What is Cantor's continuum problem? In: Benacerraf P, Putnam H (eds) Philosophy of mathematics: selected readings. Prentice-Hall, Englewood Cliffs, pp 258–273
Goldberg S (2018) To the best of our knowledge: social expectations and epistemic normativity. Oxford University Press
Goldman A (1967) A causal theory of knowing. J Philos 64(12):357–372
Goldman A (1979) What is justified belief? In: Justification and knowledge. D. Reidel Publishing, pp 1–23
Greco D, Hedden B (2016) Uniqueness and metaepistemology. J Philos 113(8):365–395
Haack S (1979) Epistemology with a knowing subject. Rev Metaphys XXXIII (2) (130):309–335
Habgood-Coote J (2019) Knowledge-how, abilities, and questions. Australas J Philos 97(1):86–104
Habgood-Coote J, Tanswell FS (2021) Group knowledge and mathematical collaboration: a philosophical examination of the classification of finite simple groups. Episteme:1–27
Hellman G (1989) Mathematics without numbers: towards a modal-structural interpretation, Oxford University Press
Hieronymi P (2006) Controlling attitudes. Pac Philos Q 87(1):45–74
Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press
Kitcher P (1992) The naturalists return. Philos Rev 101(1):53–114
Kornblith H (2012) On reflection. Oxford University Press
Kuhn T (1962) The structure of scientific revolutions. University of Chicago Press
Kusch M (1995) Psychologism. Routledge

Lackey J (2016) What is justified group belief? Philos Rev 125(3):341–396
Lackey J (2020) The epistemology of groups. Oxford University Press
Lakatos I (1976) Proofs and refutations. Cambridge University Press
Leite A (2004) On justifying and being justified. Philos Issues 14:219–253
Lewis D (1996) Elusive knowledge. Australas J Philos 74(4):549–567
Longino H (1998) Feminist epistemology. In: Greco J, Sosa E (eds) The Blackwell guide to epistemology. Blackwell Publishing, Oxford, UK, pp 327–353
Longino H (2002) The fate of knowledge. Princeton University Press
Macbeth D (2010) Diagrammatic reasoning in Euclid's elements. In: van Kerkhove B, van Bendegem JP, de Vuyst J (eds) Philosophical perspectives on mathematical practice. College Publications
Macbeth D (2021) Formal Proofs in Mathematical Practice. In B. Sriraman (Ed.), Handbook of the History and Philosophy of Mathematical Practice. Springer, pp 1–23
Mancosu P (ed) (2008) The philosophy of mathematical practice. Oxford University Press
Manders K (2008) The Euclidean diagram. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, pp 80–133
Netz R (1999) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press
Panza M (2012) The twofold role of diagrams in Euclid's plane geometry. Synthese 186(1):55–102
Pedersen NJLL, Wright C (2018) Pluralist Theories of Truth. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/win2018/entries/truth-pluralist
Peirce CS (1998) The essential Peirce. Indiana University Press, Bloomington
Pettigrew R (2021) Logical ignorance and logical learning. Synthese 198(10):9991–10020
Popper KR (1968) Epistemology without a knowing subject. In: Studies in logic and the foundations of mathematics. Elsevier, pp 333–373
Putnam H (1967) Mathematics without foundations. J Philos 64(1):5–22
Rittberg C, Tanswell F, Van Bendegem JP (2020) Epistemic injustice in mathematics. Synthese 197(9):3875–3904
Shapiro S (2007) The objectivity of mathematics. Synthese 156(2):337–381
Shapiro L, Spaulding S (2021) Embodied cognition. In: The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/win2021/entries/embodied-cognition/
Sher G (2016) Epistemic friction: an essay on knowledge, truth, and logic. Oxford University Press
Silva P Jr (2022) Knowledge-first theories of justification. In: Internet encyclopedia of philosophy. https://iep.utm.edu/kft-just/
Sinaceur H (2019) Jean Cavaillès: Philosophie Mathematique. Vrin
Singer P (1972) Famine, affluence, and morality. Philos Public Aff 1(3):229–243
Soler L, Zwart S, Lynch M, Israel-Jost V (2014) Science after the practice turn in the philosophy, history, and social studies of science (eds: Soler L, Zwart S, Lynch M, Israel-Jost V). Routledge
Sosa E (2007) A virtue epistemology: volume I: apt belief and reflective knowledge. Oxford University Press UK
Stalnaker R (1976) Possible worlds. Noûs 10(1):65–75
Tanswell F (2016) Proof, rigour & informality: a virtue account of mathematical knowledge. PhD, Philosophy, St. Andrews
Tanswell F, Kidd IJ (2020) Mathematical practice and epistemic virtue and vice. Synthese 199(1–2):407–426
Williamson T (2000) Knowledge and its limits. Oxford University Press
Wright C (1992) Truth and objectivity. Harvard University Press
Wright C (1999) Truth: a traditional debate reviewed. In: Blackburn S, Simmons K (eds) Truth. Oxford University Press
Zagzebski L (1994) The inescapability of Gettier problems. Philos Q 44(174):65–73

Agency in Mathematical Practice

Yacin Hamami

Contents

1 Introduction 2906
2 Why Take Agents Seriously 2907
3 Agency and Mathematical Activities 2909
4 Agency and Mathematical Artifacts 2911
5 Agency and Mathematical Proofs 2914
6 Agency and Mathematical Texts 2916
7 Conclusion 2919
References 2921

Abstract

A characteristic feature of the philosophy of mathematical practice is to attend to what people do when they do mathematics. But what does it mean to *do* mathematics? This question raises several issues regarding the nature of action, activity, and agency in mathematical practice. The present chapter reviews contributions in the field that have attempted to theorize about these notions. It begins with some motivations for taking agents seriously in the philosophical study of mathematical practice. The core of the chapter discusses, in turn, what it means to carry out mathematical activities, do things with mathematical artifacts, engage with mathematical proofs, and perform mathematical actions prescribed by mathematical texts. Taken together, the various lines of work reported here provide an initial, but already sophisticated, picture of what it means to do mathematics. The chapter ends with some suggestions for future research on agency in mathematical practice.

Y. Hamami (✉)
Postdoctoral Researcher FNRS, Philosophy Department, University of Liège, Liège, Belgium
e-mail: yacin.hamami@gmail.com

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_48

Keywords

Philosophy of mathematical practice · Agency in mathematical practice · Actions in mathematical practice · Mathematical agents

1 Introduction

Mathematics is something that people do, from the research mathematician trying to prove Goldbach's conjecture to the student in elementary school working on her mathematics homework. But what does it mean to *do* mathematics? One may expect philosophical accounts of the nature of mathematics to provide some clues on this question. And yet, as Kenneth Manders remarked: "[c]urrent philosophical conceptions of mathematics are predominantly agentless" (Manders 2008, p. 118). Paradigmatic examples are the main developments in the foundations of mathematics where the objects of study are the notions of mathematical theories, proofs, and computations. The branches of mathematical logic investigating these notions – model theory, proof theory, and recursion theory – do not usually require appealing to agents doing things.[1] The same observation can be made for most philosophical theorizing on the nature of mathematical objects and mathematical truths.[2] The situation seems different for epistemological considerations aiming to explain how we can possess mathematical knowledge, since knowledge needs to be attributed to an epistemic subject.[3] But here again, it is easy to set aside considerations on actions, activities, and agency and to only focus on the conditions that need to be met to attribute knowledge to an epistemic subject, without describing what a subject needs to do to acquire such knowledge in the first place.

One way the philosophy of mathematical practice[4] departs from traditional approaches in the philosophy of mathematics is precisely by taking into consideration the *active* dimension of mathematics, i.e., by attending to what people *do* when they do mathematics. This has become a leitmotif in describing the field, for instance:

> The focus has turned thus to a consideration of what mathematicians are actually doing when they produce mathematics. [. . .] These historians and philosophers agree that there is more to understanding mathematics than a study of its logical structure and put much emphasis on mathematical activity as a human activity. (Mancosu et al. 2005, p. 1)

[1] Notable exceptions are logicians that take into account the active dimension of deductive inference, e.g., Frege (1879), Gentzen (1934), Prawitz (1965), and Martin-Löf (1984).

[2] Here again, there are exceptions, for instance, works from the precursors of the philosophy of mathematical practice – e.g., Wilder (1950), Lakatos (1976), and Kitcher (1984). Also, philosophical views that take mathematical objects and facts as being the result of social constructions – e.g., Hersh (1997), Feferman (2009), and Cole (2013) – may implicitly or explicitly appeal to mathematical agents. On social construction and constitution in mathematics, see the chapter by Cantù in this handbook (Cantù 2023).

[3] On the conceptualization of epistemic subjects in the context of mathematics, see the chapter by De Toffoli in this handbook (De Toffoli 2023).

[4] For recent overviews of the field, see Carter (2019) and Hamami and Morris (2020).

> The philosophy of mathematics has experienced a renewal in recent years due to a more open and interdisciplinary way of asking and answering questions. Traditional philosophical concerns about the nature of mathematical objects and the epistemology of mathematics are combined and fructified with the study of a wide variety of issues about the way mathematics is done, evaluated, and applied [. . .]. (Ferreirós 2016, p. xi)

Once it has been acknowledged that the active dimension of mathematics is a proper object of philosophical inquiry, the next step is to figure out how to study it. This raises methodological issues on how to investigate the whole range of actions and activities present in mathematical practices as well as the different forms of agency involved.[5] It also raises theoretical questions regarding the nature of actions, activities, and agency in the specific context of mathematical practice.

The aim of this chapter is to review various ways in which the notions of action, activity, and agent have played a role in philosophical studies of mathematical practice. The focus will be on contributions that have attempted to theorize about these notions. We begin, in Sect. 2, with some motivations for taking agents seriously when investigating mathematical practice. Section 3 is concerned with the notion of mathematical activity, both as an object of study in itself and as a way of building an epistemology of mathematics where mathematical knowledge is rooted into elementary mathematical activities. In Sects. 4 and 5, we will see how the notions of action and agent have played an important role in the philosophical study of mathematical artifacts and mathematical proofs. Section 6 will be concerned with mathematical agency and mathematical texts, and more specifically with how mathematical actions can be prescribed by mathematical texts. Taken together, these contributions already provide important clues on what it means to do mathematics. Section 7 concludes the chapter with some suggestions on how to pursue further the study of agency in mathematical practice.

2 Why Take Agents Seriously

An obvious reason to take agents seriously is that virtually all the mathematical practices one may be interested to investigate would not exist without agents engaging in it at a given point in human history. As Ferreirós emphasized: "[m]athematical practice always involves communities of agents in interaction" and "there is no practice without practitioners" (Ferreirós 2016, p. 45).[6] One may then expect that the notion of agent will, one way or another, figure in any philosophical analysis of mathematical practice.

[5] The nature of actions, activities, and agency are investigated in the branch of philosophy known as the *philosophy of action* (for overviews, see Schlosser 2019; Piñeros Glasscock and Tenenbaum 2023). The theme of this chapter naturally invites connections between the philosophy of mathematical practice and the philosophy of action. So far, interactions between the two fields have been rare, but see Chemla and Virbel (2015) and Hamami and Morris (2021, 2022a) for some attempts to connect the two.

[6] On Ferreirós' view of what a mathematical practice is, see also Ferreirós' chapter in this handbook (Ferreirós 2022).

Still, one may object that the essence of a mathematical practice lies in the norms or rules that are constitutive of it, and not in its particular, contingent implementation(s). From this perspective, one could investigate a given mathematical practice independently of its implementation(s), in the same way as one could study a judiciary system or an algorithm which may never have been implemented concretely. In opposition to this view, Manders (2008) has argued that, in the context of his analysis of diagram-based geometric reasoning in ancient Greece, taking into account different *roles* for the agents engaged in this practice was important in explaining its *success*. Manders distinguishes two roles that he considers of theoretical importance (Manders 2008, p. 122): that of *asserting* the different geometrical claims made in the text and taking responsibility for them, and that of providing critical scrutiny or *probing*. One way of probing a geometrical claim is to propose a *case*, that is, a diagram which satisfies the stipulations of the claim but which is topologically different from the diagram(s) considered by the protagonist. According to Manders, proposing cases is theoretically significant because it is an essential strategy to cope with the open-ended nature of case-branching in diagram-based geometric practice, i.e., with the fact that such a practice does not possess a procedure to make sure that all topologically different diagrams that are relevant in a geometric proof have been considered. As evidence for the importance of the role of probing, Manders argues that several propositions in Euclid's *Elements* can be conceived as responses to probing claims that could potentially be raised with respect to other geometrical proofs.

Another reason to take agents seriously is highlighted by cases where the mathematics themselves are primarily about actions. Typical examples are mathematical practices geared towards geometric constructions or concrete arithmetical/algorithmic procedures. For instance, ancient Greek geometry has been described by Reviel Netz as "the study of spatial action, not of visual representation" (Netz 1999, p. 60). Some mathematical traditions, such as those from Mesopotamia or ancient China, are primarily concerned with procedures for computations which are, by definition, set of actions to be performed (see, e.g., Chemla 2012). When analyzing mathematical practices of this kind, one may naturally be led to reflect on the nature of the actions involved and on the structure of the agents performing them.[7]

But one of the main reasons to take agents seriously is that the notions of action, activity, and agent have revealed themselves theoretically useful and fruitful when addressing a number of issues that appear prominently in the research agenda of the philosophy of mathematical practice. As we shall see in the following sections, this bundle of notions occupies an important place in philosophical work on mathematical activities, mathematical artifacts, mathematical proofs, and mathematical texts.

[7] We can also mention here practices that could be qualified as mathematical such as string figure-making (see, e.g., Vandendriessche 2015) or paper folding (see, e.g., Friedman 2018), though of course whether such practices can be qualified as mathematical is a debated question.

3 Agency and Mathematical Activities

Focusing on what it means to do mathematics amounts to seeing mathematics as an activity. In this section, we will be concerned with two lines of research that naturally emerge from this perspective. The first one consists in developing an epistemology of mathematical activities. The second one explores the possibility that mathematical knowledge is rooted or grounded into elementary human activities.

In a contribution entitled "Mathematical activity," Giaquinto (2005) started from the observation that, although philosophers of mathematical practice have set for themselves the task of investigating the activities that mathematicians engage in, they have so far focused on a narrow set of activities. Giaquinto proposed then a "preliminary map" of mathematical activities worthy of philosophical investigations, recognizing that this list is not meant to be exhaustive. His list consists of the following mathematical activities: discovery, explanation, formulation, application, justification, and representation. Once such a list is established, numerous questions arise: What is the nature of these activities? Are they really different kinds of activities? What are the relations between them? Can they be decomposed into smaller activities? Giaquinto tackled some of these questions. In particular, he argues that discovering, explaining, and justifying are distinct kinds of activities which do not reduce to proving. He also proposes that, for each of the activities in the list, one can distinguish between making it, presenting it, and taking it in. For instance, in the case of discovering, one can distinguish between making the discovery, presenting the discovery to other people, and being the one who is trying to "get" the discovery, that is, who is receiving or taking in the discovery. Giaquinto conceives of his contribution as a "springboard" to pursue further philosophical investigations into the nature of mathematical activities.

Some authors have argued that the origins, grounds, or roots of mathematical knowledge are to be found in certain elementary human activities. The mathematician Saunders Mac Lane put the general idea as follows[8]:

> [M]athematics started from various human activities which suggest objects and operations (addition, multiplication, comparison of size) and thus lead to concepts (prime number, transformation) which are then embedded in formal axiomatic systems (Peano arithmetic, Euclidean geometry, the real number system, field theory, etc.). These systems turn out to codify deeper and nonobvious properties of the various originating human activities. (Mac Lane 1981, p. 463)

Mac Lane offered the following list of correspondence between elementary human activities and domains of mathematics:

Counting	: to arithmetic and number theory
Measuring	: to real numbers, calculus, and analysis
Shaping	: to geometry and topology
Forming (as in architecture)	: to symmetry and group theory

[8] See also Mac Lane (1986, chapter XII).

Estimating	: to probability, measure theory, and statistics
Moving	: to mechanics, calculus, and dynamics
Calculating	: to algebra and numerical analysis
Proving	: to logic
Puzzling	: to combinatorics and number theory
Grouping	: to set theory and combinatorics

There are different ways in which this general idea can be exploited to articulate a view of mathematics and mathematical knowledge as rooted or grounded into elementary human activities. I will now describe two of the most developed proposals in this direction, namely those of Kitcher (1984) and Ferreirós (2016).

In chapter 6 of *The Nature of Mathematical Knowledge*, Kitcher (1984) endeavors to develop an account of mathematical truth according to which mathematical statements are true in virtue of certain operations or manipulations humans can perform in the world. In the case of arithmetic, two such operations are collecting and correlating: collecting is the activity of segregating or putting objects together; correlating is the activity of matching, relating, or connecting objects. According to Kitcher, the first type of operation leads to notions like set, number, and arithmetical operations, while the second type leads to notions like that of a function. Thinking in terms of mathematical activities such as collecting and correlating is key to Kitcher's mathematical ontology: "One central idea of my proposal is to replace the notions of abstract mathematical objects, notions like that of a collec*tion*, with the notion of a kind of mathematical activity, collec*ting*" (Kitcher 1984, p. 110). Kitcher emphasizes that the operations or manipulations he has mind are not those of any given human agent but that of an ideal agent. This ideal agent should be thought as an idealized version of human agents. Kitcher conceives then of arithmetic as an "idealizing theory": "the relation between arithmetic and the actual operations of human agents parallels that between the laws of ideal gases and the actual gases which exist in our world" (Kitcher 1984, p. 109). This approach allows Kitcher to argue that arithmetic constitutes a description of the structure of reality, that is, of some operations that can be done in the physical world. In a slogan, Kitcher says that arithmetic is true: "in virtue not of what *we can do* to the world but rather of what *the world* will let us do *to it*" (Kitcher 1984, p. 108). A main goal of *The Nature of Mathematical Knowledge* is then to show how advanced mathematics can emerge from such proto-mathematical knowledge through various stages of rational transitions.

Like Kitcher, Ferreirós holds that an account of mathematical knowledge should attribute a key role to certain elementary activities or practices: "I shall argue that our knowledge of mathematics cannot be understood without emphasizing the practical roots of math, including its roots in scientific practices and technical practices" (Ferreirós 2016, p. 5). But contra Kitcher, Ferreirós' approach is not reductionist – central to Ferreirós' account is the "interplay" or "interactions" between practices, not the transitions between different mathematical practices. Another important difference with Kitcher is that Ferreirós emphasizes the cognitive dimension of elementary mathematical practices and the fact that the cognitive abilities involved must be reenacted for an individual to gain mathematical knowledge, while Kitcher situates elementary mathematical practices at the historical roots of his genealogical account.

Ferreirós also offers a precise definition of what he calls a "technical practice," namely: "a recognizable type of activity that is done—and can be taught and learned—by human agents, involving direct manipulation of objects in the world, through the use of human-made instruments" (Ferreirós 2016, pp. 40–41). The three technical practices discussed by Ferreirós (2016) are that of counting, measuring, and drawing geometrical forms. An important thesis in Ferreirós' view is that knowledge of mathematics requires the mastering of these elementary techniques.

The notion of mathematical activity, intimately connected with that of a mathematical practice, is bound to play an important role in the philosophy and history of mathematical practice. As a matter of fact, many studies in the field precisely consist in the analysis of mathematical activities. The focus in this section has been on philosophical developments that have attempted to theorize about the notion of mathematical activity itself and/or to use it in the pursuit of larger goals. Here, much is to be gained by pursuing further the inquiry initiated by Giaquinto (2005). In particular, it would be theoretically and practically useful to identify and characterize the main components of what constitute mathematical activities as well as the mathematical agents able to perform them.[9] This could not only yield methodological tools to investigate specific mathematical activities, it could also provide a perspective to compare and relate different mathematical activities – a theme central to approaches like the one of Ferreirós which emphasize the interplay between mathematical practices, and thus between mathematical activities. Possessing a rich and precise account of the nature of mathematical activities will, in turn, be directly useful for building philosophical conceptions of mathematics rooted in mathematical activities.

4 Agency and Mathematical Artifacts

Doing mathematics involves doing things with mathematical artifacts such as diagrams, symbols, and graphs. A significant part of the literature in the philosophy and history of mathematical practice is dedicated to understanding and reconstructing what agents did or do with mathematical artifacts in various mathematical practices, from the use of geometric diagrams in Ancient Greece mathematics (e.g., Netz 1999; Manders 2008) to that of commutative diagrams in contemporary homological algebra (e.g., De Toffoli 2017). The aim of this section is not to review this literature. Rather, the focus will be on contributions that have attempted to conceptualize and theorize about what it means to do things with mathematical artifacts. This line has been pursued mainly in the cases of mathematical diagrams and mathematical symbols,[10] and I will review here some representative studies in this trend. Before

[9] As an example, Carter (2008) argues that investigating the activities that mathematicians perform with mathematical structures provides a new perspective on mathematical structuralism.

[10] These two kinds of mathematical artifacts are some of the most common ones in mathematical practices past and present, which is probably why they have received most of the attention in the literature. But there are also other kinds of mathematical artifacts that invite epistemic actions, for instance abacuses, geometric instruments such as ruler and compass, and computer simulations.

that, I will present the notion of *epistemic action* from Kirsh and Maglio (1994) which has played a central role in these discussions.

The distinction between pragmatic and epistemic actions was introduced by Kirsh and Maglio (1994) in their psychological study of the video game Tetris. Pragmatic actions are physical actions performed in the world whose aim is to bring a physical system closer to a specific goal. In Tetris, the pragmatic actions are those aiming to orient the falling shapes in order to create full rows, the goal of the game being to fill up rows. Epistemic actions are also physical actions performed in the world, but their goal is different: their aim is to gain information by performing computations externally. In Tetris, the epistemic actions consist in rotating the shape in order to decide where and in which orientation to position a falling shape. The main advantages of relying on epistemic actions in Tetris is that rotating the shape in the game is less cognitively costly and more reliable than trying to mentally rotate the shape in the mind. Kirsh and Maglio argue that epistemic actions play an important role in many human activities, and that acting in the environment to gain information, knowledge, and understanding is an essential part of human cognition.

One area of contemporary mathematics where diagrams are omnipresent is knot theory.[11] De Toffoli and Giardino (2014) have conducted a detailed analysis of the use of knot diagrams in the practice of knot theory. They have argued that knot diagrams are entities that support epistemic actions such as computations and inferences. These epistemic actions require a certain cognitive faculty on the part of the agent that they called "manipulative imagination" which, they suggest, is akin to the sort of concrete manipulations one can perform on deformable objects. The moves that can be performed on knot diagrams are codified in the mathematical theory. Accordingly, agents need to be appropriately trained not only to recognize the possible or legitimate moves one can perform on a knot diagram but also to be able to perform these moves. De Toffoli and Giardino have further argued that this manipulative imagination also plays a role in the practice of low-dimensional topology (see De Toffoli and Giardino 2015). Here again, the agent must be able to recognize the permissible actions that can be performed on a given representation and so need to be appropriately trained to do so. To show this, they have conducted a detailed analysis of Rolfsen's proof of the equivalence of two presentations of the Poincaré homology sphere. In this particular case, some of the permissible actions consist of continuous transformations. According to De Toffoli and Giardino, this faculty of manipulative imagination builds on preexisting spatial and motor cognitive capacities but needs to be appropriately trained to play its epistemic function in the practice of knot theory and low-dimensional topology.

Mathematical symbols are artifacts present in virtually every branch of mathematics. De Cruz and De Smedt (2013) have made the case that a primary function of mathematical symbols is to support epistemic actions. These epistemic actions consist in typical symbol manipulations such as those related to negative and

[11] Knot theory is a branch of topology that studies knots – a mathematical knot is defined as an embedding of a circle in $\mathbb{R}^3$.

imaginary numbers. These symbol manipulations are governed by specific rules associated to the relevant symbolic systems, rules that will need to be learned by the agents. Once acquired, they will provide the agents with an efficient way to perform various forms of mathematical thinking. De Cruz and De Smedt go even further and argue that mathematical symbols support forms of mathematical cognition that would not be possible without resorting to external representations, a view that they articulate by building on the so-called extended mind thesis (Clark and Chalmers 1998; Clark 2008). Their main source of evidence for this claim comes from the history of mathematics. One of their key examples is the case of negative numbers which have received strong oppositions in the past by distinguished mathematicians such as Vieta, Pascal, and De Morgan. De Cruz and De Smedt argue that the invention of the negative numbers, which appears as counterintuitive, was made possible because of the symbolization associated to the minus sign and the operation of subtraction, a symbolic system with which one can calculate.

This perspective naturally raises the question of what form of cognition underlies thinking and acting with mathematical symbols. Landy et al. (2014) have proposed a cognitive theory of symbolic reasoning – called the *perceptual manipulations theory* – that attributes central roles to perception and action. According to this theory, external symbols and notations are treated like physical objects by the cognitive system, that is, objects that can be perceived and manipulated. Symbolic reasoning would then rely on a wide range of sensorimotor abilities such as affordance learning, pattern-matching, object tracking, symmetry detection, etc. This theory could explain the importance of careful and effective design of mathematical notations. More specifically, a well-designed notational system is, from this perspective, one that can take advantage of the sensorimotor system capacities in order, for instance, to encourage valid manipulations and refrain invalid one, or to facilitate the application of structurally similar rules from one domain to another.

Giardino (2018) has advanced an encompassing framework aiming to account for different types of mathematical artifacts, including both mathematical diagrams and mathematical symbols. Her proposal is to conceive mathematical artifacts as representational cognitive tools whose main characteristic is to play the double function of representation and instrument. As instrument, the main function of representational cognitive tools is to carry out inferences, i.e., epistemic actions. Giardino offers an account of how epistemic actions operate on representational cognitive tools by building on the notions of *material anchor* from Hutchins (2005) and *affordances* from Gibson (1979). Thinking of representational cognitive tools as material anchors amounts to considering them as material entities which have been designed so that constraints associated to their epistemic functions, as well as to what they represent conceptually, are built in the artifacts themselves. The idea of affordances is that the epistemic actions than can be performed on a given representational cognitive tool are those that are afforded by the tool. These affordances depend on the context or practice in which the cognitive tool is used, which means that agents need to be trained to recognize these affordances, a necessary prerequisite to be able to perform legitimate epistemic actions with the tool.

What these different contributions show is that accounting for what we do with mathematical artifacts in mathematical practice naturally leads to considerations of mental and epistemic actions, and thus of mental and epistemic agency, and requires paying attention to the cognitive underpinnings of the mental activities under consideration. There is thus room on these issues for fruitful interactions between the philosophy of mathematical practice and related developments in the philosophy of mind, philosophy of action, and cognitive science.

5 Agency and Mathematical Proofs

Interacting with mathematical proofs[12] is an essential part of what mathematical agents do. This includes evaluating, verifying, communicating, explaining, understanding, and reframing mathematical proofs, among others. But mathematical proofs themselves can also be conceived as being primarily about action. Hamami and Morris (2021) proposed the correspondence displayed in Fig. 1 between the static notions of deductive step and mathematical proof and the dynamic notions of deductive inference and proof activity which belong to the realm of action. That deductive inferences are first and foremost actions of an epistemic nature has been emphasized by several logicians and philosophers (see, e.g., Sundholm 2012; Prawitz 2012; Boghossian 2014; Wright 2014). Hamami and Morris introduced the term *proof activity* to refer to the sequence of deductive inferences corresponding to a mathematical proof. In this section, we will review contributions that propose an action-based perspective on mathematical proofs, focusing in turn on the notions of deductive inference and proof activity.

Larvor (2012) has argued that, in order to better understand the nature of informal proofs, it is not sufficient to approach them as bodies of propositions connected in specific ways. Rather, one should see a proof as a sequence of dynamic transitions, that is, one should focus on *inferential actions*. This change of perspective is particularly fruitful because there is no reason to think of inferential actions has

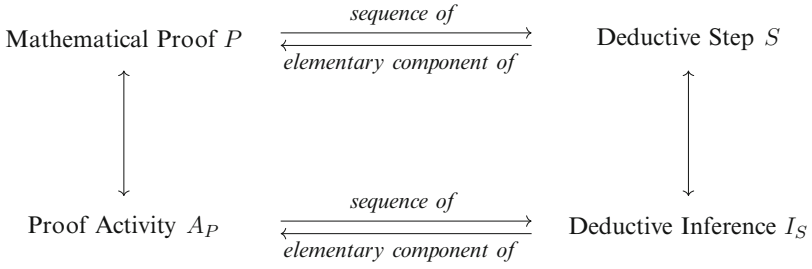


Fig. 1 Mathematical proof, proof activity, deductive step, and deductive inference

[12] By "mathematical proofs" I mean here the kind of proofs one commonly finds in mathematical practice, also commonly referred to as "informal proofs."

being restricted to linguistic representations – inferential actions can operate on different kinds of representations such as diagrams, symbolic expressions, physical models, computer models, etc.[13] Larvor builds on this observation to argue that the validity or invalidity of informal proofs can depend on their content. The main reason being advanced is that inferential actions are not always possible in all domains; only inferential actions that consist of logical inferences are. This means that the validity of these inferential actions must depend somehow on features of the domain under consideration. Larvor argues that inferential actions that are context-dependent can perfectly be rigorous insofar as they come with means of control which govern the inferential actions that are possible and legitimate in a given domain. In this respect, Manders' analysis of Euclid's diagram-based geometric practice (Manders 2008) is an archetypal example of a study aiming to make explicit the means of control governing inferential actions operating on text and diagram in the context of Euclid's geometric proofs. Larvor points out that this perspective shapes and organizes a research program on informal proofs in the philosophy of mathematical practice: the task is to identify the various types of inferential actions, to isolate the features of the domain on which they operate that are responsible for their validity, and to identify the means of control that govern them.[14]

If it is particularly fruitful to approach individual deductive steps as inferential actions, then it must be equally fruitful to approach entire mathematical proofs as sequences of inferential actions, i.e., as *proof activities* (see Fig. 1). Hamami and Morris (2021) have undertaken a detailed analysis of proof activities and the form of agency underlying them. Their investigations start from the observation that proof activities have two noticeable characteristics, namely they are goal-directed and temporally extended activities. They are goal-directed in the sense that any proof activity is always directed towards a specific goal, namely to prove or establish the mathematical proposition at hand. They are temporally extended in the sense that any action in a proof activity depends on what happened before and constraints what will happen next, that is, any action is fully integrated into the temporal structure of the activity taken as a whole. In this respect, proof activities are similar to many of our ordinary activities such as travelling or cooking which are obviously goal-directed and which unfold over time. Now, the philosopher of action Michael Bratman (1987) has observed that, for human agents, the realization of goal-directed and temporally extended activities most often requires a form of planning agency. This means that the agents engaged in goal-directed and temporally extended activities are guided by a plan that they construct, revise, and execute over time. In the case of proof activities, this means that the sequence of actions does not come from nowhere but is the result of the execution of a plan that has been constructed rationally. This has some interesting consequences for the epistemology of mathematical proofs in practice. First, it means that mathematical proofs possess what Poincaré (1908) called an *architecture* or a

[13] There is here a direct connection with the topic of agency and mathematical artifacts discussed in the previous section.

[14] This line is further pursued in Larvor (2019).

unity that can be grasped as a whole (see also Detlefsen 1992). This is due to the fact that a mathematical proof is the result of the execution of a plan produced by a mathematical agent, in the same way as a building is the result of executing the plan of an architect. Second, because proof plans are rational constructs, this also means that mathematical proofs produced by rational planning agents possess what Mac Lane (1935) called a *rational structure* – a mathematical proof in practice is not a "mere" sequence of deductive steps. Hamami and Morris (2022a) have argued that this notion of rational structure can be fleshed out by identifying the norms of rationality governing the construction of proof plans. Third, if we accept proof plans as meaningful epistemological entities, then we may expect that some of the activities carried out with proofs will have to do with proof plans. Indeed, several authors have pointed out an intimate connection between proof plans and proof understanding (e.g., Poincaré 1908; Robinson 2000; Folina 2018). Following along this line, Hamami and Morris (2022b) have advanced an account of proof understanding according to which understanding a mathematical proof amounts to being able to rationally reconstruct the proof's underlying plan. Finally, Hamami and Morris (2021) have also argued that proof plans play an important role in the activities of presenting and communicating mathematical proofs.

The action-based perspective on mathematical proofs fits very naturally with the general idea of approaching mathematics as a human activity. As we saw in this section, the static and agentless notions of deductive steps and mathematical proofs can profitably be analyzed through their active counterpart, namely the notions of deductive inferences or inferential actions and proof activities. This perspective is particularly suited to study the spectrum of activities carried out in interaction with mathematical proofs. As an illustration, we saw here that Larvor's view on inferential actions naturally leads to an account of what it means to evaluate the validity of mathematical proofs, while Hamami and Morris' analysis of proof activities can be used to shape an account of what it means to understand, present, and communicate mathematical proofs. But these are only some of the many activities related to mathematical proofs. An obvious research program to foster this action-based perspective will be to pursue and extend the analysis of these different activities. This, in turn, can shed light on the nature of agency involved when interacting with mathematical proofs.

6 Agency and Mathematical Texts

Mathematical texts are a privileged vehicle for the transmission of mathematical knowledge.[15] If we approach the nature of mathematics through the lens of mathematical practice, a key issue is then to understand how mathematical agents can

[15] This is not to say that mathematical knowledge cannot be transmitted by other means. Agent-to-agent communications through oral transmissions, eventually accompanied by gestures and external representations, have been historically dominant in some mathematical cultures and are certainly very important in present-day mathematical practice. Thanks to José Ferreirós for suggesting this clarification.

realize mathematical actions and activities on the basis of mathematical texts. In this section, we will review studies in the field that address exactly this question. As we shall see, such investigations can yield important insights on key questions in the philosophy of mathematical practice.

Computing and proving are two archetypical mathematical activities. The former is central to certain mathematical traditions such as those of Mesopotamia, ancient China, and the Indian subcontinent, while the latter has been privileged in other traditions such as that of ancient Greece (Chemla 2012). Some commentators have downplayed the activity of computing as compared to that of proving for several reasons (e.g., Hacking 2000), a main one being that computations can be carried out blindly, in a step-by-step fashion, without any form of mathematical understanding. Chemla (2015) argues against this view by undertaking a detailed investigation of how texts of mathematical procedures in ancient China prescribe mathematical actions. The focus is on two key mathematical texts, namely the *Writings on mathematical procedures* and *The Nine Chapters on Mathematical Procedures*. Her analysis yields two important results. The first one is that procedures in these texts cannot be executed in a step-by-step fashion because the specification of certain steps require to attend to other instructions appearing later on in the sequence. This means that a certain form of knowledge is required to "circulate" – Chemla's term – within these texts of procedures. This circulation is necessary to infer from the text the actual sequence of actions to be carried out in specific cases or situations. Chemla argues that this feature witnesses to the generality of the procedures – i.e., the fact that the same procedure can be applied to a wide range of different cases and situations. It is thus essential that the reader be able to infer from the text which actions to take in her specific situation, a competence that goes beyond a blind step-by-step following of the procedure. The second result is that, to infer the sequence of actions to be carried out, the reader needs a certain understanding of the reason(s) why certain steps are carried out. This follows from a meticulous analysis of what it takes to interpret steps in procedures that involve the term "likewise," that is, steps that require an understanding of what exactly is to be replicated. Taken together, this analysis of the way actions are prescribed in texts of mathematical procedures challenges, according to Chemla, the view that computations are mathematical activities that are carried out blindly without any understanding.

Tanswell (forthcoming) has offered an analysis of imperatives in mathematical proofs – introduced by terms such as "let," "assume," "solve," "observe," etc.[16] – which, he argues, yields important insights into the nature of informal proofs. Imperatives issue instructions or commands to carry out mathematical actions. Tanswell identifies three kinds of imperatives in mathematical proofs. The first kind are imperatives which directly refer to the activity to be carried out, for instance, "solve" this equation, "differentiate" or "integrate" this function, "multiply" these two matrices, etc. The second kind are imperatives corresponding to standard

[16] See the chapter by Inglis and Tanswell in this Handbook (Inglis and Tanswell 2022) for a corpus analysis of imperatives and instructions as they occur in mathematical texts.

instructions, but for which some information is left implicit, such as an "assume" clause without stating exactly what the goal is (e.g., to establish a conditional, to reach a contradiction, to do a reasoning by cases, etc.). The third kind are imperatives whose formulation does not correspond exactly to what is to be done and, for this reason, require background and expertise to be interpreted appropriately. As an example, Tanswell mentions a situation where one is told to "using the Axiom of Choice, for each *n* choose an enumeration," while "the whole point of the Axiom of Choice is that we cannot actually go about choosing infinitely many times" (Tanswell forthcoming, p. 7). Acknowledging the importance of imperatives in mathematical proofs motivates, according to Tanswell, what he calls the *recipe model* of informal proofs.[17] The model is introduced through a number of analogies between proofs and ordinary cooking recipes. It is noted that proofs, like cooking recipes, (1) employ the imperative mood, (2) are secondary to the associated activities, and (3) involve a clear distinction between the process the author(s) went through to produce them and the way it is being used by the readers or consumers. Tanswell argues that this recipe model of informal proofs is particularly adapted to account for the role of diagrams in proofs for the main reason that diagrams are used to provide instructions in many contexts, and so there is no particular reason to see diagrams as more problematic than texts when it comes to yield instructions.

The studies by Chemla (2015) and Tanswell (forthcoming) show that much is to be learned by conducting detailed analyses of how mathematical texts can lead to mathematical actions and activities. Here, we have seen that this can provide evidence for revising received views on the relation between computation and proof as well as on the nature of informal proof.[18] But, more generally, such investigations are a privileged way to better understand mathematical activities themselves. In fact, for many (most?) mathematical traditions, mathematical texts are all we have. A key task for the historians of mathematics is thus to reconstruct past mathematical activities on the basis of the mathematical texts that came down to us. As Chemla's study illustrates here, this should not only consist of identifying the actions constituting these activities, it should also amount to characterizing the knowledge, competence, and understanding that mathematical agents need to possess to carry out the relevant mathematical activities. From this perspective, the issue of mathematical agency necessarily arises whenever one undertakes to analyze or reconstruct mathematical activities, and more generally mathematical practices, on the basis of mathematical texts.[19]

[17] See Weber and Tanswell (2022) for further developments of this model and its applications to issues in mathematics education.

[18] In this respect, the studies by Chemla and Tanswell fit perfectly within the action-based perspective on mathematical proofs discussed in the previous section.

[19] Here, mathematical texts should be understood in the broad sense encompassing all the different kinds of representations used in mathematical communication, namely texts but also diagrams, symbols, graphs, etc. There is then a direct connection between the theme of this section and the ones developed in Sect. 4 on agency and mathematical artifacts.

7 Conclusion

What does it mean to *do* mathematics? Doing mathematics obviously involves doing a wide range of different things, and so the only way to progress on this question is to decompose the problem into chunks amenable to philosophical investigations. As we saw in this chapter, philosophers and historians of mathematical practice have made significant progress in this direction by investigating what it means to carry out mathematical activities, do things with mathematical artifacts, engage with mathematical proofs, and perform mathematical actions prescribed by mathematical texts. But this may only be the tip of the iceberg, and it is likely that the question of agency in mathematical practice covers a vast research territory that remains to be uncovered and explored. In this conclusion, I will suggest potential avenues for future work in this direction.

Perhaps one of the most pressing issues is to reflect on what mathematical agents are, a problem that should be addressed for every specific mathematical activity and practice. One of the few authors who have tackled this question directly is Ferreirós (2016), the notion of mathematical agent being central to the philosophical view he advances.[20] For Ferreirós, mathematical agents are first and foremost human agents with ordinary cognitive abilities and limitations, and immersed in a physical, social, and cultural world. Ferreirós argues that all those aspects are relevant to his account of mathematical knowledge and mathematical practices. Taken together, the contributions reviewed in this chapter already highlight many salient aspects of mathematical agents. In addition to the features identified by Ferreirós, mathematical agents are also planning agents (Hamami and Morris 2021, 2022a), powered with a faculty of manipulative imagination (De Toffoli and Giardino 2014, 2015, 2016), capable of performing inferential actions (Larvor 2012, 2019), grasping and understanding the reasons behind mathematical actions (Chemla 2015), and following recipes (Tanswell forthcoming; Weber and Tanswell 2022). A challenge for the philosophy of mathematical practice is to develop an account of mathematical agents capable of articulating these different dimensions in a systematic way.

Research on agency in mathematical practice has so far focused mainly on individual human agency, but several other forms of agency are present in mathematical practice – the most obvious ones are social agency, computer agency, and extended agency. Social agency is present whenever mathematical agents are doing things together, typically when several mathematicians are collaborating to prove a theorem. This raises the question of what it means for a *group* of mathematical agents to do mathematics *together*. As philosophers of action have shown,[21] social agency most often require subtle mechanisms of coordination and interaction. Investigating these

[20] In Ferreirós' words: "The thesis that multiple practices coexist and are interrelated, in a way that is crucial for the conformation of mathematical knowledge, has one interesting consequence. It means that my analysis of mathematical knowledge has to be crucially centered on the agents" (Ferreirós 2016, p. 59).

[21] See Roth (2017) for a review of the literature on social agency in the philosophy of action.

mechanisms in the context of mathematical practice will certainly lead important insights into the social dimension of mathematical practice. Additionally, it may be interesting to investigate the agency displayed by the mathematical community or subgroups thereof, for instance, when attributing credits and rewards (see, e.g., Jaffe and Quinn 1993; Rittberg et al. 2020). Similarly, the increasing role of computers in mathematical practice raises the questions of what it means *for* computers to do mathematics, if at all, and what it means to do mathematics *with* computers.[22] These issues pop up concretely whenever one investigates cases of proving or discovering where computers are involved in an essential way – e.g., in the computer-assisted proof of the four color theorem. Finally, the many studies on mathematical artifacts in mathematical practice raises the question of what exactly the agent who does mathematics consists in. When a mathematical activity is carried out by a human agent relying on some mathematical artifacts, what is doing mathematics is a sort of integrated system composed of the coupling of the agent and the artifacts, in which case the mathematical agent may better be conceived as an extended agent. Such an approach could be developed, for instance, in the line of the cyborg conception of agency proposed by Clark (2003), following the famous extended mind thesis much discussed in the philosophy of mind (Clark and Chalmers 1998). Another challenge for philosophers and historians of mathematical practice is thus to identify and characterize the different forms of agency at play in mathematical practice.

The question of agency in mathematical practice is both a theoretical and a methodological issue for the philosophy of mathematical practice. It is a theoretical issue in the sense that any account of what it means to do mathematics may be expected to say something about the agents doing mathematics. But it is also a methodological issue in the sense that our underlying conception of mathematical agency guides, in part, our investigations into specific mathematical activities and practices. More specifically, it provides us with a "template" to address questions such as: What does this mathematical activity or practice consists in? What does it take for an agent to be able to properly engage in it? These two aspects are intimately intertwined: progress on the theoretical front may yield new tools and perspectives to investigate specific mathematical activities and practices;[23] in turn, these investigations may yield new empirical data to constrain and revise our theoretical conceptions. Given that the nature of agency is one of the outstanding issues spanning the Humanities, progress in the case of mathematical agency can certainly be made by recruiting concepts and resources from other fields, both within philosophy – especially from the philosophy of action, philosophy of mind, epistemology, and the philosophy of science – but also from other areas of the social sciences such as sociology and anthropology. Understanding the nature of agency in mathematical practice is thus likely to require the full interdisciplinary perspective that was claimed as a characteristic feature of the philosophy of mathematical practice.

[22] See Avigad (2008) for some investigations in this direction.

[23] A perfect illustration of this is provided by the studies in Chemla and Virbel (2015) which build, in part, on the theoretical framework developed by Virbel and colleagues (Grandaty et al. 2000; Virbel 2000).

References

Avigad J (2008) Computers in mathematical inquiry. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 302–316

Boghossian P (2014) What is inference? Philos Stud 169(1):1–18

Bratman ME (1987) Intention, plans, and practical reason. Harvard University Press, Cambridge, MA

Cantù P (2023) The social constitution of mathematical knowledge: objectivity, semantics and axiomatics. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer, Cham

Carter J (2008) Structuralism as a philosophy of mathematical practice. Synthese 163(2):119–131

Carter J (2019) Philosophy of mathematical practice: motivations, themes and prospects. Philos Math 27(1):1–32

Chemla K (ed) (2012) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge

Chemla K (ed) (2015) Proof, generality and the prescription of mathematical action: a nano-historical approach to communication. Centaurus 57(4):278–300

Chemla K, Virbel J (eds) (2015) Texts, textual acts and the history of science. Springer, Cham

Clark A (2003) Natural-born cyborgs: minds, technologies, and the future of human intelligence. Oxford University Press, Oxford/New York

Clark A (2008) Supersizing the mind: embodiment, action, and cognitive extension. Oxford University Press, Oxford/New York

Clark A, Chalmers D (1998) The extended mind. Analysis 58(1):7–19

Cole J (2013) Towards an institutional account of the objectivity, necessity, and atemporality of mathematics. Philos Math 21(1):9–36

De Cruz H, De Smedt J (2013) Mathematical symbols as epistemic actions. Synthese 190(1):3–19

De Toffoli S (2017) 'Chasing' the diagram: the use of visualizations in algebraic reasoning. Rev Symb Log 10(1):158–186

De Toffoli S (2023) The epistemological subject(s) of mathematics. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer, Cham

De Toffoli S, Giardino V (2014) Forms and roles of diagrams in knot theory. Erkenntnis 79(3): 829–842

De Toffoli S, Giardino V (2015) An inquiry into the practice of proving in low-dimensional topology. In: Lolli G, Panza M, Venturi G (eds) From logic to practice: Italian studies in the philosophy of mathematics, Boston studies in the philosophy and history of science, vol 308. Springer, Cham, pp 315–336

De Toffoli S, Giardino V (2016) Envisioning transformations—the practice of topology. In: Larvor B (ed) Mathematical cultures: the London meetings 2012–2014. Birkhäuser, Cham, pp 25–50

Detlefsen M (1992) Poincaré against the logicians. Synthese 90(3):349–378

Feferman S (2009) Conceptions of the continuum. Intellectica 51(1):169–189

Ferreirós J (2016) Mathematical knowledge and the interplay of practices. Princeton University Press, Princeton

Ferreirós J (2022) What are mathematical practices? The web-of-practices approach. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer, Cham

Folina J (2018) Towards a better understanding of mathematical understanding. In: Piazza M, Pulcini G (eds) Truth, existence and explanation. Springer, Cham, pp 121–146

Frege G (1879) Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Louis Nebert, Halle a. S

Friedman M (2018) A history of folding in mathematics: mathematizing the margins, Science networks. Historical studies, vol 59. Springer, Cham

Gentzen G (1934) Untersuchungen über das logische Schließen I & II. Math Z 39(176–210):405–431

Giaquinto M (2005) Mathematical activity. In: Mancosu P, Jørgensen KF, Pedersen SA (eds) Visualization, explanation and reasoning styles in mathematics. Springer, Dordrecht, pp 75–87

Giardino V (2018) Tools for thought: the case of mathematics. Endeavour 42(2):172–179
Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, Boston
Grandaty M, Debanc C, Virbel J (2000) Evaluer les effets de la mise en page sur la compréhension et la mémorisation de textes procéduraux (règles de jeux) par des adultes et des enfants de 9 à 12 ans." PArole (special issue "Langage et Cognition"). 13:3–38
Hacking I (2000) What mathematics has done to some and only some philosophers. In: Smiley T (ed) Mathematics and necessity: essays in the history of philosophy. Oxford University Press, Oxford
Hamami Y, Morris RL (2020) Philosophy of mathematical practice: a primer for mathematics educators. ZDM 52(6):1113–1126
Hamami Y, Morris RL (2021) Plans and planning in mathematical proofs. Rev Symb Log 14(4): 1030–1065
Hamami Y, Morris RL (2022a) Rationality in mathematical proofs. Australas J Philos
Hamami Y, Morris RL (2022b) Understanding in mathematics: the case of mathematical proofs. Unpublished manuscript
Hersh R (1997) What is mathematics, really? Oxford University Press, New York/Oxford
Hutchins E (2005) Material anchors for conceptual blends. J Pragmat 37(10):1555–1577
Inglis M, Tanswell FS (2022) The language of proofs: a philosophical corpus linguistics study of instructions and imperatives in mathematical texts. In: Sriraman B (ed) Handbook of the history and philosophy of mathematical practice. Springer, Cham
Jaffe A, Quinn F (1993) "Theoretical mathematics": toward a cultural synthesis of mathematics and theoretical physics. Bull Am Math Soc 29(1):1–13
Kirsh D, Maglio P (1994) On distinguishing epistemic from pragmatic action. Cogn Sci 18(4): 513–549
Kitcher P (1984) The nature of mathematical knowledge. Oxford University Press, New York
Lakatos I (1976) Proofs and refutations: the logic of mathematical discovery. In: Worrall J, Zahar E (eds) . Cambridge University Press, Cambridge
Landy D, Allen C, Zednik C (2014) A perceptual account of symbolic reasoning. Front Psychol 5(275):1–10
Larvor B (2012) How to think about informal proofs. Synthese 187(2):715–730
Larvor B (2019) From Euclidean geometry to knots and nets. Synthese 196(7):2715–2736
Mac Lane S (1935) A logical analysis of mathematical structure. Monist 45(1):118–130. Oxford University Press
Mac Lane S (1981) Mathematical models: a sketch for the philosophy of mathematics. Am Math Mon 88(7):462–472
Mac Lane S (1986) Mathematics: form and function. Springer, New York
Mancosu P, Jørgensen KF, Pedersen SA (eds) (2005) Visualization, explanation and reasoning styles in mathematics. Springer, Dordrecht
Manders K (2008) The Euclidean diagram (1995). In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 80–133
Martin-Löf P (1984) Intuitionistic type theory. Bibliopolis, Naples
Netz R (1999) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press, Cambridge
Piñeros Glasscock JS, Tenenbaum S (2023) Action. In: Zalta EN, Nodelman U (eds) The Stanford encyclopedia of philosophy, Spring 2023. Metaphysics Research Lab, Stanford University
Poincaré H (1908) Science et Méthode. Ernest Flammarion, Paris
Prawitz D (1965) Natural deduction. A proof-theoretical study. Almqvist / Wiksell, Stockholm
Prawitz D (2012) The epistemic significance of valid inference. Synthese 187(3):887–898
Rittberg CJ, Tanswell FS, Van Bendegem JP (2020) Epistemic injustice in mathematics. Synthese 197(9):3875–3904
Robinson JA (2000) Proof = guarantee + explanation. In: Hölldobler S (ed) Intellectics and computational logic, vol 19:277–294. Applied Logic Series. Kluwer Academic Publishers, Dordrecht

Roth AS (2017) Shared agency. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Summer 2017
Schlosser M (2019) Agency. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Winter 2019. Metaphysics Research Lab, Stanford University
Sundholm G (2012) "Inference versus consequence" revisited: inference, consequence, conditional, implication. Synthese 187(3):943–956
Tanswell FS (forthcoming) Go forth and multiply: on actions, instructions and imperatives in mathematical proofs. In: Bueno O, Brown J (eds) Essays on the philosophy of Jody Azzouni. Springer, Cham
Vandendriessche E (2015) String figures as mathematics? In: Studies in history and philosophy of science, vol 36. Springer, Cham
Virbel J (2000) Un type de composition d'actes illocutoires directifs et engageants dans les textes de type "consigne". PArole (special Issue "Langage et Cognition") 11–12:200–221
Weber, Keith, and Fenner Stanley Tanswell. 2022. "Instructions and recipes in mathematical proofs." Educ Stud Math 111 (1): 73–87
Wilder RL (1950) The cultural basis of mathematics. Proc Intl Congr Math 1:258–271
Wright C (2014) Comment on Paul Boghossian, "what is inference". Philos Stud 169(1):27–37

The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts

Fenner Stanley Tanswell and Matthew Inglis

Contents

1 Introduction 2926
2 The Recipe Model of Proofs 2927
3 Background on Mathematical Language 2929
3.1 Mathematics Education and Philosophy 2930
3.2 Formalised Mathematics 2933
4 A Corpus Linguistics Study 2935
4.1 Corpus Linguistics 2935
4.2 Hypotheses 2937
4.3 Methods 2938
4.4 Results 2940
4.5 An Example 2946
5 Conclusion 2949
References 2951

Abstract

A common description of a mathematical proof is as a logically structured sequence of assertions, beginning from accepted premises and proceeding by standard inference rules to a conclusion. Does this description match the language of proofs as mathematicians write them in their research articles? In this chapter, we use methods from corpus linguistics to look at the prevalence of imperatives and instructions in mathematical preprints from the arXiv repository. We find thirteen verbs that are used most often to form imperatives in proofs, and that these show up significantly more often within proofs than in the surrounding

F. S. Tanswell (✉)
Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Brussels, Belgium

Department of Mathematics Education, Loughborough University, Loughborough, UK
e-mail: F.Tanswell@lboro.ac.uk

M. Inglis
Department of Mathematics Education, Loughborough University, Loughborough, UK
e-mail: M.J.Inglis@lboro.ac.uk

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_50

mathematical writing. We also show that there are many more verbs used to form a diverse selection of instructions in proofs. These findings are at odds with the view of proofs as sequences of assertions. Instead, we argue in favour of the *recipe model of proofs*: that proofs are like recipes, giving instructions for mathematical actions to be carried out.

Keywords

Proof · Mathematical language · Corpus linguistics · Instructions · Imperatives

1 Introduction

One of the central principles of the philosophy of mathematical practices is to embrace, analyse, and understand the real (and often messy) features of mathematics as a living subject. A project like this is necessarily interdisciplinary, relying on history, sociology, anthropology, psychology, cognitive science, education, argumentation studies, etc. to identify and study the salient features of mathematics. Here, we will use methods from *corpus linguistics* to study the language used in modern mathematics, and especially proofs.

Despite the central importance of attending to the realities of practices, our understanding of what proofs are really like is, in the existing literature, typically only informed by modern and historical case studies, and often by anecdote and personal familiarity. While we do not dispute the value of these, they can only provide part of the picture. In this chapter, we instead perform a systematic, large-scale, computational analysis of mathematical texts. Specifically, we will analyse texts from the *arXiv*, a preprint server for mathematics research articles.

We will investigate the relationship between the language used in written, textual mathematics, and how to best understand what proofs are. In particular, we will look at the use of instructions in proofs. One way that instructions are issued in mathematics texts is in sentences in the imperative mood. A sentence in the imperative mood instructs or commands that some action be carried out, such as "Go to your room"; "Eat your greens"; and "Do as you'd be done by." The use of the imperative mood can also be found in mathematics, such as "Suppose that the inductive hypothesis holds"; "Let I be a proper ideal of A"; and "Notice that x is a member of y." While the presence of instructions and imperative mood sentences in mathematical writing has been observed and discussed to some extent, there has been no systematic investigation of their relative frequency in modern mathematical proofs, and only to a limited extent the philosophical implications they have.

The central questions we will address through corpus linguistics concern the *frequency* and *diversity* of the instructions given in proofs in modern mathematical writing. In other words, we are asking how common various imperatives are, and how broad the range of permissible imperatives is.

There are several reasons to look specifically at the use of the imperative mood in mathematics. First of all, it contributes to philosophical debates on the nature of

proof and rigour, since it tells us about how written proofs are currently presented in research articles. Secondly, it is relevant to the epistemology of mathematics, since the potential ways that knowledge is gained from imperatives, which instruct us to carry out actions, might be importantly different to knowledge from other kinds of sentences. Thirdly, the existence of imperatives in mathematical writing points us to a different class of speech acts being used in proofs, and thereby raises questions about what the function of written proofs might be. Fourthly, it contributes to an overall picture of mathematical style and the language used in professional mathematics, which has impacts in mathematics education on how students are taught to read and write proofs. Finally, it adds to the practical efforts to embed computer assistants into mathematical practice, by showing what computers need to be able to read and produce in order to engage in mathematics alongside mathematicians.

The structure of this chapter is as follows. In Sect. 2, we outline the recipe model of proofs, under which a proof provides a recipe of instructions for the reader to follow. In Sect. 3, we look at the background of existing literature discussing the use of instructions and imperatives in mathematics. In Sect. 4, we present our corpus study: Sect. 4.1 explains corpus linguistics more generally and the body of texts we are working with; Sect. 4.2 offers our hypotheses to test based on the recipe model and the existing literature; Sect. 4.3 describes the methods; and Sect. 4.4 goes through the various results we obtained. Section 5 will offer an empirically informed discussion of the philosophical questions in light of the results.

2 The Recipe Model of Proofs

In this section, we describe the *recipe model of proofs*, as defended by Tanswell (forthcoming). This model stands in contrast to the *standard model of proof*, on which proofs are seen as sequences of assertions, which are stepwise inferences from premises to a conclusion, and all of which are declarative statements. From that perspective, all of the activity is reduced to acts of inferring, and the steps in the proof represent the results of those inferences. The standard model of proof has many virtues; however, it also often mischaracterises both what proofs look like in practice and how they are engaged with by practitioners. The recipe model of proof is intended to more faithfully account for these aspects of mathematical proofs. This alternative model develops an analogy that mathematical proofs are akin to cooking recipes. This section will fill out the important aspects of this analogy. Our summary of the recipe model will focus on four key points: (1) the instructions given by the text; (2) the language used to issue them; (3) the relationship between the text and the activity it tells you to carry out; and (4) the epistemic significance of proving. In this section, we discuss these four points.[1]

[1] As we are interested in written, purely textual proofs here, we set aside a fifth point by Tanswell (forthcoming) concerning the use of diagrams to issue instructions.

On point (1), the reason to compare proofs to recipes is that this emphasises the ways in which proofs are used to issue instructions, either explicitly or implicitly, for us to carry out a series of mathematical actions. On the recipe model, we propose that the actions that a proof instructs the audience to perform are far more diverse than inferring alone, and can include instructions for constructing useful mathematical objects, for manipulating objects and equations, for hypothetically considering mathematical assumptions, for fixing notations, for directing the readers' attention, for structuring their reasoning, and for many more possible epistemic actions.

For point (2), the emphasis of the analogy to recipes is on the similarity of the language used to issue instructions. There are, of course, large variations in the presentation of both cooking recipes and mathematical proofs. However, in cooking recipes the predominant mood is usually the imperative mood, directly listing instructions to follow to carry out the recipe.[2] The recipe model thus predicts that proofs will similarly include imperative sentences to offer instructions to the reader. However, two disanalogies are salient. First of all, recipes might sometimes include elements describing the state of the food, but in mathematics there is reason to find more declarative content because we are likely to see some aspects of the traditional view of proofs, where they do contain declarative statements with inferences from one to the next. Secondly, as a matter of style, cooking recipes tend to use more purely imperatival language, while in maths we can also make use of the "We [verb]" constructions. Nonetheless, in both disanalogies the instructions are still there: in the latter case explicitly, but in the case of inferring from statement to statement, the proof is still implicitly instructing the reader to undertake those mathematical acts of inference.[3] This means the recipe model is not dependent on there being many imperatives but does certainly look more plausible if there are. We expect that there are plenty of proofs without imperatives, but if there are very few proofs that contain any imperatives at all, that would be a bad sign for the recipe model generally.

The next point (3) is about how both proofs and recipes should be understood in terms of the relationship between a written text and the actions they describe. A recipe is intrinsically linked to the activity of making, and by this analogy we should understand proofs to be the same. A proof is a device for recording and communicating how to carry out some piece of mathematics, but it is the correct performance of the relevant actions that constitutes mathematical practice. On this model, therefore, we prioritise the doing of mathematics over the written artefact of a proof itself. Putting it another way: the recipe model makes proofs secondary to proving. Here,

[2] Detailed considerations of the language used in cooking recipes in different contexts can be found in Gerhardt et al.'s (2013) edited volume.

[3] What will "implicit instructions" look like? It will not be the case that all verbs are used to give implicit imperatives. Rather, the idea is that declarative sentences can carry with them instructions for, say, inferences (or other things) to be carried out. It may be that one needs a certain level of mathematical expertise and know-how to understand what these implicit instructions are exactly. Analysing the verbs in certain ways, as we will do, is one way to access these instructions, but we don't assume that all verbs correspond to instructions.

the activity of proving is construed broadly, such that the authors of the proof have proved the theorem, then use a proof to communicate to their readers how to carry out the mathematical activity, and thereby establish the truth of the theorem. Similarly, we expect the author of a recipe to have cooked it themselves, then use the recipe to enable the audience to reproduce it.

The final point (4) concerns epistemology: one of the central functions that proofs play in mathematics is epistemic: a correct and rigorous proof is meant to establish a theorem to be known with certainty.[4] But an important question for the standard model is: how does a presented proof link to knowledge at all? What is connecting the proof on the page (or wherever) with the knowledge in your head? A plausible story would already have to be recipe-like, where the proof tells you exactly which inferences to carry out. But as soon as the proof has any inferential gaps (Fallis 2003; Hamami 2014; Andersen 2020), which many proofs do, an additional account is needed. The recipe model's answer to this question is more direct: that knowledge is gained through the mathematical activity that we are instructed in by proofs, with this activity acting as a key intermediary to explain the epistemic functioning of proofs. Furthermore, the recipe model also expands the traditional conception of mathematical knowledge, which is largely about propositional knowledge, to also include mathematical knowledge-how. In order to cook from a recipe, one needs to *know how* to perform the relevant actions; likewise, to carry out a proof one needs to *know how* to perform the mathematical actions it tells you how to do. In fact, the mixture of imperative and declarative forms suggests that the knowledge gained from proving will be a combination of knowledge-how and knowledge-that. (For a longer development of this epistemic picture see Tanswell (2016, Chap. 4) and Tanswell (forthcoming, Sect. 6).)

The recipe model of proofs gives us concrete predictions on our two main areas of investigation, concerning the frequency and diversity of imperatives, namely that they should be common and diverse. They should be common because the issuing of instructions is one of the primary functions of proofs. They should be diverse because the possible space of mathematical activities that one could be instructed to carry out is large. These hypotheses are set out in Sect. 4.2 and tested in Sect. 4.4.

3 Background on Mathematical Language

In this section, we survey previous discussions of the use of instructions and imperatives in mathematics. We group these into two broad categories, based on the sources of the discussion: (1) mathematics education and philosophy; and (2) formalised mathematics.

[4] This is again a rather strong traditional conception of the epistemic function of proofs, and we agree with De Toffoli (2020) that this needs to be tempered to properly reflect practice.

3.1 Mathematics Education and Philosophy

Researchers in mathematics education and philosophy have both, at times, noted the use of the imperative mood in mathematical proof texts. However, in many cases this has just been done parenthetically or tangentially. We will set out a number of examples from these literatures chronologically.

David Pimm, in *Speaking Mathematically* (1987), analysed the kinds and effects of the language used in the mathematics classroom. As an aside, he remarked on mathematical imperatives:

> There is also the equally common use of the imperative in mathematical discourse (e.g. *consider*, *suppose*, or *define*) [...]. Another form is 'Let x be the number of . . .' Who is giving permission for this to be done? The mathematical imperative [. . .] is a topic worthy of considerable attention. (Pimm 1987, p. 72)

However, Pimm does not go on to give them any such attention, merely suggesting that the impersonal style allows mathematics texts to suppress and omit the agent who is carrying out the actions. This is a clear case of a parenthetical mention of the imperative mood in mathematics.

The work of Brian Rotman (1988, 1993) included the unusually careful examination of imperatives in mathematics. His broader project is one of semiotic analysis of mathematical signs, which is the analysis of how signs come to have meaning in mathematical texts. He considers imperatives to be a major feature of mathematical writing, and in establishing the shared meanings used in mathematics:

> But proof in turn involves the idea of an argument, a narrative structure of sentences, and sentences can be in the imperative rather than the indicative. [. . .] Mathematics is so permeated by instructions for actions to be carried out, orders, commands, injunctions to be obeyed — 'prove theorem T', 'subtract from y', 'drop a perpendicular from point P onto line L', 'count the elements of set S', 'reverse the arrows in diagram D', 'consider an arbitrary polygon with k sides', and similarly for the activities specified by the verbs 'add', 'multiply', 'exhibit', 'find', 'enumerate', 'show', 'compute', 'demonstrate', 'define', 'eliminate', 'list', 'draw', 'complete', 'connect', 'assign', 'evaluate', 'integrate', 'specify', 'differentiate', 'adjoin', 'delete', 'iterate', 'order', 'complete', 'calculate', 'construct', etc. that mathematical texts seem at times to be little more than sequences of instructions written in an entirely operational, exhortatory language. (Rotman 1988, p. 8)

He argued that imperatives in mathematics can be separated into two distinct kinds: *inclusive* imperatives, which are for creating shared mathematical discourse (by introducing shared standards, notations, referents etc.), and *exclusive* imperatives, which take a shared mathematical discourse for granted and operate within it. Rotman takes "consider," "define," "prove," and synonyms of these to be inclusive, constructing a shared domain of meaning, while other mathematical verbs merely tell us to operate on the shared domain of meanings that is already established. For example, "consider the metric space *(X,d)*" tells us to use X as a set with a distance measure d obeying the definitions of a metric space, constructing shared meanings for the author and reader for the coming mathematical text, while "add the distances d_1 and d_2" merely instructs the reader to carry out an operation

within the space, where meaning for the variables has already been established. This distinction is important for understanding the semiotics of a mathematics text and how some imperatives work to create shared mathematical discourses, but we don't think the distinction coincides with the epistemic and pragmatic functions of different kinds of imperatives. We also note that while Rotman listed many potential mathematical instructions, there is no study as to how common these actually are. We will carry out such a study below.

Besides the creation of shared meaning in mathematical texts, another feature of the use of imperatives in mathematical texts is to mark who is part of the mathematical community, something discussed by Candia Morgan (1996). She says:

> One common characteristic of academic mathematics texts (and some school texts) is the conventional use of imperatives such as *consider, suppose, define, let* x *be*. Like the use of we, these implicate the reader, who is addressed implicitly by the imperative form, in the responsibility for the construction of the mathematical argument. The use of imperatives and of other conventional and specialist vocabulary and constructions characteristic of academic mathematics marks an author's claim to be a member of the mathematical community which uses such specialist language and hence enables her to speak with an authoritative voice about mathematical subject matter. At the same time it constructs a reader who is also a member of the same community and is thus in some sense a colleague (although the nature of this relationship may vary according to the type of action demanded). (Morgan 1996, p. 6)

Her analysis was primarily focused on how these issues of community membership work for students of mathematics who are not yet members of the community but are in the process of being inducted and enculturated into it. However, she did not further concern herself with imperatives in particular.

Paul Ernest has also discussed imperatives in mathematics in multiple places. For example, in his book *Social Constructivism as a Philosophy of Mathematics* (1998), he stated:

> Thus mathematical texts comprise specific assertions and imperatives directed by the writer to the reader. (Ernest 1998, p. 170)

We are broadly sympathetic to this view of proofs as, at least in part, functioning as communicative devices to transfer knowledge of mathematical reasoning from author to reader. In more recent work, citing Rotman, Ernest went so far as to say:

> Imperatives are orders that instruct or direct actions either inclusively, such as: let us ..., consider ..., or exclusively, such as: add, count, solve, prove, etc. Imperatives occur more frequently in mathematics than in any other academic school subject. (Ernest 2018, p. 191)

We note that this is a direct empirical claim, offered without evidence, and one that could potentially be refuted by other disciplines which use imperatives in their written work such as chemistry, philosophy, or geology (cf. Swales et al. 1998). Below we carry out one study of the frequency of imperatives in proofs in mathematical research articles.

The work of Rotman was also extended and challenged by Roy Wagner (2009, 2010) in his detailed semiotic study of Gödel's proofs of the first incompleteness

theorem. Wagner discussed the imperatives both explicit and implicit in the text, and incorporated these into a nuanced picture of the mathematical subject, agency, and the relationship these have to a mathematical text. One major difference from our findings below is that Wagner finds very few explicit imperatives:

> First, there are very few imperatives in these texts. Perhaps the imperative mood was considered ill-suited for civilised written communication in the Vienna and Princeton circles. Perhaps it never occurred to writers in these circles to dominate a mathematical text with the imperative mood. Instead of imperatives, we have the frequent use of the indicative mood in both active and passive voices, often attributed to the character we. (Wagner 2009, p. 48)

It is notable, but maybe not surprising, that the modern writing style is different in some ways to Gödel's in the 1930s. Wagner's work also provides a nice contrast for what can be achieved with different methodological approaches: Wagner uses detailed analysis of a single mathematician's writing, while the corpus linguistics approach below looks for patterns across thousands of papers by thousands of authors.

A notable difference related to action-based vocabulary was also found by Mejía-Ramos and Inglis (2011). On a proof-evaluation task with undergraduate students, they found that changing the question from "Does the argument prove the claim?" to "Is the argument a proof of the claim?" had a significant impact on how the students evaluated a specific visual argument. The students were more willing to agree that the argument proved the claim than that it was a proof for the claim. The researchers built the case that the verb form was more connected to judgments of conviction, while the noun form was more connected to judgments of validity. They related this to issues of semantic contamination: where the everyday meanings of words affect our understanding of them in more formal and technical settings like mathematics (cf. Tanswell 2018, Sect. 4).

As a final example, Lew and Mejía-Ramos (2019) investigated the norms of mathematical writing by testing the reactions of mathematicians and students to cases where those norms were flouted. One of their stimuli was an ill-formed imperative: "Suppose $(S \cdot R)^{-1}$ such that [...]." They found that professional mathematicians especially responded very negatively to the ungrammatical phrasing of this. For the mathematicians, imperatives must be used correctly and grammatically, and they appear to have strong conventions on how that is achieved. In contrast, students often did identify that the construction didn't sound right but could not clearly express why (Lew and Mejía-Ramos 2019, p. 134).

From a different direction, Paul Halmos (1970) provided a classic style guide for mathematics and discussed both imperatives in maths and the use of the first-person plural "we." Regarding imperatives, he said: "A frequently effective and time-saving device is the use of the imperative. 'To find P, multiply q by r.' 'Given p, put q equal to r.'" (Halmos 1970, p. 141).[5] He also proceeded to explain using the first-person plural to convey instructions on how a proof goes:

[5] Somewhat amusingly, the immediately prior sentence is: "Most (all?) mathematical writing is (should be?) factual; simple declarative sentences are the best for communicating facts." (Halmos 1970, p. 141).

> There is nothing wrong with the editorial "we", but if you like it, do not misuse it. Let "we" mean "the author and the reader" (or "the lecturer and the audience"). Thus, it is fine to say "Using Lemma 2 we can generalize Theorem 1", or "Lemma 3 gives us a technique for proving Theorem 4". (Halmos 1970, p. 141)

Indeed, it is discussions of the first-person plural "we" that are found most often in mathematical style guides. For example, Krantz likewise stated that "The custom in modern mathematics is to use the first person plural, or "we." It stresses the participatory nature of the enterprise, and encourages the reader to push on." (Krantz 1997, p. 33). A similar point was made by Knuth et al. (1989, p. 2).

3.2 Formalised Mathematics

There is a large area of research dedicated to formalised mathematics, where the central goal is to provide formalised counterparts of informal proofs that can be formally verified in one of the proof-checkers, such as Lean, Coq, Mizar, or Isabelle. Standardly, these come with their own language, software, and libraries of completed formal proofs that can be called on in later proofs. There are two obvious goals to this formalisation project: firstly, to check the existing mathematical literature, and secondly, to develop software that can assist mathematicians in new work.

Especially for the second goal, it is important that the formal language used by the computer is accessible to human mathematicians. For this reason, there is some research on the language of everyday mathematics focusing on how to formalise it. Oddly enough, researchers here sometimes seem to downplay the significance of imperatives in proofs. We look at three examples of work in this area, and what they say with regard to imperatives in mathematical writing.

The oldest we consider is the work of de Bruijn (1994). His goal was to set out the "mathematical vernacular," which is meant to be the native language of mathematics, described as "[...] the very precise mixture of words and formulas used by mathematicians in their better moments" (de Bruijn 1994, p. 865). This is immediately contrasted to the "official," purely formal language of mathematics. De Bruijn then claimed that this mathematical vernacular is in essence the same as the rigorous part of mathematics:

> Roughly speaking, the [mathematical vernacular] part of a piece of mathematics will be the rigorous part. (ibid., p. 867)

This is then used to exclude various common parts of mathematical texts from the mathematical vernacular as he sees it, such as historical remarks, textual signposting, references to the syntactical form of the presented equations, and indications of how to reproduce omitted material. Included in this list is "Commands, like `show that'." (ibid., p. 867). Hereby de Bruijn excludes imperatives from the domain of rigorous mathematics.

This exclusion is also present in the book *The Language of Mathematics* by Mohan Ganesalingam (2010). The purpose of the book was to analyse the language used in

mathematical texts in keeping with a broadly generative tradition, to give a formal, objective, and precise syntactic structure and associated semantics. Ganesalingam said that he will restrict himself to the language used in "more rigorous, careful textbooks" (ibid., p. 7) and only the part in the *formal mode*, where "[a] sentence is in the formal mode if and only if it consists only of assertions about mathematical objects and about mathematical facts, so that its semantic content is purely truth-conditional and can be completely captured in an appropriate logic" (ibid., p. 7). This clearly precludes imperatives, since they are not assertions, nor do they have a purely truth-conditional semantics.[6] Ganesalingam thus stated that:

> [T]he language of mathematics consists purely of assertions about mathematical objects. As a result, textual mathematics predominantly uses the third person singular and third person plural, to denote individual mathematical objects (or propositions) and collections of mathematical objects (or propositions) respectively. The first person plural ('we') is also used in a more restricted capacity, with a limited and potentially closed class of verbs, typically to refer to the mutual intent of the author and reader. (ibid., p. 21)

This seems to allow constructions similar to imperatives (e.g., "We suppose," "We assume," "We obtain," etc.) but not the imperative forms ("Suppose," "Assume," "Obtain"). Of course, if we restrict ourselves to the formal mode, which by definition is exclusively about mathematical assertions, then the absence of imperatives must follow. Nonetheless, this doesn't justify these bolder claims about *the language of mathematics* being limited to this class of sentences, nor even claims about rigorous textbooks.[7]

Below we will demonstrate that imperatives are present in mathematical writing as found on the *arXiv*. Whether any given article on the *arXiv* is part of rigorous mathematics or not must surely be debatable, but there is no doubt that such a corpus must represent a collection of paradigm examples of the language of mathematics.

In contrast to de Bruijn and Ganeslingam's claims about rigorous mathematics, formalised mathematics does not need to exclude imperatives. Here is an example of a simple proof in Lean[8]:

---

[6] The formal mode is contrasted with the *informal* mode: "The informal mode consists of commentary on the ongoing mathematics. Individual sentences in the language of mathematics tend to be either entirely formal or entirely informal, though there are exceptions. It is worth noting that mathematicians have very strong intuitions about the distinction between these two modes." (p. 7). Below we show that imperatives are a part of normal mathematical language, but it bears emphasising that no evidence is provided for the existence or reliability of these intuitions, and even if mathematicians do have robust intuitions about the difference between doing the mathematics and commentary on it, that does not necessarily align with the characterisation he gives of the difference between the formal and informal modes.

[7] A similar criticism can be levelled concerning another exclusion: that of *modality* in mathematics. Ganesalingam said: "[...] textual mathematics is much more restricted than natural language. [...] Mathematical truths are true in all possible worlds, and therefore there is no modality." (pp. 22–23). However, Hodges (2013) demonstrated that there are many uses of modals in mathematical writing but that these don't refer to metaphysical necessity and possibility.

[8] This example is taken from https://leanprover.github.io/introduction_to_lean/ written by Jeremy Avigad, Gabriel Ebner, and Sebastian Ullrich (Accessed 11/02/2021).

```
theorem even_add : ∀ m n, even m → even n → even (n + m) :=
  take m n,
  assume ⟨k, (hk : m = 2 * k)⟩,
  assume ⟨l, (hl : n = 2 * l)⟩,
  have n + m = 2 * (k + l),
  by simp [hk, hl, mul_add],
  show even (n + m),
  from ⟨_, this⟩.
```

The details of how the system works are not important here; all we need is to see the imperatives at work in the proof. What is remarkable is that the imperatives serve the dual functions of both acting as imperatives in a proof in the sense outlined above for the human mathematician, and acting as imperatives as commands for the computer (in the imperative programming tradition).

The system Naproche-SAD (De Lon et al. 2020) similarly incorporates many imperatives into its "controlled natural language," a fact notable because of the system's explicit aim to emulate normal mathematical language. Here is an example of Naproche-SAD text taken from Frerix and Koepke (2018, p. 3):

Theorem 1 *Assume f is holomorphic and the domain of f is a region. If f has a local maximal point then f is constant.*

Proof *Let z be a local maximal point of f. Take ε such that Bε(z) is a subset of Dom(f) and |f[w]|≤|f[z]| for every element w of* $B_\varepsilon(z)$*. Let us show that f is constant on* $B_\varepsilon(z)$*. Assume the contrary. Then* $f[B_\varepsilon(z)]$ *is open. We can take δ such that* $B_\delta(f[z])$ *is a subset of* $f[B_\varepsilon(z)]$*. Therefore there exists an element w of* $B_\varepsilon(z)$ *such that* $|f[z]|<|f[w]|$*. Contradiction. end. Hence f is constant.*

Again, observe the natural use of imperatives to construct the proof here. The creators of *Naproche* are amongst a relatively small set of researchers who also use linguistics to study mathematical language, and we hope that our empirical work below will support its further development.

4 A Corpus Linguistics Study

4.1 Corpus Linguistics

The driving idea underlying the philosophy of mathematical practice is to ensure that our philosophical claims about mathematics are true and accurate of real mathematics, not just some idealised version of it. This means that we need to give empirical grounding to any factual claims about mathematical practice that we rely on (e.g., Aberdein and Inglis 2019). In this chapter, we are focused on the language of mathematics, and especially on the language used in proofs. In our review of the literature, however, the main evidence used by previous authors is their own experience as mathematicians (e.g., Ganesalingam 2010) or the detailed analysis of isolated

texts (e.g., Wagner 2009, 2010). Here, we take a very different approach: we examine a large number of mathematics texts using the tools of corpus linguistics.

A corpus in this sense is a substantial body of texts that are machine-readable to allow for computational analysis. In our case, we are using the corpus assembled by Alcock et al. (2017) and used by Mejía-Ramos et al. (2019) in their study of language relating to explanation in mathematics. The corpus is made up of all of the papers uploaded to the arXiv in the first 4 months of 2009 to the *Mathematics* category. The arXiv is an online repository for texts from mathematics, physics, computer science, and various related disciplines, which acts as the main preprint server for large sections of research mathematics, especially in English-speaking mathematics communities. Many mathematicians upload copies of their academic papers as preprints to make them freely accessible online. There is a level of selection before papers can appear on the arXiv, partially moderated by volunteered and partially by an automated system, and the assigned category for the paper does have social implications, as described by Reyes-Galindo (2016).

Texts on the arXiv are predominantly written in the TeX/LaTeX mark-up language, then compiled into pdf format. In order to process these into a form useable for computational linguistics (plain texts without extraneous code), the texts were processed to strip out the TeX/LaTeX code while preserving the structure of the written language, as described by Mejía-Ramos et al. (2019, p. 249–250).[9] This posed the challenge of what to do with inline mathematics, such as "Let f: X →Y be a bijection." The problem is that the same mathematics can be coded in many different ways in LaTeX, which speaks against leaving the code as is because the linguistics software would not be able to identify the different codes as representing the same syntactic unit. However, simply removing it would interfere with some of the linguistics tools (measuring word proximities, for example). As such, they opted to replace all inline mathematics with the tag "inline_math." For the above example, this gives us "Let inline_math be a bijection." The result of this processing is sufficient for our needs too because we are interested in the natural language components and not the mathematical content.[10]

In total, the corpus then contained 6,988 files and 30,892,695 words of mathematics (with 231,400 distinct words). As a useful subset of this for our purposes, we also have a corpus that is made up of only the proofs found in these papers, containing 3,268 files and 4,973,892 words of mathematics (with 48,025 distinct words).[11] This will allow us to focus specifically on the mathematical writing found

[9] This process was automated, and Chris Sangwin has made the code available online at https://github.com/sangwinc/arXiv-text-extracter (accessed 23/02/2021).

[10] The decision of how to treat the mathematical content of the papers does affect the kinds of results we get. If we were to instead convert mathematical content to verbalised counterparts (i.e., x^2 + 5 becoming "x squared plus five"), this would lead to different frequencies, and "dilute" the natural language parts in the corpora. However, there is certainly no unique or uncontroversial algorithm for this process of verbalisation, so our approach here is the pragmatically obvious one.

[11] Selected automatically, using everything between "\begin {proof}" and "\end{proof}." This is the most common way to write proofs but may have left out a small minority of written proofs coded differently.

in written proofs in these papers. Simple subtraction then also lets us compare the proof-only corpus to the non-proof parts of the mathematical texts. The non-proof part therefore contains 25,918,803 words. Henceforth, we refer to the main corpus as the *All-math* corpus, and the proof only part as the *Proof-only* corpus, and the result of subtracting the *Proof-only* from the *All-math* will be called the *Non-proof* corpus.

Something important to be sensitive to in analysing these texts is how representative they are of textual mathematical proofs more generally. Since the arXiv is the standard repository for preprint papers for many mathematical researchers, we do take it that the corpus is fairly representative. However, this must come with some caveats. The first is the most obvious: we are conducting this research in English, and so our results apply to mathematical texts written in English. They may still be applicable to nearby languages, where we believe many similar linguistic conventions apply for mathematical writing, but the current study would not reveal that. Secondly, language and other factors will in turn affect who submits their papers to the arXiv, and there are obvious demographic and geographic biases likely to arise in whose work we are seeing.[12,13] Thirdly, since these are preprint articles primarily from professional mathematicians, the mathematical language we are analysing is theirs and not, say, that of mathematics students or teacher, nor the language used in proofs in maths textbooks or other media (such as spoken proofs, or proofs on blackboards in seminars). Finally, the papers are all from 2009 and so represent a snapshot of proof writing from that year.[14] We do not believe there has been any substantial change since then, but writing and presentational styles are always changing to some degree. Obviously, the results of our study will not immediately generalise to historical mathematical writing either.

4.2 Hypotheses

Based on the previous sections, we can form various hypotheses about the prevalence of instructions and imperatives in our corpus.

In the literature from mathematics education and philosophy, we saw that the researchers like Pimm (1987), Rotman (1988, 1993), Morgan (1996), and Ernest (1998, 2018) broadly claimed:

[12] The arXiv themselves do have various statistics about submissions and downloads at https://beta.arxiv.org/help/stats (accessed 23/02/21), but these do not include demographic or geographic data.

[13] For more on how social features of mathematics can involve epistemic exclusion, see Rittberg et al. (2018) and Tanswell and Rittberg (2020). Indeed, Reyes-Galindo (2016) examines how the automated filtering of the arXiv is partially based on linguistic analysis, and can lead to the exclusion of certain researchers.

[14] They are also from only the first 4 months of 2009. We do not believe that the time of year makes a significant difference to writing styles, though it is possible it affects the demographics of who is submitting their papers to some small degree.

(H1): Imperatives are common in mathematical proof texts.

The recipe model of proof also predicts that (H1) will hold.

In the works of De Bruijn (1994) and Ganesalingam (2010), however, the opposite claim was made regarding imperatives. They saw rigorous proofs as only containing assertions about mathematical objects. So, the opposite hypothesis can be considered:

(H2): There are no imperatives in mathematical proof texts.

The recipe model of proof will predict that (H2) will not hold.

Next, consider the first-person plural in mathematics, which we saw was mentioned in various style guides for mathematics. This can be used with the same verbs to form "We [verb]..." sentences, which can potentially also be used to give instructions for the reader to follow. We will test how common this is as the following hypothesis:

(H3): The construction "We [verb]..." is common in mathematical proof texts.

The recipe model of proof again predicts that (H3) will hold.

Ganesalingam did note the use of the "We [verb]..." constructions (like "We suppose...," "We assume..." etc.) but states that they are only used for a "limited and closed class of verbs." Generalising to cover imperatives too, and putting this as a hypothesis:

(H4): The verbs used in imperatives and the "We [verb]..." constructions in mathematical proof texts are drawn only from a limited and closed class of verbs.

On the recipe model of proofs, proofs contain many instructions (be they explicit or implicit), and therefore the more diverse the selection of verbs the better for the model. As such, the recipe model of proofs predicts that (H4) does not hold.

In sum, (H1) and (H2) are about the *frequency* of imperatives in mathematical texts, while (H3) is about "We [verb]..." constructions. The hypothesis (H4) is about the *diversity* of verbs used to form mathematical imperatives. We believe that both of these can be used in mathematical texts to give instructions so will investigate both below.

4.3 Methods

We used Laurence Anthony's *AntConc* software (Anthony 2020) and Yasu Imao's *CasualConc* software (Imao 2018) to analyse the corpora.[15]

[15] The *AntConc* software is available online at https://www.laurenceanthony.net/software/antconc/ (accessed 24/02/21) and *CasualConc* is available at https://sites.google.com/site/casualconc/ (accessed 06/03/21).

We had two primary targets for investigation from our hypotheses: explicit imperatives and "We [verb]" constructions. Both of these have textual markers that we could use. For explicit imperatives, we could perform a case-sensitive search for capitalised verbs, as this is a reasonable and conservative operationalisation that will capture many imperatives (e.g., "Let...," "Assume...," "Suppose...," and "Construct..." all are formed with a capitalised verb to start the imperative sentence.) A drawback of this is that it will miss any imperatives constructed in different ways, such as "Firstly, suppose...," but we did not have a systematic way to predict the idiosyncratic ways sentences may be constructed, and we believe that missing these did not alter our findings. For "We [verb]" constructions, the software allowed us to search for these strings directly.

For both explicit imperatives and "We [verb]" constructions, we needed to identify the specific verbs we will investigate in the first place. We required some way to select these systematically, but the verbs most likely to be used in mathematical imperatives are typically mathematical keywords; we shouldn't expect, for instance, that common mathematical imperatives would use the same verbs as imperatives in general English, so couldn't use that as a guide. To solve this, we produced a list of the 500 most common words in the *All-math* corpus and selected all of them that can ever be verbs in English. This gave a total of 91 possible verbs, listed in Table 1. One obstacle we faced was that many of the words have multiple meanings, for example, the word "set" is ambiguous between the verb form and the noun form and is used in both senses in the corpora. The same held for many of the terms for mathematical objects, such as "group," "ring," "field," and "structure." A second obstacle is that some of these words were extremely unlikely to be used as verbs in the current context, such as "stable," "like," and "hand." Finally, some of the verbs have existing functions in mathematics papers, such as "claim," "step," and "remark."

To separate out the verbs used as imperatives in the corpus, we generated a file for each of the 91 verbs containing all capitalised instances of that verb in the *All-math*

Table 1 The 91 potential verbs among the 500 most common words in the *All-math* Corpus

Action	Construct	Free	Lie	Order	Satisfy	Step
Apply	Curve	Function	Like	Pair	Say	Structure
Assume	Define	Further	Limit	Part	Second	Study
Bound	Denote	Get	Map	Point	See	Sum
Bundle	Do	Give	Measure	Present	Sense	Suppose
Call	Estimate	Graph	Model	Prime	Sequence	Take
Check	Exist	Group	Need	Process	Set	Term
Choose	Factor	Hand	Norm	Prove	Show	Time
Claim	Field	Have	Note	Rank	Side	Type
Complete	Figure	Hold	Number	Recall	Smooth	Use
Conclude	Find	Image	Observe	Remark	Space	Will
Condition	Fix	Introduce	Obtain	Respect	Stable	Work
Consider	Form	Let	Open	Ring	State	Write

Table 2 The 91 verbs divided into three categories

Imperatives		Mixed		Non-Imperatives			
apply	note	claim	number	have	hand	respect	structure
assume	observe	complete	order	action	hold	ring	term
call	obtain	do	pair	bound	image	satisfy	time
check	prove	estimate	present	bundle	lie	second	type
choose	recall	factor	process	condition	like	see	will
conclude	say	find	rank	curve	limit	sense	
consider	show	form	set	exist	need	sequence	
construct	suppose	get	study	field	norm	side	
define	take	give	sum	figure	open	smooth	
denote	use	group	work	free	part	space	
fix	write	map		function	point	stable	
introduce		measure		further	prime	state	
let		model		graph	remark	step	

corpus and their contexts (which shows six words either side of the verb). For example, an entry from the file for the verb "choose" is:

```
     [circle of a copy of inline_math.  Choose a base point
inline_math on inline_math,]
```

By examining these files, we separated the 91 verbs into three classes: (1) those where the capitalised form was used exclusively (or almost exclusively) as an imperative; (2) those where there was a mixture of imperatives and other uses; and (3) those where there were no imperatives. Separating the verbs into the three lists using the context files gave us the three categories reported in Table 2.

From here, we performed a selection of analyses using the corpus linguistics software, reported below.

4.4 Results

4.4.1 Comparing Imperatives in the *Proof-Only* and *Non-Proof* Corpora

To test (H1) and (H2) on explicit imperatives, we compared the frequency (per million words) of the capitalised verbs in our *Proof-Only* and *Non-Proof* corpora. We used only the verbs from the first category of Table 2 so that our operationalisation of using capitalised verbs to track sentences in the imperative mood will be accurate. Recall that the third category contained no imperatives, and the second category contained a mixture of imperatives and non-imperatives, which would interfere with the goal of the study. The results are displayed in Table 3. What we see is that the total frequency per million words for our selected verbs is 9406 in the *Proof-only* corpus and 6854 in the *Non-proof* corpus, a highly significant difference, Fisher's exact test, $p < 0.001$.

These results demonstrate that imperatives using common mathematical verbs appear more commonly within proofs than outside of proofs. If mathematical texts

Table 3 Comparing imperative frequencies per million words in the *Proof-only* and *Non-proof* corpora

Verb	Proof-only	Non-proof
Let	4523	4035
Suppose	944	512
Note	929	681
Consider	570	314
Assume	556	339
Recall	304	265
Define	272	167
Fix	255	106
Denote	218	145
Observe	213	92
Choose	199	45
Take	178	49
Write	117	39
Apply	53	6
Use	28	9
Call	14	11
Introduce	11	9
Construct	8	4
Say	7	7
Show	3	10
Check	2	2
Prove	1	4
Obtain	1	0
Conclude	0	0
TOTAL	9406	6854

outside of proofs, as in the *Non-Proof* corpus, are a reasonable baseline for imperatives in mathematical writing, then these results provide evidence for (H1) and against (H2).

The results here also show major variation in the prevalence of imperatives using different verbs. By far the most common is the imperative using "Let," followed by "Suppose" and "Note," then by "Consider" and "Assume." The imperatives using "Recall," "Define," "Fix," "Denote," "Observe," "Choose," "Take," and "Write" were also relatively common. Other capitalised verbs we searched for were all present to some degree but were less common overall. The only three capitalised verbs that showed up as a higher frequency per million words in the *Non-Proof* corpus were "Show," "Prove," and "Conclude."

4.4.2 Comparing "We [verb]..." Constructions in the *Proof-Only* and *Non-Proof* Corpora

To test (H3) concerning the "We [verb]..." constructions, we took all 91 of the (possible) verbs from the top 500 most common words in the *All-math* corpus (i.e., all of the verbs found in Table 1) and looked for the "We [verb]" string in the

Table 4 Comparing "We [verb]" construction frequencies per million words in the *Proof-Only* and *Non-Proof* Corpora.

Verb	Proof-Only	Non-Proof	Verb	Proof-Only	Non-Proof
We apply	51	18	We model	0	0
We assume	59	73	We need	92	39
We bound	3	1	We norm	0	0
We bundle	0	0	We note	61	75
We call	12	85	We number	0	0
We check	11	3	We observe	30	20
We choose	55	27	We obtain	65	35
We claim	231	34	We open	0	0
We complete	5	2	We order	1	1
We conclude	137	41	We point	2	11
We condition	0	0	We present	4	17
We consider	95	108	We process	0	0
We construct	21	16	We prove	128	69
We define	96	140	We recall	32	70
We denote	77	158	We remark	14	35
We do	18	35	We satisfy	0	0
We estimate	13	3	We say	9	147
We factor	1	0	We see	40	22
We find	12	16	We set	71	40
We fix	37	24	We show	105	74
We form	1	1	We smooth	0	0
We further	8	10	We state	1	9
We get	53	21	We structure	0	0
We give	20	42	We study	3	23
We group	1	0	We sum	1	1
We have	666	348	We suppose	9	15
We hold	0	0	We take	37	21
We introduce	14	33	We term	0	0
We let	31	27	We use	121	83
We like	0	0	We will	318	387
We limit	0	0	We work	3	7
We map	0	0	We write	47	52
We measure	0	0			
			TOTAL	2923	2522

Proof-Only and *Non-Proof* corpora. We made this case-sensitive to only look at instances where the "We" is capitalised, thus indicating the start of a sentence. The results are reported in Table 4.[16]

[16] We have omitted all the verbs that returned no results. As such, the zero entries that do appear in the table are the result of rounding.

The results here are more mixed than in the previous section. The two biggest entries are from "We have" and "We will," used to form the past and future tenses, so they are not in and of themselves giving implicit instructions, but could be used that way. Especially "We have" has a diverse range of potential uses: it could describe carrying out an action in the past tense, as in "We have applied...," or can be used to summarise as in "We have seen...." It can be used in the present tense to emphasise tools, facts, or mathematical objects at our disposal at that point in the proof, as in "We have that...." It can also be used to point forwards to set a shared target, as in "We have to [achieve some goal/carry out some procedure/etc.]." These constructions can also be used in the abstracts of the papers, describing what the authors do in the paper, in which case the "we" is not being used to include the reader, unlike within a proof. Examining the corpus revealed examples of all of these constructions.

While the totals overall again find more "We [verb]" constructions in the *Proof-Only* corpus than the *Non-Proof* corpus (and this difference is again significant, Fisher's exact test, $p < 0.001$), most of the difference can be accounted for just by the "We have" construction, which we just noted can be used in a variety of ways, and so does not tell us much about instructions being issued.

Looking more at the specific verbs, the more common imperatives we found above do not appear to be anywhere as common in the "We [verb]" form: none of the verbs used in the thirteen most common imperatives appear in this form more than 100 times per million words in the *Proof-Only* corpus. Nonetheless, several other verbs are more common in this form, especially "We claim," "We conclude," "We prove," "We show," and "We use." What is notable about these constructions is that they appear to be more structural than the verbs we found as imperatives, in that they are used to structure where the proof is aiming to get to, or (in the case of "We use") what method or fact will be employed to get there.

With regard to (H3), we did find plenty of "We verb" constructions in the *Proof-Only* corpus, suggesting that they are not uncommon, but we also didn't find them to be overall more common than in the more general mathematical texts, as represented by the *Non-Proof* corpus. The lack of more substantial finding here may be because of the restriction to the start of a sentence, as identified with starting capital letters. We will return to this concern in Sect. 4.5 and look at these constructions in an example from the corpus.

4.4.3 Imperative Frequency in *Proof-Only* Files

In Sect. 4.4.1, we tested (H1) and (H2) by looking at the frequency of imperatives in proofs relative to their frequency in mathematical texts outside of proofs. To provide another angle on how common imperatives are, we now investigate the percentage of files they appear in. Recall that our *Proof-Only* corpus consists of 3,268 files. A file in our corpus contains all of the proofs from one file uploaded to the arXiv. Barring a minor complication, we can thus take this to give us a representation of

what percentage of papers that contain proofs at all also contains a given imperative somewhere in those proofs.[17]

Once again, to look for imperatives we searched for the capitalised verbs from the first category in Table 2 using a case-sensitive search such that, for example, it will count "Consider" and "consider" separately.

The results are found in Table 5 in the left columns. To put the percentages in some context, we have included an arbitrarily selected word which appears in a similar percentage of the papers, listed on the right.

The results agree with the findings from Sect. 4.4.1, with the same set of verbs serving most often as imperatives in proofs. The results also show that these

Table 5 Number of and percentage of files in the *Proof-Only* corpus containing the capitalised verb, alongside other keywords appearing at a roughly similar frequency for reference

Verb	Number of files	% of files	Nearby word	Number of files	% of files
Let	2692	82.4%	then	2706	82.8%
Note	1486	45.5%	function	1477	45.2%
Suppose	1250	38.3%	thus	1274	39.0%
Consider	1186	36.3%	So	1186	36.3%
Assume	1085	33.2%	bounded	1053	32.2%
Recall	844	25.8%	know	843	25.8%
Define	712	21.8%	simple	722	22.1%
Fix	606	18.5%	action	598	18.3%
Choose	541	16.6%	length	553	16.9%
Denote	538	16.5%	precisely	544	16.7%
Take	459	14.1%	always	465	14.2%
Observe	447	13.7%	less	455	13.9%
Write	343	10.5%	derivative	348	10.7%
Apply	173	5.3%	block	173	5.3%
Use	104	3.2%	convenient	122	3.7%
Call	57	1.7%	congruent	54	1.7%
Introduce	44	1.4%	Lagrangian	42	1.3%
Construct	34	1.0%	walk	32	1.0%
Say	27	0.8%	adopt	27	0.8%
Check	7	0.2%	Nash	7	0.2%
Show	5	0.2%	Method	5	0.2%
Prove	5	0.2%	functionally	5	0.2%
Obtain	2	0.1%	unpruned	2	0.1%
Conclude	1	0.0%	Hamiltonianly	1	0.0%

[17] The minor complication is that if a paper was uploaded as multiple files, then it will remain as multiple files in our corpus. However, only files containing proofs at all will be included in the corpus, which reduces the frequency of this. Only a small minority of the files were uploaded this way, and this is conservative such that the effect on our results is that the percentage of files we report is a slight underestimate of the percentage of papers a given word appears in.

imperatives are not only common relative to non-proof mathematical texts, but also common in general, appearing in a substantial proportion of the files of the corpus, and therefore in a substantial percentage of the papers containing proofs.

4.4.4 Imperative Diversity

Our final hypothesis (H4) concerned the diversity of verbs used in imperatives and "We [verb]" constructions in mathematical proofs, alluding to Ganesalingam's claim that they are drawn only from a limited and closed class of verbs. Our more fine-grained results from the previous sections already address this hypothesis to some degree.

Of course, the truth of (H4) depends on what that class is. For example, the English language is finite, so on one reading the claim could be trivially true. We propose that a natural reading of Ganesalingam (2010) in light of our results would be to have the "limited and closed class of verb" be those that we found to be most common in Sects. 4.4.1 and 4.4.3, namely, "let," "consider," "assume," "denote," "note," "define," "suppose," "recall," "write," "take," "choose," "fix," and "observe." From our results in the previous sections, then, it is correct to say that a large number of the instructions in proofs are given with these verbs. This is clearer in the case of imperatives from Sects. 4.4.1 and 4.4.3, but also occurs in the "We [verb]" constructions.

However, the hypothesis of (H4) states that these verbs are the *only* ones used in instructions given by imperatives and "We [verb]" constructions. Above, we found many other instances of verbs being used both as imperatives and in "We [verb]" constructions. Here, we offer two more pieces of evidence for there being a diverse number of instructions issued in mathematical proofs.

First of all, in Sect. 4.3 we set aside 22 of our 91 verbs because in the *All-Math* corpus they were used in their capitalised form both as imperatives and in some other ways, which would have obscured the above studies. These were the verbs listed in the middle column of Table 2. Returning to these verbs, we searched for them in their capitalised forms in the *Proof-Only* corpus. Of these, 16 returned results that were clearly imperatives.[18] For each of these, we selected a clear example, displayed in Table 6. Given that these examples sound like normal written mathematics, this demonstrates further diversity of the verbs used as imperatives in the *Proof-Only* corpus.

The second way we chose to look at the diversity of verbs used as imperatives in proofs was using the mathematical verbs listed by Rotman in the quote in Sect. 3.1. We omit all of those that overlapped with the previous analysis. Like in Sect. 4.4.1, we compared the frequency per million words of the capitalised versions of the verbs between the *Proof-Only* and the *Non-Proof* corpora. The results are displayed in

[18] We avoid giving the numbers of imperatives because not all of the cases were clear, and mostly the numbers were low. That is not important, though, as here we are interested in diversity of imperatives rather than the frequency.

Table 6 Examples of imperatives from the *Proof-Only* corpus for the verbs found in the middle column of Table 2

Verb	Example
Set	Set the total degree equal to the sum of the bidegrees.
Group	Group the non-constant linear generators of inline_math so that [. . .]
Map	Map this product to a neighborhood inline_math of the torus knot as follows [. . .]
Form	Form the commutative cube in which the front and back faces are pullbacks, so that [. . .]
Order	Order the eigenvalues of inline_math so that the inline_math eigenvalue of the spectrum of inline_math is [. . .]
Number	Number the elements of inline_math.
Get	Get rid of the self-intersection as before to obtain as before.
Give	Give the knot inline_math an orientation inline_math.
Find	Find a partition of unity inline_math on inline_math subordinate to the cover inline_math.
Sum	Sum the estimates in the previous corollary.
Do	Do the above coarsening process for every element inline_math with inline_math.
Complete	Complete this elimination process, ending with inline_math.
Pair	Pair these elements and return to inline_math and repeat the process, skipping over any letters already paired.
Estimate	Estimate the difference on the right hand side of [. . .] by the triangle inequality to find [. . .]
Study	Study the case inline_math.
Factor	Factor inline_math as a product of prime ideals inline_math in inline_math.

Table 7. These verbs appeared significantly more often per million words in proofs than outside of proofs, Fisher's exact test, $p = 0.006$.

One striking feature here is that the verbs suggested by Rotman are orders of magnitude less common than those selected by our method above, suggesting that while these may be mathematical words broadly speaking, they are not particularly common ones. More importantly for (H4), we did find that many of them were present as imperatives in our *Proof-Only* corpus. Once again, this result suggests that many more verbs are amenable to being used as imperatives in mathematical proof texts. We thus reject (H4), as the class of verbs does not appear to be closed or limited.

4.5 An Example

It is easy to lose oversight of what these results look like in practice. To address this, and provide some reflection on the methodological choices made above, we will now demonstrate the use of imperatives and instructions in our corpus with an illustrative example. We will look at the preprint of the paper by Mazzeo and Rowlett (2015), selected as the chronologically first paper in our corpus to contain at least one proof.

Table 7 Comparing the Rotman verb imperative frequencies per million words in the *Proof-Only* and *Non-Proof* Corpora

Verb	Proof-only	Non-proof
Add	2.01	2.28
Adjoin	0	0.08
Assign	1.81	1.12
Calculate	0.80	1.50
Compute	2.41	6.29
Connect	2.41	0.58
Count	0	0.15
Delete	0.60	0.81
Demonstrate	0	0.04
Differentiate	0.40	0.58
Draw	0.60	0.77
Drop	0	0.23
Eliminate	0	0.12
Enumerate	2.21	1.00
Evaluate	0.20	0.62
Exhibit	0	0.00
Integrate	3.62	0.50
Iterate	1.01	0.31
List	0.40	2.16
Multiply	7.84	0.93
Reverse	0.20	0.58
Specify	0	0.46
Subtract	1.21	0.12
TOTAL	27.74	21.22

Here is an excerpt from the introduction section of the paper. This contains three imperatives ("For simplicity, suppose," "Let," and "Note") of which our analysis using capitalisation would have found two. There are also two "We [verb]" constructions ("We now explain" and "We assume"), of which we would have found one. Interestingly, in the context of an introduction, the "We [verb]" phrases are not acting to include the reader in epistemic actions but instead are simply describing what the authors are doing. What is also noteworthy is that this paragraph is relatively high in instructional language, but is also *proof-like* in the sense that it is describing a piece of mathematical reasoning. Looking further through the paper, this seems to be the case for a large number of the imperatives being used. Direct imperatives appear in theorem statements, lemma statements, numbered remarks, and in the long paragraphs of description of mathematical reasoning. Since our analysis placed all of those in the *Non-proof* category, the results we have described are likely to understate the extent to which imperatives are characteristic of proofs in research papers, since this style of proof writing can spill over into the surrounding text (Fig. 1).[19]

[19] Our thanks to a referee for pushing us to be explicit about this important point.

We now explain the desingularization more precisely. For simplicity, suppose that Ω_0 and Ω_ϵ all lie in some slightly larger ambient open surface $\widetilde{\Omega}$, and that the metrics g_ϵ on Ω_ϵ are all extended to metrics (still denoted g_ϵ) on this larger domain. We assume that this family of metrics converges smoothly on $\widetilde{\Omega}$. Let p be a vertex of Ω_0 and consider the portion of Ω_ϵ in some ball of fixed size around p, $B_c(p) \cap \Omega_\epsilon$. Our main assumption is that the family of pointed spaces $(B_c(p) \cap \Omega_\epsilon, \epsilon^{-2} g_\epsilon, p)$ converges in pointed Gromov-Hausdorff norm, and smoothly, to a noncompact region $Z \subset \mathbb{R}^2$ with smooth boundary, such that at infinity, ∂Z is asymptotic to a cone with vertex at 0 and with opening angle α, the same angle as at the vertex p in (Ω_0, g). Note that this is actually pointed Gromov-Hausdorff convergence for the ambient space $(\widetilde{\Omega}, g_\epsilon, p)$.

Fig. 1 An excerpt from the introduction of Mazzeo and Rowlett's (2015) preprint

Proof of Theorem 1.6: We first construct a particular family of parametrices for the heat kernel on Ω_ϵ. For any $0 \leq \epsilon < \epsilon_0$, decompose

$$\Omega_\epsilon = \Omega_{\epsilon,1} \cup \Omega'$$

where $\Omega_{\epsilon,1} = \Omega_\epsilon \cap B_1(p)$, and $\Omega' = \Omega_\epsilon \setminus (\Omega_\epsilon \cap B_1(p))$. Note that Ω' is independent of ϵ. Lemma 2.2 shows that

$$H^{\Omega_\epsilon}(t,z,z) = \chi_1(z) H^{\epsilon Z}(t,z,z) + \chi_2(z) H^{\Omega_0}(t,z,z) + K(t,z), \qquad (3.1)$$

Fig. 2 An excerpt from the proof of Theorem 1.6 of Mazzeo and Rowlett's (2015) preprint

The first substantial proof is of Theorem 1.6, which spans over several pages. The opening lines of this proof can be seen in Fig. 2.

The proof contains eight direct imperatives (using the verbs "decompose," "note," "write," "set," "choose," and "let"), of which our analysis would have found five. In fact, we have not looked for imperatives using the verb "decompose" above, so this reinforces our finding of diversity of imperative phrases. The proof also contains ten instances of using "we" with a verb to indicate activities (using the verbs "construct," "denote," "see," "defer," "examine," "write," "prove," "appeal," and "find"). Surprisingly, only two of these were in the exact "We [verb]" form, and most were embedded in much more complicated sentence structures. This is coherent with our analysis that we did find "We [verb]" phrases in both the *Proof-only* and the *Non-proof* corpora but suggests that further analysis might reveal a richer picture. Indeed, this proof also contains several phrases that suggest epistemic and mathematical actions in other ways: "For that, we may as well replace the upper limit of integration by [...]," "To finish the proof, we must analyze the behaviour of [...]," and "The decay of each of term as $|\lambda| \to \infty$ is straightforward, so we can write [...]."

These three all make use of modal language, something examined by Hodges (2013). Computational corpus linguistics may be a suitable method for capturing these more complex constructions, but would require more advanced tools than we have used here.

We also note that at the time of writing, Keith Weber (submitted) has carried out a smaller scale analysis of the use of instructions and imperatives, examining the entirety of Kunen's classic textbook *Set theory: An introduction to independence proofs* (1980). This will provide a useful piece of methodological triangulation for our results, something which Löwe and Van Kerkhove (2019) have argued is centrally important for empirical approaches to mathematical practice. While our computational analysis allows us to examine trends across a large number of papers, Weber's approach does not rely on automation and thus allows him to find and analyse instructions our methods wouldn't catch. Indeed, Weber is able to carry out an interesting investigation of the different kinds and functions of instructions he finds, providing different insights into the instructional language of proofs. Nonetheless, Weber's results largely confirm our results, finding both a high frequency and broad diversity of instructions and imperatives in Kunen's text. There are some differences in the exact distribution of the imperatives he finds, but this is easily explained by the fact that this is a single author with certain stylistic preferences. In light of Weber's work, we can feel more confident that the various small methodological choices we have made above have delivered reliable results.

5 Conclusion

Our analysis of the language used in a large number of mathematics papers allows us to make empirically informed contributions to the philosophy of mathematical practice, and especially questions concerning the nature of real, modern mathematical proofs. We now relate our findings to the philosophical debates discussed earlier in the chapter.

Recall that we were interested in the recipe model of proofs, which tells us that proofs are akin to recipes, commanding that we undertake various kinds of mathematical actions to perform a piece of reasoning. These actions were not limited to inferential actions, but also include constructing and manipulating mathematical objects, fixing notations, entertaining hypotheticals, structuring their reasoning, and more besides. Viewing proofs in this way has important implications for understanding the nature of proofs, the epistemology of proofs, the functions of proofs, how to best educate students to engage with proofs, and how to best integrate formal proof checkers into mathematical practice. As such, while it may appear to be a mere stylistic variation, we hold that the specifics of how proofs can be written is of substantial importance in the philosophy of mathematical practice.

Based on the recipe model, we predicted that imperatives and instructions are both common and diverse in mathematical proof texts. The results of our study certainly lend support to the recipe model because we did indeed find that imperatives and

instructions are used often and use a broad range of verbs. However, there were a core set of more common imperatives using the verbs "let," "consider," "assume," "denote," "note," "define," "suppose," "recall," "write," "take," "choose," "fix," and "observe." This core set of verbs make up the bulk of the imperatives we found in our study, suggesting that these imperatives are used in fairly standardised ways for particular mathematical instructions. Nonetheless, we found many different verbs used as imperatives in our corpus. While many of these were used less frequently, it does indicate that there is also the flexibility in written proofs to deploy a diverse selection of instructions using imperatives.

Finding broad empirical support for the recipe model of proofs is also beneficial for the epistemic picture underlying it. If proofs are giving us instructions to carry out, then mathematical knowledge is more intimately tied to knowledge-how than is standardly recognised.

Our results should also be of interest to the formal mathematics community. If part of the aim of formal proof checkers is to gain widespread use among mathematicians by making the formal tools better at reading and producing natural-sounding proofs, then we will need accurate information on how those proofs do and ought to look. The results in this chapter suggest that such natural-sounding formally-verifiable proof should contain imperatives and "We [verb]" constructions, especially those most common verbs.

We should re-emphasise a note of caution from above. Our results are for papers from the arXiv and so represent a particular kind of mathematical text. It is thus not clear whether they would generalise to other contexts, such as proofs in textbooks, classrooms, or seminars. We can speculate that journal proofs may be at the more formal end of mathematical language, and therefore we might expect greater flexibility in terms of the verbs used to form instructions in other contexts. Research articles are also a context where one assumes a great deal of competence from the reader, so we might also hope that proofs aimed more at teaching will include more imperatives as a communicative aid.

To conclude, we have made progress in investigating real mathematical proofs in a systematics and large-scale way. If we truly want our philosophy of mathematics to attend to the realities of practice, then these methods should be integral to uncovering what those realities are.

Acknowledgments Work by the first author was funded by the Research Foundation – Flanders (FWO) – Project G083620N: *The Epistemology of Data Science: Mathematics and the Critical Research Agenda on Data Practices*. We are grateful to Lara Alcock, Kristen Lew, Pablo Mejía-Ramos, Paolo Rago, and Chris Sangwin for the assembly and processing of the corpus with which we conducted this work. We would also like to thank Laurence Anthony for the creation and free distribution of the *AntConc* software with which we carried out our main analysis. Likewise, we thank Yasu Imao for the creation and free distribution of *CasualConc*, with which we carried out some subsidiary analyses. We thank audiences in Prague and online via Copenhagen for helpful feedback. We are grateful to Valeria Giardino and Roy Wagner for the careful refereeing of this chapter. We also had comments and suggestions from a large number of friends and colleagues, to whom we also extend our thanks.

References

Aberdein A, Inglis M (eds) (2019) Advances in experimental philosophy of logic and mathematics. Bloomsbury, London

Alcock L, Inglis M, Lew K, Mejía-Ramos JP, Rago P, Sangwin C (2017) Comparing expert and learner mathematical language: a corpus linguistics approach. In: Weinberg A, Rasmussen C, Rabin J, Wawro M, Brown S (eds) Proceedings of the 21st annual conference on research on undergraduate mathematics education. RUME, San Diego, pp 478–484

Andersen LE (2020) Acceptable gaps in mathematical proofs. Synthese 197:233–247

Anthony L (2020) AntConc (Version 3.5.9) [Computer Software]. Waseda University, Tokyo. Available from https://www.laurenceanthony.net/software

De Lon A, Koepke P, Lorenzen A (2020) Interpreting mathematical texts in Naproche-SAD. In: Benzmüller C, Miller B (eds) CICM 2020, LNAI, vol 12236, pp 284–289. https://doi.org/10.1007/978-3-030-53518-6_19

De Toffoli S (2020) Groundwork for a Fallibilist account of mathematics. Philos Q:1–22. https://doi.org/10.1093/pq/pqaa076

Ernest P (1998) Social constructivism as a philosophy of mathematics. State University of New York Press, New York

Ernest P (2018) The ethics of mathematics: is mathematics harmful? In: Ernest P (ed) The philosophy of mathematics education today. Springer Nature, Cham, pp 187–216

Fallis D (2003) Intentional gaps in mathematical proofs. Synthese 134:45–69

Frerix S, Koepke P (2018) Automatic proof-checking of ordinary mathematical texts. In: Hasan O, Kaliszyk C, Naumowicz A (eds) Proceedings of the workshop formal mathematics for mathematicians (FMM), Hagenberg, Austria, 13–Aug 2018, published at http://ceur-ws.org

Gerhardt C, Frobenius M, Ley S (eds) (2013) Culinary linguistics. The chef's special. John Benjamins Publishing Company, Amsterdam

Halmos P (1970) How to write mathematics. Enseign Math 16:123–152

Hamami Y (2014) Mathematical rigor, proof gap and the validity of mathematical inference. Philos Sci 18:7–26

Hodges W (2013) Modality in mathematics. Log Anal 56:5–23

Imao Y (2018) CasualConc (Version 2.1.0) [Computer Software]. Osaka University, Osaka. Available from https://sites.google.com/site/casualconc/

Knuth DE, Larrabee T, Roberts PM (1989) Mathematical writing, MAA notes number 14. Mathematical Association of America, Washington, DC

Krantz SG (1997) A primer of mathematical writing. American Mathematical Society, Providence

Kunen K (1980) Set theory: an introduction to independence proofs. Elsevier, Amsterdam

Lew K, Mejía-Ramos JP (2019) Linguistic conventions of mathematical proof writing at the undergraduate level: Mathematicians' and students' perspectives. J Res Math Educ 50:121–155

Löwe B, Van Kerkhove B (2019) Methodological triangulation in empirical philosophy (of mathematics). In: Aberdein A, Inglis M (eds) Advances in experimental philosophy of logic and mathematics. Bloomsbury, London, pp 15–37

Mazzeo R, & Rowlett J (2015) A heat trace anomaly on polygons. Math Proc Camb Philos Soc 159: 303–319. https://arxiv.org/pdf/0901.0019.pdf

Mejía-Ramos JP, Inglis M (2011) Semantic contamination and mathematical proof: can a non-proof prove? J Math Behav 30:19–29

Mejía-Ramos JP, Alcock L, Lew K, Rago P, Sangwin C, Inglis M (2019) Using corpus linguistics to investigate mathematical explanation. In: Fischer E, Curtis M (eds) Methodological advances in experimental philosophy. Bloomsbury, London, pp 239–264

Morgan C (1996) 'The language of mathematics': towards a critical analysis of mathematics texts. Learning Math 16:2–10

Pimm D (1987) Speaking mathematically: communication in mathematics classrooms. Routledge & Kegan Paul, London

Reyes-Galindo LI (2016) Automating the Horae: boundary-work in the age of computers. Soc Stud Sci 46:586–606

Rittberg CJ, Tanswell FS, Van Bendegem JP (2018) Epistemic injustice in mathematics. Synthese 197:3875–3904. https://doi.org/10.1007/s11229-018-01981-1

Rotman B (1988) Towards a semiotics of mathematics. Semiotica 72:1–35

Rotman B (1993) Ad Infinitum... The Ghost in Turing's Machine. Stanford University Press, Stanford

Swales JM, Ahmad U, Chang Y, Chavez D, Dressen-Hammouda D, Seymour R (1998) 'Consider This. . .': the role of imperatives in scholarly writing. Appl Linguis 19:97–121

Tanswell FS (2016) Proof, rigour and informality: a virtue account of mathematical knowledge. PhD thesis, University of St Andrews. https://research-repository.st-andrews.ac.uk/handle/10023/10249

Tanswell FS (2018) Conceptual engineering for mathematical concepts. Inquiry 61:881–913

Tanswell FS (forthcoming) Go forth and multiply: on actions, instructions and imperatives in mathematical proofs. In: Brown J, Bueno O (eds) Essays on the philosophy of Jody Azzouni. Springer, Cham

Tanswell FS, Rittberg CJ (2020) Epistemic injustice in mathematics education. ZDM 52:1199–1210. https://doi.org/10.1007/s11858-020-01174-6

Wagner R (2009) S(zp, zp): post-structural readings of Gödel's proof. Polimetrica International Scientific Publisher, Milan

Wagner R (2010) Who speaks mathematics: a semiotic case study. Philos Perspect Math Pract 12:205–234

Weber K (submitted) Instructions and constructions in set theory proofs (unpublished manuscript)

Section XIV

Semiology of Mathematical Practice

Juan Luis Gastaldi

Introduction to the Semiology of Mathematical Practice

Juan Luis Gastaldi

Contents

1 Introduction 2956
2 What Do You Mean, "Semiology"? 2956
3 Semiology in History and Philosophy of Mathematics: Contributions to This Section 2961
4 Conclusion 2968
5 Cross-References 2969
References 2969

Abstract

The philosophy and history of mathematical practices have brought the study of mathematical language and signs to the forefront of contemporary mathematical thought. However, despite the fruitfulness of this research trend, a comprehensive and unified account of its various aspects and the diverse approaches taken to explore it remains elusive. Recognizing this gap, we have undertaken the task of editing the present section of the Handbook of the History and Philosophy of Mathematical Practice as a much-needed remedy. Before providing an overview of the various contributions to the section, this introduction provides some context for the subject matter and a few conceptual clarifications.

Keywords

Semiology · Semiotics · Signs · Peirce · Saussure · Mathematics · Mathematical Practice · Language · Textuality

J. L. Gastaldi (✉)
ETH Zurich, Zurich, Switzerland
e-mail: juan.luis.gastaldi@gess.ethz.ch

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_130

1 Introduction

The purpose of the following pages is to present the section "Semiology of Mathematical Practices" in this Handbook of the History and Philosophy of Mathematical Practice (Sriraman 2020). The section provides a comprehensive account of the philosophical and historical approaches to the study of mathematical *signs* and *language* from the viewpoint of mathematical practice. Conceived as a whole, the series of contributions that make up this section demonstrate the consistency and diversity of these approaches, offering a unified yet plural perspective on this important dimension of the history and philosophy of mathematical practices.

The section is composed of nine chapters written by leading specialists, providing a rich account of the state of the art in most areas of this research field. The initial chapters present the philosophical and historical underpinnings of a practice-driven theory of mathematical signs. They are followed by compelling case studies, which demonstrate the fruitfulness of the semiological approach in analyzing the historical and conceptual aspects of mathematical practices. Finally, the section offers fresh insights into current and upcoming challenges where a theory of mathematical signs can play a critical role.

The authors of this section have done an exceptional job of enriching their unique perspectives with a comprehensive introduction to the respective topics, along with a presentation of the necessary context and conceptual tools. Consequently, attempting to provide a detailed introduction to the semiology of mathematical practices here would lead to redundancy. Instead, the goal of these pages is to offer a concise presentation of all the contributions, accompanied by a few elements of context and some conceptual clarifications that could only be addressed from the general standpoint enabled by the completed work.

2 What Do You Mean, "Semiology"?

An initial clarification is in order concerning the term "semiology." While some may find the term either unknown or outdated, it addresses a precise object of inquiry that no other alternative seems to capture entirely. At its core, semiology is the *theory of signs*. The study of signs has a rich history, stretching back to ancient medical practices where signs and symptoms were essential in diagnosing ailments. Building upon this foundation, the Stoics further developed the concept of the sign within the context of logical thought. In modern philosophy, John Locke drew upon this tradition to characterize one of the three divisions of science as "σημειωτική" (*semeiotiké*), concerned with "the nature of signs the mind makes use of for the understanding of things, or conveying its knowledge to others" (Locke 2013, p. 720). In contemporary thought, semiology has two primary roots that have significantly shaped its present understanding. One of these roots can be traced back to Charles Sanders Peirce (1839–1914) within the tradition of Boolean logic. Peirce retained the term "semiotics" or "semeiotics" to refer to the philosophical study of the semiosis process, whereby a sign, a signified object, and an interpretant

are related. Parallelly, within the linguistic tradition, Ferdinand de Saussure (1857–1913) introduced the term "semiology" as a key component of the ideas leading to the structuralist revolution in the field. At the opening of his influential *Course in General Linguistics* (Saussure 1959), Saussure stated:

> *A science that studies the life of signs within society* is conceivable; it would be a part of social psychology and consequently of general psychology; I shall call it *semiology* (from Greek *sēmeîon* 'sign'). Semiology would show what constitutes signs, what laws govern them. Since the science does not yet exist, no one can say what it would be; but it has a right to existence, a place staked out in advance. Linguistics is only a part of the general science of semiology; the laws discovered by semiology will be applicable to linguistics, and the latter will circumscribe a well-defined area within the mass of anthropological facts. (Saussure 1959, p. 16)

As Saussure explains in the passages following these lines, postulating semiology as a scientific discipline was crucial for successfully "assigning linguistics a place among the sciences" (p. 16). The reason is that a semiological approach to language seeks to determine the "true nature of language" by examining its similarities and differences with other sign systems, rather than deferring the task to the study of individual behaviors or social institutions. In this way, semiology affords the theory of language independence from fields such as psychology or sociology.

It is this specific feature that our use of the term "semiology" would like to emphasize.[1] More precisely, our choice seeks to highlight the idea that there exists a consistent region of phenomena *akin to language*, which cannot be fully understood by relying solely on current linguistic analysis without missing crucial aspects of them. Yet, attempting to approach those phenomena from extra-linguistic perspectives can lead to even worse reductions to the principles of other disciplines such as psychology, sociology, logic, biology, physics, or even mathematics. If we agree to call "*signs*" those nonlinguistic entities with language-like characteristics, the term "semiology" refers to the idea that these signs, irrespective of their connections to other domains of knowledge, are *things* that deserve independent and thorough study *in and of themselves*.

A less pedantic way of expressing the same idea is to say that the perspective adopted here considers semiology as a form of *generalized linguistics*, encompassing languages other than natural language. Hence its relevance to addressing the problem of what is usually referred to, for lack of a better expression, as *mathematical language*. Throughout its documented history, mathematical practice has employed all kinds of semiotic artifacts and principles to express the specific contents at stake. Those artifacts involved natural language in significant ways. Yet, crucially, they also transcended those resources in an equally essential manner. In its minimal expression, the idea of a semiology of mathematics is that all those expressive means and their relationship to their respective contents should be studied

[1] For a more comprehensive approach, the reader can consult Barthes (1988); Sebeok (2001); Robering et al. (1996).

as such from a perspective that leaves room to explore the similarities and differences with other linguistic phenomena.

A second point of clarification concerns the distinction between the terms "semiology" and "semiotics." While these terms are often used interchangeably, our focus on the idea of a generalized linguistic in this section motivates a slight preference for "semiology," as evidenced by our choice of the section title. Other arguments stemming from the evolution of the structuralist tradition could be also summoned. In particular, Louis Hjelmslev, following Saussure's program, offered explicit definitions precisely characterizing the relation between the two notions. In Hjelmslev (1975), we can find the following definitions for both terms:

> A *Semiotic* [. . .] is a Hierarchy, any of whose Components admits of a further Analysis into Classes defined by mutual Relation, so that any of these classes admits of an analysis into Derivates defined by mutual Mutation. (Df. 24, p. 11)

> A *Semiology* [. . .] is a Metasemiotic that has a Non-Scientific Semiotic as an Object Semiotic (Df 47, p. 15)

Although Hjelmslev's definitions may appear obscure – in particular, each capitalized expression has been previously defined in his treatise – his idea is simple: the term "semiotic" is reserved to refer to any sign *system*, while "semiology" denotes those sign systems that take other semiotics as their objects of analysis. These definitions have at least the merit of replacing the alternative between both terms with a joint and coherent characterization. In our case, this would mean characterizing mathematics as a semiotic, and our theories about mathematical signs as semiologies.[2]

However, although these and other arguments could be put forth to justify the choice of either term, it is crucial to keep in mind that our intention from the outset of this project was to take a minimalistic and nonprescriptive stance. Above all, our primary aim for this section was to enable a pluralistic approach to the study of mathematical language and signs, granting authors the freedom to expose their unique perspectives on this fundamental aspect of mathematical practices. Accordingly, we intentionally refrained from any standardization in the use of the terms "semiotics" and "semiology," nor did we encourage specific conceptual choices or theoretical preferences that the use of one term or the other might implicitly entail. Hence, as a rule, both terms are to be read as synonyms across the different chapters of this section, and any variations should be attributed to the authors' reflective choices or individual tastes.

[2] The fact that semiologies are defined as having "non-scientific" semiotics as objects should not lead us to assume that mathematics cannot be studied as a science from a semiological perspective. For Hjelmslev, a "scientific semiotic" is merely a semiotic organized as a specific analytical procedure. If mathematics were to fall into this category, its study as a sign system would be termed a "meta-(scientific semiotic)" according to Hjelmslev's view (cf. Hjelmslev 1975, Df. 41,46).

Notwithstanding the plurality of perspectives represented in this section, there is one significant feature setting the semiological standpoint apart from other philosophical approaches to language, such as logicism, or generative linguistics, with which all our authors seem to agree to some extent. This crucial aspect is the key role attributed to *practices* in the study of signs. This will come as no surprise in a Handbook on the history and philosophy of mathematical practices. However, from a semiological viewpoint, the relationship between signs and practices is not an accidental feature. Given its significance, it might be helpful to provide some conceptual insight on this point.

The emphasis on practices is a component of both the Peircean and the structuralist traditions. However, it is the latter that has given a more prominent role to it, as evidenced by the development of Saussure's semiological project within various disciplines such as anthropology, psychology, sociology, and history, each one addressing specific dimensions of human practices. Within this tradition, the notion of practice is consubstantial to that of sign. We have seen how Saussure's semiological project placed language as part of "the life of signs within a society," determining "a well-defined area within the mass of anthropological facts" (Saussure 1959, p. 16). Yet, a more systematic account of the internal connection between signs and practices is provided by Hjelmslev and Uldall. In Hjelmslev and Uldall (1936), the authors write:

> A language consists of three concentric parts [. . .]: A central part, the *system*, i.e. the elements arranged in a pattern of mutual relations; the *norm*, i.e. a set of rules based on the system and fixing the limit of variability for each element; the *usage*, i.e. a set of rules based on the norm and fixing the limit of variability tolerated in a given community at a given time. It is necessary to distinguish between language as an institution, and the use to which it is put by the individual speaker; this is called the *practice* (Saussure's *parole*). (Hjelmslev and Uldall 1936)

Figure 1 offers a possible illustration of Hjelmslev and Uldall's (1936) views. In their text, the authors propose the example of the English phoneme /r/, which, as an element of the English linguistic *system*, is defined through a complex array of oppositions with respect to all the other phonological units of the system. Yet, as well specified as this phoneme may be, it leaves significant room for variability within those limits, thus defining a *norm* for English determined by all the possibilities of pronouncing the letter *r* without it being mistaken for another element of the phonetic system. Specific *usages* of English, such as the Scottish, Victorian, or Jamaican, exploit only a restricted region of that norm. Finally, individual *practices* represent singular manifestations of those collective usages, like the unmistakable accent of Bertrand Russell. These conceptual distinctions can be easily imported into the analysis of mathematical language. An unimaginative example is provided by arithmetic, where numbers, defined as elements of a system, can be expressed in a myriad of ways without violating elementary requirements of the assumed arithmetical system. However, different historical and geographic numeral systems (Roman, Arabic, binary, etc.) restrict the variability of that norm to only a few writing usage principles, which will then be practiced by individuals in their own style.

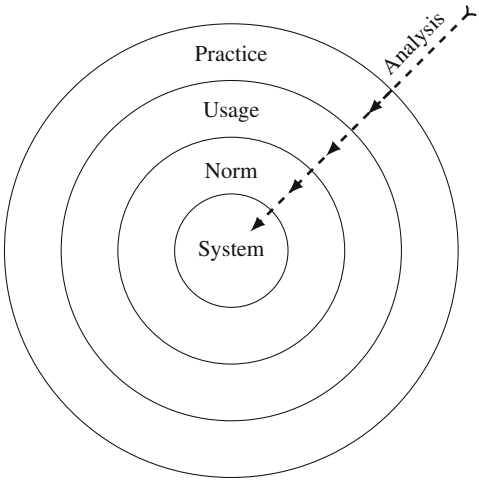


Fig. 1 Illustration of Hjelmslev and Uldall's (1936) view of semiological analysis through the various dimensions of language

Interestingly, Hjelmslev and Uldall affirm that the true object of a theory of language and signs is the exhaustive description of a language *as a system*. However, the key point is that the elements and rules composing semiotic systems *are not observable as such*. In their most fundamental nature, *languages and signs are only given in individual practices*. As long as the analysis of signs is to remain empirical, its task is to infer systematic features from idiosyncratic and constantly evolving practices. As Uldall and Hjelmslev say:

> The [. . .] system is found inductively through a series of ascending abstractions: an empirical study of the practice leads to the recognition of the usage, a study of the usage to the recognition of the norm, and a study of the norm to the recognition of the system. (Hjelmslev and Uldall 1936)

A semiology of mathematics is, therefore, inseparable from the analysis of mathematical practices, even as it aims to reveal the underlying system. Nonetheless, a semiological approach must also characterize the distinctiveness of mathematics compared to all other sign systems. In this regard, it is not impossible that part of what makes mathematical practice different from most semiotic practices is its recurrent endeavor to explicitly articulate systematic features of its own language. Admittedly, these features are themselves conveyed through signs and are, as such, the object of individual practices. Peano axioms can certainly be taken as a characterization of the *system* of arithmetic. But this cannot erase the fact that such a system is the result of multiple systematization practices (including but not restricted to Peano's), involving a complex interplay with norms and usages. What is more, the stabilization of one system cannot entirely preclude the emergence of new and nonentirely equivalent systematizations in the future. The chapters in this section provide numerous examples of such a phenomenon. However, the explicit pursuit of systematization as a critical aspect of mathematical practices can motivate alternative

original trajectories in the semiotic space, distinct from the ones traditionally conceived for the study of other sign systems. These and other specificities of mathematical practices indicate the potential contribution a semiology of mathematics can offer to a general theory of signs.

3 Semiology in History and Philosophy of Mathematics: Contributions to This Section

In 1902, Bertrand Russell reproached Frege for not having "clearly disentangled the logical and linguistic elements of naming" (Russell 2010, p. 519). Half a century later, the young Noam Chomsky opposed any attempt to understand natural language as logical systems such as Caranap's:

> But in the case of the artificial 'languages' investigated by Carnap in his logical laboratory, there is little if any antecedent reason for regarding these as in any way comparable to the actual languages of the outside world. (Chomsky 1955, p. 43)

These cross-accusations involving the most prominent figures in the fields of logic and linguistics serve as a compelling illustration that despite the so-called "linguistic turn," insofar as the philosophy of mathematics has been governed by a classic logical approach, the true problems of language have been kept at arm's length. The tutelary figure of the late Wittgenstein only confirms this circumstance, as it demonstrates how a multitude of new problems linked to mathematical signs emerge once the strict logical perspective is subject to critical examination.

As a consequence, strictly semiological approaches have occupied a relatively marginal position in contemporary philosophy of mathematics. As the new century emerged, alternative approaches to the language of mathematics, distinct from strictly logical accounts, were limited to only a few rather disconnected and not necessarily converging individual works originating from both philosophy and history (cf., for instance, Rotman 1988, 2000; Netz 1999; Vinciguerra 1999; Herreman 2000; Serfati 2005; Kvasz 2008; Macbeth 2005; Chemla 2004). However, as a result of the "practical turn" in the philosophy of mathematics over the past decades, numerous studies have been pursued to explore various facets of mathematical knowledge as products of human practices rather than solely arising from abstract logical properties. Accordingly, a significant portion of research within the field involves the analysis of various mathematical corpora. These corpora encompass a wide range of materials, including actual mathematical texts, experimental data, historical and contemporary works, whether authored individually, institutionally, or collectively, and comprising both published and unpublished sources. The problem of mathematical signs – inscriptions, symbols, marks, diagrams, expressive means, representations, etc. – has thus regained prominence in the field. Drawing on insights from classical semiological theories, important aspects of mathematical knowledge have been shown to be intricately related to regularities and emergent patterns identifiable at the level of signs. From a philosophical viewpoint, these

works have contributed to bridging the traditional gap that separated mathematical from natural language by demonstrating the many ways in which mathematical signs, classically assumed to be purely formal, are conditioned by empirical and historically determined human practices. This shift in perspective has further facilitated the avoidance of anachronistic and ethnocentric biases in the philosophical and historical study of mathematical knowledge.

The philosophy and history of mathematical practices have thus brought the study of mathematical language and signs to the forefront of contemporary mathematical thought. However, despite the fruitfulness of this research trend, a comprehensive and unified account of its various aspects and the diverse approaches taken to explore it remains elusive. Recognizing this gap, we have undertaken the task of editing the present Handbook section as a much-needed remedy.

As emphasized earlier, our primary aim was to prioritize the richness of diverse perspectives rather than impose artificial uniformity and systematicity in the practice-driven study of mathematical language. To achieve this goal, we sought to include a wide array of viewpoints and topics in our selection of contributions. Despite the varied nature of approaches, the resulting series of chapters exhibit a cohesive structure. The section commences with papers that tackle the problem of mathematical signs from a broad and overarching perspective, both philosophically and historically. These pieces set the stage for the exploration of case studies that vividly demonstrate the fruitfulness of a semiological approach in analyzing various historical and conceptual dimensions of mathematical practices. Finally, the section concludes by offering fresh perspectives that confront new challenges, paving the way for further developments that can build upon and expand the rich insights showcased in this section.

The first chapter of this section is **David Waszek**'s *Signs as a Theme in the Philosophy of Mathematical Practice*. This contribution is conceived as a detailed introduction to the whole section. Waszek provides a comprehensive *survey* of the diverse works that delve into the problem of mathematical signs within the tradition of the history and philosophy of mathematical practice. His insightful analysis explores the reasons behind this apparent convergence in the field toward the study of signs, revealing that it conceals a diversity – and sometimes even a disparity – of perspectives and motivations, each with its own agenda.

Waszek identifies three main trends within this landscape. The first trend aims to reveal the norms of informal mathematical practices, suggesting that mathematics is characterized by specific manipulations of signs rather than solely relying on classic logical inference. The second trend focuses on exploring the boundaries of rigorous norms, such as ambiguity and polysemy, to underscore the openness of mathematical practices to sociohistorical conditions. Lastly, the third orientation investigates the cognitive aspects associated with the manipulation of signs by individuals, raising questions about the cognitive conditions that shape mathematical knowledge and the design choices that can enhance or hinder the efficiency of mathematical practices.

Notably, the three trends identified by Waszek align, at least to some degree, with the three external layers of semiotic systems presented earlier: abstract norms, sociohistorical usages, and individual practices. One can hardly conceive of a better

account of how the current research on the semiology of mathematical practices is distributed across the landscape of a theory of signs, encompassing multiple trajectories.

After Waszek's comprehensive survey, **Roy Wagner** proposes a theoretical introduction to the topic in the chapter titled *Structural Semiotics as an Ontology of Mathematics*. Wagner's specific approach revolves around the *structuralist and poststructuralist theories of signs*, building upon the foundational works of Saussure, Lévi-Strauss, and Derrida. His analysis conceptually reconstructs the historical evolution of this major current of semiological thought, skillfully extracting the elements and principles relevant to the study of mathematical practices. In particular, Wagner centers his account around the *commutation test*, which allows to identify semiotic units through correspondences between signifier and signified substitutions. Wagner's reconstruction is both precise and accessible, making the chapter an exceptional resource for readers seeking an initial foray into the semiology of mathematics.

However, Wagner's contribution goes beyond the sole presentation of this theoretical framework; he also puts the corresponding conceptual tools to the test through a historically informed mathematical example – the history of the numeral 0. This example not only showcases the strengths of the approach but also highlights its limitations. Through this elaboration, Wagner advances a crucial thesis: the semiological perspective in question can offer more than just methodological tools or epistemological insights to the history and philosophy of mathematics; *it can offer an ontology*. In other words, the analysis of mathematical signs involves a way of answering the question of what mathematics *is* or what it is *about*.

Crucially, Wagner avoids falling into a naïve position holding that everything is a sign. Instead, he crafts a formula that is rich with subtleties: everything is *also* a sign. This implies that everything, including everything mathematics is supposed to be about, is at any time subject to the displacement operations necessary to determine it as a sign through the commutation test. Departing from classic ontologies of mathematics centered around fixed referents and concepts, Wagner concludes, supported by various examples, that this perspective offers an insightful image of mathematics as a series of signs determinable only through the regularities found amidst multiple changes of context. Consequently, mathematical knowledge becomes inherently reliant on practices in a central and inescapable manner.

The next chapter reminds us that the role of history in fostering a renewed awareness of the significance of signs for the study of mathematical practices cannot be overstated. Notably, Reviel Netz's *The Shaping of Deduction in Greek Mathematics* (1999) stands as a landmark and an exemplar in this regard. The reader will certainly recall Netz's celebrated "two lanes" thesis, positing that the two components of deduction in Greek mathematics – necessity and generality – were shaped through a complex articulation between two semiotic devices: the lettered diagram, and the formulaic language. While the semiotic aspects of these "cognitive tools" were expounded with remarkable detail and erudition, the methodological and epistemological underpinnings of the analysis ended up appearing scarce in the face of the increasing fruitfulness of the approach.

The significant impact of this historical study led to a growing need for further theoretical and methodological elaboration guiding the replication of the same historiographical gesture in other contexts. In the third chapter of this section, **Reviel Netz** willingly embraced the challenge of revisiting his celebrated work after almost a quarter of a century since its publication. In *Shaping, Revisited*, Netz embarks on a remarkable endeavor of retroactive methodological and epistemological reflection, directly addressing the semiological dimension of this and subsequent works. Reviewing a rich series of historical cases, he probes the attributes that can be attributed to strictly semiotic properties in each case. This reflective exercise reveals that the study of mathematical signs, from a strictly semiological perspective, is inextricably intertwined with two other dimensions: the "culturally and historically malleable" cognitive apparatus and, more crucially, a materialist approach to the history of culture. Thus, we see the name of Marx joining that of Lévi-Strauss to indicate the plausibility of a materialist-structuralist perspective in the history of mathematics, yet only as a provisional orientation contingent on the empirical conditions of particular historical explanations.

The intricate connection between textual practices and culture in the historical study of mathematical practices has been a focus of an intense research program led by Karine Chemla and her colleagues for over two decades (Chemla 2004). In the fourth chapter of this section, titled *Mathematical Practices and Written Evidence: General Reflections Based on a Historian's Experience*, **Karine Chemla** returns to this question through a theoretical lens, offering critical insights into the *nature of textuality* in the history and philosophy of mathematical practice.

Beginning with a clear distinction between texts, documents, and inscriptions, Chemla builds on a comprehensive survey of recent works on the topic, drawing from a historically grounded theoretical discussion on how these forms of textual evidence can inform our understanding of mathematical practice. She first challenges the prevailing notion that the primary function of texts is communication. While this may be true for many published texts, historical investigations reveal that communication is not an immutable property of textuality but rather a situated practice that demands proper contextualization. Yet, the universality of communication can also be questioned at a deeper level, particularly when the notion of textuality is extended beyond published sources. Historical records offer numerous instances where texts serve other purposes than communication, such as exploration, learning, or computation.

By meticulously surveying these diverse cases, in which cultural variability plays a pivotal role, Chemla convincingly demonstrates how different conceptual and philosophical perspectives on the nature of writing can shape our understanding of texts. Emphasizing the importance of looking beyond published, communicative documents and considering the materiality of texts, she reveals a wealth of mathematical practices and sheds light on the social conditions that enabled their existence.

The first four chapters of the section offer an overarching perspective on how philosophical considerations about signs and language can ground the analysis of concrete, historical mathematical practices, while simultaneously showcasing how historical analyses of practices can inform and shape philosophical perspectives. With **Lucien Vinciguerra**'s contribution, *What Happens, From a Historical Point of*

View, When We Read a Mathematical Text?, the section takes a step further by presenting a series of three contributions that exemplify this virtuous circle in action, combining conceptual elaborations with detailed case studies.

Vinciguerra's chapter is a testament to his rich and original contributions, where he extends Foucault's archaeological program to the history and philosophy of mathematics (cf., in particular, Vinciguerra 1999). Here, he provides a compelling example of how such a perspective can yield critical tools for a theory of mathematical signs, emphasizing the inseparability of philosophy and history in this context. The central focus is on the act of *reading* mathematical texts, which becomes pivotal in addressing the classic Husserlian question concerning the stability of mathematical truth within an ever-evolving history of mathematical objects and concepts.

Intriguingly, as entangled as philosophy and history may be, the nature of mathematics itself introduces a tension between both, which Vinciguerra embraces as productive without the need of taking a definitive stance. He proposes to view texts as continuous but divisible spaces, where partitionings are effectively enacted through reading practices. Specifically, Vinciguerra explores three distinct ways in which the act of reading can redistribute the lines within a text: partitioning what is true and false, partitioning what is text and what is image, and partitioning what is mathematical and what is not in a text.

Mobilizing these critical tools, the second part of the chapter presents a meticulous case study centered around the reading of equations in Descartes' *Geometry*, skillfully drawing insights from other Cartesian texts like *The World* and the *Dioptrics* to reveal hidden partitioning lines. Vinciguerra's original focus on the philosophical conditions of reading historical texts uncovers essential differences between the historicity of mathematics and that of experimental sciences.

Moving forward in history, **Ladislav Kvasz** proposes a stimulating perspective in his chapter *Symbolic Algebra as a Semiotic System*, by delving into the *dynamic effects* that sign systems can exert on mathematical practices. Building on his prior work (Kvasz 2008), which associates crucial dimensions of the evolution of mathematical knowledge with different features of mathematical language, Kvasz's contribution focuses on the invention of symbolic algebra during the sixteenth and seventeenth centuries. He contends that understanding the changes occurring during this period requires attention to two dimensions of algebraic symbols: compositionality and referentiality, which characterize pivotal components of any semiotic system. To elucidate both aspects, Kvasz draws upon the works of Frege and Wittgenstein, respectively. Both references are surprising in their own way, determining the originality of Kvasz's approach. The first one because, unlike classic analytic approaches, Frege is here summoned to study the functioning of signs, and not of logical properties. Kvasz's use of Wittgenstein's philosophy, on the other hand, is doubly innovative. At odds with the usual references to the late Wittgenstein in this area of research, Kvasz returns to the early Wittgenstein. Yet, instead of relying on the *Tractatus*'s logical apparatus, the referential dimension of mathematical language is elaborated through the pictorial theory of signs. The result is a thought-provoking chapter proposing a framework wherein analytical properties of mathematics emerge from historical developments.

Concluding the series of case studies and shifting the focus to the nineteenth century, **Anna Kiel Steensen**'s chapter titled *Reading Mathematical Texts with Structuralist Semiotics* adopts the principles and tools of structuralist semiology, building on Wagner's presentation in this section, to conduct a *close reading* of a historical mathematical text. Steensen's proposes a proper operationalization of the commutation test, which enables her to identify semiotic units by examining substitutions within mathematical expressions and their corresponding contents. The specific text under scrutiny is Dedekind's *Über die Composition der binären quadratische Formen*, initially published in 1871. With exceptional precision, Steensen's analysis of this famous piece tackles the genesis of the mathematical notion of "ideal" from the internal perspective of the text.

Through the lens of structuralist semiology, Steensen convincingly shows that the progressive specification of expressions within the text's narrative structure engenders a series of productive ambiguities requiring the intervention of the reader. The irreducible underspecification of certain signifiers thus revealed allows her to suggest two intriguing properties of mathematical signs which could be extended beyond this particular text. First, a mathematical text can – and sometimes must – be read before its content is rigorously established. Second, in this process, important dimensions of mathematical content are determined as the effect of the interaction between strictly expressive means. Looking ahead, while Steensen's analysis remains circumscribed to a single text, her contribution can have broader implications in the use of distributional analyses for the study of mathematical text and corpora.

The last two chapters of the section redirect the attention from historical analysis to current and future challenges in mathematical sign practices. In the first of these chapters, *The Design of Mathematical Language*, **Jeremy Avigad** addresses the question of the *design choices* that arise when conceiving mathematics as a *semi-formal language*. Retrieving the theme of the previous chapter, Avigad observes that traditional formal languages used to represent mathematical content often face the issue of being underspecified in comparison to ordinary mathematical language. While the control and precision offered by formal methods are arguably essential for meeting the inferential requirements of mathematics, other features that are absent in formal specifications are equally necessary or desirable when considering mathematical language as the object of collective practices. Such features encompass, among others, principles and tools for effective error checking by a community, fostering creativity, ensuring efficiency, and maintaining reliability.

In addition to drawing on examples from the mathematical literature – particularly textbooks – Avigad's exploration of this problem crucially incorporates insights from the design of computational proof assistants and their libraries. This constitutes one of the principal originalities of Avigad's contribution, as it brings contemporary mathematical practices into play to reflect on the nature and conditions of mathematical language. While his approach is not necessarily limited to contemporary mathematics, Avigad's focus in this case centers around contemporary mathematical language. The outcome is an impressive survey of various features of mathematical language that are subject to design choices, ranging from the different categories

used to organize mathematical objects to various abstraction principles, theory-building tools, and other aspects of mathematical language, such as diagrams, algorithms, and heuristics.

Avigad's *catalogue raisonné* is an insightful and useful tool for analyzing past practices and making informed decisions in current and future contexts. As computer-assisted mathematics gains prominence, Avigad's contribution provides essential guidance for effectively leveraging the benefits of formal methods while preserving the richness and flexibility of mathematical language as a dynamic and evolving tool for collective mathematical practice.

Finally, in *How to Do Maths With Words. Neural Machine Learning Applications to Mathematics and Their Philosophical Significance*, **Juan Luis Gastaldi** tackles one of the most pressing issues facing the mathematical community today – the widespread proliferation of *machine learning applications to mathematical knowledge*, stimulated by the development of Deep Neural Networks (DNNs) and Artificial Intelligence (AI) in the past decade. Gastaldi's primary goal is to demonstrate that these applications hold profound implications worthy of the attention and interest of historians and philosophers of mathematics, despite the natural skepticism that may arise toward current AI branding. Gastaldi bases his claim on two compelling reasons: the potential impact of these new technologies on mathematical practices and, more profoundly, the emergence of a new perspective on mathematical knowledge bearing a new and surprising relationship with the principles of natural language.

After providing a concise presentation of DNNs, Gastaldi conducts a meticulous survey of the work in this booming research area, unveiling its main identifiable orientations. Significantly, those trends tend to be spontaneously organized according to the AI researchers' implicit assumptions as to what particular practice characterizes the production of mathematical knowledge, namely: finding proofs, manipulating objects, acquiring skills, or supporting heuristics. Through this survey, Gastaldi highlights that despite the pronounced differences among various machine learning applications, they all share a distinct and novel perspective on mathematical language, at odds with the conventional view that understands mathematical expressions as simple notations for existing contents or as an arbitrary syntax for a predefined semantics. The reason is that the successful performance of tasks requiring the manipulation of mathematical content is grounded on the processing of textual data alone. Yet, this does not bring neural models closer to a formalist approach to mathematical knowledge and language, because the language implied both in the training data and in the learning models is no other than natural, informal language. Challenging a clear-cut distinction between natural and artificial languages, natural language datasets and models are thus mobilized to address tasks involving mathematical content in a way that no existing philosophy of mathematics could have foreseen.

The implications of this new role for natural language raise numerous open questions and challenges that demand attention from a pluralistic, critical, and historically informed theory of mathematical language – precisely like the one outlined in this entire section.

Inevitably, every publication comes with its share of *regrets*. Despite our best efforts, some approaches representing significant contributions to a practice-driven theory of mathematical signs and language could not find a place in this section. Four of them deserve to be briefly mentioned here, thus inviting the reader to explore these topics further and fill in the gaps left by their absence in this work. The first one concerns the study of mathematical language and other semiotic artifacts from the perspective of *ethnomathematics*. The work of Barton (2008) offers a beautiful example of what this approach can contribute to the study of mathematical language. The stimulating work of Vandendriessche (2015) also provides original insights in this regard. Second, it barely deserves to be mentioned how much a philosophy of mathematical *notations* has to offer to a theory of mathematical signs. While Waszek's survey addresses some of this work, this section would have greatly benefited from a chapter dedicated to this important subject matter. We refer the interested reader to the references in Waszek's chapter. Thirdly, the intersection of *digital humanities*, *corpus linguistics*, and mathematical practices is an area of growing importance, where the automatic processing of mathematical texts plays a fundamental role. Insights from this research orientation can be particularly insightful to articulate many dimensions of semiotic practices in mathematics. Fortunately, this Handbook features a chapter by Tanswell and Inglis providing an example of this interesting line of research. Finally, it is worth mentioning the combined efforts of the research communities gathered around the Conference on Intelligent Computer Mathematics (CICM). The results exhibited in the processing of many mathematical tasks assessed by this community – such as informal-formal translation, information retrieval, classification, formula parsing, OCR, digital edition, typesetting systems, automatic proof generation, etc. – constitute valuable resources for the study of mathematical signs within the tradition of the history and philosophy of mathematical practice. We refer the reader to the series of proceedings of the Conference (CICM 2008–2022).

4 Conclusion

The culmination of our endeavor is a comprehensive work comprising over 250 pages, skillfully crafted by confirmed and emerging specialists in the field. We are confident that the collective efforts invested in this work will significantly enhance our understanding of the intricate connections between mathematical knowledge and semiotic practices. By providing a diverse, consistent, and historically informed account of this important dimension of the history and philosophy of mathematical practices, we aspire to contribute to unraveling the distinctiveness of mathematics as a human practice. Ultimately, the aim of this work is to forge innovative pathways that bridge the realms of formal and natural sciences with the rich domain of the social sciences and the humanities, fostering a harmonious integration of knowledge across diverse disciplines.

5 Cross-References

- ▶ Avigad, J., The Design of Mathematical Language
- ▶ Chemla, K., Mathematical Practices and Written Evidence: General Reflections Based on a Historian's Experience
- ▶ Gastaldi, J. L., How to Do Maths with Words: Neural Machine Learning Applications to Mathematics and Their Philosophical Significance
- ▶ Kiel Steensen, A., Reading Mathematical Texts with Structuralist Semiotics
- ▶ Kvasz, L., Symbolic Algebra as a Semiotic System
- ▶ Netz, R., Shaping, Revisited
- ▶ Vinciguerra, L., What Happens, From a Historical Point of View, When We Read a Mathematical Text?
- ▶ Wagner, R., Structural Semiotics as an Ontology of Mathematics
- ▶ Waszek, D., Signs as a Theme in the Philosophy of Mathematical Practice

Acknowledgments I want to thank Jeremy Avigad, Karine Chemla, Ladislav Kvasz, Reviel Netz, Anna Kiel Steensen, Lucien Vinciguerra, Roy Wagner and David Waszek for their generosity, patience, and invaluable contribution. I am deeply thankful to all the anonymous reviewers for their careful work, dedication, and constructive advice. Last but not least, I wish to express my gratitude to Bharath Sriraman for his constant support, comprehension and trust.

Funding Information This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 839730

References

Barthes R (1988) The semiotic challenge. Basil Blackwell, Oxford
Barton B (2008) The language of mathematics: telling mathematical tales. Springer US, Boston
Chemla K (2004) History of science, history of text. Kluwer Academic, Boston
Chomsky N (1955) Logical syntax and semantics: their linguistic relevance. Language 31(1):36–45
CICM (2008–2022) Intelligent computer mathematics. Springer International Publishing, Cham. https://link.springer.com/conference/mkm
Herreman A (2000) La topologie et ses signes: éléments pour une histoire sémiotique des mathématiques. L'Harmattan, Paris
Hjelmslev L (1975) Résumé of a theory of language. Travaux du Cercle linguistique de Copenhague, Nordisk Sprog-og Kulturforlag, Copenhagen
Hjelmslev L, Uldall HJ (1936) Synopsis of an outline of glossematics, Levin og Munksgaard, London
Kvasz L (2008) Patterns of change. Birkhäuser, Basel
Locke J (2013) An essay concerning human understanding. The Clarendon Edition of the Works of John Locke. Oxford University Press
Macbeth D (2005) Frege's logic. Harvard University Press, Cambridge, MA
Netz R (1999) The shaping of deduction in Greek mathematics. Cambridge University Press, Cambridge
Robering K, Posner R, Sebeok TA (eds) (1996) Semiotik, no. Bd. 13 in Handbücher zur Sprach- und Kommunikationswissenschaft, W. de Gruyter, Berlin
Rotman B (1988) Toward a semiotics of mathematics. Semiotica 72(1–2):1–36

Rotman B (2000) Mathematics as sign: writing, imagining, counting. Stanford University Press, Palo Alto
Russell B (2010) Principles of mathematics, Routledge classics. Routledge, London
Saussure F (1959) Course in general linguistics. McGraw-Hill, New York
Sebeok TA (2001) Signs: an introduction to semiotics, Toronto studies in semiotics and communication, 2nd edn. University of Toronto Press, Toronto
Serfati M (2005) La révolution symbolique: La constitution de l'écriture symbolique mathématique. Éditions Pétra, Paris
Sriraman B (ed) (2020) Handbook of the history and philosophy of mathematical practice. Springer International Publishing, Cham
Vandendriessche E (2015) String figures as mathematics?: An anthropological approach to string figure-making in oral tradition societies. Springer, Cham
Vinciguerra L (1999) Langage, visibilité, différence: histoire du discours mathématique de l'âge classique au XIXe siècle. J. Vrin, Paris

Signs as a Theme in the Philosophy of Mathematical Practice

David Waszek

Contents

1 Introduction 2972
2 The Ideal of Signless Mathematics 2973
3 Proof Beyond Logic: The Norms Governing Nonverbal Signs 2976
3.1 A Historical Example: Euclid's Geometry 2976
3.2 Norms of Nonverbal Signs in Contemporary Mathematics 2979
3.3 Philosophical Ramifications 2980
3.4 The Problem of Reliability and Diagrammatic Formal Systems 2982
4 The Open-Endedness of Signs and the Historical Development of Mathematics 2985
5 Mathematics for Finite Agents: From Psychology to Understanding 2987
5.1 Mathematical Cognition 2988
5.2 The Broader Issue of Computational Limitations 2990
5.3 The Philosophical Stakes 2992
6 Conclusion 2995
7 Cross-References 2996
References 2996

Abstract

Why study notations, diagrams, or more broadly the variety of nonverbal "representations" or "signs" that are used in mathematical practice? This chapter maps out recent work on the topic by distinguishing three main philosophical motivations for doing so. First, some work (like that on diagrammatic reasoning) studies signs to recover norms of informal or historical mathematical practices that would get lost if the particular signs that these practices rely on were translated away; work in this vein has the potential to broaden our view of the specific normativity of mathematics beyond logical proof. Second, some work focuses on signs to highlight the open-ended and "nonrigorous" aspects of mathematical practice, and the role of these aspects in the historical development of mathematics. The third motivation is based on the following observation: the reason differences in

D. Waszek (✉)
ITEM, Ecole Normale Supérieure, Paris, France

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_66

signs matter in mathematics is that humans are finite agents, with cognitive and computational limitations. Accordingly, studying signs can be a way to study how human computational constraints shape the mathematics that we do, and to tackle the topic of mathematical understanding.

Keywords

Representations · Signs · Notations · Diagrams · Mathematical proof · Heuristics · Computational limitations

1 Introduction

That mathematical texts do not solely consist of arguments conducted in natural language is obvious: Geometrical diagrams and symbolic formulas routinely serve as emblems for the discipline. Yet, in a sense, much of twentieth-century philosophy treated this fact as *inessential*, assuming that mathematics could be recast in some uniform mold – say, first-order logic, itself often seen as little more than a regimentation of natural language – and that nothing philosophically crucial about the nature of mathematical knowledge would be lost in the process. In contrast, recent years have seen a surge of philosophical interest in the various "signs" or "representations" (more on terminology below) actually employed by mathematicians. My goal in this chapter is to explain why: Philosophically, what is there to gain by focusing on mathematical signs?

First, more needs to be said about the mathematical and philosophical context that made the variety of signs actually used in mathematics seem unimportant. One aspect of this context, namely the rejection of diagrams in "rigorous" proofs, has been much discussed – and debated – over the past thirty years. However, the attitude toward diagrams is but one part of a broader story: As we shall see, there has been a concomitant tendency to also treat symbolic notations as a superficial matter. Sect. 2 offers a brief and tentative sketch of this background.

The chapter then presents recent work that takes mathematical signs seriously by distinguishing three broad reasons, or philosophical agendas, for doing so. Some work looks to the signs (and in particular to diagrams) to recover norms of informal or historical mathematical practices; attempts to draw philosophical conclusions from such efforts often look to the (broadly speaking Peircean) thesis that what makes mathematical knowledge special is not logical inference in the traditional sense, but the reliance on the construction and manipulation of signs (Sect. 3). Some work studies signs as a way to highlight the heuristic, or more generally the "nonrigorous," aspects of mathematical practice, emphasizing how the ambiguity and openness to reinterpretations of mathematical signs is crucial to the historical development of the discipline (Sect. 4). Finally, the work surveyed in Sect. 5 focuses on what may, at first, appear to be psychological or pragmatic features of mathematical signs: the way they leverage human cognitive abilities, the way they make certain tasks more or less efficient. One reason why such a study matters

philosophically, as we shall see, is that questions of computational efficiency – and hence, the signs we use – shape our mathematics quite deeply, constraining which mathematical objects end up being salient for us.

With respect to the existing literature and to other papers in this volume, my focus here is distinctive in two ways. On the one hand, I restrict myself to external, *nonverbal* signs (like diagrams and symbolic notations), excluding mere written text. While this choice reflects recent trends, and makes sense as a reaction to the philosophical context alluded to above, one should keep in mind that there is no clear dividing line between verbal and nonverbal signs – between, say, introducing new terminology and introducing a new piece of symbolic notation. The language of mathematical texts, even when it does not rely on specific symbols and looks just like, say, English, ancient Greek, or Chinese, can in fact be quite specialized, constrained, and in a sense "artificial," as emphasized by historians (see Netz 1999; Chemla 2006); indeed, the close parallel between questions of terminology or concepts on the one hand, and questions of notation or representation on the other, will come up repeatedly below. Accordingly, some authors choose a broader scope from the outset, discussing "mathematical language" (Avigad, ▶ "The Design of Mathematical Language") or the "semiotics of mathematics" (Wagner, ▶ "Structural Semiotics as an Ontology of Mathematics") without separating the seemingly verbal from the seemingly nonverbal (for a short general introduction, see also Colyvan 2012, chap. 8). On the other hand, as mentioned already, I focus on nonverbal signs *in general*, including symbolic notations; many surveys are more restricted, covering only visualization (Mancosu 2005; Giaquinto 2020) or diagrams (Shin et al. 2013; Giardino 2017) (Giardino, "Diagrammatic Proofs in Mathematics").

Before getting started, a terminological remark is in order. How should one refer, generically, to diagrams, graphs, formulas, matrices, and so on – in short, to the many nonverbal inscriptions used in mathematical practice? An obvious choice is to call them *representations*, which is in line with well-established usage in philosophy of science.[1] It is not without problems, however; among other things, the term already has technical uses in mathematics. In keeping with the title of this section of the *Handbook*, on the *semiology* of mathematics, I decided to call them *signs* instead, though no substantive position should be taken to hinge on this.

2 The Ideal of Signless Mathematics

It would be hopeless to survey the entirety of twentieth-century thinking about mathematics here, and preposterous to claim that all of it was marked by a dismissal of the role of signs: there certainly have been major countervailing tendencies (some varieties of formalism, Wittgenstein's late philosophy of mathematics, and much of "post-structuralist" continental philosophy come to mind). Nevertheless, it is a crucial background for the rest of this chapter that some of the main currents of

[1] See, e.g., Frigg and Nguyen (2020).

thought of the period *have* made the issue of mathematical signs decidedly marginal. To bring home this point, I highlight two suggestive elements of context, one philosophical and one mathematical.

In the philosophy of mathematics, it is the conjunction of two widely held views – inherited from the intense work on the foundations of mathematics in the first half of the twentieth century – that presents the most significant obstacle to the study of mathematical signs. The first view is that mathematical proofs, and hence mathematical knowledge, can be adequately captured as derivations in some formal system. The general notion of a formal system, of course, is very broad: It includes systems that encode the legal moves in chess, rewriting systems that encode algebraic computations, the lambda calculus and its derivatives, and indeed the systems that encode the construction and manipulation of diagrams that we'll meet in Sect. 3. But this first view is often conjoined with a second one, namely that the formal systems appropriate for mathematics are *logical*, roughly in the sense that their derivation steps regiment, model, or capture logical inferences. The prototypical examples of such systems are the various formulations of first-order logic (in which set theory is standardly phrased), typical inference rules of which are *modus ponens* and universal generalization. Such rules are meant to be general rules of reasoning, which could be formulated in words just as well as in symbols.

In combination, these two theses yield a view of mathematics in which, *in principle*, any proof could be replaced by a formal derivation in some logical system, and this formal derivation in turn by a piece of reasoning conducted in natural language. It is often conceded that such a reduction is not practical, but the impression remains that, as far as the nature of mathematical knowledge goes, little is lost *philosophically* by focusing on natural language arguments and idealizing all the rest away. The result is not just to exclude diagrams and other ostensibly nonlinguistic signs from serious consideration, but also to downplay the distinctiveness of symbolic calculi and computations by assimilating them to verbal reasoning.

While some philosophers of mathematical practice might be inclined to reject both theses at once, notice – anticipating the next section – that the first thesis, formalizability, does not on its own preclude a substantial investigation of the variety of mathematical signs. Nothing in the concept of a formal system, at least as understood in mathematical logic or computer science, immediately prevents its application to ostensibly nonlinguistic signs like diagrams, or more broadly to sign manipulations of any kind that do not obviously track verbal proof steps. In particular, we shall encounter a range of efforts at formalizing diagrammatic reasoning in Sect. 3; tellingly, many have come from researchers trained in mathematical logic or computer science rather than from philosophers.

The second element of context are attitudes, prevalent in parts of pure mathematics in the twentieth century, according to which the discipline should become "conceptual" in some sense, working from intrinsic definitions of its objects rather than relying on notations and computations. As a signpost to broader developments, let me focus on a few suggestive remarks by members of the influential Bourbaki group in the middle of the century.

At the heart of Bourbaki's outlook (as presented in their manifesto *The Architecture of Mathematics*[2]) is what they call the "axiomatic method," which invites identifying and separating out, axiomatically, the elementary structures that are normally mixed together in any given mathematical theory or problem. Their ideal is that, once the right structures have been isolated and the crucial ideas properly encoded in the theory, the proofs should flow easily, in what may be called a purely "conceptual" way – that is, without recourse to opaque calculations or geometric intuitions not adequately captured by the axioms. As they famously phrased it (taking up a slogan from Dirichlet that is typically associated with a purported nineteenth-century German school of "conceptual mathematics"[3]), an overarching goal of theirs was to "substitute *ideas* for *calculations*";[4] in the retrospective words of Chevalley, one of the founding members of the group, "anything that was purely the result of a calculation was not considered by us to be a good proof."[5]

This is where the connection with signs arises: From such a perspective, the real mathematical action ought to take place at the conceptual level, so that, once the appropriate structures have been worked out, the particular signs used ought to be inessential. An offhand remark by Weil, another founding member of Bourbaki, is particularly telling in this respect. In a famous (and famously nasty) polemical reply to Unguru – a historian of mathematics who had insisted that, to obtain a proper understanding of ancient Greek mathematics, it was crucial to refrain from transcribing ancient results into modern algebraic symbolism and, instead, to respect ancient phraseology and notations – Weil wrote, as though it were self-evident, that "the content of a theorem does not change greatly, whether it is expressed in words or in formulas; the choice, as we all know, is mostly a matter of taste and of style."[6] In Weil's view, what marks out real mathematicians is precisely their competence and ability to *see beyond the signs*.

Again, the two trends just discussed – in academic philosophy, a focus on logical formalization; among some mathematicians, a preference for "conceptual" methods – should not be taken to represent the entirety of thinking about mathematics in the twentieth century. Nevertheless, they played a major role in making the diversity of signs actually used in the practice of mathematics seem peripheral. Explaining how and why the tide has now turned, at least in some parts of the philosophy of mathematics, is the goal of what follows.

[2] Bourbaki (1950) = Bourbaki (1948).

[3] For a first introduction, see Ferreirós (2007, 24–38); but see Haffner (2014, chap. 1) for a good overview of the difficulties involved in identifying such a school, or in tracing the origins and meaning of this slogan.

[4] Bourbaki (1950, 231) = Bourbaki (1948, 47); emphasis in the original French.

[5] Guedj (1985, 22).

[6] Weil (1978, 92).

3 Proof Beyond Logic: The Norms Governing Nonverbal Signs

Mathematical practice is a peculiarly constrained activity: There seem to be rigid norms – more rigid, at least, than in most other human pursuits – governing which statements and justifications practitioners take to be correct. At first sight, formalizations in logical systems offer an explanation for this rigidity: A mathematical proof is correct, one might say, if and only if it can be routinely formalized in a suitable logical system. However, such formalizations are far removed from mathematics as it is actually found in the classroom or in research: Proofs there might contain diagrams, say, or algebraic calculations; even when they look like a purely logical argument, they are not usually couched in a formal language like that of first-order logic, nor are they formally correct – and formalization can prove quite difficult. And yet it is clear that generations of pupils have found their mathematics teachers more exacting and rigid in their expectations than any others. Prima facie, it is not clear how formalizations that are neither used, nor even accessible in practice, can nevertheless account for the manifest rigidity of practitioners' judgments.

Accordingly, one motivation for looking at the particular signs used by mathematicians is to describe the *surface* norms, so to speak, that govern their practices. One possible goal can be to provide a *bridge* to underlying formalizations: If we take formalizability as the criterion for correctness of mathematical proofs, a full account of mathematical practice requires explaining how *informal* proofs are actually judged to be correct, and how such informal correctness judgments manage (at least in many cases) to track formalizability. We shall discuss efforts of this kind with respect to both a much-studied historical example, namely elementary plane geometry in the style of Euclid, and more contemporary uses of diagrams.

A more radical goal, as we shall see, can be to account *directly* for the rigidity and distinctiveness of mathematics, bypassing formalization accounts altogether: Work on diagrams (and on nonverbal signs in general) has the potential to broaden our vision of proof beyond logical inferences as standardly construed, or at least – if one prefers to keep the usual meaning of "proof" fixed – to broaden our view of the distinctive normativity of mathematics beyond (logical) proof. One can end up arguing that what makes mathematics unique is not its link to logical deductions, but the fact that it is based on strictly normed sign manipulations.

3.1 A Historical Example: Euclid's Geometry

First, consider the historical example – much discussed in the recent literature – of elementary plane geometry, with its *diagram-based* proofs that the angles at the base of an isosceles triangle are equal, that the angles in any triangle sum to two right angles, and the like. Under various guises, such proofs have been a mainstay of mathematics in the West at least since Euclid's *Elements*, often serving until recently as *the* exemplar of mathematics. Paradoxically, though, their use of diagrams makes

them difficult to accommodate for contemporary philosophical views, a discrepancy that has of late spawned a fairly large literature.[7]

Let us first clarify the problem. Roughly speaking, from a contemporary point of view, elementary geometry studies what can be deduced logically from fundamental, verbally formulated assumptions about space ("axioms") – where, as a rough approximation, one can think of logical deduction as corresponding to the use of a handful of simple rules like *modus ponens* and universal generalization. But, though Euclid's text does appear axiomatic – with "postulates" and "common notions" that, at first sight, play the role of axioms, and an apparent commitment to prove everything on the basis of these – his proofs come with diagrams and *crucially rely on them*, in the straightforward sense that if the diagrams are removed, some claims in the proofs no longer appear supported by preceding ones.[8] From the vantage point of contemporary mathematics, we can only see this as a flaw, since it is not clear whether or how those steps that rely on the diagram could be deduced logically from the hypotheses, postulates, or common notions. Hence, insofar as we accept Euclid's theorems or proofs as correct at all, it can only be because *we* are able to *patch them up* according to our standards, eliminating any use of diagrams in the process – or so it seems to the modern observer. While one could try to argue that our patched-up proof is, in some sense, "essentially" the same as the ancient one, it seems clear that such "corrections" import later tools and standards into the earlier proofs instead of assessing them on their own terms.

This perspective on Euclid, based on modern views of proof, raises two crucial problems. First, the use of diagrams obviously does not mean that "anything goes." In ancient Greece – just as today – geometry was seen as a distinctively rigorous enterprise, with uniquely little room for disagreement among practitioners: It is clear that ancient geometers had their own norms as to what counted as a correct proof, which we cannot hope to describe by invoking standards inaccessible to them. Second, even admitting that we need to patch up Euclid's proofs before we can accept his theorems as correct, how is it that we are able to do this for pretty much *all* of Euclid, without stumbling upon clear mistakes? We need to explain this fact, unless we are prepared to conclude that ancient geometers were incredibly lucky, or that later mathematicians were so committed to making the *Elements* come out as true that they found creative ways of reinterpreting and correcting them so as to *make*

[7] Of course, there is no unitary practice spanning the two millennia and counting from Euclid to today's classrooms; for the sake of precision, the literature has mostly focused on the first books of the *Elements* themselves, usually without attending too closely to the details of the text.

[8] See Netz (1999, esp. chap. 5, 175–182) for a careful investigation of the various ways diagrams are used and relied upon in Euclid's proofs. In some cases, the diagram only serves as a kind of shortcut, in the sense that unproblematic chains of logical inferences could easily be added to the text to "fill in" the apparent justificatory blank left by the removal of the diagram; in others, such a "filling in" would require adding postulates; in others still (proposition 35 of book I is a simple example), Euclid's proposition seems general, as does the beginning of its proof, but at a certain point the proof turns out to be about a particular case displayed by the diagram (i.e., about particular relative positions of the relevant points and lines), so that, absent the diagram, some statements in the proof no longer follow from preceding ones at all – the proof would have to be reorganized.

them correct (but see Vinciguerra, ▶ "What Happens, from a Historical Point of View, when we read a Mathematical Text?").

These two problems motivate much recent work, the goals of which are, accordingly, the following: first, to describe the norms that governed Euclidean geometric practice as the practitioners themselves understood them; and second, to account for the fact that these norms, old-fashioned though they may be, largely produce conclusions that are *correct by contemporary standards*. The most philosophically sophisticated discussions along these lines are those by Manders (2008a, b) and by Panza (2012) (see also Netz 1999 for a philological perspective, and Ferreirós 2016, chap. 5 for a helpful introduction).

On the picture of Euclid that emerges, diagrams are not – as the modern perspective would have it – mere illustrations, sometimes used spuriously in the course of proofs, of a subject matter that could be specified axiomatically independently of them. Instead, they are a full-fledged part of the practice and are subject to precise norms, much like the pieces of reasoning conducted in Euclid's text.

The norms governing the diagrams are, in the first instance, codified in Euclid's postulates, which turn out not to be axioms in the modern sense: instead, they serve to codify how diagrams can be *constructed* and *enlarged* – if one insists on an analogy with modern logic, they are more akin to *formation* and (to a certain extent) *transformation* rules than to axioms. Once the diagrams are constructed, one can draw conclusions based on their appearance. While codified nowhere explicitly, this, too, turns out to be done only in very constrained and controlled ways. Equalities of lengths or angles, for example, are not (and indeed could not) be deduced from the appearance of diagrams. One *can*, however, rely on other features of diagrams, for instance on the fact that two circles that are partly inside each other have a point of intersection (like in Euclid's infamous first proposition) – features that turn out to be such that they can be reliably reproduced by different practitioners, working on different and imperfect diagrams. Other aspects of the practice also seem designed to secure agreement among practitioners: The division of the text into propositions whose proofs are kept relatively short by a systematic reliance on previously proved propositions, for instance, ensures that diagrams remain simple and uncluttered enough to be reliably reproduced, checked, and inspected.

In sum, a close inspection of Euclid's practice shows that there are clear and substantial norms governing how diagrams can be constructed, transformed, and exploited, and that these norms seem designed to ensure – as far as possible – univocal agreement among experienced practitioners.

A question then remains: Since Euclid's diagram manipulations are not justified by reference to other principles – in the sense that the diagrams are not illustrations of a subject matter that Euclid had means to define independently of them, and that, therefore, the norms governing their use could not be reduced to properties of this independent subject matter – what is Euclid's geometry *about*? Historians emphasize that it is a theory of figures (line segments, triangles, squares, circles, etc.) and *not*, as today's geometry, a theory of space.[9] But then, what exactly are these figures that

[9] On the transformation of geometry from a theory of figures to a theory of space, see, for instance, De Risi (2015).

cannot be defined independently of the diagrams yet are not reducible to them? Netz (1999, Sect. 2.3.2) explains the problem well: On the one hand, Euclid's geometry – full of expressions like "let a circle have been drawn," "let the given bounded straight line be AB" – appears to be *about the diagrams themselves*, or at least to identify the objects it talks about with the concrete lettered diagrams that are referred to within the text. On the other hand, this cannot be quite right, as the diagrams clearly do not exactly have, and are not meant to exactly have, the properties ascribed by the text to its objects.

Panza (2012) (▶ "Platonism, *De Re*, and (Philosophy of) Mathematical Practice") articulates and defends the most straightforward answer: Euclid's geometry is about abstract objects that are *individuated*, and some properties of which are *determined*, by the concrete diagrams. (This makes these objects *quasiconcrete* in the sense of Parsons, like the "expression types" familiar from the type-token distinction for languages: They "have an intrinsic relation to the concrete," in the sense that they "are determined by their concrete embodiments."[10]) In other words, Euclid's constrained practice with concrete diagrams gives rise to a *sui generis* domain of abstract objects.

Instead, Manders (2008b) suggests focusing on the inferential norms governing the practice and doing away with any talk of geometrical objects, arguing in particular that Euclid's use of "impossible" diagrams in proofs by contradiction – like the diagram with squished circles that is used to prove, in proposition 10 of book III, that two circles cannot cut in more than two points – puts too much pressure on accounts of diagrams as representations for them to be viable. (For an analysis of this argument, see Waszek 2022.)

3.2 Norms of Nonverbal Signs in Contemporary Mathematics

In contrast to Euclid, contemporary mathematicians typically do possess resources to define their subject matter independently of diagrams (e.g., by defining its objects axiomatically or constructing them from set theory). Nevertheless, controlled uses of diagrams do still make their way into widely accepted proofs; as we shall see, this seemingly threatens common contemporary views about mathematical rigor.

Here are some examples from the recent literature. De Toffoli, partly in collaboration with Giardino (De Toffoli and Giardino 2014, 2015a; De Toffoli 2017, 2023), explores the use of various kinds of diagrams in knot theory and more broadly in low-dimensional topology, as well as in homological algebra: knot diagrams, polygons representing surfaces, arrow diagrams, etc. Giardino and Patras (n.d.) examine the role of Feynman diagrams in a proof from a recent research paper straddling algebra and mathematical physics.

[10] Parsons (2008, 33, and the entirety of §7).

In these cases, diagrams may be eliminable *in principle*, but in practice they are crucial to the way practitioners assess proofs as correct or incorrect. They may even, at least in the Feynman diagrams case study, be in some sense *more fundamental* than any nondiagrammatic definition of the relevant subject matter: While there is no doubt that an appropriate nondiagrammatic definition of "Feynman diagrams" could be provided, it would be cumbersome and, more importantly, it seems that the standard by which its *adequacy* would be judged would precisely be the diagrams themselves. For contemporary (informal) proofs too, then, attending to the signs actually used allows one to describe *surface* norms governing the practice.

Such diagram-based practices from contemporary mathematics put pressure on what is sometimes called the "standard view" of mathematical rigor, most carefully spelled out by Hamami (2022): On this view, (1) an informal mathematical proof is *rigorous* if it can be *routinely translated* into a formal proof; (2) practitioners *judge* whether informal proofs are rigorous by recognizing what one might call "syntactical shortcuts," essentially higher-order inference rules; which (3) ensures that practitioners' rigor judgments track actual rigor, that is, *track translatability into formal proofs*. If in fact practitioners – at least sometimes – judge proofs on the basis of sign manipulations that do not straightforwardly correspond to logical inferences or sequences thereof, one will have to offer another account of whether and how informal proofs that are judged correct by practitioners manage to track formal proofs.

One might hold that correct informal proofs do not, in fact, track formal proofs at all; see for instance Rav (1999, 2007). Alternatively, one might maintain that formalizability holds, but articulate other accounts of the relationship between informal proofs and formal proofs. Burgess and De Toffoli (2022) map out some possible options. One well-known proposal comes from Azzouni (2005, 2009) (► "The Algorithmic-Device View of Informal Rigorous Mathematical Proof"). He introduces the notion of "inference packages," of which the use of diagrams in elementary geometry is the prototypical example: According to him, these function as "black boxes," allowing practitioners to extract consequences *reliably* from sets of assumptions while hiding from them the precise logical relations linking hypotheses and conclusions.[11]

3.3 Philosophical Ramifications

Let us take stock: Practices prototypically called "mathematical" sometimes – perhaps even frequently – rely not only on logical inferences in the usual sense, but also on controlled manipulations of nonverbal signs (prominently including diagrams). What should we make of this? On the one hand, historians or sociologists

[11] For an attempt to apply this idea to a contemporary example, see again Giardino and Patras (n.d.) on Feynman diagrams.

of mathematics would do well to adopt a broad notion of "proof" that can accommodate such phenomena. This is precisely the move made by Chemla (2012), who suggests that a misplaced philosophical obsession with absolute certainty and hence logical proof has truncated and distorted our perception, not only of contemporary mathematics, but also of its ancient counterparts:

> By the term 'proof' [. . .] we simply mean texts in which the ambition of accounting for the truth of an assertion or the correctness of an algorithm can be identified as one of the actors' intentions.[12]

On the other hand, philosophically, doing this leaves us with a puzzle: If any justification appearing in a mathematical context can be called a proof, what – if anything – distinguishes "mathematical" practices from other practices? If the peculiar rigidity of mathematics does not come from a specific method of demonstration, where does it come from?

One might answer that the question is misguided. Perhaps so-called mathematical practices are only special insofar as they rely on logical inference, which they normally do to a partial extent at best; as soon as one would leave behind the firm ground of logical proof, there would be nothing but a continuum of practices with no natural boundary for what should properly count as a mathematical justification. In particular, the various diagram-based practices discussed above would be nothing but more or less imperfect relatives of logical deduction.

The work just surveyed, however, suggests a different answer, one that has been most explicitly articulated by Peirce: Perhaps the peculiar character of mathematics comes not from logical inference but from controlled sign manipulations. Peirce's account has, for this reason, undergone a revival in the philosophy of mathematical practice (see in particular Carter 2014, 2020). For him, mathematics is about drawing "necessary consequences" from hypotheses, but in contrast to more standard deductivist positions, this drawing of necessary consequences does not proceed by logical inference in the usual sense, but by the construction and manipulation of signs:

> [I]n drawing those [necessary] consequences, the mathematician uses what, in geometry, is called a "construction," or in general a diagram, or visual array of characters or lines. Such a construction is formed according to a precept furnished by the hypothesis. Being formed, the construction is submitted to the scrutiny of observation, and new relations are discovered among its parts, not stated in the precept by which it was formed, and are found, by a little mental experimentation, to be such that they will always be present in such a construction. Thus the necessary reasoning of mathematics is performed by means of observation and experiment, and its necessary character is due simply to the circumstance that the subject of this observation and experiment is a diagram of our own creation, the conditions of whose being we know all about. (Peirce 1931–1958, §3.560)

[12] Chemla (2012, 18).

In fact, for Peirce, logical inference is *itself* a species of the broader genus of reasoning through the construction and observation of signs – a broader genus which Peirce variously calls *deduction* (in a broad sense) or *diagrammatic reasoning*, it being understood that his use of "diagram" includes such things as algebraic formulas and the schematic sentences of syllogistic logic.[13]

Further discussion of Peirce's views goes beyond the scope of this chapter; for a helpful introduction written from the perspective of mathematical practice, see Carter (2020).

3.4 The Problem of Reliability and Diagrammatic Formal Systems

Views of the normativity of mathematics that, like Peirce's, center on sign manipulations face an apparent roadblock in the oft-repeated (though distinctively modern) idea that diagrams are a source of fallacies. Indeed, it is often supposed that mathematics is made distinctive by its certainty and reliability – so is it not counterintuitive, from a modern point of view, to model one's *overall* account of mathematics, including logical inferences and simple algebraic computations, on something so reputedly unreliable as diagrammatic reasoning?

On the face of it, the work discussed above does show that even geometrical diagrams can be used in controlled and (at least to some extent) reliable ways. As for nongeometrical diagrams, many of those used in contemporary mathematics (like the arrow diagrams used in homological algebra or category theory) do not seem to threaten reliability at all. Might it nevertheless be the case that the use of diagrams, or at least of some of them, is inherently less safe than purely deductive reasoning?

In Euclid's geometry – to again use it as an example – one faces the so-called *generality problem*: One proves a *general* statement by constructing and inspecting a *particular* diagram, but is the particular diagram adequate to the general claim, or are there other, substantially different diagrams in which the claim might be wrong, or for which a further proof would be required? Applying Peirce's remarks, quoted above, to the case of Euclidean diagram-based proofs (which were, for him, a prototypical example), one needs to ensure "by a little mental experimentation" that the conclusions drawn from the constructed diagram are "such that they will always be present in such a construction." This, however, cannot be done according to mechanical rules; It requires tact and care, and leaves one exposed to the risk of an oversight – a risk that, in the common image of mathematics, bulletproof logical deductions are meant to avert.

There are two complementary strategies to mitigate such worries. On the one hand, one can emphasize that run-of-the-mill, nondiagrammatic proofs *too* are fallible (though often very reliable); De Toffoli, for instance, was led – after her

[13] He writes that "all deductive reasoning, even simple syllogism, involves an element of observation" and that, for instance, the "syllogistic formula, All M is P, S is M ∴ S is P" is "a diagram of the relations of S, M, and P" (Peirce 1885, 182).

work on diagrams, though her motivations are broader – to a fallibilist account of mathematical knowledge (De Toffoli 2021). On the other hand, one can bolster diagrams' claim to reliability by attempting to *formalize* (instances of) diagrammatic reasoning, which by now has been done in a number of cases, including for the simplest parts of a reputedly difficult case, namely Euclid's geometry (for surveys, see Mancosu 2005; Shin et al. 2018; Giaquinto 2020).

The first strategy ties in with broader discussions about the nature of informal proofs. Above, I wrote as if diagrammatic proofs raised a unique challenge to the "standard view" of mathematical rigor (Hamami 2022) – i.e., as if there were questionable proofs using diagrams on the one side and rigorous deductive proofs on the other, the latter basically being similar to formal proofs, just with *gaps* that correspond to the application of higher-order inference rules and hence can routinely be filled. However, much recent work has argued that informal mathematical proofs are often *not* like that, especially in research contexts. This is not to say that informal proofs are unreliable. Rather, the disagreement concerns where the reliability of informal proofs comes from. On the "standard view" as articulated by Hamami, informal proofs reliably track formal proofs because they are *sketches* of formal proofs and are *recognized as such* by practitioners; their reliability then comes directly from the unimpeachable solidity of the underlying formal proofs. In contrast, the argument made in papers, such as Avigad (2021) and De Toffoli (2021), is that the reliability of informal proofs – the fact that they *do* track formal derivations – comes from a host of features of mathematical practice that allow for effective communal error checking. On such views, informal proofs do not straightforwardly inherit their solidity from the formal derivations they track; the relationship between the two can be quite complex.[14]

From such perspectives, the fallibility or defeasibility that some diagrammatic proofs might display comes to be seen as typical, but it is compatible with high reliability: The remaining question is merely one of degree. In Manders's vocabulary, does a certain practice afford us enough *control* – by way of error-checking and dispute-resolution mechanisms – to avoid a fall into *disarray*, with recurrent errors and endless controversies? As Mancosu (2005) already concluded, this will have to be adjudicated on a case-by-case basis, whether there are diagrams or not. Euclid-like diagram use, for instance, certainly seems to offer sufficient control for elementary geometry, if not in more advanced settings; in many contemporary cases, diagrams do not seem to threaten reliability at all, and in fact might sometimes *increase* it by introducing helpful error-checking mechanisms (see Avigad 2021, 7392–7393 for such an argument).

The second strategy is to argue for the reliability of (some uses of) diagrams by developing *diagrammatic formal systems*. These come in two broad flavors, which face different challenges.

[14] See again Burgess and De Toffoli (2022) for a survey of the various ways of understanding this relationship.

Systems of the first kind have a syntax that is truly diagrammatic, in the sense that it is *not* made up of strings of symbols; instead, the formation and transformation rules instruct one to draw, erase, or transform geometric objects like lines or curves. Systems of this kind were mostly developed around Barwise and Etchemendy's diagrammatic reasoning program (see Allwein and Barwise 1996 for an overview), and include the first attempt at a diagrammatic formal system for plane geometry (Luengo 1996); the best known of them is Sun-Joo Shin's (1994) system for Venn diagrams. By explicitly codifying how diagrams can be constructed and manipulated, such systems are meant to allow – as usual formal systems do – conducting consistency or soundness proofs, as well as performing "move by move" checks of the correctness of purported proofs.

In systems of this kind, however, issues like the generality problem are still lurking. Setting them up or reasoning about their properties requires being able to survey the totality of possible diagrams that could be constructed using certain rules. But, in contrast to standard formal systems – whose inductive formation rules give us a sufficient grasp on possible formulas to reason about them – diagrams may have nonobvious spatial properties: We may have difficulty picturing the full range of possible cases, and may make mistakes by having only simple and unrepresentative examples in mind.[15] It does seem that genuinely diagrammatic systems, despite their appearance of perfect codification, raise problems of control that do not arise with standard formal systems.

Diagrammatic formal systems of the second kind – for instance, those developed by Miller (2007), Mumma (2006, 2019), and Avigad et al. (2009) for Euclid's geometry – have a more familiar syntax that is made up of strings of symbols, but is meant to *model* the diagrams and the reasoning conducted on them. Since their syntax is not itself diagrammatic, these systems are no less reliable than the standard ones found in logic textbooks. Instead, the challenge they face is one of *goodness of fit*: Do they closely track the preexisting diagrammatic practice, or are they more akin to *reconstructions* that eliminate the difficulties involved in diagram use? Consider again the case of Euclid. Rabouin (2015, note 33) points out that the extant systems for formalizing the *Elements* all end up rejecting at least a few of Euclid's proofs. The problem comes from those *reductio* proofs that require drawing apparently "impossible" diagrams (like the impossible diagram for proposition 10 of book III, with two "circles" that cut in four points): Formalization, because it requires stipulating exactly what kind of diagrams can legally be drawn, is liable to block at least some of these proofs.

A thorough assessment of diagrammatic systems lies beyond the scope of this survey. However successful such formalizations may be, and whether or not one ends up accepting a Peircean view of mathematics, one conclusion is clear.

[15] Shin's original system turned out to contain a mistake of this kind: It had a transformation rule permitting the erasure of any curve in a Venn diagram, which overlooked the fact that in complex diagrams (having more than four curves), though not in simpler ones, such a transformation can, for topological reasons, lead to a diagram that is not well formed (see Scotto di Luzio 2002; Howse et al. 2002).

Diagrammatic formal systems demonstrate that formalizing does not necessarily mean logicizing, at least not in the traditional sense: One can formalize norms of mathematical sign manipulations that do not correspond to logical inferences as usually construed. Note, though, that one can express the same conclusion in a different vocabulary: Instead of talking of sign manipulations that are not logical, one can attempt to redefine logic so as to cover diagram manipulations, as suggested in particular by Barwise, Etchemendy, and Shin on the basis of a semantic view of logical consequence (see for instance Barwise and Hammer 1994 and Shin 2004).

4 The Open-Endedness of Signs and the Historical Development of Mathematics

Let us now turn to work that, instead of looking to mathematical signs to recover *norms* of informal or historical mathematical practices, does – if anything – the very opposite, emphasizing the ambiguity and incidental features of signs, their openness to reinterpretation and change, and the way these drive the historical development of mathematics. The main thrust of such work, which often goes hand in hand with discussions of mathematical discovery and of the historical growth of mathematical knowledge, is to combat a vision of mathematics as an isolated and inward-looking realm governed by strict rules, and instead to emphasize its plasticity and open-endedness. My aim here is to give a quick feel for the relevant literature; the topic is covered in much more detail in Wagner's survey (this section).

The simplest point to be made will be familiar to everyone who has studied mathematics: It is not unusual for definitions, theorems, or proofs to be motivated by a visualization or a heuristic computation, only to be rigorously formulated or established in a different way. Such bifurcations between demonstration and heuristics – where the heuristics often rely on diagrams or other sign manipulations – go at least as far back as the work of the ancient Greek mathematician Archimedes (see, e.g., Detlefsen 2008, Sect. 2 for a historical survey), and reflect the norms in force as to what counts as an acceptable proof. In the recent literature, Carter (2010) and Starikova (2012) nicely document examples of diagram-driven discovery in contemporary mathematical research (see also Giaquinto 2020, Sect. 4 for a survey).

Perhaps most interesting, however, are cases in which the line between the demonstrative and the heuristic – between what is a legitimate sign manipulation, sanctioned by clear rules, and what is an unintended representational artifact – is not so clear-cut.

As a first example, consider the derivatives of a function f. In principle, it should not matter much whether they are written f', f'', etc. (Lagrange's notation) or $\frac{df}{dx}, \frac{d^2f}{dx^2}$, etc. (a descendant of Leibniz's notation); there may be "pragmatic" differences, of a kind we shall return to in Sect. 5 – the Lagrangian notation may be more economical, and the Leibnizian one better suited for remembering the chain rule, say – but the rules of the differential calculus themselves should remain the same whether they are couched in one notation or the other. From this point of view, the fact that differential operators $\frac{d}{dx}, \frac{d^2}{dx^2}$ appear separately from the function f in the Leibnizian case is merely

an artifact of the notation. Heuristically, though, this apparent artifact matters: It invites officially illicit, but sometimes productive, algebraic manipulations of the differential operator itself. This seems to have given rise historically to the collection of methods called the "operational calculus," in which differential operators *are* used algebraically. While the manipulations in question do not follow any preestablished norms, they depend on particular features of the signs previously in use, by creatively exploiting analogies these signs suggest or opportunities they afford.

Note that caution is required here: The idea that a mere artifact of notation is what led to the operational calculus is a caricature. Rarely if ever does it happen that an incidental feature of a notation adopted for unrelated reasons is solely responsible for a major breakthrough, however tempting the idea may be. Nevertheless, this caricature gets something right: Notations can matter in exploratory contexts. For a cautious attempt at seeing how much of a role notational changes did play in the earliest attempts, by Leibniz and Johann Bernoulli, at an algebraic treatment of differential operators, see Waszek (n.d.).

Importantly, there is already more at play in this example than a disjunction between discovery and justification; it is not merely a case where something was discovered at some point based on incidental features of previously used signs, but was then justified by a different route. Methods associated with the "operational calculus," defined by their algebraic use of differential symbols, ended up being used in a *systematic fashion* over extended periods of time, often with no more than partial and piecemeal justifications available. (On the history of the operational calculus, see Koppelman 1971 and Lützen 1979; for an engaging and mathematically sophisticated survey of various instances in which sketchy algebraic manipulations turn out to be surprisingly and systematically helpful, see Cartier 2000. Feynman diagrams are a well-known analogous case, but with diagrams rather than symbolic formulas; see, e.g., Kaiser 2005.)

The operational calculus – especially methods developed in the late nineteenth century by the electrical engineer Oliver Heaviside[16] – is one of the favorite examples of Wilson (2006, in part. chap. 8; 2021), who emphasizes that scientific and mathematical progress often comes by way of exploratory sign manipulations that, at first, cannot be justified nor fully understood. His broader point, which is pragmatic in spirit, is that it often takes a lot of scientific developments before we are able to understand why and how our successful methods work, and that insisting prematurely on perfect rigor is counterproductive.

Crucially, though, Wilson sees the operational calculus merely as an extreme and obvious instance of a phenomenon that is pervasive throughout scientific practice. He argues through numerous examples that signs (and more broadly, concepts) tend to shift in meaning in subtle and often unrecognized ways depending on context – even in core scientific work that is not considered heuristic by practitioners – and that this allows for successful and efficient methods whose underpinnings we only *think* we understand.

[16] On these, see Lützen (1979).

Work by Grosholz (2007) and Wagner (2017) on the role of ambiguity in mathematics bolsters such claims. Wagner's study of generating functions (2017, chap. 4) provides us with a particularly clear example coming from contemporary mathematical practice. Generating functions are series of the form $\sum a_n x^n$ that are used to solve combinatorial problems. What Wagner shows is that the various techniques commonly used in this field require interpreting the symbol x in different ways: Sometimes x has to be a real or complex variable, sometimes a mere placeholder defining a formal power series, and sometimes a (transcendental) element of an extension of the field to which the coefficients belong. While any particular technique – save perhaps for delicate cases at the frontier of research – can be justified satisfactorily, the interpretations and mappings between interpretations that would be required to do this are not usually spelled out, and often could not be except by sophisticated practitioners. In effect, under a veneer of rigor, the field thrives on ambiguity.

More radically, Wagner and collaborators emphasize that signs permit circulations, not just between different mathematical contexts, but also between mathematical and nonmathematical ones. (For ease of exposition, I take for granted that there is a clear boundary between the mathematical and the nonmathematical, but of course the point of such work is to call this into question.) Wagner (2010, 2017, chap. 2) focuses on Italian abbacist treatises of the sixteenth century; he argues that the way they conceptualized and handled (what we would now call) negative numbers, fractional numbers, and square roots was closely related to *economic* practices of currency conversion. In other work (2009; 2017, chap. 4), he studies a specific combinatorial problem that, in the mathematical literature, is often phrased in terms of marriages; he argues that, through such choices of terminology and notations, gender stereotypes can end up shaping how the problem is framed and what kind of questions are asked about it.

In sum, and in contrast to the image of perfect rigor that mathematics often enjoys, the work we surveyed in this section argues that the open-endedness of signs – their ability to function ambiguously, to be experimented upon beyond what is allowed in principle, to be reinterpreted, and to allow circulations from contexts apparently far removed from mathematics – plays a crucial role in the practice and historical development of the discipline.

5 Mathematics for Finite Agents: From Psychology to Understanding

Our last motivation for studying the signs actually used in mathematics is to explore how finite human beings, with their specific cognitive makeup and computational limitations, can practice mathematics and attain mathematical knowledge.[17] Indeed, while in principle, one could practice arithmetic using the symbols of set theory or do

[17] As such, it is an instance of what Carter (2019) has called the "agent-based" strand of the philosophy of mathematical practice.

geometry without any diagrams, both are effectively impracticable: For humans, the particular signs used matter considerably.

It may seem that such differences between signs are merely "psychological" or "pragmatic," in the sense that they do not impact the *justification* or *content* of mathematics: They are irrelevant – one might think – to whether a theorem is true or to what kind of objects (if any) it is about. It is true that part of the literature rests content with fitting mathematical knowledge into a reasonable account of cognition and ignores ontological questions altogether. But as we shall see, the stakes are higher than this. Some recent work clearly suggests that human computational limitations, and the representational tools that allow us to make do with these limitations, deeply shape our mathematics, including in its justifications and its ontology.

Note that similar questions have long been discussed in relation to the philosophically central topic of *infinity*. Indeed, one way of formulating Hilbert's foundational program (at least in its mature formulations, from the second half of the 1920s) is as aiming to justify all talk of the actual infinite on the grounds that it ultimately provides an *efficient* (though in principle eliminable) way of talking about the finite (see Zach 2023, Sect. 3), thus seemingly grounding our mathematical ontology on concerns of pragmatic accessibility. Later on, questions about how different notations impact practical feasibility have been raised in the context of discussions of strict finitism (see Dean 2018). Early efforts at developing an explicitly semiotic account of mathematics were also largely motivated by the problem of infinity (Rotman 1993, 2000; on this, see also Van Bendegem 1996). Despite its obvious relevance, however, the question of infinity has not had a central role in recent work on mathematical signs; moreover, it would deserve an extended discussion. This topic is therefore best left for elsewhere. (For a starting point, see Easwaran et al. 2021.)

5.1 Mathematical Cognition

Let us start from a few examples of psychologically oriented research that explores how human cognitive abilities constrain, and take advantage of, the various signs used in mathematics. Landy, Goldstone, and collaborators have shown experimentally that perceptual cues, like spacing, impact accuracy and performance in symbol manipulation tasks – e.g., that $a + b*c + d$ is easier than $a+b * c+d$ to parse (properly) as $a + (b * c) + d$ – and, more broadly, that fluency in parsing symbolic expressions relies on a kind of *perceptual training*.[18] Wege et al. (2020) show that even the most superficial aspects of mathematical signs, like symmetry, can impact performance. Experiments by Hamami et al. (2021) suggest that human reasoners typically assess the validity of elementary geometric claims through a diagram-based search for counterexamples. Finally, in a more phenomenological (rather than

[18] See Landy and Goldstone (2007); Landy et al. (2014); Marghetis et al. (2016).

experimental) vein, De Toffoli and Giardino (2014) argue that the practice of topology often relies on the use of sophisticated visuospatial skills to imaginatively manipulate diagrams of three-dimensional objects.[19]

Such work reintegrates mathematics into cognition and, at its best, has the potential to reveal unexpected ways in which mathematics taps into the strengths of the human sensorimotor apparatus. Its core message is that the practice of mathematics involves the body – the eye and the hand – in many ways. As such, its target is less traditional philosophy of mathematics than, on the one hand, a certain cultural image of mathematics as abstract and disembodied, and, on the other hand, broad theories of cognition that conceive of thinking as a computational process taking place "inside the head" – showing, instead, how much mathematics relies on the body, with its specific sensory apparatus, as well as on external signs.[20] In this context, mathematics is discussed less for its own sake than as a strategic battleground in broader debates on the embodiment of cognition – a position mathematics owes to the fact that it is typically seen as *the* paradigmatic instance of abstract thought.

However, it is clear that differences between signs are not just a matter of psychology. Think of algebraic symbolism: Details of the perceptual system may determine, in part, what kind of symbolic expressions (including spacing, the shapes of symbols, the spatial encoding of hierarchical structure, and so on) are easy for humans to process. Nevertheless, the fact that the calculus used in modular arithmetic, say, makes the solution of many divisibility problems much more accessible than run-of-the-mill algebra, which is symbolically quite similar, cannot have much to do with psychology at all.

Admittedly, the boundary between the psychological and the nonpsychological is sometimes subtle. For instance, given three sets A, B, and C such that $B \supseteq A$ and $C \cap B = \emptyset$, noticing that $A \cap C = \emptyset$ may be easier on an Euler diagram (which represents the sets A, B, and C as circles, with circle A located inside circle B and circle C disjoint from the others) than on the symbolic relations just given. Presumably, this is partly due to our visual ease at perceiving the containment and noncontainment of shapes (though even this is tricky to make precise: it might be due to human ease at processing spatial relations rather than to anything specifically visual, as Euler diagrams can be helpful even if perceived through touch). But even there, our visual abilities can only come into play because of an *antecedent*, *non*-psychological fact, namely that Euler circles systematically make set inclusion and set disjointness accessible in the form of geometric containment and noncontainment patterns, independently of the fact that we are good at visually picking up the latter – while conversely, a list of various relations such as $B \supseteq A$ and $C \cap B = \emptyset$, which could be phrased in a multitude of other ways, does not make inclusion and disjointness systematically accessible as single patterns.

[19] See also De Toffoli and Giardino (2015a, b).

[20] For general discussions of mathematical signs (diagrams as well as symbolic notations) in this spirit, see, for example, Giardino (2016); Vold and Schlimm (2020).

5.2 The Broader Issue of Computational Limitations

Wherever the boundary lies between factors that do and factors that do not depend on the peculiarities of the human sensorimotor apparatus, any mathematics practiced by *computationally limited agents* will be sensitive to the particular signs used, because these impact what is easily accessible. A study of how sign systems systematically reorganize accessibility – in theorem proving and problem solving, as well as exploration – can (and should) also be conducted independently of psychology, although both psychological and nonpsychological investigations will have to come together for a full account of mathematics as practiced by humans.

In this spirit, there are a number of recent case studies describing the advantages of particular mathematical signs in particular contexts, both historical and contemporary. To take just a few examples, De Toffoli (2017, 2023) discusses the use of commutative diagrams in homological algebra, and of arrow diagrams for representing closed connected surfaces; Feferman (2012) explores the use of some diagrams that contain ellipses (i.e., "...") to indicate that they could be continued indefinitely, including a commutative diagram used in homological algebra and a diagram often used to prove the Cantor-Bernstein theorem in set theory; Carter (2010) discusses the role an ad hoc diagram played in the discovery of particular concepts and later of a particular proof; and, in a more historical vein, Manders (1999) (an unpublished but widely shared manuscript) compares traditional geometry in Euclid's style, based on diagrams and verbal arguments, with Descartes' addition of algebraic equations.

So far in this survey, logic has mostly appeared as a foil – a provider of uniform symbolic languages that risk erasing the diversity of mathematical signs. However, logical notations too are designed for specific purposes (e.g., the perspicuous decomposition of sentences into constituents or of deductions into elementary inferences; the effective checking of the correctness of a deduction; the effective discovery of proofs; etc.). They are subject to psychological and computational constraints and can be analyzed in such terms. For example, Schlimm (2018) offers such an analysis of Frege's *Begriffsschrift* notation (or concept script); Waszek and Schlimm (2021) revisit the contrast between the notations of Boole and Frege, with an eye to their respective purposes and, relatedly, pragmatic advantages. There has also been work in this spirit on Peirce's many notations for logic, including his diagrammatic "existential graphs" (e.g., Schlimm and Waszek 2020; Bellucci and Pietarinen 2016).

The points made in the case studies just mentioned, be it in mathematics or in logic, usually boil down to the following: Using certain signs rather than others can make it more *efficient* for agents to perform certain tasks. The tasks in question are most often related to the discovery (or reconstruction) of proofs (or of solutions to problems). For instance, to borrow an example discussed by Feferman (2012) and De Toffoli (2017, 2023), using commutative diagrams rather than long lists of (symbolically formulated) commutation relations can make it considerably easier to reconstruct the proofs of many lemmas in homological algebra.

However, it is frequent to find such advantages phrased in a different idiom, namely in terms of gains in "understanding." In broad terms, one can make sense of such terminology if one explicates understanding in terms of the possession of *abilities* to perform certain tasks (to construct proofs, etc.); such an account is often endorsed at least implicitly in discussions of signs or representations (see, e.g., Feferman 2012), and is most explicitly articulated in Avigad (2008). To tighten the connection between understanding and the efficiency advantage of the signs used, however, one needs to clarify *which* tasks one has in mind and what these are good for.

Upon closer inspection, the precise tasks that certain signs make easier can be quite specific and varied. For example, they can make it easier to formulate a high-level sketch of the proof; to locate a particular step in such a sketch; to identify which steps one could possibly take next in a proof, and guess which is likely to help; etc. Intuitively, all of these are related to "understanding what one is doing" in a more precise sense than the mere ability to prove theorems efficiently. For attempts at identifying the most relevant tasks in particular examples, and at elaborating a suitable theoretical vocabulary, see Manders (1999, 2012) and Keränen (2005).

Trying to generalize, as just discussed, about the *kinds* of efficiency advantages that are most relevant to mathematical practice is one way of making case studies like the above more systematic. Another avenue of progress for going beyond the descriptive character of case studies lies in the following question: What exactly does it mean to say that, for certain purposes (solving a certain kind of problem, say, or proving a theorem), using certain signs offers advantages in efficiency? To give a clear answer to this question, one needs a *systematic way of comparing* different approaches, relying on different signs, to one and the same task. Though I cannot enter into details here, two broad strategies for making such comparisons more precise have emerged in the literature.

The first strategy to compare different signs, following Larkin and Simon (1987), is based on a computer model: problem solving is modeled by computer programs, and the signs or representations used in the problem-solving process are modeled by the data storage schemes used by these programs – in computer science terminology, by their "data structures." The idea is that different ways of storing the same data may lend themselves to different kinds of processing; in Larkin and Simon's terms, they may be "informationally equivalent" but "computationally different." Two ways of solving the same mathematical problem using different signs would then be modeled by two programs using different data structures, and their efficiency would be compared by counting the number of elementary operations they require before reaching a solution. (For a general discussion of this strategy, see Waszek 2023; for examples of recent attempts to apply this idea to mathematical signs, see Schlimm and Neth 2008 and Carter 2018.)

The second strategy boils down to treating different signs as the *syntax* of different *logical systems*, in order to quantify their efficiency for particular purposes by counting proof steps. In the case of diagrams, such efforts rely on diagrammatic formal systems as discussed in Sect. 3 (see Stapleton et al. 2017; Stapleton et al. 2018 for examples).

Both strategies rely on quantifying the efficiency of using certain signs by *counting* – elementary operations in one case, proof steps in the other – and both face a similar challenge, in that getting results for *particular* problems may seem of limited interest (and certainly has not much interested logicians), while getting *general* results, about the comparative efficiency of certain signs in a broad range of cases, is tricky. For instance, an obvious candidate for producing such general results would be computational complexity theory, but this approach requires identifying a class of similar problems of *arbitrarily large* size and yields only asymptotic results, which is of dubious relevance to practical concerns (see Dean 2019). Identifying classes of problems that are large enough to yield general results while still being relevant to the practice is difficult.

5.3 The Philosophical Stakes

Why should we care *philosophically*? Aren't the kinds of differences between signs discussed above, if not merely psychological, then merely "pragmatic" – in the sense that they are a matter of convenience or efficiency rather than substance? I believe that the very formulation of this question is wrong: For limited beings like us, considerations of efficiency end up shaping the substance. In particular, the signs we use for (at least partly) pragmatic reasons end up shaping what our mathematical objects and theorems are. Though research on this topic remains underdeveloped, two suggestive examples, one on notations and one on diagram use, will make this point clearer.

Before turning to examples, though, it may be helpful to notice that the parallel claim for *concepts* or *definitions* is easier to accept. Think of the central concepts in abstract algebra, say those of group, ring, or field. A main rationale for adopting them is surely efficiency – so much so, in fact, that the Bourbaki group famously referred to their axiomatic method, centered around isolating and studying such concepts, as the "'Taylor system' for mathematics"[21]: for instance, instead of studying related phenomena separately for numbers, permutations, geometric transformations, and so on, the theory of abstract groups allows focusing on them only once and then applying them to many cases (for a deeper analysis of this and related advantages, see Avigad 2020). But evidently, these concepts come with a change of subject matter: abstract algebra, even if applicable to number theory or geometry, is profoundly different from them – it is ostensibly about different objects, and its theorems are not reducible to their instances in other fields. Accordingly, the philosophical importance of conceptual or definitional choices is perhaps more evident than that of notational choices;[22] the Bourbaki group, one of whose main members (Weil) we have seen downplaying the role of notations, certainly staked its entire project on the importance of organizing mathematics around the right

[21] See Bourbaki (1950, 227) = Bourbaki (1948, 42).

[22] For discussion on this topic, see in particular Tappenden (2008a, b) and Lange (2017, chap. 9).

(structural) concepts. Nevertheless, the examples below suggest that choices of signs and choices of concepts or definitions lie on a continuum; both are perhaps best seen as species of the same genus. In practice, conceptual shifts often go hand in hand with changes in (nonverbal) signs, and both can end up reorganizing the subject matter. (For a further articulation of this parallel, see Waszek n.d.)

As a rough metaphor, think of people engaged in mathematical problem-solving, theorem proving, or mere exploration as travelers making their way through rough terrain. The signs or representations they use – and more broadly, their concepts and general organization of knowledge – trace paths and roads across the landscape, which privilege certain itineraries and make certain places *more accessible* to them than others. Nothing stops them from traveling cross-country once in a while (though some ravines may prove impassable without exceptional skill), it's just that taking the highway is more straightforward and efficient. But this "pragmatic" efficiency advantage ends up structuring most travelers' understanding of the land: which features of it are most salient, which places everyone visits, and which remain largely unknown. Of roughly this sort, I believe, is the impact of mathematical signs.

To make this more concrete, my first example comes from Avigad and Morris (2016), who discuss a famous theorem of number theory, first proved by Dirichlet: Any arithmetic progression whose terms do not all share a common factor contains an infinity of primes.[23] The proofs found in modern textbooks rely on the notion of *character* of a group; they involve summing over characters, the relevant sums being evaluated term by term based on the properties of the indexing characters. Dirichlet's original proof, however, does not refer to characters at all: His sums are indexed by intricate data and, while they contain what we can retrospectively recognize as the explicit algebraic expressions of the relevant characters, these are not the explicit focus of attention. In fits and starts, subsequent presentations of the proof progressively extricated what we now call characters, introducing specific *notations* for them and isolating their properties, thus making the proof much more surveyable.[24]

Dirichlet's original proof is just as valid as its successive reformulations, up to modern, character-based versions; in fact, there is an intuitive sense in which all of them are essentially the same. From this perspective, as Avigad and Morris note, it is tempting to dismiss the differences as "merely pragmatic." Yet in the historical arc they survey, these "pragmatic" differences had, as end result, a clear conceptual and even *ontological* change: The concept of character emerged, and characters became accepted as objects in their own right.[25] Notice the continuity, in this story, between

[23] For more historical detail, see also their companion paper, Avigad and Morris (2014).

[24] Among other things, introducing characters allows one to hide the complexity of the underlying explicit formulas, making the indexation of the sum easier to grasp, and making the proof more focused by spinning off the general properties of characters to separate lemmas.

[25] I should emphasize that the story is more complex than this sketch: treating (what we now call) characters as full-fledged objects, rather than just shorthands for parts of complex calculations, relies on a shift – associated with broader evolutions in nineteenth-century mathematics – from seeing them as algebraic expressions to seeing them as functions defined by their input-output behavior. Avigad and Morris (2016) deserves a full reading.

apparently simple *notational* abbreviations for characters and full-fledged conceptual developments.

My second example comes from Manders (1996), who uses the history of geometry to argue against the idea that "individuation of content, and hence *content*, is representation-neutral": in other words, he argues that the very content of our mathematical statements ultimately depends upon the representations (or signs) we use. His strategy is to contrast traditional geometry in the style of Euclid with *projective* geometry – more precisely, the form of it practiced by the nineteenth-century-mathematician Poncelet.

Euclid, Manders explains, proved two theorems that amount to the following: "when two lines through a point O meet a circle, say, one in A and B and the other in C and D, the rectangles (products) OA by OB and OC by OD are equal."[26] But Euclid did not state, and in fact did not have the means to handle, the result in this general form. Instead, he proved it separately for O inside and for O outside the circle.[27] This is because Euclid-style proofs always come with, and rely on, a diagram, and Eulidean diagrams *have to* settle such questions as whether a point is inside or outside a circle. Poncelet's projective geometry, on the other hand, does have means to state and prove all of these cases uniformly. So is there a single theorem here, or two? In Euclid-style geometry, the best we can say is that there are two somewhat analogous theorems with different proofs, one about "point inside circle" and one about "point outside circle." If instead we follow Poncelet, we can state and prove a single theorem about the undifferentiated content "point and circle," which Euclid cannot handle.

Before further analysis, two remarks will help forestall misunderstandings. First, there obviously *are* cases in which Poncelet's geometry goes beyond Euclid's. Manders's choice to focus on a case in which Poncelet and Euclid's theorems are intertranslatable is strategic: It is intended to bracket other phenomena so as to highlight the reorganization effected by the representational differences between them. Second, while Manders chose an example in which the focus is on the *diagrammatic* means of representation, similar points have also been made about the *conceptual* (or *definitional*) differences between Euclidean and projective geometry, highlighting again the continuity between representational and definitional differences. For instance, many of the unifications permitted by projective geometry turn upon accepting, on top of normal points, *points at infinity* at which parallel lines intersect – in other words, they turn upon redefining the

[26] Manders (1996, 396). In modern terms, these theorems are about what is sometimes called the "power of a point to a circle."

[27] The case with O inside is proposition 35 of book III of Euclid's *Elements*. The case with O outside is proposition 36, but with a further complication: the theorem still holds if one or both of the lines is *tangent* to the circle (e.g., if the first line is tangent, it will have a single intersection point A with the circle, and the theorem holds with the square on OA instead of the rectangle OA by OB), which means that there should be three cases to consider, depending on whether two, one, or none of the lines are tangents. Euclid proves only the case with one tangency, presumably because the others immediately follow.

concept of point. (The value of this redefinition is that any two lines, including parallel ones, will then intersect somewhere, which allows many theorems mentioning intersecting lines to also cover the case where the lines are parallel.) In practice, the concepts and the signs (or means of representation) are intertwined: The change in the concept of point is tied to a new way of seeing geometrical diagrams as invariant under projective transformations, some of which can send a run-of-the-mill point to infinity.

Returning to Manders's point, why does it matter if there is a single theorem (as in Poncelet) rather than two (as in Euclid)? Is it not the same knowledge, just packaged differently? Above all, the reorganization Manders analyzes matters because it alters what is *salient* for mathematicians in contexts of research and problem-solving. If using diagrams like Poncelet, one will, for instance, be attuned to any situation involving a point and a circle, irrespective of their relative positions, and may expect further theorems to also hold generally for points and circles. In other words, when faced with a problem, the signs one is in the habit of using – those in which one's theorems are phrased – determine what one pays attention to, what expectations one has, and what potential moves one thinks of first.

The metaphor of the traveler introduced above should be clearer by now: Much like concepts, the signs used trace privileged paths across the mathematical landscape, in a way that impact what kind of theorems, and even what kind of objects, are seen as structuring the land.

As a last point, note that these considerations have interesting connections to a different literature – one that did not originate in philosophy of mathematics, but in the *historiography* of mathematics. The question at issue is whether historians can legitimately use "anachronistic" mathematical notations and concepts to make sense of historical mathematics. Our discussion above – about the wide-ranging consequences that "pragmatic" differences between signs can have – should make the potential dangers of such a procedure clear: Summarily translating an ancient source into modern algebraic symbolism is like trying to retrace the winding itinerary of an ancient traveler by following the contemporary highway system; while convenient and doubtlessly useful for orientation, it is often inadequate in the details.

6 Conclusion

In recent years, a large literature has developed around "representations" or (as I have put it here) "signs" in mathematics, the aims of which – beyond that of offering precise descriptions of aspects of mathematical practice – are diverse and sometimes hard to discern. This survey attempted to clarify what is at stake.

Part of the literature discussed above is not much concerned with philosophy of mathematics in the traditional sense. Some authors investigate mathematical cognition; others mainly aim at altering our perception of mathematics, say by shifting the emphasis away from certainty, rigor, or abstraction and toward open-endedness, creativity, or embodiment.

For the rest, we encountered two main ways in which work on signs bears on more strictly philosophical questions. First, work on the norms governing nonverbal signs in various mathematical practices has the potential to revive views of the nature of mathematics that center on sign manipulations rather than on logical inference, perhaps along Peircean lines (Sect. 3.3). It is fair to say that such views remain marginal today, and are insufficiently discussed head-on. However, they will likely keep proving attractive to researchers wishing to accommodate the variety of mathematical practices without giving up on the specificity of mathematics. From an ontological perspective, discussions of how reference to abstract objects can emerge from sign manipulations – like those of Parsons for symbols and of Panza for Euclidean geometricals – will be relevant here (see Sect. 3.1).

Second, a focus on signs (including concepts, though my emphasis here has been on nonverbal signs like symbols and diagrams, where the issues can often be formulated more sharply) allows approaching the way human computational limitations shape the mathematics that we do, and in particular determine what our salient mathematical objects and theorems are (Sect. 5, esp. 5.3). Moreover, studying signs is a promising avenue for tackling in a precise way the topic of mathematical understanding, to which – under various names and guises – a large portion of the philosophy of mathematical practice literature is devoted.

7 Cross-References

- ▶ Diagrammatic Proofs in Mathematics: (Almost) 20 Years of Research
- ▶ Platonism, *De Re*, and (Philosophy of) Mathematical Practice
- ▶ Structural Semiotics as an Ontology of Mathematics
- ▶ The Algorithmic-Device View of Informal Rigorous Mathematical Proof
- ▶ The Design of Mathematical Language
- ▶ What Happens, from a Historical Point of View, When We Read a Mathematical Text?

Acknowledgments For comments on previous drafts that helped me improve this survey considerably, I would like to thank Jeremy Avigad, Paola Cantù, Jessica Carter, Silvia De Toffoli, José Ferreirós, Valeria Giardino, Yacin Hamami, Nicolas Michel, Dirk Schlimm, Jean Paul Van Bendegem, and Roy Wagner, as well as my section editor, Juan Luis Gastaldi.

References

Allwein G, Barwise J (eds) (1996) Logical reasoning with diagrams. Studies in logic and computation 6. Oxford University Press, New York/Oxford

Avigad J (2008) Understanding proofs. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford/New York, pp 317–353

Avigad J (2020) Modularity in mathematics. Rev Symb Log 13(1):47–79

Avigad J (2021) Reliability of mathematical inference. Synthese 198(8):7377–7399

Avigad J, Morris R (2014) The concept of "character" in Dirichlet's theorem on primes in an arithmetic progression. Arch Hist Exact Sci 68(3):265–326
Avigad J, Morris R (2016) Character and object. Rev Symb Log 9(3):480–510
Avigad J, Dean E, Mumma J (2009) A formal system for Euclid's elements. Rev Symb Log 2(4): 700–768
Azzouni J (2005) Is there still a sense in which mathematics can have foundations? In: Sica G (ed) Essays on the foundations of mathematics and logic. Polimetrica, Monza, pp 9–47
Azzouni J (2009) Why do informal proofs conform to formal norms? Found Sci 14(1–2):9–26
Barwise J, Hammer EM (1994) Diagrams and the concept of logical system. In: Gabbay DM (ed) What is a logical system? Studies in logic and computation 4. Clarendon Press, Oxford, pp 73–106
Bellucci F, Pietarinen A-V (2016) Existential graphs as an instrument of logical analysis: Part I. Alpha. Rev Symb Log 9(2):209–237
Bourbaki N (1948) L'Architecture des mathématiques. In: Le Lionnais F (ed) Les grands courants de la pensée mathématique. Cahiers du Sud, Marseille, pp 35–47
Bourbaki N (1950) The architecture of mathematics. Am Math Month 57(4):221–232. Transl. of Bourbaki (1948)
Burgess JP, De Toffoli S (2022) What is mathematical rigor? APhEx 25:1–17
Carter J (2010) Diagrams and proofs in analysis. Int Stud Philos Sci 24(1):1–14
Carter J (2014) Mathematics dealing with 'hypothetical states of things'. Philosophia Mathematica (III) 22(2):209–230
Carter J (2018) Graph-algebras – faithful representations and mediating objects in mathematics. Endeavour 42(2–3):180–188
Carter J (2019) Philosophy of mathematical practice – motivations, themes and prospects. Philosophia Mathematica (III) 27(1):1–32
Carter J (2020) Logic of relations and diagrammatic reasoning: structuralist elements in the work of Charles Sanders Peirce. In: Reck EH, Schiemer G (eds) The prehistory of mathematical structuralism. Logic and computation in philosophy. Oxford University Press, Oxford/New York, pp 241–272
Cartier P (2000) Mathemagics. In: Planat M (ed) Noise, oscillators and algebraic randomness from noise in communication systems to number theory. Lectures of a school held in Chapelle des Bois, France, April 5–10, 1999. Lecture notes in physics 550. Springer, Berlin/Heidelberg, pp 6–67
Chemla K (2006) Artificial languages in the mathematics of ancient China. J Indian Philos 34(1–2):31–56
Chemla K (2012) Prologue. Historiography and history of mathematical proof: a research programme. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge/New York
Colyvan M (2012) An introduction to the philosophy of mathematics. Cambridge University Press
De Risi V (2015) Introduction. In: De Risi V (ed) Mathematizing space the objects of geometry from antiquity to the early modern age. Trends in the history of science. Birkhäuser, Basel, pp 1–13
De Toffoli S (2017) 'Chasing' the diagram – the use of visualization in algebraic reasoning. Rev Symb Log 10(1):158–186
De Toffoli S (2021) Groundwork for a Fallibilist account of mathematics. Philos Q 71(4):823–844
De Toffoli S (2023) Who's afraid of mathematical diagrams? Philos Impr. https://doi.org/10.3998/phimp.1348
De Toffoli S, Giardino V (2014) Forms and roles of diagrams in knot theory. Erkenntnis 79(4): 829–842
De Toffoli S, Giardino V (2015a) An inquiry into the practice of proving in low-dimensional topology. In: Lolli G, Panza M, Venturi G (eds) From logic to practice. Italian studies in the philosophy of mathematics. Boston studies in the philosophy and history of science 308. Springer, Cham, pp 315–336

De Toffoli S, Giardino V (2015b) Envisioning transformations – the practice of topology. In: Larvor B (ed) Mathematical cultures. The London meetings 2012–2014. Trends in the history of science. Birkhäuser, pp 25–50
Dean W (2018) Strict finitism, feasibility, and the sorites. Rev Symb Log 11(2):295–346
Dean W (2019) Computational complexity theory and the philosophy of mathematics. Philosophia Mathematica (III) 27(3):381–439
Detlefsen M (2008) Proof: its nature and significance. In: Gold B, Simons RA (eds) Proof & other dilemmas. Mathematics and philosophy. Spectrum series. The Mathematical Association of America, pp 3–32
Easwaran K, Hájek A, Mancosu P, Oppy G (2021) Infinity. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Winter 2021 edn. Metaphysics Research Lab, Stanford University, Stanford. https://plato.stanford.edu/archives/win2021/entries/infinity/
Feferman S (2012) And so on. . .: reasoning with infinite diagrams. Synthese 186(1):371–386
Ferreirós J (2007) Labyrinth of thought. A history of set theory and its role in modern mathematics, 2nd edn. Birkhäuser, Basel/Boston/Berlin
Ferreirós J (2016) Mathematical knowledge and the interplay of practices. Princeton University Press, Princeton/Oxford
Frigg R, Nguyen J (2020) Scientific representation. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Spring 2020 edn. Metaphysics Research Lab, Stanford University, Stanford. https://plato.stanford.edu/archives/spr2020/entries/scientific-representation/
Giaquinto M (2020) The epistemology of visual thinking in mathematics. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Spring 2020 edn. Metaphysics Research Lab, Stanford University, Stanford. https://plato.stanford.edu/archives/spr2020/entries/epistemology-visual-thinking/
Giardino V (2016) Behind the diagrams: cognitive issues and open problems. In: Krämer S, Ljungberg C (eds) Thinking with diagrams. The semiotic basis of human cognition. De Gruyter Mouton, Boston/Berlin, pp 77–101
Giardino V (2017) Diagrammatic reasoning in mathematics. In: Magnani L, Bertolotti T (eds) Springer handbook of model-based science. Springer, Dordrecht, pp 499–522
Giardino V, Patras F (n.d.) Proving with graphs: the mathematician's toolkit. (Forthcoming)
Grosholz ER (2007) Representation and productive ambiguity in mathematics and the sciences. Oxford University Press, Oxford/New York
Guedj D (1985) Nicholas Bourbaki, collective mathematician. An interview with Claude Chevalley. (Transl. by Gray J.). Math Intell 7(2):18–22
Haffner E (2014) The "science of numbers" in action in Richard Dedekind's works. Between mathematical explorations and foundational investigations. PhD thesis, Université Paris VII Diderot, Paris. https://hal.science/tel-01144626
Hamami Y (2022) Mathematical rigor and proof. Rev Symb Log 15(2):409–449
Hamami Y, Mumma J, Amalric M (2021) Counterexample search in diagram-based geometric reasoning. Cogn Sci 45(e12959):1–25
Howse J, Molina F, Shin S-J, Taylor J (2002) On diagram tokens and types. In: Hegarty M, Meyer B, Narayanan NH (eds) Diagrammatic representation and inference. Second international conference, diagrams 2002. Lecture notes in artificial intelligence 2317. Springer, Berlin/Heidelberg, pp 146–160
Kaiser D (2005) Drawing theories apart. The dispersion of Feynman diagrams in postwar physics. The University of Chicago Press, Chicago/London
Keränen JPM (2005) Cognitive control in mathematics. PhD thesis, University of Pittsburgh
Koppelman E (1971) The calculus of operations and the rise of abstract algebra. Arch Hist Exact Sci 8(3):155–242
Landy D, Goldstone RL (2007) How abstract is symbolic thought? J Exp Psychol Learn Mem Cogn 33(4):720–733
Landy D, Allen C, Zednik C (2014) A perceptual account of symbolic reasoning. Front Psychol 5(275):1–10

Lange M (2017) Because without cause. Non-causal explanations in science and mathematics. Oxford studies in philosophy of science. Oxford University Press, Oxford/New York
Larkin JH, Simon HA (1987) Why a diagram is (sometimes) worth ten thousand words. Cogn Sci 11(1):65–100
Luengo I (1996) A diagrammatic subsystem of Hilbert's geometry. In: Allwein G, Barwise J (eds) Logical reasoning with diagrams. Studies in logic and computation 6. Oxford University Press, New York/Oxford, pp 149–176
Lützen J (1979) Heaviside's operational calculus and the attempts to rigorise it. Arch Hist Exact Sci 21(2):161–200
Mancosu P (2005) Visualization in logic and mathematics. In: Mancosu P, Jørgensen KF, Pedersen SA (eds) Visualization, explanation and reasoning styles in mathematics, Synthese library, vol 327. Springer, Dordrecht, pp 13–30
Manders K (1996) Diagram contents and representational granularity. In: Seligman J, Westerståhl D (eds) Logic, language and computation, CSLI lecture notes 58, vol 1. CSLI Publications, Stanford, pp 389–404
Manders K (1999) Euclid or Descartes? Representation and responsiveness. Unpublished but widely circulated manuscript, August 1999
Manders K (2008a) Diagram-based geometric practice. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford/New York, pp 65–79
Manders K (2008b) The Euclidean diagram (1995). In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford/New York, pp 80–133
Manders K (2012) Expressive means and mathematical understanding. Unpublished draft, May 2012
Marghetis T, Landy D, Goldstone RL (2016) Mastering algebra retrains the visual system to perceive hierarchical structure in equations. Cogn Res Princ Implic 1(25):1–10
Miller N (2007) Euclid and his twentieth century rivals diagrams in the logic of Euclidean geometry. In: Studies in the theory and applications of diagrams. CSLI Publications, Stanford
Mumma J (2006) Intuition formalized. Ancient and modern methods of proof in elementary geometry. PhD thesis, Carnegie Mellon University. http://johnmumma.org/Writings_files/Thesis.pdf
Mumma J (2019) The Eu approach to formalizing Euclid. A response to "On the inconsistency of Mumma's Eu". Notre-Dame J Formal Log 60(3):457–480
Netz R (1999) The shaping of deduction in Greek mathematics. A study in cognitive history. In: Ideas in context 51. Cambridge University Press, Cambridge
Panza M (2012) The twofold role of diagrams in Euclid's plane geometry. Synthese 186(1):55–102
Parsons C (2008) Mathematical thought and its objects. Cambridge University Press, Cambridge/New York
Peirce CS (1885) On the algebra of logic: a contribution to the philosophy of notation. Am J Math 7(2): 180–196. Repr. in Peirce (1931–1958), vol 3, §359–403
Peirce CS (1931–1958) Collected papers. 8 vols. The Belknap Press of Harvard University Press, Cambridge, MA
Rabouin D (2015) Proclus' conception of geometric space and its actuality. In: De Risi V (ed) Mathematizing space. The objects of geometry from antiquity to the Early Modern Age. Trends in the history of science. Birkhäuser, Basel, pp 105–142
Rav Y (1999) Why do we prove theorems? Philosophia Mathematica (III) 7(1):5–41
Rav Y (2007) A critique of a formalist-mechanist version of the justification of arguments in mathematicians' proof practices. Philosophia Mathematica (III) 15(3):291–320
Rotman B (1993) Ad infinitum. The ghost in Turing's machine: taking god out of mathematics and putting the body back in. Stanford University Press, Stanford
Rotman B (2000) Mathematics as sign. Writing, imagining, counting. Stanford, Stanford University Press
Schlimm D (2018) On Frege's *Begriffsschrift* notation for propositional logic: design principles and trade-offs. Hist Philos Log 39(1):53–79

Schlimm D, Neth H (2008) Modeling ancient and modern arithmetic practices. Addition and multiplication with Arabic and Roman numerals. In: Love BC, McRae K, Sloutsky VM (eds) Proceedings of the 30th annual conference of the Cognitive Science Society, CogSci 2008 (Washington, DC. July 23–26, 2008). CD-ROM, archived online. Cognitive Science Society, Austin, pp 2097–2102. http://csjarchive.cogsci.rpi.edu/Proceedings/2008/pdfs/p2097.pdf

Schlimm D, Waszek D (2020) Multiple readability in principle and practice: existential graphs and complex symbols. Logique Anal 251:231–260

Scotto di Luzio P (2002) Patching up a logic of Venn diagrams. In: Vermeulen K, Copestake A (eds) Algebras, diagrams and decisions in language, logic and computation. CSLI lecture notes 144. CSLI Publications, Stanford, pp 119–134

Shin S-J (1994) The logical status of diagrams. Cambridge University Press, Cambridge/New York

Shin S-J (2004) Heterogeneous reasoning and its logic. Bull Symb Log 10(1):86–106

Shin S-J, Lemon O, Mumma J (2013) Diagrams. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Fall 2013 edn. Metaphysics Research Lab, Stanford University, Stanford. https://plato.stanford.edu/archives/fall2013/entries/diagrams/

Shin S-J, Lemon O, Mumma J (2018) Diagrams. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Winter 2013 edn. Metaphysics Research Lab, Stanford University, Stanford. https://plato.stanford.edu/archives/win2018/entries/diagrams/

Stapleton G, Jamnik M, Shimojima A (2017) What makes an effective representation of information: a formal account of observational advantages. J Log Lang Inf 26(2):143–177

Stapleton G, Shimojima A, Jamnik M (2018) The observational advantages of Euler diagrams with existential import. In: Chapman P, Stapleton G, Moktefi A, Perez-Kriz S, Bellucci F (eds) Diagrammatic representation and inference. 10th international conference, diagrams 2018. Lecture notes in artificial intelligence 10871. Springer, Cham, pp 313–329

Starikova I (2012) From practice to new concepts: geometric properties of groups. Philosophia Scientiæ 16(1):129–151

Tappenden J (2008a) Mathematical concepts and definitions. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford/New York, pp 256–275

Tappenden J (2008b) Mathematical concepts: fruitfulness and naturalness. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford/New York, pp 276–301

Van Bendegem JP (1996) The strange case of the missing body of mathematics. Rev Brian Rotman 112(3–4):403–413

Vold K, Schlimm D (2020) Extended mathematical cognition: external representations with non-derived content. Synthese 197(9):3757–3777

Wagner R (2009) Mathematical marriages: intercourse between mathematics and semiotic choice. Soc Stud Sci 39(2):289–308

Wagner R (2010) The natures of numbers in and around Bombelli's *L'algebra*. Arch Hist Exact Sci 64(5):485–523

Wagner R (2017) Making and breaking mathematical sense: histories and philosophies of mathematical practice. Princeton University Press, Princeton

Waszek D (2022) Are Euclid's diagrams representations? On an argument by Ken Manders. In: Zack M, Schlimm D (eds) Research in history and philosophy of mathematics. The CSHPM 2019–2020 volume. Annals of the Canadian Society for History and Philosophy of Mathematics. Birkhäuser, Cham, pp 115–127

Waszek D (2023) Informational equivalence but computational differences? Herbert Simon on representations in scientific practice. Mind Mach. https://doi.org/10.1007/s11023-023-09630-4

Waszek D (n.d.) Notational differences, exploration, and discovery in mathematics: a case-study of notation-driven discovery. (Forthcoming)

Waszek D, Schlimm D (2021) Calculus as method or calculus as rules? Boole and Frege on the aims of a logical calculus. Synthese 199(5–6):11913–11943

Wege TE, Batchelor S, Inglis M, Mistry H, Schlimm D (2020) Iconicity in mathematical notation: commutativity and symmetry. J Numer Cogn 6(3):378–392

Weil A (1978) Who betrayed Euclid? (extract from a letter to the editor). Arch Hist Exact Sci 19(2): 91–93

Wilson M (2006) Wandering significance. Clarendon Press, Oxford/New York

Wilson M (2021) Innovation and certainty. Cambridge elements in the philosophy of mathematics. Cambridge University Press, Cambridge

Zach R (2023) Hilbert's program. In: Zalta EN, Nodelman U (eds) The Stanford encyclopedia of philosophy, Spring 2023 edn. Metaphysics Research Lab, Stanford University, Stanford. https://plato.stanford.edu/archives/spr2023/entries/hilbert-program/

Structural Semiotics as an Ontology of Mathematics

Roi Wagner

Contents

1 Introduction 3004
2 Structural Semiotics 3005
3 A Mathematical Example 3009
4 Post-Structural Semiotics 3014
5 Semiotics as Ontology 3018
6 Examples for Semiotics as Ontology 3022
7 Conclusion 3028
8 Cross-References 3029
References 3029

Abstract

This chapter introduces a minimalist framework for structuralist and post-structuralist semiotics based on the work of de Saussure, Lévi Strauss, and Derrida. It then applies this framework to mathematical case studies. Finally, it explains what it means to consider semiotics as an ontology of mathematics and provides examples of this approach.

Keywords

Philosophy of mathematical practice · Structuralist semiotics · Post-structuralism · Ontology of mathematics

R. Wagner (✉)
ETH, Zurich, Switzerland
e-mail: roy.wagner@gess.ethz.ch

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_131

1 Introduction

As explained in the introduction to this section of the handbook, the terms semiotics and semiology are used in a variety of ways, loosely grouping together various extensions of linguistic methods beyond the context of natural languages. In the context of history and philosophy of mathematics, self-proclaimed semiotic approaches are few and far between, so a careful delineation of strands of semiotic approaches in this field is probably premature (an exception to this is the field of mathematics education, where the discussion of semiotics is much more advanced; e.g., Ernest (2006, 2018), and Presmeg et al. (2018)). Waszek's chapter in this handbook (▶ "Signs as a Theme in the Philosophy of Mathematical Practice") covers two important strands in the semiotics of mathematics: the use of diagrams in rigorous mathematical argumentation and the not necessarily rigorous, but creative use of signs to suggest new mathematical explorations.

It seems that most often, semiotic approaches to mathematics explore the interactions of several layers of signs – either layers of different kinds of signs, or layers within an apparently uniform regime of signs. The goal is usually to understand how such linguistic layers interact, formally or informally, coherently or ambiguously, to form new meanings.

Some prominent examples for this approach include Reviel Netz' (1999, see also his ▶ "Shaping, Revisited" in this handbook) discussion of how text and diagrams interacted to form the new discourse of classical Greek geometry, set in a cognitive history framework, as well as Vinciguerra's Foucauldian analysis of the interaction between the sayable and the visible in early modern mathematics (1999). Other attempts include O'Halloran's (2005) analysis of various semiotic modalities and their interactions across mathematical history, couched in the framework of systemic functional grammar, as well as Kvasz' chapter in this handbook (▶ "Symbolic algebra as a semiotic system") discussion of the differences between the expressive capacities of geometric and algebraic languages. Friedman (2021) opens up a general exploration of the notion of translation between mathematical languages, and Barton (2008, 7–11) presents an exciting case study of what happens to the meaning of number words when they are translated between English and the Maori language.

Looking for linguistic layers inside an apparently uniform linguistic system, we can site Rotman's discussion of the collapse of meta-language into language in mathematics (1987) and the way that different grammatical moods generate a division of labor between mathematical subject positions (2000, chs. 1–2; see also Wagner 2009, ch. 1). Gastaldi tracks down the distinction between algebraic expressions and logical propositions in Boole's work (2022), and between names and propositions in de Morgan's (forthcoming). Grosholz (2007) explores the notion of "productive ambiguity," where a single sign is interpreted with respect to two different, not necessarily coherent, systems of signification.[1]

[1] Wittgenstein's discussion of mathematical language(s) (1976, 1981) can be read as an attempt to tease layers of language that we tend to mix together around mathematics; however, Wittgenstein never considered himself, as far as I know, to be doing "semiotics."

Most semiotic approaches to mathematics point to Peirce as a theoretical source (although sometimes only superficially). Very little semiotic analysis of mathematics is done in an explicitly structuralist (or post-structuralist) framework. Brown (2001) delineates a general structuralist and post-structuralist semiotic framework for mathematics, but Herreman's (2000) Hjelmsevian approach is clearly the most developed. It breaks down mathematical languages into "planes of content" and explores how their interactions produce new mathematical ideas. Wagner (2009, 2017, chs. 3–4) explores shifts of meaning that occur when different layers of interpretation of the same signs interact, building on post-structural inspirations. Steensen's doctoral dissertation (2022) develops both approaches further, and applies her enriched theoretical framework to analyze nineteenth century algebra. Finally, Gastaldi's contribution to this handbook ("Neural Machine Learning Applications to Mathematics") points the way to a digital analysis of mathematical texts based on structuralist methods.

Here I will try to outline a semiotic approach to the analysis of mathematical text. I will first present a minimalist structuralist analytical framework, which can be extended or embedded into more elaborate ones. Next, I will apply it to a mathematical test case, that of the numeral 0. I will go on to highlight the instabilities embedded in this framework, in the vein of post-structuralist semiotic thinking. Then I will argue that the semiotics of mathematics affords a new approach to the ontology of mathematics, in a manner that fits with the interests of the philosophy of mathematical practice, rather than those of more classical ontological debates. I will conclude with some examples of such ontological descriptions.

2 Structural Semiotics

Let's start with an "origin myth" for signs, inspired by de Saussure (2011, Part I, ch. 1, part II, chs. 2–4). The pre-Saussurian origin myth of signs tells us that there are objects (such as trees), and that we can make articulate sounds (such as the word "tree"). By connecting one to the other, conventionally or by means of some hidden logic, a signifying relation is formed. De Saussure objected to this story, because it assumes that words and objects are unproblematically there in advance. But why would a tree be an objective unit, rather than the bottom half of the trunk and the surrounding patch of earth accompanied by a warm, pleasant feeling? And if words are unproblematically given, how come languages are so different, that the speaker of one language can hardly figure out where, in another language, one word ends and another begins?

To answer these questions, let me present a different origin myth, more Saussurian in nature, to explain how signs emerge. According to this myth, we are exposed to continuous, unarticulated streams of experience. These include sensory streams (like perceived sounds) and cognitive-experiential streams (like our thought-experience). These streams do not have constitutive units; their flows resemble that of water. However, such streams often correlate. We may, e.g., hear similar things while seeing similar things and undergoing similar experiences. These

conjunctions divide the streams into correlated units, which, a-posteriori, may end up being described, in the case of one such possible conjunction, as "bird-chirping sound," "the view of the sun rising," and the experiential-cognitive "concept of morning."

If we are in a linguistic community, this correlation may be further correlated with another articulation, this time an articulation of the stream of sounds of spoken language. The word "morning" may emerge from this stream and correlate with the above experiences. A sign is born: the correlate articulation of portions of two or more streams of sensory-cognitive experience. Note that the sign is neither the word (the signifier "morning") nor the experience (the signified or concept of morning), but that which binds them together as co-articulated units of the respective streams. The streams are now streams of signifiers and signifieds co-articulated into signs.

I call this an "origin myth" not just because of its idyllic setting, which connotes a biblical "let there be light" with a naïve and optimistic picture of waking up (this choice is mine, not de Saussure's). I call this an origin myth primarily because it is not at all clear how these correlations turn into a co-articulation of units and what are the metaphysical or empirical conditions for this to take place. For example, the notion "similar" in the above narrative is not well defined, and a definition risks forcing pre-fabricated units on the supposedly unarticulated experiential streams.

De Saussure is, in fact, careful enough so as not to provide this or any other origin myth. Instead, he gives us a metaphor: Just as the unarticulated flow of air and unarticulated flow of water interact to generate discrete waves, so do the flows of perceived sound and lived experience interact to generate articulated signs that bind words and concepts (de Saussure 2011, 112). De Saussure does, moreover, provide us with ways to operationalize the notion of sign articulation that he introduces. This is the test of commutation (below I will subsume all of structural semiotics under this test; this is a substantial simplification and conceptual stretch, but would facilitate the exposition).

Commutation helps identify the scope of a sign by attempting to replace a portion of one of the correlated streams and observe its effect on the other stream. For example, the phoneme "r," according to de Saussure, is not defined as such in any natural way. Different speakers of the same language (or the same speaker at different times) may pronounce the sound "r" very differently. Such differences, which are insignificant in one language, which may not even be perceived as different by speakers of a given language, may correspond to completely different phonemes in another language. Think, e.g., of the distinction between "r" and "l" in English, its collapse to a single phoneme in some Chinese accents, and its expansion into five different distinct phonemes in a language like Malayalam.

What defines the English "r" from this point of view is therefore not some physical-physiological constraints, but rather the fact that some ways of expressing "r" make a difference on the signified side, while others do not. The "r" of two different native English speakers is considered the same, because commuting one for the other does not break the bond of their different pronunciations with a common signified (e.g., the word "war" means the same, whether the "r" is rhotic or non-rhotic). However, someone will be said to have a speech impediment or a

confusing accent, if what they commute for an "r" sound may lead to bind their pronunciation of "war" with the signified normally associated with "wall." The English "r" would then be defined as all those sounds which, when substituted for Eliza Doolittle's or Arnold Schwarzenneger's "r" sound when they say "war," would still be linked to their intended signified. The phenomenon is similar on the signified side. If someone says that "this is a bush" and "that is a shrub," but I cannot figure out why I can't commute this for that and still call it a bush, then my articulation of the concept of bush is different from theirs.

A sign, in structuralist terms, is that which binds a certain part of one experiential flow with a certain part of another, and the features of those "parts" can be identified by commutation tests. A sign, therefore, is not defined positively, by its content, but negatively, by the fact that some changes in one flow (e.g., sound) make a difference in another (thought experience or concepts). Those productive differences constitute the *value* of a sign. Signs, therefore, do not have content, only values witnessed by commutations.

This is the main principle of structural semiotics: *Signs are systems of differences that are captured by commutation tests*.

Note that this principle is not restricted to the relation between spoken sounds and thought concepts. Any two streams can play the roles of signifier and signified, as long as the co-articulation of their units is captured by commutation tests. The extension of the scope of analysis from correlating speech or writing streams with concepts to correlating any two streams is precisely where structural linguistics becomes structural semiotics. This has a surprising corollary: everything is a sign. More precisely, everything is *also* a sign. Even a hammer is a sign: if I see a person holding a hammer and a person holding a pen, I tend to have different expectations about what they are going to do and who they are. I may make assumptions about their jobs and social positions. In other words, I will probably conceive them differently. This is why hammers and pens appear as meaningful signs on flags and in commercial logos.

This does not mean that commutation tests would capture everything there is to know about everything. For example, a commutation test would not capture the physical causality between smoke and fire, even if it could establish (under some constraints) that the appearance of one in the visual experience stream is correlated with the other in the heat experience stream. And one could, e.g., use a stop sign to hit someone over the head, rather than convey a commutationally-articulated meaning.[2] But structural semiotics is interested precisely in things *as* signs, that is, in what we can learn about them by means of commutation tests.

Commutation tests are in fact richer then they appear at first sight. They can be performed within a language: e.g., can I always exchange "dog" by "canine"? If so, they are perfectly equivalent; otherwise, they are not. They can also take place between languages: for instance, if commuting "I" with "you" corresponds to a

[2] Within an appropriate ontological framework, these two examples might arguably be reducible to semiotic values, but I am not interested in pursuing such a reductive trajectory here.

commutation between "je" and "tu," then the English sign relation "I : you" would be equivalent to the French "je : tu" (the presence of the French honorific "vous" complicates the picture). Such cross-linguistic commutation tests also help us understand the nuances of relations between signs and differences between languages.

There is yet another kind of commutation test. For example, if, in English, I change the subject of a sentence from "I" to "he," this often entails adding the suffix "s" somewhere else in the sentence – a change that contributes to the articulation of the grammatical notion of "verbs" and of their conjugation. Similarly, if I consider what can replace the word "dog" in the sentence "I see a dog" while maintaining grammatical correctness, I am on my way to articulate a certain class of so-called nouns. Therefore, syntactic rules are also captured by some kind of commutation tests.[3]

Note that the last two examples impose a coarser re-articulation on the already articulated stream of words (here, into verbs and nouns). This articulation is not about linking words with concepts, but about linking words with other words. The analysis focuses on the level of the signifier (words), rather than on their relation to the signified (concepts). In the development of structural semiotics, this shift of focus was indeed the trend. In fact, as structural semiotics evolved, the relation between the signifier and signified became less and less obvious. For Lévi-Strauss (1987), not every articulated signifier-word corresponds to an articulated signified-idea. The development of language is actually all about re-articulating the units of signifying matter with respect to signified experience – an adaptation that can never be perfect. For Lacan, semiotics is characterized by "the incessant sliding of the signified under the signifier" (2007, 419), and only occasionally do the two adhere to each other in a stable manner.

As a result, the structural semiotics developed with an increasing emphasis on the formation of relations between signs, rather than between signifiers and signifieds. The work of Lévi-Strauss is exemplary in this respect (Lévi-Strauss 1955 provides self-contained, accessible examples for the processes I mention below). One process of sign formation that he identifies is mediation. When an exclusive binary opposition in language is to be used in a non-binary world of experience, humans may come up with intermediary signs. For example, the opposition between life and death may be mediated by hunting, which is killing in order to live, or by ghosts.

Another process that Lévi-Strauss identifies is analogy: by relating a challenging binary opposition to a more mundane opposition, one that we are more comfortable with, we may feel that we can make sense of the former. For example, the life-death opposition may be considered analogous to sleep-wakefulness. This may help us understand the relatively rare and traumatic transition of the former in terms of an everyday occurrence. It may help us come to terms with the kind of absence that death entails, and even suggest a hope for revival.

[3] In contemporary computational settings, such binary commutation tests are usually replaced by a statistical analyses of a concrete textual corpus (see Gastaldi 2021).

Yet another process of sign formation is recombination: re-arranging the components of a sign relation in order to form new relations, which may complement them and enrich our expressive capacities. For example, if life is associated with motion and reproduction, whereas death is associated with stillness and decay, what happens when we recombine motion and decay on the one hand and stillness and permanence on the other?[4] I leave it to the reader to consider what physical, theological, or social concepts can emerge from this recombination, and how they relate to the original opposition between life and death.

This kind of sign-forming, according to Lévi-Strauss, is the universal of human thought. Even modern scientists look for phenomena at the border of life and death (between complex proteins and viruses, or among organisms that can survive a long suspension of all metabolic processes). They also use analogies, trying to come to terms with the biological life-death opposition by recasting it in chemical and physical terms, and conceptual recombinations also abound (e.g., artificial life). Of course, different regimes of knowledge scrutinize these procedures differently (a mythical analogy between life : death and sleep : wakefulness depends on a very different justification then the recasting of life : death in terms of some opposition of chemical processes), but, according to Lévi-Strauss (1966, 262–269), they share the same set of underlying semiotic moves.

3 A Mathematical Example

Let us try to explore a mathematical example for the application of semiotics. This example serves two purposes. First, it is meant to elucidate and concretize the above discussion. But it will also serve to pinpoint its weaknesses. For rhetorical reasons, I will first suppress these weaknesses, only to address them later on. The reader is asked to play along and reserve their critique until the subsequent section.

Consider the symbol 0. Let's focus on positive integer arithmetic in standard decimal place value notation. In this context, the signifier 0 is almost entirely commutable with any other digit – but only almost. We do not put zero at the leftmost decimal place of a number (i.e., the form 058 would usually be considered faulty or at least redundant). This imposes a syntactic distinction on 0, which sets it apart from other signs, even before we consider its meaning.

Does this syntactic difference make a conceptual difference? There is at least a possible difference: unlike other numerals, which represent numbers or quantities, in the above context, 0 can be conceived as a placeholder. A decimal place that is there in a number, but is unoccupied by a numerical quantity, is held by a 0. Note, however, that a place that does not belong to the range of a number, such as the place of the hundreds or thousands in the number 58, need not be "held" – it is

[4] This is a take on Lévi-Strauss's so called canonic formula (Lévi-Strauss 1955, 442). Since Lévi-Strauss himself was quite cavalier about it (Sperber 1985, 65–66), I allow myself to improvise around the original, rather than attempt to capture the original nuance.

simply not there. 0 serves as the mark of the existence of a relevant, but otherwise unmarked space.[5]

What kind of evidence can support or clarify this interpretation? One test is that of inter-linguistic commutation, namely, translation to other systems of representation. Indeed, in some closely related systems, such as the Chinese rod numerals or some Cuneiform notations (Katz 2007, 196; Schilmm and Skosnik 2011), 0 is replaced by a visible gap in the sequence of numerals. Another commutation test that can strengthen this interpretation is the presentation of unknown quantities in rule of three problems. In such contexts, 0 has a history of being used to hold the place of the unknown answer (instead of a blank space, a verbal designation, or modern algebraic letters). We can see examples of this in very diverse cultures, such as in the Indian *Bakhshali manuscript* (probably late first millennium; Hayashi 1995), in some Hebrew manuscripts of Abraham Ibn Ezra's *Book of Number* (originally twelfth century; Katz et al. 2016, 231–232) and in Simon Stevin (*La Pratique d'Arithmetique*, sixteenth century, see Coquard 2021, 467 for the extract). The numerical value of the answer has nothing to do with the modern quantitative notion of 0, so 0 is clearly a placeholder (or, in semiotic terms, commutable with empty but relevant positions), rather than a quantity.

Yet another way to explore the relationship between the syntactic distinction of 0 and its special mathematical meaning is provided by constructing closely related systems that do not have such a syntactic distinction. This would be a sort of "meta-commutation" test, comparing systems that are "almost" the same, but differ with respect to the syntactic rule we wish to analyze. This comparison would help to isolate the impact of the specific syntactic rule. Consider the place value number representation where there is no 0, and to the symbols 1,2,. . .,9 we add the new numeral X, representing 10 units. In this representation the number 100, for example, will be presented as 9X ($= 9 \times 10 + 10 \times 1$), 101 will become X1 ($= 10 \times 10 + 1 \times 1$), 110 will be XX ($10 \times 10 + 10 \times 1$), and 111 will be 111. Here, X will be syntactically commutable with any other numeral, and is clearly not a place holder (Foster 1947).[6] This lends plausibility to the relation between 0's syntactic distinction and its semantic distinction, serving as a placeholder rather than a quantitative numeral.

Let us continue our exploration of alternative representations in order to study the relation between the syntactic exception of 0 in standard decimal notation and its meaning. To do that, we will consider similar systems that do contain 0, but where it manifests no syntactic exception. Consider, for example, modern computer-oriented contexts where the number of digits is fixed in advance, and in order to maintain this

[5] Of course, this may change when this practice is considered within richer contexts, such as real numbers or temperature measurement, but a change of context often changes sign relations, as we will see below.

[6] I owe my exposure to this system to its independent development by Dirk Schlimm in his unpublished "A decimal place-value system without zero" (2012), which lists earlier publications of this recurrent idea.

number, a string of zeros is required to fill the empty decimal places on the left. Here, there is no longer a syntactic difference between 0 and the other numerals.

In this case, I claim, the entire set of semiotic relations changes, and the meaning of 0 with it. This is witnessed semiotically by the arithmetic procedures and identities applied to such strings (note that an identity, e.g., between 2 + 3 and 5, allows to commute one for the other, so it is a commutational-semiotic relation). In the fixed-length representation, some arithmetic identities must undergo some changes. For example, this representation requires handling the possibility of overflow, as in the case of 50...0 + 50...0. Moreover, the way a computer would perform these operations may be different, involving all sorts of logical operations on strings that are divorced from our handwritten decimal arithmetic algorithms, so as to adapt the code to the relevant data structure or hardware. This introduces new operations, new identities, and therefore new semiotic relations. It turns out that, in this case, commuting one materiality for another (electronic circuits for pen and paper) is correlated with differences in semiotic relations that are greater than they initially appeared.

Moreover, in the same computational context, the translations of the above fixed-length representations to the standard handwritten decimal representation (or, semiotically, their cross-linguistic commutation) is often under-determined, as in floating-point representations (where the string 0123 could mean .00123, 1.23, etc.).[7] We see that the slightly different notation, with its slight difference concerning the positions available to 0, is actually embedded in a very different notation, subject to very different operations, and therefore entering different intra and cross linguistic commutation relations.

In such a context, the gained symmetry between 0 and other numerals comes at the cost of a more formal, rather than numerical, evaluation of digits, where the arithmetical meaning of strings, numerals and even places is much less stable. This is most salient when we say things like "computers express everything by strings of 0's and 1's," considering both 0 and 1 as arbitrary and exchangeable, not necessarily numerically evaluated symbols. What appeared to be a slight syntactic difference ended up being correlated with many other differences, which change the meaning not only of 0, but of all other digits. One may even draw the conclusion that the two

[7] We have noted that the semiotic values of signs involve not only intra-linguistic commutations (e.g., the relation between 0 and other digits), but also cross-linguistic ones (e.g., the relation between different systems of number representation). This is an important aspect of mathematical signs. Historically, we can witness it in the case of negative numbers. In one context, they evolved through binomials: subtraction of two terms. Some of the earliest acknowledgments of negative numbers are in the form "0 less a number," and the rules for applying arithmetical operations to these numbers are analogous to the rules for applying operations to combinations of the form "number less number" where the result is positive (Wagner 2017, ch. 2; La Nave and Mazur 2002). However, the articulation of these new negative numbers also depended on their cross-linguistic translation to debts. Neither kind of relation determined negative numbers alone. Similarly, in the case of complex numbers, sets of rules (again, by analogy to other binomials and their roots) were usually not enough for mathematicians to endorse them as meaningful. Geometric representations also played a part, even before the emergence of the modern complex plane (Wagner 2010).

systems are not commensurable enough for the purposes of a fine-grained semiotic comparison.

The last example is actually meant as a sort of cautionary tale: semiotic comparisons that appear simple may actually depend on completely different materialities and practices, which render the comparisons possibly misleading. Can we, instead, try to isolate only the "internal" or "syntactic" difference (0 is/isn't allowed at the beginning of a number) and consider its impact alone, without introducing all the other changes involved in the computer oriented practice of the above fixed-length representation? Note that while isolating a difference in this way is tempting, it is not always practically possible. As we saw, actually attested changes in one syntactical aspect may be associated with a change in functionality and materiality (this is also attested in a remarkable manner, in the context of 0, in Herreman 2001). Trying to pinpoint a specific difference that does not occur in isolation in any "real-world" system might therefore entail uncontrollable speculations, which could lead us everywhere or nowhere. The fact that a specific semiotic difference is not actually included or excluded in isolation in "real-world" semiotic systems, without triggering other semiotic differences, *might* (but does not have to!) mean that it makes no sense to consider it separately from those other semiotic differences.

However, in our case we do have a real life close analogue that captures the relevant syntactic difference. Consider positive real number decimal representations.[8] We can easily attest the representations 0.2, 0.200, and 0.2000... in actual use. This appears to be precisely what we want to capture, only applied to the right side of the number instead of its left-hand side. What happens to 0 here? Is 0 still a place holder? The equivalence of the different decimal representations, with and without additional 0s on the right, suggests that fractional decimal places are there whether 0 holds them or not (as opposed to the standard representation, where 058 would be deviant, suggesting that the hundreds place is simply not there, and therefore should not be held by 0). The empty places to the right of the decimal point are part of the number, even if they are empty, and they do not require 0 to hold their place. If 0 is not required to hold the place, but may still be there, we may suspect that 0 is not strictly a placeholder, but has some other function – perhaps it even becomes a numerical value.

So regardless of what happens to the concept of 0 in this context, something significant has changed: all decimal places are always there, as evidenced by the potential commutations of strings with and without additional 0s on the right. Every decimal representation implies an infinity of other equivalent representations and even an implied infinite representation. As a result, there is something infinite in any real number – its infinite set of decimal places. If we allow ourselves to take the risk and speculate (in the absence of historical evidence or appropriate cognitive experiments), returning to positive integers, but now allowing arbitrary 0 signs to be added

[8] Note that I insist on positive numbers, so that the use of 0 by itself, which implies that it is a numerical value, is not available as evidence in our discussion. This is, I admit, an artificial move, as will become apparent in the next section.

on the left, the implicit infinity associated with infinite decimal representations of real numbers might end up being projected on the finite integers as well. Regardless of the meaning of 0, something might change about our notions of number. Given such a change, mathematicians may even come up with the idea of "extended natural" numbers, allowing for infinitely many digits.[9]

To clarify: I do not claim that in the original representation (standard decimal representation of positive integers) 0 is *only* a place holder, and that any attempt to tinker with its syntactic distinction implies that it is *not only* a place holder. We do not have enough evidence to claim that, and, as we will see in the next section, the question is not necessarily even well posed. The point here was to give examples of semiotic explorations of 0, building on its semiotic distinction in one system, by means of applying various kinds of commutation tests.

So let's move on. We discussed in the previous section some processes for generating new signs from old ones. What would, for example, a Lévi-Straussian "mediation" look like in a mathematical context? Consider the strict binary opposition between absence and presence of a numerical value, that is between zero and non-zero. Mathematicians have been toying with ways of bridging the gap in many different cultures and contexts well before the explosion of infinitesimals in early modern Europe (e.g., Archimedes' method for area and volume calculations, see Netz 2022, 187–193; Bar Hiyya's triangulation of the circle, see Katz et al. 2016, 306–307; Chinese volume calculations in commentaries on the *Nine Chapters*, see Katz 2007, 256–259, 261–263). But we will be interested in the example of Euler, for whom an infinitesimal was a zero ("an infinitely small quantity is simply an evanescent quantity and therefore actually equal to zero," from *Institutiones Calculi Differentialis*, quoted in Ferraro 2004, 45).

The zero-ness of infinitesimals was expressed by the fact that they were additively neutral: for Euler, a finite magnitude plus an infinitesimal strictly equals (rather than "converges to" or "is indistinguishable from") the original finite magnitude. However, without introducing a difference, the concept of infinitesimal will be strictly identical, according to the rules of commutation tests, with the old 0. So what is the difference?

The difference occurs in the contexts of division. While before infinitesimals, 0/0 would be simply meaningless – an indication that our mathematical description or calculation went wrong – in the context of infinitesimals, 0/0 (in equivalent forms such as $dx/5dx$ or ddx/dx) would have determinate values that depend on the specific infinitesimal, or, in different words, on the specific kind of zero. For Euler, quantity signs were not reducible to numerical values, not because they might lack values, but because they could be more than numerical values – their function depended on the context of their generation as well (Ferraro 2004, 42, 47–48). In this context, infinitesimals were zeros, but zeros emerging from different circumstances, and thus manifesting different relations to each other (e.g., $d(x^2)$, dx and $d \log x$ are all

[9] As my students occasionally did in the context of discussions of real number representation and naïve infinite set theory at the early stages of an introductory logic and set theory course.

zeros, but have different relations to each other). Zeros were no longer restricted to the roles of absence or place holding; as infinitesimals, they had a relative quantification (ratios to each other and to other quantities) that was different from that of finite magnitudes.

This was an instance of a Lévi-Straussian mediation. What about analogy? Euler's interpretation is not the only – nor indeed the main – interpretation of infinitesimals. There were many other approaches (Boyer 1959, chs. 4–6). One interesting approach is that which denies the existence of the infinitesimal itself, but acknowledges that there is an infinitesimal ratio. In this context, neither *dx* nor *dy* are meaningful entities, but their ratio, *dy*/*dx*, can be identified with a ratio between numbers (this was philosophically elaborated by Hegel, see Sommers-Hall (2010), but was not unique to him). In this way, we approach the question of what is an infinitesimal by analogy ("I don't know what they are exactly, but they relate to each other like . . . relates to . . .").

The story of zero is far from exhausted by these examples. Zero changes its semiotic functions (behavior in commutation tests) in as many ways as the systems in which it is embedded. When one moves from natural numbers to negative or real numbers, the notion of zero changes. Later on, when 0 is conceived as the additively neutral element in commutative groups or fields, or as the empty set in a canonical set-theoretic model of the natural numbers, it changes even further.

The bottom line is that the semiotics of mathematics studies mathematical signs and their meanings by exploring all kinds of commutation tests. We ask which signs or combinations of signs are commutable or equivalent to others in the same language, how a change of sign in one expression correlates with changes of other signs somewhere else in that expression, and how a change of sign in one system of presentation translates to changes of related signs in other systems of presentation. The ensuing characterization of signs captures their relations or values as signs, and the system of all such relations is the proper object of structural semiotics and its vehicle for discussing questions of meaning.

4 Post-Structural Semiotics

Much of what I wrote above may appear confused and sketchy. This is partly the cost of trying to condense a theoretical approach and some not-completely-trivial examples into the span of a few pages – some sketchy arguments and approximate formulations are inevitable.[10] But there are more substantial grounds to object to the exposition above. In fact, a certain kind of objection is an inherent feature of the theory.

Consider, e.g., the case of the pronunciation of "r" in English. People with different indigenous British accents may pronounce the sound "r" in "war" in

[10] At any rate, the analysis above is only meant to show how the theory works, not to make proper historical claims. The latter may be found in analyses such as Rotman (1987) and Herreman (2001).

recognizably different ways, while still maintaining the bond with the same signified. In that sense, their "r" sounds are commutable, and form the same signifier. But the difference is not meaningless. It can provide clues for where the person might be from or indicate their social position. So, if the context of signification is sensitive to such matters (e.g., "people from whatavashire actually mean 'a playful exchange' when they say 'war'," or "only if a person from the somewhere-isles threatens war, should you worry that they actually intend to deliver"), the two "r" sounds will form different signifiers belonging to different signs. This may be even be the case where the two variants sound the same to any human ear, but can be distinguished by some sensitive voice analysis instrument!

The point is that commutation tests are always context sensitive. Tests that rely on inter-linguistic translation depend on the kind of translation we are looking for and its purposes (legal? poetic?). Tests that consider intra-linguistic relations between signs are also never perfectly determined. I can claim that the word "dog" in the sentence "I see a dog" is commutable with the word "greenly," if my corpus includes what is usually referred to as non-grammatical sentences (say, in poetry, or in a recording of a language learning class, or in a corpus of examples of grammatically deviant sentences).

The same sensitivity applies to the analyses of mathematical examples, as in the case of 0 above. In fact, throughout the analysis, I pretended that I could zero in on a limited portion of mathematical language (such as positive decimal integer arithmetic). But in real life, we use mathematics in a rather fluid manner, considering at the same time various representations and uses. I may mix up notations such "504" and the spoken word "five hundred and four" (where 0 is absent); I may combine standard calculation algorithms with some ad-hoc maneuvers (using various tricks); I may apply numerological associations, and may even throw in some non-mathematical statements ("zero is like the nothingness of being"). Which manipulations and representations are and are not relevant when considering commutation tests is thus precisely a question of context.

However, the problem is actually worse. Even if one pins down a context, it is usually impossible to impose it on one's research. In a linguistics lab, there are all sorts of tricks for suppressing all sorts of semiotic contexts that one would like to exclude, but these are only ever partially successful, and there might always be some unintended context lurking in the subject's mind that the researcher neglected to suppress. In a historical setting, your sources are safe from the intervention of modern notations introduced in a later period, but you cannot always rule out practices and representations that leave little or no material evidence (e.g., fingers calculation or drawing lines in the sand), and it's terribly difficult to suppress anachronistic notations in the historian's head (or notepad). All these uncertainties can affect the intended context and outcomes of commutation tests.

In fact, this indeterminacy of context is unavoidable not only in practice, but also in principle. As Derrida observed, "by virtue of its essential iterability, a written syntagma can always be detached from the chain in which it is inserted or given without causing it to lose all possibility of functioning. . .. No context can entirely enclose it" (Derrida 1988, 9). The condition of possibility for a sign to work as a sign

is that it can be extracted from its context and copied somewhere else. Something entirely singular and immobile, like Stonehenge, is not a sign, because it cannot be extracted and commuted for other signs. However, we did say that everything is *also* a sign. This is the case to the extent that Stonehenge is copied, reproduced, photographed, painted, disseminated. The possibility of copying (a possibility that is there even before the sign is actually copied, when one only "copies" its image in one's mind) turns it into a sign that can be commuted and can attain a semiotic value; and Stonehenge is indeed used as a symbol of certain cultures. Therefore, considering something as a sign amounts to the ability to take it out of its singular context.

This entails that attempting to establish a precise context is paradoxical. When one says "consider the use of 0 only in arithmetical positive integer decimal place value notation," one assumes that one could, at least in principle, isolate this context. But the signs used *in* this context have a long history of – indeed depend on, by the fact that they are signs – being taken out of context. When using 0 in *this* context, I am using it *as something transferrable between* contexts (language learning, math schooling, a zero-shaped mirror on the wall, hearing someone being called "a big dumb zero"). The context is, therefore, not really closed, but connects with all those contexts that are involved in the fact that the sign is a sign because it can be taken out of/into them.

Even when learning a totally new sign, the context of learning is not entirely identical to the context of application, so already then more than one context is involved. Even recognizing a new sign *as a sign* assumes a certain de-contextualization for this very act of recognition ("Oh, it is a sign, like those letters I learned about last year"). To consider a sign *only* in a given context is to consider it as less than a fully functional – that is iterable – sign. The ideal sign depends on actual and potential context-crossing for its existence, so even an ideal context hosts signs that, by the fact that they are signs, echo foreign contexts, which might affect the semiotic analysis. The results of any commutation test, in any context, are therefore always somewhat open, even under idealized circumstances.

In order to try to pinpoint a specific context, one can stipulate rules of usage, and exclude any commutation test where these rules are violated. However, even in those (rare and usually artificial situations) where precise, complete rules can be stipulated, they must be stipulated in a natural language, where it is difficult to totally exclude new interpretations of the rules. Moreover, using the same signs in new or unexpected circumstances may require a new interpretation of the previously given rules – as Wittgenstein affirmed, rules are not really "rails invisibly laid to infinity" (Wittgenstein 2009, §218).

This is the principle of post-structural semiotics: *contexts are never completely determined; therefore, commutation tests are always at least somewhat open ended*.[11]

[11] Some scholars consider post-structuralism as a reaction against structuralism. I tend, like others (e.g., Maniglier 2006) to see it as a continuation and acknowledge that some of the principles most strongly identified with post-structuralism are already there in earlier structuralist thought. What changed is, I think, mostly the emphasis and goals, not the basic toolkit.

This is not a bug or a weakness of the theory. This openness is a necessary condition for our possibility to adapt our signs to a changing world. Trying to bracket, ignore or suppress this principle does not lead to a better scientific understanding of the meaning of signs, but to an essential misunderstanding. If a semiotic analysis provides a completely decisive answer, then you should know that it is wrong. A good semiotic analysis should follow the signs as they change, and study their potential for change as contexts flow and transform, rather than simply determine their meanings (for mathematical examples, see Herreman (2001) and Rotman (1987) in the analysis of 0; for other examples see Grosholz 2007; Wagner 2017, chs. 3–4). To understand 0 is not simply to decide whether, in a given context, it is a place holder, a quantitative value or a formal entity; rather it is to understand how these different roles interact across the contexts that should contain them, leading to the emergence of new meanings of 0. What may appear at first sight as a weakness or even failure of the theory, is actually its strength.

This opens up two further dimensions for discussion: materiality and power. As we said, signs can always, by the fact of being signs, be taken out of context. This means that their contexts and meanings are always somewhat open. But this does not mean that they always actually are taken out of any specific context. They may be harder or easier to carry (clay tablet vs. paper); they may survive for a longer or shorter time (clay tablet vs. lines in the sand); they may facilitate or hinder various context-transitions (pen and paper calculation facilitates moving from the context of calculation to that of verification, whereas calculating with tokens on a counting board does not).

Moreover, carrying signs between contexts is not a solipsistic, perfectly voluntary venture. Physical and metaphorical walls may be erected to prevent taking signs out of their intended contexts. Prisons may be built for people who use certain signs in unorthodox ways, or such people may form pacts of secrecy. A reality of power relations can slow or expedite the way signs move between certain contexts. A semiotic analysis is therefore not an extraction of the pure meaning or unrestricted possibilities of signs, but of the meaning and possibilities of signs in material contexts where power relations apply.

Material contexts and power relations also enable to impose some objectivity on what may seem like a theory leading to purely relativistic "anything goes." Indeed, when Wittgenstein wrote that rules were not "rails invisibly laid to infinity" (Wittgenstein 2009, §218), he meant that no amount of explanations and examples could preclude new or unorthodox interpretations of even the simplest of rules. He strongly acknowledged, however, that in some circumstances, only few explanations and examples are enough to get a large bunch of people to agree for a long while on the right way to follow a rule. This, according to him, was not due to some metaphysical conditions, but to concrete practical-cognitive circumstances (e.g., Wittgenstein 1976, 168), and, we might add, concrete power relations.

The stability or reproducibility of a semiotic analysis is, therefore, a practical issue. It is never absolute, but it can sometimes be effectively achieved. Semiotic analysis, then, even in a post-structuralist setting, can be objective, provided we understand objectivity in a practical and moderate sense. One way is to understand objectivity as the effective reproducibility of the results of an analysis among a

community of critical but well-trained experts, while acknowledging that what an objective judgment should capture changes across time and scientific contexts (Daston and Galison 2010). Alternatively, and more in line with the ethical goals of post-structuralism, we may consider objectivity as something that is achieved only in a responsible mediation of the analyses performed by a variety of relevant stakeholders – especially those who will suffer the most from the consequences of the analysis (e.g., Harraway 1988; Harding 1992; Barad 2007).

5 Semiotics as Ontology

Classically, the debates concerning the ontology of mathematics have revolved around what kind of objects serve as mathematical *referents*. This is a tricky issue, because the referents of mathematics have to be rigid enough to account for the stability and precision of mathematical knowledge and at the same time be accessible enough to account for our success in knowing and applying mathematics. If mathematical objects live in some independent realm of abstract, ideal objects, it is not clear how we know about them and why they are useful in describing the real world; if they exist in empirical reality, then they are subject to the kind of change and decay that would clash with the stability and precision of mathematical knowledge. There is a long history of attempts to find middle grounds. This middle ground could be found in logic, model theoretic structures, constructs of (pure Kantian) intuition, psycho-cognitive objects, or simply written symbols, to quote the most prominent approaches (cf. Fried's chapter in this handbook, ▶ "Ontology in the History and Philosophy of Mathematical Practice: An Introduction"). However, some philosophers completely reject the project of trying to determine object-like referents for mathematics.

A prominent instance is Wittgenstein. Wittgenstein has repeatedly insisted that talking about something in the same way as we talk about empirical objects does not always mean that it makes sense to think of them as actual objects. One example is pain (as discussed across Wittgenstein 2009). A common approach to pain is to try to locate it physically (in the body, in the brain) and come up with objective ways of observing it (by medical imaging devices). However, this is a notoriously difficult task – pain is extremely elusive. Private experiences, such as pain, cannot live up to the standard governing the discussion of inter-subjective empirical objects. It is therefore best, according to Wittgenstein, not to try to force them into the ontological template of such objects. Instead, when we are talking about pain, we should see ourselves as engaging in normative communication of demands and expectations ("he is in pain, try not to bother him"), rather than in describing or asserting the presence of some object.[12]

[12] Note, however, Wittgenstein's statement: "'But you will surely admit that there is a difference between pain behaviour with pain and pain-behaviour without pain.' – Admit it? What greater difference could there be?" (Wittgenstein 2009, §304). Wittgenstein does not deny the private experience, he only objects to its object-ification.

Wittgenstein follows a similar path in the context of mathematics. Instead of trying to figure out what kind of object "2" or a "triangle" or a "Hilbert space" is, he insists that mathematical statements impose grammar-like normative rules on descriptions without implying a reference to any real or ideal object (I will not try to justify this approach here; for a summary see Wagner 2020). My proposal is to follow Wittgenstein in denying mathematics an objective referent, but to complement his grammatical-normative approach by collapsing the sign-referent or tool-object distinction. I suggest we articulate *the ontology of mathematics as semiotic*, namely, as a collection of co-articulated mathematical signs (where, recall, everything, not just formal symbols, is *also* a sign), without an external, non-signifying referent.

Before I explain what a semiotic ontology means, let me contrast it with another proposal, which may sound similar: the formalist position, which claims that mathematics is about symbols in formal systems. According to this approach, a mathematical statement simply claims that there is a succession of symbols (a formal proof) that starts with the given axioms and hypotheses, and, following the given rules of inference, ends with the stipulated statement. So, mathematics is a body of claims about symbols. Saying, however, that mathematics is made of signs, in the semiotic sense of the term, is quite different.

First, a Saussurian sign is formed by the mutual co-articulation of symbolic material (perceived sounds or inscriptions) and thought-experience (concepts), so it is at least as cognitive as symbolic. If one prefers to discuss signs that are co-articulated against other material realities (like water and air co-articulate waves in de Saussure's metaphor) or against some ideal realm, then signs would involve the real or the ideal as well. So talking about signs does not necessarily reduce the discussion to the level of formal mathematical symbols.

Second, formalism assumes that symbols are given as such, and that, in mathematics, they are manipulated only according to the rules of a specific given context: a formal language with its axioms and rules of inference. We have already seen that this has nothing to do with the semiotic understanding of signs. Semiotically, signs are never given as such, but come into existence via the co-articulation of two or more streams, and cannot be rigidly restricted to a single context, not even in principle. It is never just a set of formal rules that determines the value of signs. The values of signs carry with them the echoes of the contexts through which these signs were and are still shaped.

Recall further that we claimed that everything is *also* a sign. The positions of planets are signs, to the extent that they enter into commutation relations with other signs (drawings, tables or astrological charts; using the naked eye or precision instruments). Food divided between workers is a sign, to the extent that the different shares can be commuted between workers or for equivalent commodities. This is precisely what happens when the positions of stars or the food allocated to workers are embedded in mathematical practice. We can therefore think about mathematics as a system of translations between various semiotic modalities (commodities, planet positions, geometric diagrams, symbolic calculations). A semiotic ontology of mathematics, then, does not reduce the scope or purport of mathematics to a limited formal subset. Instead, it covers everything that is (1) considered according to its

semiotic value and (2) is "mathematized" in the sense of being co-articulated with respect to the signs that we recognize as belonging to the core of mathematics. This core, we should emphasize, is historically and culturally contingent.

To say that the ontology of mathematics is semiotic is to say that mathematics is a collection of signs whose values are captured by the regularities in the practices of their commutation[13] – regularities that are embedded in materialities and power relations and are never completely determined. Since the regularities that define sign values are those witnessed in *practices* of commutation (rather than in some arbitrary speculation about them), a philosophy of mathematics grounded in semiotics is a *philosophy of mathematical practice*.

To say that the ontology of mathematics is semiotic also displaces mathematical concepts. From that point of view, concepts (e.g., the concept of a number as a multitude of units, as an intuitive succession, as a Fregean set, or as a symbol controlled by the axioms of Peano arithmetic) no longer dominate mathematics, but are co-articulated with practices involving the relevant mathematical signs. Instead of telling stories about how changes in mathematical concepts change mathematics, we recall that mathematical concepts have material-signifying aspects (formal or informal definitions, practices that express these concepts) and are therefore one more link in the net of co-articulated signs that is the ontology of mathematics. They are no longer *the essence* of mathematics.

The distinction between this and more traditional kinds of ontology can be highlighted by a reference to Rabouin (2017). Rabouin shows that a continuity of "style" (used to denote specific linguistic-symbolic practices, such as the classical Greek geometric style, the Cartesian style of so-called analytic geometry, the Weierstrassian style of the calculus, or the modern formal-axiomatic style) does not necessarily align with a continuity of concepts or of traditional epistemic-ontological positions. People may use the same style while making very different explicit statements on how we know mathematics and what it is about.

While a traditional ontology may aim to extract the referents of mathematics from its style, or, oppositely, to capture them by abstracting the style away, semiotics as ontology captures both the style of mathematics and its practical relations to implicit and explicit, theoretical or practical, philosophical articulations. These are all part of a connected network of correlated signs with more and less stable components. Instead of stating that mathematics is about real, ideal, pure-intuitive, logical, formal *or* other objects, we ask how mathematical signs are co-articulated with respect to reality, ideality, pure intuition, logic *and* formal systems. Instead of deciding between classical ontological proposals, we study how they relate to each other in practice via sign articulations.[14]

[13] It might sound more sensible to characterize mathematical signs in terms of operations rather than commutations, but, as we noted above, an operation is something that allows commuting the array of signs that forms the operation and the array of signs that forms the result of the operation. Moreover, this is not the only relevant kind of commutation, as we saw in the examples above.

[14] Exemplary analysis in non-European contexts, which relate the semiotic-material and conceptual aspects of mathematics include Chemla (2018, 2020) and Keller (2015).

Put this way, a semiotic ontology for mathematics allies with approaches that are already established in the context of science studies. Latour (1999, in the chapter *Circulating Reference*) describes scientific practice as forming chains of translations from one materiality to another, relating them by semiotic co-articulation, making specific manipulations (or commutations) on one side of the signifying-matter chain correspond to specific differences on the other. If science is understood entirely in terms of such co-articulations (which may be more than Latour would assert), then it can be captured by a semiotic ontology (for a mathematical case study, to which I owe this connection with Latour, see Steensen et al. forthcoming; for my elaboration of mathematical abstraction in this framework see Wagner 2019).[15]

Note that this perspective provides the term "reference" with a new meaning. In such a framework, to say that a scientific representation "refers" to some natural phenomenon is to establish rigorous enough chain of co-articulations that links one signifying materiality to another in a stable manner across the chain. In this framework, it is the stability, mobility and replicability of such chains that would render the reference "objective." Reference then circulates in a chain of signs, rather than connects signs to an ontologically distinct referent.

A related approach from a different philosophical tradition is that of Nancy Cartwright (2019, part I), who asserts that the formal "laws of nature" are never successfully universal. Rather, real contexts always require skillful tinkering with material artifacts and contextual adaptations to match real situations with laws (whereas laws that are designed to require no tinkering and no exceptions end up having an empty scope). This skillful material tinkering is something that a semiotic ontology of mathematics takes seriously – mathematics is the skillful, contingent, and context dependent fitting together of various kinds of signs, including, but not restricted to, those that belong to formal systems.[16] Our knowledge of mathematics and its applicability then derive from our involvement in crafting its formal and informal signs from the materiality of the real world (this raises the question of the stability of mathematics, which I defer to the conclusion).

The last few paragraphs show that semiotics as ontology is not necessarily unique to mathematics. However, reducing the ontology of empirical sciences to semiotics may raise objections, citing the commitment of empirical sciences to something beyond signs (or more precisely, to things not seen *only* as signs). I think that mathematics, however, can embrace the semiotic ontology proposal more readily, even without the hard work carried out in attempting to counter the former objection.[17] Even keenly empiricist views of mathematics often view mathematics as the *language*

[15] This convergence between Latour and semiotics is no accident – Latour's thought emerged, at least in part, in the context of French structuralist and post-structuralist semiotics.

[16] Isabelle Stengers (2010–2011) can be read as (but should not be reduced to) someone who incorporates Latour's and Cartwright's approaches, dwelling on the various weaknesses of scientific chains of reference without thereby dismissing them.

[17] Karen Barad (2007) goes farthest in attempting to design a full-fledged feminist post-structuralist ontology that insists on the inseparability of materiality and semiotics, and on the fact that materiality is given objectively only in semiotic co-articulations.

in which the book of nature is written. A semiotic ontology of mathematics gives both signification and materiality the respect they deserve in grounding mathematics, while focusing on the relationality that mathematics is usually considered as capturing.

6 Examples for Semiotics as Ontology

Let's move on to some mathematical examples of semiotics as ontology. The author whose work I will appropriate for providing examples of semiotic ontologies of mathematics is Reviel Netz (see his ▶ "Shaping, Revisited" in this handbook for a brief overview), even though he might have some objections to this cooptation of his work.[18] I will follow his analysis of two domains: Greek arithmetic and Greek geometry.

In *Counter Culture* (2002) Netz explains that Greek arithmetic is about counters. One of the written Greek representations of numbers is similar to the well-known Roman numeral representation, using Greek rather than Latin letters. This representation, in fact, reflects the primary representation of numbers in Greek culture: as counters on a counting board. A (Greek equivalent of a) number such as CCXXXVIIII is a description of the following situation. You have a surface divided into columns. The rightmost columns is that of units, and has four counters. The next column is the fives, and contains one counter. The next is the ten, and contains three counters. The next column, that of fifties is empty. The following column that of hundreds has two counters in it. The subsequent columns to the left (if there are any) are empty.

This notation of numbers was superseded later by a different alphabetic number representation, where the numbers one through nine, multiples of these numbers by ten, and multiples of these numbers by a hundred each had its own Greek letter. Calculation by means of counters on a counting board, however, remained the most salient means for performing non-mental calculations. Calculations (from the Latin *calculus*, or little stone – a counter!) is literally about moving counters around. Adding is grouping the counters of the added numbers together, replacing large groups by a counter in the next column. Multiplication is literally about replacing each counter by many or performing repeated addition, combined with position shifts. Subtraction and division are just as concrete. All the difficulties supposedly associated with calculations involving Roman (or Greek) numerals turn into intuitive manipulations on the counting board.

Such counter-calculations prevailed in many places in the world before and after the Greeks. The relations between counters and symbolic numeration goes at least as far back as ancient Mesopotamia (Robson 2008; Overmann 2019), and the Chinese rod numerals (Katz 2007, 194–201) are a variant of the same practice. But if we are to take semiotics as ontology seriously, we should consider the specific material

[18] In terms of Netz's paper, the scope of "semiotic" here covers both his "structural" and "material," so this is not the problem. But the commitment to a semiotic ontology for mathematics might still invoke other kinds of objections.

commutations as actualized in Greek cultural practices, rather than suggest an abstract universalization.

Where else do we find counters in Greek culture? First and foremost, Netz explains, in the newly created coin based economy, where counters (coins) are used for trade (but also in religious rituals, such as placing obols on the eyes of the deceased). Counters played a role in political and legal life as well. The identification of people, the drawing of lots for serving as jury or representatives, the votes in courts – all involved various kinds of tokens or counters. But counters were also used in games, and in playful theoretic manipulations, such as the Pythagorean geometric representation of numbers (triangular numbers, square numbers, etc.), based on arranging a given number of counters into a geometric shape.

This review is not meant to suggest a restriction of the ontology of Greek arithmetic to counters. Semiotic relations involving arithmetic operations, whether by means of counters or one of the written number representations, extended to many other contexts (architecture, engineering, various crafts). The co-articulations between the core arithmetic representations (counters, written numbers) and other signs (in economy, politics or crafts) establish cross-domain commutation relations, which all belong to the semiotic network that forms the ontology of Greek arithmetic. Indeed, if we take seriously the reciprocity of co-articulation, then anything that is co-articulated with arithmetic representations participates in the determination of the values of arithmetic signs, and is therefore part of the ontology of Greek arithmetic.

The prevalence of counters in specific social, economic and cultural domains in Greece facilitated the co-articulation between the latter and counter-calculations. This means that these domains of Greek life were more strongly "arithmetized" than they might have been otherwise. A curious hiatus, Netz emphasizes, occurs between record keeping and calculation. Since the letter representations of numbers, which dominated book keeping, are not amenable to calculation algorithms, the medium of calculations (counters on a counting board) and the medium of record keeping (letters inscribed on stone or papyrus), were different. This does not mean that record keeping was not arithmetical, but that a material barrier weakened the ability of signs to move between these two contexts, forming a weaker semiotic integration of calculation and record keeping than we would expect today, when digital or written numerals serve both purposes (for more on Greek record keeping see Cuomo 2013).

The ontology of mathematics changes dramatically, when the medium of calculation changes to place-value numerals on erasable dust-boards and later still on paper. I will not go into the details of the relevant Indian, Arabic or European systems, I only note that this change of medium weakens the relation between calculation and the actual manipulation of counters or tokens. The numeral loses its immediate relation to a collection of units ("4" is no longer primarily four tokens on the units column), and turns into an element that might participate in different kinds of combinatorial commutation relations (symbolic arithmetic). This opens up a path toward what is usually described as a more formalized kind of mathematics (once again, see Herreman 2001).

When we think about a number today, our immediate association is no longer the feeling of counters between our fingers or on a board, but a visual numerical symbol.

This transition echoes Jacob Klein's (1968; for a recent overview see Baratelli 2021) famous thesis on the difference between Greek known and unknown numbers and early modern algebraic ones. The former, according to Klein, refer to collection of units, probably imagined as counters; the latter refer to the symbolized concept of a specific or general quantity, such as three-ness, confounded with an instance of this concept, such as three things. But instead of restricting ontology to such presumed *referents*, a semiotic ontology would focus on the practices whereby sign-chains weaken or strengthen links between numerals, tokens, collections of objects and terms/concepts such as "three-ness." Indeed, both Rotman (1987) and Herreman (2001) offer such semiotic takes with interesting possible relations to Klein's thesis.

We further note that the transition to written place-value numerals places paradigmatic arithmetical signs on paper and in books, at the hands of the bookkeeper, rather than on counting boards, at the hands of the merchant. The relations between this shift and the rise of capitalism have been tentatively explored by several authors (e.g., Swetz 1987; Hadden 1994). The semiotic-ontological shift may thus be correlated with changes in social structures and the distribution of economic power.

I will briefly consider another example, that of geometry. If we consider semiotics as ontology, then, as established by Netz (1999), classical Greek geometry was organized around very strict signifying material: the geometric language and the diagrams. The language consisted of very restricted vocabulary, repeated combinations of words (which Netz calls "formulas," but are not to be confused with modern algebraic formulas) and inferential transitions. The diagrams were poorly drawn sketches, but (outside counterfactual situations) strictly differentiated straight and curved lines and were true to topological relations such as inside/outside and intersecting/non-intersecting. Letters indexed diagrammatic objects of interest, binding diagram and text. While in the philosophy of mathematics the Platonic and Aristotelian ontologies are often discussed, many Greek mathematicians (especially in the generation of Archimedes) were uninterested in philosophy – they didn't need an ontology beyond their carefully regulated diagrams and texts: "the diagram . . . acted, effectively, as a *substitute* for ontology" (Netz 1999, 54–57; for the changing attitudes of mathematicians to philosophy, see Netz 2022).

Of course, geometry was involved in various other sciences (like astronomy and optics) and practical applications (like architecture and engineering), but the strict language and specific way of using diagrams helped segregate the scholarly mathematical context from other contexts, and slowed down (but not closed off) the communication and co-articulation between the scholarly geometry and more applied contexts. This seems to have been a deliberate choice: the aristocratic creators of classical Greek geometry deliberately wanted to segregate themselves from the lower class people who made a living by means of practical geometry. In particular, in scholarly classical geometry, geometric objects were never measured by numbers (although the ratios between geometric magnitudes could equal ratios between numbers), and no counter calculations were involved (Netz 1999, ch. 6; Asper 2008 – in Euclidean-style number theory, even numbers were represented by lines, rather than tokens). The classical Greek geometers wanted to isolate an aristocratic, exceptionally stable truth, applied to their idiosyncratic realm of signs

(Netz 1999, ch. 7, Sect. 3). The zealous warding of their distinct semiotic context was an expression of power.

Other pre-modern and modern geometrical cultures, however, depended on different semiotic systems. By the eighteenth century, the classical Greek rigid linguistic structure had lost its footing and was mingled together with algebraic representations (analytic geometry). Classical diagrams were used interchangeably with new projection based graphical representations (in mapping, art, projective geometry and descriptive geometry). The latter representations could interchange straight and curved lines and distort topological relations (e.g., parallel lines intersecting at the vanishing point on the canvas or paper). Moreover, early modern and modern mathematicians, associated with a sort of intellectual or professional bourgeoisie rather than an aristocracy, were no longer as interested in segregating theory from practice, and integrated numerical measurement with geometric arguments.

The semiotic ontology of geometry changed from a co-articulation of classical language and classical diagrams that capture the properties of specific shapes to a co-articulation of numerical calculations, loosely regimented geometrical languages, algebraic representations, topologically diverse pictorial representations, and practical contexts where space itself was "sign-ified" or articulated by measurements (De Risi 2016a). A more classical view would focus on the "conceptual change" or change in the ontological referent: from shapes to space. Here, however, I would like to view the process as a change in the entire semiotic system of geometry, where referents (shapes or space) form only a part – not necessarily a dominant part – of the system of co-articulated representations that formed around geometry.

What were the implications of this change? I claim that this ontological shift was substantial enough to change the status of the parallel postulate[19] from "true but difficult (although probably not impossible) to prove" to "not necessarily true" (cf. Manders' analysis as reviewed in Waszek's chapter in this handbook, ▶ "Signs as a Theme in the Philosophy of Mathematical Practice"). To understand this claim, let's start with the classical conception of the parallel postulate. I should first clarify that classical Greek geometry was not axiomatic in the modern sense. It did, of course, have axioms (or more precisely, *common notions* and *postulates*), but there was never a stable list of axioms (De Risi 2016b), and the axioms did not play the same role as axioms in a modern formal theory. Indeed, they were never meant to control all possible inferences. Many inferences in Euclid's *Elements* derive from observing the topological properties of diagrams, and others are built into the rigid language of classical Greek geometry without an axiomatic underpinning. The axioms capture only some of the necessary principles or rules of inference. Moreover, attempts to prove axioms were quite common, and not restricted to the case of the parallel postulate.

[19] In its original form (postulate V of the Euclid's *Elements*) or in any of its modern forms, e.g.: for any straight line and any point outside it, a unique parallel to this line passes through this point.

An example of an inference that was not grounded in axioms is the ability to move a shape while preserving angles and straight lines, which is implicit in Euclid's *Elements* I.4 (the first triangle congruence proposition). Similar maneuvers came up in many of the attempted proofs of the parallel postulate (e.g., in Ibn al-Haytham's proof, see Rosenfeld 1988, 59–64). Standard historiography says that such purported proofs relied on unjustified implicit assumptions. This is true, but this is just as true for many proofs of Euclid himself! There are no unambiguous, non-anachronistic criteria for deciding whether Ibn al-Haytham's assumptions about straight-line and angle preserving motion follow or depart from those embedded into Euclid's proof of proposition I.4 (or, in other words, if one is justified in commuting one for the other within the context of the semiotic articulations of Euclidean geometry).

So, it is wrong to say that all past proofs of the parallel postulate are wrong – some of them were just as ambiguously justified as Euclid's own propositions. The persistent, unresolved disputes around this issue, which spanned over two millennia, demonstrate this fact. Instead of saying that the proofs were wrong, it would be less anachronistic to say that the semiotic articulations that specified what counts as a valid proof in classical geometry were insufficient for deciding whether such proofs should be accepted. The kind of semiotic articulations that enable a clear-cut decision on these matters (which depend on some kinds of syntactically formalized language) belongs to the nineteenth and twentieth centuries, and cannot be projected backwards. Indeed, formal axiomatized geometry depends on a very different semiotic network, where language and syntax have a dominant role, and this recent semiotic system is not the only possible historical continuation of the classical geometry that gave rise to the problem. Nevertheless, regardless of the disputes concerning *proofs* of the parallel postulate, we have no hints of a doubt as to the *truth* of the parallel postulate well into the eighteenth century. While the existence of a valid proof of the parallel postulate was a contentious issue, the truth of the postulate was evident to all – it was visually obvious.

What happened in the nineteenth century that allowed this attitude to change, even *before* formal axiomatized geometry became a dominant representation? I believe that this has to do with the semiotic-ontology shift described above. I will follow here Hodgkin's (2005, ch. 8) simple but effective summary of the main secondary literature, with some support from Rodin (2010). They both explain how Lobachevsky argued for the possibility of a non-Euclidean hyperbolic geometry, where parallels to a given line through a given point are not unique.

First, the very algebraization of geometry (stemming from Viète's and Descartes' so-called analytic geometry, capturing geometric objects and relations by algebraic ones) was a preliminary condition. Second, the new functional analysis (especially Euler's mid-eighteenth century treatment of exponentials) allowed mathematical descriptions that were crucial for the quantitative analysis of Hyperbolic geometry but were not available before. Indeed, Lobachevsky's claim that hyperbolic geometry was consistent was based on a sort of "isomorphism" between the functional equations that capture spherical trigonometry and those that capture hyperbolic trigonometry. Lobachevsky argued that since the former is consistent and

"isomorphic" to the latter, and since these equations capture the entirety of geometric relations, the latter must also be consistent.

But this was not enough for a wide acceptance of the new geometry. A crucial ingredient was the use of topologically and linearly distortive diagrammatic representations of geometry. Whereas in the classical Greek system, diagrammatic practice did not allow drawing a curved line for a straight one, the new cartographic and artistic geometric projections could commute one for the another. This meant that a straight line no longer had to *appear* straight, while a curved line could *appear* straight. Moreover, after the integration of numerical measurement practices into the semiotic network of theoretical geometry and the establishment of the nineteenth century's empirical approach to measurement (reflected paradigmatically by Gaussian geodesic science, which was re-articulated as a definition of spatiality by Riemann), it was no longer its apparent shape, unstable under a change of representation, that determined the straightness of a line. Instead, the straightness of a line was re-articulated with respect to practices of measurement. New semiotic co-articulations affected the ontology of straight lines.

For the Greeks, an arc of a great circle on a sphere could never be conceived as a straight line, simply because it was curved, and straight and curved lines were not commutable.[20] By the late nineteenth century, however, such an arc could be referred to as a straight line in spherical geometry. Indeed, the new semiotic ontology allowed a straight line to be straight even if it clearly looked curved, as long as it spanned the shortest distance between its ends according to a well-defined practice of measurement. The models of hyperbolic and elliptic geometry (e.g., Beltrami's model), which were taken as proofs of the consistency of the new geometries, depended on such curved representation of so-called straight lines.

The transition from the classical semiotic-ontology, which focused on rigid language and a specific practice of diagrams, and which made an effort to segregate itself from other semiotic contexts, made the parallel postulate obviously true, even if it was difficult to agree on its proof. The modern ontology, which included algebra, new diagrammatic representations, and empirical measurement, gave up on the rigidness of language and classical diagrammatic standards. The new representations turned the parallel postulate into a falsifiable claim. The Greeks would reject these developments because they were simply unacceptable in terms of their semiotic system. They depended on something different – not merely a different concept, but a different network of co-articulated signs: a different semiotic ontology. The co-articulation of the new geometry in terms of formal axiomatics added an extra layer that solidified this development, rather than form its ground or origin.

[20] There are other arguments for claiming that arcs of great circle on a sphere are not straight line segments – for example, the fact that two such segments may intersect in two points. But the fact is that the question was never even raised (if it had been raised, some creative tinkering might have resolved it). The above argument did not need to arise, because considering an arc as a straight line, as far as I can tell, was simply out of the question.

7 Conclusion

I hope that the above examples clarify what it means to consider semiotics as an ontology of mathematics. Mathematics is thereby conceived as the practice of co-articulating certain kinds of signs (inscriptions on paper, cultural artifacts, natural-phenomena-as-signs) with respect to each other. Given such an ontological view, the philosophical analysis would turn to the analysis of the values of signs in various contexts, their limitations, and potentials. The philosophical normative analysis would analyze how different mathematical goals are served or dis-served by various semiotic ontologies with their integrations and segregations. But the analysis would also recall that sign values are always bound with materiality and power relations. It would further assert that nothing can guarantee the purity of semiotic contexts, and therefore nothing can guarantee the perfect stability of mathematics, although various measures can be taken to strengthen or weaken it. The resulting philosophy is a philosophy of practices that value and re-evaluate signs, which stabilize or de-stabilize the semiotic systems of mathematics (those that I elsewhere referred to as constraints; see Wagner 2017, ch. 3). This philosophy analyzes and evaluates the impact of mathematical practice as a collection of semiotic, material, power-laden maneuvers.

As a philosophy of mathematics, semiotics provides a lot of flexibility for formulating solutions to philosophical challenges. Both the ontology and epistemology of mathematics become diverse and distributed, and lose their awesome mystery. The price that we pay, however, is a loss of stability and a threat to the rigor of mathematics. Indeed, how can inherently unstable semiotic systems serve as ontological grounds to something as rigorous and stable as mathematics?

My answer is that the stability of mathematics tends to be exaggerated, as witnessed by the example of the parallel postulate, where ambiguities are covered over by the anachronistic claim that all historical proofs of the parallel postulate were simply false (for a more general discussion of such ambiguities in mathematics, see Wagner 2022). Since the beginning of the twentieth century, we indeed have very stable and rigorous semiotic systems – the so-called formal systems. Earlier, we had other, relatively stable systems, like the language and diagrams of classical geometry, which were nevertheless not as stable as modern formal systems. We also had notoriously unstable systems, like early infinitesimals. Practices that stabilized the co-articulations between very stable semiotic systems and less stable mathematical semiotic systems,[21] practices that did not only consolidate, but actually changed mathematics (as argued in the latter sections of Wagner 2022), strengthened the stability of mathematics, making it a paragon consensus, at least in some of its aspects.

But mathematics is still a network of many co-articulated systems, and formal mathematics does not have an absolute hold over the entire field (Rav 2007). New

[21] This process spanned much of the twentieth century and continues, for example, in the co-articulation of computer verified proofs and contemporary mathematical practice, as discussed in Avigad and Harrison (2014).

semiotic articulations and attempted context segregations help render mathematics more stable, but cannot make it entirely so. The question whether a certain formal system is the right one for doing a certain kind of mathematics, the competition between different kinds of formal-foundational systems, and the fact that formal systems are hardly ever the tool for producing new mathematics show that such systems are just components in the semiotic ontology of mathematics. They are more or less successfully co-articulated with other semiotic systems that form mathematics, but are not, in any way, the foundation of mathematics.

8 Cross-References

- ▶ Ontology in the History and Philosophy of Mathematical Practice: An Introduction
- ▶ Signs as a Theme in the Philosophy of Mathematical Practice
- ▶ Symbolic Algebra as a Semiotic System

References

Asper M (2008) The two cultures of mathematics in ancient Greece. In: Robson E, Stedall J (eds) The Oxford handbook of the history of mathematics. Oxford University Press, Oxford, pp 107–132

Avigad J, Harrison J (2014) Formally verified mathematics. Commun ACM 57(4):66–75. https://doi.org/10.1145/2591012

Barad K (2007) Meeting the universe halfway: quantum physics and the entanglement of matter and meaning. Duke University Press Books, Durham

Baratelli G (2021) How letters become symbols Jacob Klein's genealogy of formalization. In: The new yearbook for phenomenology and phenomenological philosophy, vol 18. Routledge, London, pp 407–445

Barton B (2008) The language of mathematics: telling mathematical Tales. Springer, New York

Boyer CB (1959) The history of the calculus and its conceptual development. Dover, New York. http://archive.org/details/TheHistoryOfTheCalculusAndItsConceptualDevelopment

Brown T (2001) Mathematics education and language: interpreting hermeneutics and post-structuralism. Kluwer, Dodrecht

Cartwright N (2019) Nature, the artful modeler: lectures on laws, science, how nature arranges the world and how we can arrange it better. Open Court Publishing, Chicago

Chemla K (2018) How has one, and how could have one, approached the diversity of mathematical cultures? In: Mehrmann V, Skutella M (eds) Proceedings of the 7th European congress of mathematics 2016. EMS Press, Berlin, pp 1–61

Chemla K (2020) Different clusters of texts from ancient China, different mathematical ontologies. In: Lloyd GER, Vilaça A (eds) Science in the forest, science in the past. HAU Books, Chicago, pp 121–146. https://library.oapen.org/handle/20.500.12657/47354

Coquard J-M (2021) Stevin's mathesis and number. In: The new yearbook for phenomenology and phenomenological philosophy, vol 18. Routledge, London, pp 459–487

Cuomo S (2013) Accounts, numeracy and democracy in classical Athens. In: Asper M (ed) Writing science: medical and mathematical authorship in ancient Greece. De Gruyter, Berlin, pp 255–278. https://doi.org/10.1515/9783110295122.255

Daston L, Galison P (2010) Objectivity. Zone Books, New York

De Risi V (2016a) Francesco Patrizi and the new geometry of space. In: Vermeir K, Regier J (eds) Boundaries, extents and circulations: space and spatiality in early modern natural philosophy.

Studies in history and philosophy of science. Springer International Publishing, Cham, pp 55–106. https://doi.org/10.1007/978-3-319-41075-3_3
De Risi V (2016b) The development of Euclidean axiomatics. Arch Hist Exact Sci 70(6):591–676. https://doi.org/10.1007/s00407-015-0173-9
de Saussure F (2011) Course in general linguistics (eds: Meisel P, Saussy H, trans: Wade B). Columbia University Press, New York
Derrida J (1988) Limited Inc. Northwestern University Press, Evanston
Ernest P (2006) A semiotic perspective of mathematical activity: the case of number. Educ Stud Math 61(1/2):67–101
Ernest P (2018) A semiotic theory of mathematical text. Philos Math Educ J 33. https://education.exeter.ac.uk/research/centres/stem/publications/pmej/pome33/Ernest%20A%20Semiotic%20Theory%20of%20Mathematical%20Text.doc
Ferraro G (2004) Differentials and differential coefficients in the Eulerian foundations of the calculus. Hist Math 31(1):34–61. https://doi.org/10.1016/S0315-0860(03)00030-2
Foster JE (1947) A number system without a zero-symbol. Math Mag 21(1):39–41. https://doi.org/10.2307/3029479
Friedman M (2021) On mathematical towers of babel and 'translation' as an epistemic category. Math Intell 43(4):62–73. https://doi.org/10.1007/s00283-020-09969-x
Gastaldi JL (2021) Why can computers understand natural language? Philos Technol 34(1):149–214. https://doi.org/10.1007/s13347-020-00393-9
Gastaldi JL (2022) Boole's Untruth Tables: The Formal Conditions of Meaning before the Emergence of Propositional Logic. In Logic in Question, edited by Jean-Yves Béziau, Jean-Pierre Desclés, Amirouche Moktefi, and Anna C Pascu, 119–49. Studies in Universal Logic. Cham: Birkhäuser
Gastaldi JL (forthcoming) De Morgan's Laws: The Placeof Duality in the Emergence of Formal Logic. In Duality in 19th and 20th Century Mathematical Thinking, edited by R. Krömer, E. Haffner, and K. Volkert. Birkhäuser. https://hal.science/halshs-03995323v1
Grosholz ER (2007) Representation and productive ambiguity in mathematics and the sciences. Oxford University Press, Oxford
Hadden RW (1994) On the Shoulders of Merchants: Exchange and the mathematical conception of nature in early modern Europe. SUNY Press, Albany
Haraway D (1988) Situated Knowledges: the science question in feminism and the privilege of partial perspective. Fem Stud 14(3):575–599. https://doi.org/10.2307/3178066
Harding S (1992) Rethinking standpoint epistemology: what is 'strong objectivity?'. Centennial Rev 36(3):437–470
Hayashi T (1995) The Bakhshālī manuscript: an ancient Indian mathematical treatise. Egbert Forsten, Groningen
Herreman A (2000) La topologie et ses signes : éléments pour une histoire sémiotique des mathématiques. L'Harmattan, Paris
Herreman A (2001) La mise en texte mathématique. Une analyse de l'"Algorisme de Frankenthal". Methodos. Savoirs et textes 1. https://doi.org/10.4000/methodos.45
Hodgkin L (2005) A history of mathematics: from Mesopotamia to modernity. Oxford University Press, Oxford
Katz VJ (ed) (2007) The mathematics of Egypt, Mesopotamia, China, India, and Islam: a sourcebook. Princeton University Press, Princeton
Katz VJ, Folkerts M, Hughes B, Wagner R, Berggren JL (2016) Sourcebook in the mathematics of medieval Europe and North Africa. Princeton University Press, Princeton
Keller A (2015) Ordering operations in square root extractions, analyzing some early medieval Sanskrit mathematical texts with the help of speech act theory. In: Chemla K, Virbel J (eds) Texts, textual acts and the history of science. Springer International Publishing, Cham, pp 183–218. https://doi.org/10.1007/978-3-319-16444-1_5
Klein J (1968) Greek mathematical thought and the origin of algebra. MIT Press, Cambridge

La Nave F, Mazur B (2002) Reading Bombelli. Math Intell 24(1):12–21. https://doi.org/10.1007/BF03025306
Lacan J (2007) Écrits: the first complete edition in English (trans: Fink B). W. W. Norton, New York
Latour B (1999) Pandora's hope: essays on the reality of science studies. Harvard University Press, Cambridge
Lévi-Strauss C (1955) The structural study of myth. J Am Folk 68(270):428–444. https://doi.org/10.2307/536768
Lévi-Strauss C (1966) The savage mind. University of Chicago Press, Chicago
Lévi-Strauss C (1987) Introduction to the work of Marcel Mauss. Routledge & Kegan Paul, London
Maniglier P (2006) La Vie énigmatique des signes: Saussure et la naissance du structuralisme. Editions Léo Scheer, Paris
Netz R (1999) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press, Cambridge
Netz R (2002) Counter culture: towards a history of Greek numeracy. Hist Sci 40(3):321–352. https://doi.org/10.1177/007327530204000303
Netz R (2022) A new history of Greek mathematics. Cambridge University Press, Cambridge
O'Halloran KL (2005) Mathematical discourse: language, symbolism and visual images. Continuum Press, New York
Overmann KA (2019) The material origin of numbers: insights from the archaeology of the ancient near east. The material origin of numbers. Gorgias Press, Piscataway. https://doi.org/10.31826/9781463240691
Presmeg N, Radford L, Roth W-M, Kadunz G (eds) (2018) Signs of signification: semiotics in mathematics education research. Springer, Cham. ICME-13 Monographs. https://link.springer.com/book/10.1007/978-3-319-70287-2
Rabouin D (2017) Styles in mathematical practice. In: Chemla K, Keller EF (eds) Cultures without culturalism. Duke University Press, Durham, pp 196–224. https://doi.org/10.1515/9780822373094-010
Rav Y (2007) A critique of a formalist-mechanist version of the justification of arguments in mathematicians' proof practices. Philos Math 15(3):291–320. https://doi.org/10.1093/philmat/nkm023
Robson E (2008) Mathematics in ancient Iraq: a social history. Princeton University Press, Princeton/Oxford
Rodin A (2010) Did Lobachevsky have a model of his 'imaginary geometry'? arXiv. https://doi.org/10.48550/arXiv.1008.2667
Rosenfeld BA (1988) A history of non-Euclidean geometry: evolution of the concept of a geometric space. Springer, New York
Rotman B (1987) Signifying nothing: the semiotics of zero. Macmillan Press, Houndmills
Rotman B (2000) Mathematics as sign: writing, imagining, counting. Stanford University Press, Stanford
Schlimm D, Skosnik K (2011) Symbols for nothing: different symbolic roles of zero and their gradual emergence in Mesopotamia. In: Cupillari A (ed) Proceedings of the 2010 meeting of the Canadian Society for History and Philosophy of Mathematics, vol 23, pp 257–266. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.381.2648&rep=rep1&type=pdf
Somers-Hall H (2010) Hegel and Deleuze on the metaphysical interpretation of the calculus. Cont Philos Rev 42(4):555–572. https://doi.org/10.1007/s11007-009-9120-2
Sperber D (1985) On anthropological knowledge: three essays. Cambridge University Press, Cambridge
Steensen AK (2022) Texts in progress: reading 19th century mathematics with structural semiotics. PhD dissertation, ETH Zurich, Zurich
Steensen AK, Johansen MW, Misfeldt M (forthcoming) Textual materiality and abstraction in mathematics. Sci Context
Stengers I (2010) Cosmopolitics I, II (trans: Bononno R). University of Minnesota Press, Minneapolis

Swetz F (1987) Capitalism and arithmetic: the new math of the 15th century (trans: Smith DE). Open Court Publishing, La Salle
Vinciguerra L (1999) Langage, visibilité, différence: histoire du discours mathématique de l'âge classique. Vrin, Paris
Wagner R (2009) S(Zp, Zp): post-structural readings of Gödel's proof. Polimetrica, Milan
Wagner R (2010) The geometry of the unknown: Bombelli's algebra Linearia. In: Heeffer A, Van Dyck M (eds) Philosophical aspects of symbolic reasoning in early modern mathematics. College Publications, London, pp 229–269
Wagner R (2017) Making and breaking mathematical sense: histories and philosophies of mathematical practice. Princeton University Press, Princeton
Wagner R (2019) Mathematical abstraction as unstable translation between concrete presentations. Philos Math Educ J 35
Wagner R (2020) Conatus mathematico-philosophicus. Allgemeine Zeitschrift für Philosophie 45(1):85–120. English version available in: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/408346/wittgensteinpaper.pdf
Wagner R (2022) Mathematical consensus: a research program. Axiomathes. https://doi.org/10.1007/s10516-022-09634-2
Wittgenstein L (1976) Wittgenstein's lectures on the foundations of mathematics, Cambridge, 1939. University of Chicago Press, Chicago
Wittgenstein L (1981) Remarks on the foundation of mathematics, 3rd edn. Wiley, London
Wittgenstein L (2009) Philosophical investigations, 4th edn (trans: Anscombe GEM, Hacker PMS, Schulte J). Wiley and Blackwell, Chichester

Shaping, Revisited

Reviel Netz

Abstract

In this chapter, I review the argument of "The Shaping of Deduction in Greek Mathematics," explaining its elements that may be construed as "semiological," comparing them with other work I did in the semiology of mathematics, and conclude by pointing out the scope, and limits, of semiological explanation in the history of mathematics.

Keywords

Greek mathematics · Formulaic language · Diagrams · Abacus · Materialism

When asked to reflect upon *The Shaping of Deduction in Greek Mathematics* (Netz, 1999) I find myself plagued, naturally, by self-doubt. This takes a specific form. If you've read Netz (1999) you know that I do not shun the sweeping generalization but this is usually historical rather than philosophical or historiographical in character. I am out of my depth discussing methodology and usually end up with the platitude that different historiographical tools are useful for different historiographical purposes. And yet, I understand my role in this essay to be historiographical in character. The interesting question is how the example of this particular book can serve for the purposes of studying the semiology of mathematical practice in other contexts. To try and approach this question, I will present the book, discuss some of my own attempts to extend its arguments further, and – reflecting upon the limits of such extensions – conclude with some notes on the

R. Netz (✉)
Stanford University, Stanford, CA, USA
e-mail: netz@stanford.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_63

scope, and limits, of the semiology of mathematical practice as a type of historical explanation.

And so, first of all: to *The Shaping of Deduction* itself, as an example of semiology.

It is actually not too difficult to distinguish the strictly semiological parts of the book. The first four chapters formed a description of the main formal features of Greek mathematical texts. Chapters 5 and 6 discussed consequences of those formal features for mathematical practice (chapter 5 was about "The Shaping of Necessity" – what gives rise to the sense that an argument is demonstrative, while chapter 6 was about "The Shaping of Generality" – how an argument couched in terms of a particular case is seen as valid for a general mathematical statement). Chapter 7 was my attempt to provide a historical account – in social and cultural terms – for how the formal features came to be entrenched and the Greek mathematical specific form of writing, or genre, formed. The overall structure then was to derive necessity and generality, via the formal features of texts, from a social and cultural context. This is indeed a very historical project but its central portion – and the one highlighted in the layout of the book – was the description of the formal features. Those, indeed, were not approached very historically. My descriptive strategy was to make claims concerning the Greek mathematical genre based on samples ranging across individual authors or a group of authors, but I did not try to find contrasts between authors, subgenres, or between historical eras. The underlying assumption was that there is indeed much in common between the works in the Greek mathematical genre and that we should concentrate on such generalities. As noted: I believe generalizations are valuable. Indeed, much of my work since historicized the detail of Greek mathematics further. In *Ludic Proof* (Netz, 2009) I looked for the specific motivations for mathematical writing in the Hellenistic era. In *The Transformation of Mathematics in the Early Mediterranean* (Netz, 2004) I considered the rise of commentary form in Late Antiquity (which I have since considered in a much wider historical setting in *Scale, Space and Canon in Ancient Literary Culture*, (Netz, 2020), chapter 6.5; in the same book, I consider also the specific trends of imperial era science, in chapter 6.3). A major thrust of *A New History of Greek Mathematics* (Netz, 2022) is the historical divide between the two major creative generations of Greek mathematics, the generation of Archytas and that of Archimedes. In short, Greek mathematics has a well-articulated history. Archytas' motivations differed from those of Archimedes, differing from those of Ptolemy and Pappus. Their audiences were different and their authorial values differed accordingly. What differed much less was their genre – their mode of writing – and the continuity at this level is made more, not less evident, by the consideration of the changing historical background.

The formal features surveyed in chapters 1–4 – the semiology of Greek mathematics, if you will – fell into two main parts. Chapters 1–2 described the use of diagrams (chapter 1 discussed the diagrams themselves; chapter 2 discussed the use of letters, labeling the diagram, within the main text). Chapters 3–4 described the use

of language (chapter 3 described the mathematical lexicon, chapter 4 described the use of formulaic expressions). Looking back at the reception of the book, the most surprising thing to me was the extent to which this focused on diagrams and not on language. A lot of this seems to be a bias toward the new and surprising: readers naturally focus more on what they did not expect. When considering the semiology of a text, it is obvious that one would describe its language; it was more surprising, and so more useful for readers, to learn about the place of the visual. This is especially the case because of the claim (which was made at the same time, from the philosophical point of view, by Manders (1995/2008)) that Greek mathematical diagrams are not some mere ornament, in principle eliminable, but instead serve as one of the informative vehicles of the text. This makes for a more complicated, two-lane semiology. There is more: diagrams are evidently "embodied," they subsist as graphic and ocular phenomena. To say that mathematics of a certain kind essentially relies on diagrams is to say that this mathematics is essentially embodied; this attitude, vaguely anti-Platonist, spoke to many readers (as one can see in particular in the most significant review of my book, that by Bruno Latour (2008)). The emphasis in the reception on the diagrams is thus understandable but also a source of chagrin. I did *try* to say new things about the language, as well! More seriously, and more to the point, the operation of Greek mathematics cannot be understood apart from its *two* lanes. It is precisely the combination of the visual and the verbal that makes Greek mathematics effective. The complicated semiology is not merely interesting in its added complication of a visual, put on top of the verbal: it is the interaction of the two – for which, both must be acknowledged – which matters most.

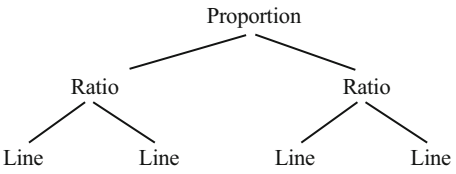
In what follows, then, I will restate my main observations; I will then proceed to discuss later additions and revisions, and then return to my tentative reflections.

To redress somewhat the balance of the reception, let me begin here with the language. The observation of chapter 3 was mostly negative (disambiguation of terms was not the only or even the central function of most introductory passages to Greek mathematical works; much of the lexicon was undefined or ambiguously defined and in general there was not a deliberate effort to regiment the actual vocabulary in use). This leads to chapter 4 where I made my more positive contribution. My argument was that the unit around which the language was organized, in use, was not the individual word but the formulaic expression. By this I refer to expressions such as ("<. . .>" mark elided words):

(1) "The (feminine article) <line> AB"
(2) "Let the <Line> AB be drawn through the <point> A parallel to <line> CD"
(3) "As the <line> AB to the <line> CD, so the <line> EF to the <line> GH".

The number of possible expressions one could compose within the Greek mathematical lexicon is astronomically bigger than the number of expressions in actual use. Typically, a Greek mathematical proposition can be broken down to a few

repeated expressions, so that even an entire treatise can be shown to use in the order of magnitude of a hundred or so. What I found most significant was the embedded tree structure of such formulaic expressions. The expression for proportion (3 above) brings together two expressions for ratio, each in turn often composed of two expressions for lines (1 above). As linguists would put this:



Now, consider a statement such as (Apollonius' *Conics* IV. 46, 72.4–8):

> Since it is: as the <square> on MΨ to the <square> on ΨI, so the <rectangle contained> by AΠB to the <rectangle contained> by ΔΠE, but as the <rectangle contained> by AΠB to the <rectangle contained> by ΔΠE, so the <square> on ΛT to the <square> on TI, therefore also: as the <square> on MΨ to the <square> on ΨI, so is the <square> on ΛT to the <square> on TI.

Such statements are the backbone of Greek mathematics: the manipulation of proportion. They appear almost impenetrable to the untrained reader, but a modern reader would immediately understand the content by being told that this is the transitivity of proportion and that the claim has the form:

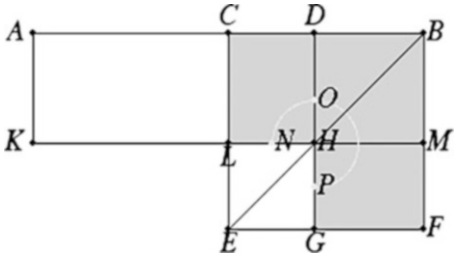
> a:b::c:d and c:d::e:f, therefore a:b::e:f

One of the key arguments of *The Shaping of Deduction in Greek Mathematics* was that (1) formulaic expressions such as those Greeks use for proportion were as transparent, to the Greek trained readers, as visual symbolic expressions are to modern trained readers, (2) that what made their manipulation so easy was the human capacity to read off the phrase structure of verbal expressions. The key operation of the Greek mathematical language was the substitution of elements within formulaic expressions. The significance of expressions such as that of proportion is that they provide, simultaneously, a substrate for substitution (the formulaic expression for proportion is composed of six lower slots – two ratios, four objects – each of which may be independently replaced) as well as a warrant for substitution (the very claim of the proportion asserts the substitutability of the two ratios).

The Shaping of a Deduction in Greek Mathematics was subtitled *A Study in Cognitive History*, not *A Study in Semiology*. In my account of the Greek mathematical language, I did provide something of a morphology of the Greek mathematical expression, a linguistic and a semiological exercise. But what interested me then was not the symbolic system itself but its interface with the human mind, making the performance of certain tasks possible.

As noted, the cognitive exercise of Greek mathematics was a two-lane process, verbal and visual. Some of the operations were supported by verbal manipulation, mediated by the phrase structure of embedded formulaic expressions. Others were supported by visual identification. Hence the role of the diagram.

The morphology of the Greek mathematical diagram is harder to provide (the semiotics of the visual – two dimensional, never a pure Piercian symbol – is always more difficult, hence the greater interest it attracts). The Greek mathematical diagram is always labeled, the combination of letters conforming with the potential combinations of geometrical objects. When lines or areas are presented in isolation, they usually are labeled with a single letter directly; otherwise labels are attached to the points at the intersections or other significant points on the lines in configuration. The purpose of the diagram is to enable readers to pick up such relations of configuration. One main function of the diagram is to support claims such as (paraphrasing Elements II.5):



[area] CH is equal to [area] FH, therefore, adding [area] DM, [area] CM is equal to [area] DF.

That is, it is through the diagram that we tell that, e.g., CH + DM = CM. This relation is not supported by any other symbolic representation. In more general terms, the diagram provides the reference for the objects mentioned by the verbal discussion and allows one to skip steps, since identities are visually preserved. For instance, in this case it is verbally asserted that the relevant lines are parallel so that we can tell that, for instance, ACLK is a parallelogram. A Greek mathematical author would not be able to tell such a property directly from the diagram (that two lines are parallel cannot be visually represented by the semiotics of Greek mathematics), but once the reference is established, one does not need to go through a separate deduction to establish that AC and KL are opposite sides, hence AC=KL. That opposite sides are equal to each other is explicitly proved, but that AC and KL are the opposite sides in the parallelogram (and not AC, AK) is *visually* established. I mention the example of parallels and the relations they engender. This is not an accidental example. In general, parallels engender, to the Greeks, similar figures. Similarity, in turn, signifies, to the Greeks, the presence of proportions. This then is the normal process of a Greek mathematical proof: embed in the visual sign the reference to parallel lines; use this to extract similarity, hence proportion; manipulate proportions via the verbal sign; this in turn may give rise to new results embedded and then manipulated via the visual sign. This

two-lane process of Greek mathematics is at the root of many creative moments in Greek mathematics. It may be that, in general, our lucky number is two. Working within a single sign system, one can do only so much that is surprising: the signs themselves suggest their horizon of applicability. Combine two sign systems together, and one no longer can tell the possible results in advance. The mathematical support for this duality, in this case, is the homology of proportion and similarity; the semiological support is the duality of language and diagram.

In the years since the publication of *The Shaping of Deduction in Greek Mathematics*, I made a few further contributions related to the semiology of Greek mathematics, which are worth mentioning before I return to sum up the relation between semiology and history.

As regard the language, two observations I made – both related more to arithmetic than to geometry – are perhaps relevant. One is in the article "Counter Culture: Toward a History of Greek Numeracy" (Netz, 2002), the other is in the article "Reasoning and Symbolism in Diophantus: Preliminary Observations" (Netz, 2012).

Both articles engage, yet again, with the theme of semiotic duality. In Netz (2002), I described the practices of calculation in the ancient Mediterranean organized around the abacus, which is simply a flat surface on which spaces are marked for specific types of numerical representations, with small tokens moved along the surfaces according to the rules of the required calculation. In the most familiar case we have the surface divided into vertical columns marked by "1," "5," "10," "50," etc. (this is most familiar as the Roman numerals). The rule is that five counters on a "1"-type column are removed and replaced by a single counter on the column immediately following; the same applies for two counters on a "5"-type column. For instance, calculate $33 + 22 = 55$ as the series of states on the abacus:

I	V	X	L	
***		***		33

I	V	X	L	
***		***		
**		**		+22

I	V	X	L	
	*		*	=55

I emphasized in Netz (2002) that such counters, moved on the surface of the abacus, form, to the Greeks, the most natural way of representing numbers. Just as,

to our mind, certain traces on paper taking shapes such as "2," "3," are the self-evident symbol for the relevant numbers, so, to the Greeks, the same numbers are self-evidently represented by the heap of counters containing the correct cardinality. The metonym of number were *psephoi* (Latin: *calculi*. To be clear: the suggestion is that the heaps of counters stood as the metonym for number – not the Greek symbolic analogues to "I," "V," "X," "L," which were more obviously mere tools for manipulating number). At the same time, the Greeks – like almost all cultures – possessed a verbal representation for numbers which, in their case, followed the base 10. (The distinction between "1" and "5" columns in the abacus is of course motivated by the limitations of the human cognitive capacity of subitization). A separate scribal convention existed to abbreviate this verbal representation. Thus, instead of writing out in full "thirty two" the Greeks could write in sequence the symbol for thirty followed by the symbol for two. The symbolism for this system – which has distinct symbols for each of the values 1, 2, . . ., 9, 10, 20, . . ., 90, 100, 200, . . . 900 – was taken from a slightly expanded alphabet. In this case, the duality of the semiotic systems – counters in the calculation of numerical values, alphabetical numbers in its record – was perhaps, if anything, counterproductive. This is because there is no two-way traffic allowing one to combine the two systems. Numbers are calculated through counters and then are entrenched in writing as alphabetical, where the process ends. The manipulation of number as a written device is the great invention of the algorithm which is, essentially, a way of turning the abacus into a written form.

Part of my interest in Netz (2012) was in the symbolism employed in the manuscripts of Diophantus' Arithmetica. (I stand by my purely semiological observations on the symbolism as such; I am much less confident of my interpretation of Diophantus offered in that article, in part because I am not longer confident that Diophantus should be considered a Late Ancient author). Diophantus explicitly develops a system of symbols based on transparent abbreviations: so, for instance, C^{u} for "cube." A special symbol is used for "number," which often is used for the number sought in the problem at hand: this then makes it functionally equivalent, in some cases, to our "x" (I will represent this as $). Remember that one may use the alphabet in a particular way, to abbreviate numerals – and, with all those abbreviations in place, many statements in Diophantus seem to be written in a quasi-algebraic format, for instance:

> [concluding a calculation] result: $A M^{o}K

This is equivalent to what we would write down as X + 20. (M^{o} stands for "monads" or the units of the number being calculated in this case. A is 1, K is 20).

My observation in Netz (2012) – which, to be honest, was not very original or controversial – was that in the manuscripts of Diophantus abbreviated and full forms are always interchangeable. In one sentence one writes down "$," in the next – "number." "K" and "twenty" may appear side by side. There seems no obvious pattern inside manuscripts, between them or across time: in all likelihood, I believe this pattern was found in Diophantus from the very beginning. The symbolic abbreviation is essentially an allograph. (This is true, to be clear, for both pairs,

"$"/"number" as well as "K"/"twenty": It should be noted that the latter allography between number abbreviation and full number word is visible everywhere in the manuscript tradition and is not specific to Diophantus). It is tempting to suggest, once again, that Diophantus was at the cusp of a semiological/cognitive breakthrough but failed to achieve it. (Because, obviously, his interests were different: he did not try to invent algebra, he tried to solve arithmetical problems). It would be only in early modern Europe that authors would start to employ abbreviated forms more consistently, creating a semiotic system in which certain observable relations between symbols may be calculated directly on the symbols as a visual system, independently of the symbols' verbal interpretation.

With both developments – the rise of visual number manipulation, the rise of algebraical symbolism – we study a similar trajectory where, ultimately, modern European mathematics would contrast with the Greek. This trajectory from geometry to algebra clearly invites a semiological account. Still, I did not pursue it much further, and only in part because of my training in ancient, and not modern mathematics. Something seems to change, I believe, in the very nature of the process, so that I no longer see how to pursue it as part of the same research program. To see why, let me add a few words on the other component of the Greek mathematics semiological system – the diagrams.

This brings us back to Greek geometry itself. The key to Greek geometry is the duality of the two semiotic systems – language and diagram, each operating based on its own cognitive lane. The two lanes, the visual and the verbal, are then most directly combined in early modern Europe by the tool of the symbolic algebraic formula, in some ways equivalent to the Greek formulaic expression but made accessible, in cognitive terms, to a more purely visual computation, bypassing the grammatical parsing of the verbal expression. In some sense, then, early modern Europe – with its Indo-Arabic numerals and its symbolic algebra – represents semiological progress.

Perhaps one reason why Greek mathematical diagrams were found, by readers of my book, more interesting than the Greek mathematical language is that there is no simple story of progress to be told and if anything there is a case to be made that the Greek mathematical use of diagrams was more subtle and sophisticated than that of contemporary geometrical pedagogy. It is often believed today that students should not rely on diagrams while making inferences, and instead rely on the discursive proof, organized in a formal manner (say, as a list of claims, each supported explicitly by axioms, previous claims, etc.). This attitude represents, at some level, a fundamental misunderstanding, as if the discursive proof is somehow semiologically innocent, as if the signs of the text and its symbolism are not a medium in and of itself. But at another level this represents a valid concern: it is essentially impossible for diagrams to convey certain types of information precisely. Roughly put, the distinction is between metrical information (the measurements of the elements of the geometrical configuration) and topological information (the contacts, containments, and intersections of the configuration itself). Greek mathematics relies on diagrams in a prudent and sophisticated way: allowing to read off the diagram the topological information but conveying all metrical information via the language.

This use of the diagram seems to be reflected in the very drawing of the diagram. In our manuscript evidence we see less metrical fidelity than one typically sees in

modern diagrams. There is absolutely no effort at any painterly effects (solid objects often appear entirely flat and indeed are often hard to interpret), and different elements, of different shapes and sizes, are often represented by the same schematic representation: the same acute isosceles triangle, whose base is parallel to the base of the page, every time a triangle needs to be represented; unequal lines often represented as equal. The only thing "correctly" represented are the relation of configuration required for the proof itself.

The transmission of ancient Greek texts is mediated, with few exceptions, through Medieval manuscripts, copies of copies of the original. Philological techniques make it possible to judge, with a great deal of certainty, how close our texts are to the authorial source. It seems clear that there is a certain amount of interpolations as well as simple mistakes contaminating our texts but that this does not disrupt our basic understanding of the language. Our semiology is safe in this regard. The matter of diagrams is more difficult. Until recently, there was not even a consistent effort to try and recover the evidence of the Medieval manuscripts concerning the diagrams (many philologers, especially in the nineteenth century, made the significant semiological choice to edit text, but to freely recreate diagrams). The observations noted above concerning the shape of Greek mathematical diagrams were made very recently, and I contributed to them in my edition of Archimedes (Netz, 2004, 2017, and see the reflections in Netz, 2012b). The best starting point is Saito and Sidoli (2012).

Since the evidence is so mediated, there is room for reasonable doubt. Carman, in articles starting in 2018, made the case that the more schematic diagrams seen in Medieval manuscripts could result from the "broken telephone" game of continuous copying. I argued against this in Netz (2020), especially from the perspective of the actual chains of transmission. But the issue should not be considered as settled.

Perhaps I can reiterate another argument made in Netz (2020), of direct relevance to semiology. If indeed we believe that schematic diagrams are authorial, how to account for this schematic character? One obvious response might be that Greek mathematicians used schematic diagrams because these are mathematically effective, for the reasons explained above. However, I find this unlikely. The argument concerning metrical properties, topology, and the logic of diagrams is complicated and was only very recently articulated by philosophers of mathematics such as Manders. (Poincaré did anticipate this argument in 1913: but few had thought as deeply on topology as he did). Even more to the point, there are no traces of any such discussions in Greek antiquity – and, had there been such discussions, how likely were they to result in a consensus? These were Greeks, after all. More likely, in my view, is that Greek geometers lucked into a semiotic device that happened to be useful for certain purposes. Indeed, this can be supported by the consideration of the scribal practices of ancient papyrus. This is, in general, a remarkably "flat" semiological practice. Formal, bookroll writing in Greek antiquity involved a single font (even nothing like our division into upper case or lower case: it was all, essentially, upper case), a single size (practically no bigger initials or titles), a single color (all the same dark brown, charcoal-based ink), no punctuation, and – most remarkable to a modern reader – no space between words. The rule can be stated simply: the scribal

code of the ancient papyrus involved zero variation because it involved zero *interpretation*. This makes sense as we consider the function of the ancient bookroll (for which see Johnson, 2000, 2010). While the ancients did read silently, they also read primarily the masterpieces of performative literature – Homer, tragedy, comedy, rhetoric, and dialogue. While reading, one always tried to project, imaginatively or even in out-loud performances, the original voice and embodied presence of the text in its full real-world power. The ability to perform a written text in this manner was the ultimate achievement of a full education: one first learned the letters themselves, the *grammata*, in a grammar school, to matriculate into a rhetorical education whose main product was the ability to declaim texts in an elegant way. Ancient education was about the ability to interpret letters. Hence interpretation, left entirely to the reader. Since the same scribes who produced texts for the general public also produced (very occasionally, never as a major part of their occupation) the texts of Greek mathematics, the unadorned, repetitive, unpainterly, and, in general, uninterepted diagrams make sense as part of the same scribal system. These suggest to us how diagrams were to be read: by a heavy interpretative investment by their readers. It is this, above all, which modern diagrams eschew.

Is this final observation at all semiological? It seems as if it should be: It is directly about the nature of a signifying system. But it is useful to note that our understanding of this practice cannot be based, in this particular case, on the symbols alone. My argument does not rely on the structural properties of Greek mathematical diagrams but instead embeds them within a cultural interpretation: I bring in a comparative structure – that of the scribal bookroll – and I interpret this, in turn, following on Johnson (2000, 2010), in terms of the ancient culture of reading as a whole but also – following especially Cribiore (2001), in terms of the practices of ancient education. Is this semiology, or history?

In general, the stories I have been retelling in this brief retrospective have three protagonists. One is the semiological system itself. This is in principle not a historical or even a specifically human system and we study it essentially in the same way we study logical or mathematical systems. The second is the cognitive apparatus with which the semiological system is engaged. This is of course human, and to some extent culturally and historically malleable (everyone can learn to parse automatically the syntactic structures of a Greek mathematical formula – to some extent, I did – but it takes a specific learning which was the hallmark of authors in a certain tradition, and not others). The third is the historical context giving rise to semiotic systems, and, as just explained, influencing the cognitive apparatus itself. I am mostly interested in the fact that people do different things in different times and places. This divides into two questions: *how* do they do that – what makes this possible? And *why* do they do that then – and not in other times and places? Often, I find situations where the *how* question has to do with the interface of semiotics and cognition – a particular semiotic system makes a certain kind of cognition possible; the *why* question has to do with the interface of both with a historical context. I thus try to build a through line from a historical context, via semiotic and (responsive) cognitive practices, to the achievement of different tasks, in this way making the history of science flow from a more fundamental historical context.

The underlying impulse for this historiography is, roughly speaking, a materialist approach to the history of culture. What is it that people actually *do* that serves to explain their more intellectual achievements?

A materialistic impulse leads one to look for conditions and presuppositions of practice that are, indeed, beyond the actors' control. I emphasized, for instance, that I doubt that Greek mathematicians simply *decided* to use schematic diagrams. I proposed an alternative account that they had such diagrams, so to speak, forced upon them given the underlying context of scribal practices as a whole, in turn driven by the underlying context of Greek writing in a world of performative literature and rhetorical education. The impulse to look for materialist accounts does not make any materialist argument correct. The question is empirical: For instance, if a lost ancient work comes to light, making the explicit case that one should rely on schematic diagrams, this would considerably reduce the likelihood of my materialist account of schematic diagrams. I tried to make the case, empirically, where possible. For instance, I argue that the set of formulaic expressions was not explicitly planned by anyone and, in this case, there are handful of historical sources of evidence we may cite: This is in part why I needed to study the Greek practice of definitions (which does not, in fact, cover the use of formulaic expressions). A particular point is that the most self-evident use of formulaic language is in the way it structures the proposition as a whole. (One of those expressions – in Latin acronym, "Q.E.D." – is so well known to have become, effectively, an English phrase!) It was therefore important for this argument to show that when the Late Ancient commentator presents an explicit schema, structuring the typical Greek mathematical proposition in terms of six parts, he does not repeat an ancient formulation but instead present what was in all likelihood the produce of later commentary (Netz, 1999b). In short, I argued in *The Shaping of Deduction* that the Greek mathematical semiotic system emerged organically, in response to a particular historical setting, and I still believe this is true. But there are of course historical examples of semiotic systems being proposed and adopted not through an organic, mostly unreflective process, but through the actors' own deliberate actions. Authors might make the case for a semiotic innovation, others might choose to follow, all based on a particular higher-order theoretical argument. I say "higher order" in the sense that such a process is no longer materialist: in such a case, what we study is the history of ideas, pure and simple. In general, we will probably find that we need to pay attention, in varying degrees, to *both* the material and to the higher order.

This is already straying beyond my field of expertise but, at any rate, it does seem that much of the history of mathematics, later than Greek mathematics itself, involves paying more attention to explicit semiotic innovations. I do not think there is a great mystery why this is the case. Greek mathematics, quite simply, presents a very clearly structured genre. It just is very evidently a semiological system. Once one wishes to reflect upon it so as to transform it, it becomes natural to propose changes that involve the rules of the genre as such, hence to propose semiotic innovations: which is obviously part of the story of the transformation of Greek mathematics in early modern Europe.

In this way, perhaps, there is something unique about the constitutive moment, as a literary genre is formed for the first time. Of course, the very semiotics of the genre are likely to emerge, in this case, more organically. Once the genre is established, however, its later evolution is described by a more complex interplay of explicit reflection and material conditions. Materialist historians may well find then that their effort is best spent on explaining aspects of the practice that are less strictly semiotic.

The point is perhaps even more basic. Why do I find materialist history compelling? Because I am, in general, interested in nonobvious explanations. There is something quite frustrating, to my mind, about arguments such that "the Copernican worldview ultimately replaced the Ptolemaic worldview because evidence gradually accumulated in support of the earlier and against the latter." The problem is not quite that this is the wrong explanation (at some level, this is obviously correct) but that this is not an explanation at all. As long as you point out such obvious, surface level observations, you do not yet explain a historical development but merely describe it. The question has to be what were the conditions that made such accumulation possible, and compelling, at a given historical moment. In other words, we are looking to explain the rise of new ideas, through some underlying conditions that are categorically distinct from the ideas themselves. This approach to intellectual and cultural history is materialist or, perhaps, structuralist.

I pursue such approaches myself but I am not keen to offer general historiographical advice, for reasons similar to those pointed out concerning the scope, and limits, of the semiotic approach in the history of mathematics. It is telling that I do not even find it important, at this point, to distinguish between materialist and structuralist accounts. Whatever works, for a given explanation, is itself an empirical question, depending on what the actors were actively considering (which varies historically) and depending on what, in the surrounding environment, was decisive (which once again changes through time).

There is nothing very controversial about a preference for materialist or structuralist explanations. From the mid-twentieth century, through several decades, such an approach became dominant (this coincided with an overall intellectual optimism associated with the expansion of public universities and a general explosion in scholarship in the humanities and social sciences). I suspect that at some point the good intuition, underlying such approaches – that explanations should be *explanations* – was marred by a bad intuition, as if there can be a single universal key, a particular type of materialist or structuralist parameter explaining all of history. Several decades ago, such universal claims began to seem – rightly – hollow. The response of scholars was to give up not only the universalist claims themselves but the very project of explanation of culture. I guess it was hard to imagine, at this point, the baby aside from its bathwater (can one imagine Marx, or Levi-Strauss, apart from their claims of universality?). Perhaps we should. One should be a materialist, or a structuralist, as the occasion demands: looking for different material or other conditions in each different case. Occasionally, then – but only occasionally – semiology, itself, may serve as historical key.

References

Carman CC (2018) Accounting for Overspecification and indifference to visual accuracy in manuscript diagrams: a tentative explanation based on transmission. Hist Math:217–236
Cribiore R (2001) Gymnastics of the mind: Greek education in Hellenistic and Roman Egypt. Princeton University Press, Princeton
Johnson WA (2000) Toward a sociology of Reading in classical antiquity. Am J Philol:593–627
Johnson WA (2010) Readers and reading culture in the high Roman empire: a study of elite communities. Oxford University Press, Oxford
Latour B (2008) Review essay: the Netz-works of Greek deductions. Soc Stud Sci 38:441–459
Manders, K. 2008 [circulated 1995]. The Euclidean diagram, in P. Mancosu (ed.), The philosophy of mathematical practice: 80–133. Oxford University Press: Oxford
Netz R (1999) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press, Cambridge
Netz R (1999b) Proclus' division of the mathematical proposition into parts: how and why was it formulated? Class Q:282–303
Netz R (2002) Counter culture: towards a history of Greek numeracy. Hist Sci:321–352
Netz R (2004) The transformation of mathematics in the early Mediterranean world: from problems to equations. Cambridge University Press, Cambridge
Netz, R. 2004b. Archimedes: translation and commentary, volume I: the two books on sphere and cylinder. Cambridge University Press: Cambridge
Netz R (2009) Ludic proof: Greek mathematics and the Alexandrian aesthetic. Cambridge University Press, Cambridge
Netz R (2012) Reasoning and symbolism in Diophantus: Preliminary observations. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 327–361
Netz R (2012b) The texture of Archimedes' writings: through Heiberg's veil. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 163–205
Netz R (2017) Archimedes: translation and commentary, Spiral Lines, vol II. Cambridge
Netz R (2020) Scale, space and canon in ancient literary culture. Cambridge University Press. Cambridge
Netz R (2022) A new history of Greek Mathematics. Cambridge
Poincaré (1912 [1963]) Mathematics and science: last essays. Dover, New York
Saito K, Sidoli N (2012) Diagrams and arguments in ancient Greek mathematics: lessons drawn from comparisons of the manuscript diagrams with those in modern critical editions. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 135–162

Mathematical Practices and Written Evidence: General Reflections Based on a Historian's Experience

Karine Chemla

Contents

1 Introduction 3048
2 Texts Viewed as Communicating Knowledge 3048
2.1 Texts as Acts of Communication and Their Shaping 3049
2.2 A Text Communicates at Various Scales 3050
2.3 Contexts of Communication 3051
2.4 Approaching Texts in the Context of Mathematical Cultures and the Issue of Interpretation 3053
3 Texts Written to Work With 3055
3.1 How Various Uses of the Page Highlight Different Dimensions of the Practice ... 3056
3.2 The Private and the Public 3057
3.3 Intertwinement Between Exploring and Communicating Knowledge 3058
3.4 Another Case of Intertwinement Between Exploring and Communicating Knowledge 3060
4 Contexts and Environments 3062
4.1 Learning Mathematics with Texts 3062
4.2 How Texts Reflect Material and Social Environments 3064
4.3 Ephemera 3065
4.4 Inscriptions as Spatial and Dynamic Entities 3067
5 Conclusion 3068
References 3069

Abstract

Historians and philosophers of mathematics rely primarily on written evidence to discuss the knowledge and practices produced or put into play by actors. The point of this chapter is to offer a critical reflection on the *nature* of the written evidence on which historians and philosophers *qua* observers draw to this effect.

K. Chemla (✉)
SPHERE (UMR7219, CNRS & Université Paris Cité), Paris, France
e-mail: chemla@univ-paris-diderot.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_75

Keywords

Texts as communicating knowledge · Texts as tools to work with · Drafts · Documents · Ephemera · Inscriptions · Mathematical knowledge · Mathematical practice

1 Introduction

Historians and philosophers of mathematics rely primarily on written evidence to discuss the knowledge and practices produced or put into play by actors – for sure, we should not forget that they also sometimes rely on objects, like calculating machines, but including these in the scope of the essay would lead me too far here. The point of this chapter is to offer a critical reflection on the *nature* of the written evidence on which historians and philosophers *qua* observers draw to this effect.

To follow the discussion of this issue that I offer here, some terminological distinctions will be useful. I will thus tentatively introduce indications about some terms I use, even though I am aware of the complexity of the matter, and of the preliminary character of my suggestions. In what follows, then, I will use the term "text" to designate that which we read in our written evidence (with reference to both discourse and non-discursive elements). Texts are available to us through "documents," a term I use to refer to a material object (e.g., a papyrus, a clay tablet, or bamboo slips) on which a text under consideration features. Texts have also been written by means of ephemeral "inscriptions" (using, e.g., calculating rods, tokens, or chalk), and in such cases, due to its very nature, the written evidence must be restored from clues.

This essay surveys recent work in the history of mathematics to offer a theoretical discussion of the various ways in which texts, documents, and inscriptions reflect knowledge and practices, my aim being to highlight the various dimensions of mathematical practice to which they attest. In my view, such a clarification helps us apprehend how our sources might be used to work on the activities undertaken by our actors. It can also inspire general reflections on what writing has meant – an issue to which I return in conclusion.

2 Texts Viewed as Communicating Knowledge

A common tacit assumption that historians and philosophers of mathematics make about the mathematical texts they use as sources is that these texts were written to *communicate* and more specifically to communicate pieces of knowledge – data, concepts, results, procedures, reasoning, theories, thoughts, techniques, and so on – to their intended readers. We can perceive the effect of this assumption in the way these observers engage with their sources: they seek to account for *what* is being *communicated* and formulate an evaluation of *the way* communication is *performed*

(i.e., whether the exposition is clear or ambiguous, whether it has gaps or not, and so on.)

To be sure, letters meant to convey information to an addressee and further to a wider readership when it was common practice that they be reproduced to circulate more broadly. Articles were intended for readers and, to begin with, in modern times, for referees. Books were composed with specific readerships in mind, even when they turned out to be tapped by other collectives of readers and for uses that their authors or their editors had not foreseen. However, can we assume more generally that this motivation governed the production of any text that we use as a source? Before addressing, in the next sections, the problems with this assumption, I will first analyze what recent work in the history of mathematics tells us about how we might apprehend communication.

2.1 Texts as Acts of Communication and Their Shaping

Take Euclid's *Elements* or the Chinese first-century CE *The Nine Chapters on Mathematical Procedures* (*Jiuzhang suanshu* 九章算術, hereafter *The Nine Chapters*). It seems clear that the former delivers propositions with proofs, organized according to an axiomatico-deductive order, while the latter communicates, in the form of "procedures" (*shu* 術), lists of mathematical actions to be carried out to solve problems. If we extend to written texts the classification of speech acts inaugurated by John Austin (1962), we might provisionally suggest that the former work is of the assertive type, whereas the latter seems rather of the directive type.[1]

Reviel Netz (1999) goes one step further, asking not only *what* the text of Euclid's *Elements* communicates, but *how* the act of communication was carried out. As I will argue, the "how?" question is as complex as the "what?" question. Netz's book illustrates some of the ways in which this question can be interpreted and addressed. His study of the lettering of the diagram attached to each of the propositions of Euclid's *Elements* leads to several conclusions on these issues.

First, Netz establishes that the point of theorems or problems put forward in a proposition, as well as the reasoning attached to them and the diagram on which the reasoning was based, were fully formed before the related text in Euclid's *Elements* was composed. In other words, the text did not derive from a reflection that unfolded while writing, but aimed to present propositions, after they had been acquired. This conclusion might be interpreted as supporting the claim that the intention behind the composition of Euclid's *Elements* was indeed to communicate knowledge.

Second, the diagram was drawn, and letters were added to it, following an order that combined the process of geometric construction and the flow of the argument, before the text of the argument was written up (p. 85). Paying attention to the writing practices involved in the making of Euclid's *Elements* thus enables Netz to document the writing act that produced the text.

[1] On this extension, see Chemla and Virbel (2015).

2.2 A Text Communicates at Various Scales

One could address the question of what and how the text communicates from several other angles.

For instance, a work like Euclid's *Elements* seems to communicate not only propositions but also an organization thereof, thereby conveying a meaning at a global level and hence calling for a perusal that goes, so to speak, from cover to cover. Was *The Nine Chapters*, with its 246 problems and related procedures, intended to communicate meanings of a similar nature – that is, meanings at the level of the textual organization – and therefore did the work require a global reading of this kind? This does not appear to be the case to the observers that we are. However, what did the ancient actors think about these questions? A method of addressing the issue in a historical fashion is to examine ancient readers' approach to the text. *The Nine Chapters* was handed down with a selection of ancient commentaries that were composed on the classic – that is, the one completed by Liu Hui 劉徽 in 263 as well as the one composed under Li Chunfeng's 李淳風 supervision and presented to the throne in 656. These readers do not seem to have considered that *The Nine Chapters* should be perused going simply from beginning to end. By contrast, for them as well as for us when we deal with Euclid's *Elements*, the classic required forms of global reading. In their view, at a meso-level, the chapters conveyed a mathematical meaning *qua* chapters (Chemla et al. 2018). Moreover, the set of nine chapters pointed out fundamental operations to which all mathematical procedures could be reduced: a commentator's interpretation thus aimed to establish the list of such operations using a specific reading of the procedures brought together in *The Nine Chapters* (Chemla 2008). Whatever the case might be, neither Euclid's *Elements* nor *The Nine Chapters* deal with these elements of knowledge in a discursive way. Nor do they address the kind of handling of the text required for a successful uptake. The question of how the authors or the editors of these works intended to convey meaning is thus not as straightforward as it might appear at first sight. Nor consequently is the issue of the interpretation of our written evidence.

Problems with respect to our "how?" and "what?" questions also arise at a much more local level. What did terms for operations, such as those occurring in the text of a procedure in *The Nine Chapters*, refer to? Høyrup (1990) has raised doubts about the basic answer that earlier historians had given to this issue for the case of cuneiform tablets.[2] For them, the sentences of a procedure in which these terms occur referred directly to arithmetical operations and to the operands to which they should be applied. Why, then, Høyrup asked, do we have different terms referring to what for us would be the same arithmetical operation? In contrast with earlier interpretations, Høyrup put forward the thesis that these terms referred not only to an arithmetical but also to a geometric operation: In his view, if different terms referred to the same arithmetical operation, geometrically they designated different

[2] Høyrup (2002) gives a more comprehensive treatment of this issue.

types of action. Such an interpretation of these terms entails a wholly different reading of the text of a procedure: For Høyrup, the text not only communicated the prescription of a sequence of computations but at the same time referred to a geometric configuration and its successive transformations, which made the meaning of the computations clear and thus formulated an argument for the correctness of the procedure. Contrary to interpretations that considered the text of a procedure as a simple sequence of directives giving readers the means to fulfil a task without conveying any understanding – which is how texts of procedures from the ancient world have frequently been assumed to make sense – Høyrup proposed an analysis that read proof and algorithm interwoven into a single text. This work has opened the way to a reinterpretation of whole corpora of ancient sources (Chemla 2012). One could not explain more clearly what is at stake in securing a correct apprehension of the communicative acts that mathematical sources carry out. In the same way as the questions of what and how a speech act communicates have given rise to much theorization, the same questions need to be attended to with respect to mathematical texts.

2.3 Contexts of Communication

The study of speech acts has emphasized the importance of taking into account the context of communication in order to analyze the way in which a given act was successfully performed. To investigate the nature of the communicative act carried out by given mathematical texts, taking into consideration the context in which they were produced also appears as essential. However, the nature of the contexts that are meaningful in this case raises yet another theoretical problem, which is of extreme difficulty and still awaits research in the history and philosophy of mathematics. Let me offer here some preliminary reflections on this issue.

Geoffrey Lloyd (1990) has already given the notion of context of communication pride of place. He found in this notion resources to disqualify explanations that accounted, for example, for the emergence of the axiomatico-deductive reasoning in ancient Greece, and not in China, in terms of mentality. In particular, Lloyd suggested that the latter difference could be understood as correlated with differing contexts of communication in these two parts of the ancient world, which owed to contrasting political backgrounds. In ancient Greece, unlike early imperial China, Lloyd argues, adversarial debate was common and, in specific contexts, "adjudicated by a lay audience" (p. 126). In Lloyd's view, such an environment was conducive to the development of critical analyses of argumentation and to the shaping of textual forms of reasoning such as those found in works like Euclid's *Elements*. Lloyd's argument points to a more general thesis with respect to our written evidence: it suggests how and why textual forms of communication adopted in mathematical writings may be intimately related to textual practices at play in a broader "textual culture" – in this case, argumentative practices familiar to the citizen in the context of classical Greece, owing to his or her political and social experience. However, it is interesting to note that not all Greek mathematical texts fall into this category. This

warns us against the temptation to generalize too quickly in any part of the reasoning. In particular, the identification of textual cultures appears as a potentially fruitful project, but also as a challenging one.[3]

Indeed, this is not the only case where a textual form employed in mathematical writings echoes a form common in the broader textual cultures in relation to which mathematical texts were composed. A telling case in point is that of Sanskrit mathematical "treatises" – or mathematical chapters in "treatises" on astral sciences – that are written in the form of *sūtras,* as are "treatises" in many other fields of inquiry from South Asia (Plofker 2009, Appendix A). It is beyond the scope of this chapter to discuss a definition of the *sūtra* from a textual viewpoint, assuming that such a definition that would hold for all times and places exists (Keller 2015). Suffice it to say that *sūtras* are as a rule versified and characterized by their utmost concision. Accordingly, their formulation has often led observers to emphasize their ambiguities and underline that they are "not always easy to understand" (see, e.g., Plofker 2009: 307–308, 66, respectively). The key question for us is to capture how actors came to grips with these types of text, whether the texts dealt with mathematics or with any other field of inquiry. Unfortunately, as a rule there remains no evidence that would enable us to address this question for the precise context in which *sūtras* were produced. However, "treatises" written in *sūtra* form were in general handed down with commentaries, which offer a linguistic analysis of the *sūtra* in a commentator's interpretation – making clear the meaning of the terms and the grammatical structure of the sentences. Moreover, in cases when the base text dealt with mathematics, commentaries also offered mathematical developments based on the *sūtra* as well (Keller 2006; Plofker 2009: 67). These remarks have two crucial consequences.

First, taking the *sūtras* as that which should be understood implies that one considers the "treatise" independently from the commentaries with which it was handed down. However, for the transmission of the text, actors have shaped the textual act of communication as the combination of the *sūtras* and an associated commentary. A source of this kind attests to an actor's uptake of the *sutras* – it materializes how for this actor the *sutras* communicated meanings and makes explicit the meaning he read into them. Probably, the term "treatise," commonly used in Indology to refer to a collection of *sūtras*, misleadingly creates expectations that the *sūtras* form a stand-alone text and that they carry out an act of communication different from what the actors understood (Keller 2010). Interpreting written evidence requires that we identify what for actors was the textual act.

Second, each commentary embodies a way in which actors processed a collection of *sūtras*. Observing how they drew on the text to make sense of it gives us evidence about the kind of text *sūtras* were for these actors. One example will suffice to illustrate that commentators apply ways of reading that are quite unexpected for modern readers. When the seventh-century commentator Bhāskara I reads *sūtra* 19 of chapter 2 of the *Aryabhatiya* (499), from the list of operations that composes

[3] On the problems and promises attached to various notions of cultures, see Chemla and Keller (2017).

the first part of the *sūtra*, he derives four different procedures on series through using different approaches to the wording, each of which omits a different section of the list (Keller 2006, vol. 1: xlvi). It would not make sense to consider that his reading gives the final word on the *sūtra*. However, the observation of his mode of handling the text of this *sūtra* as well as that of other commentators clearly demonstrate how misleading a modern reading of the *sūtra* would be (Keller 2015). Not all texts communicate in the same way. Applying a decontextualized approach to a text is what might lead to our perception of the text as "ambiguous" or hard to understand. By contrast, the unpredictability for us of Bhāskara I's apprehension of the text shows how essential an approach of reading from the perspective of historical anthropology is for understanding the type of text *sūtras* constituted for actors interpreting them in different given contexts. This approach is after all, here as in any other case, what can help us grasp the mathematical knowledge and practices to which a given text attests. What is more, the fact that *sūtras* deal with mathematics might give us specific resources to describe modes of reading that would then be relevant for other types of *sūtras*.

These remarks lead to a general conclusion valid far beyond the case of *sūtras*. The interpretation of mathematical writings requires that they be contextualized *qua* texts. I have suggested elsewhere that the appropriate contexts were mathematical cultures (Chemla and Keller 2017). I return to this point below. The uptake of textual acts necessitates in particular that the actual texts with which, for actors, acts were carried out be determined and that texts be approached with operations with which actors in a given context dealt with them. For the case of *sūtras*, a given commentary epitomizes a context of interpretation of this kind, which will be passed on to readers who will approach the *sūtras* within this commentary. This observation further highlights that, in addition to the context in which our written evidence was produced, the contexts in which it was appropriated and used are equally meaningful for the history and philosophy of mathematics. The last remark is meant to refer to the practices of reading and interpretation at play in the contexts of appropriation. However, in the case of mathematical writings, it extends also to actors' mathematical knowledge and practices in these contexts (Goldstein 1995; Chemla 1999; Ehrhardt 2017). Textual communication needs to be not only contextualized, but also historicized.

2.4 Approaching Texts in the Context of Mathematical Cultures and the Issue of Interpretation

Placing proofs of a Euclidean type into a wider context of communication allows Lloyd to argue for the thesis that this context might – at least partly – account for how textual forms and practices took shape in mathematics. The case of mathematical *sūtras* highlights another issue at stake in this type of contextualization: the focus here is on interpretation, an issue that is meaningful for any work written in *sūtra* form, although it might take on specific features for mathematical writings. By contrast, the works of Netz and Høyrup mentioned above highlight facets of

mathematical written evidence that are specific to "mathematical cultures," in the sense of "shared ways of doing mathematics" that can be observed in given local contexts and over given periods of time. Lettered diagrams in Greek documents and texts of procedures in cuneiform tablets are arguably facets of this kind. For Netz (1999: 60–61), lettered diagrams emerged as an artefact central to mathematical practice in some Greek contexts and became a crucial part of written documents for some (but not all) practitioners. Like Lloyd, Netz focuses on the question of the historical shaping of elements making up mathematical texts, while in Høyrup's work the emphasis is on how modern historians should interpret the lists of operations in our written evidence from Old-Babylonian Mesopotamia.

In brief, we see how mathematical texts are composed with resources taken from broader contexts as well as with resources shaped by collectives of actors practicing mathematics (Chemla and Keller 2017). In this respect, the types of text used by practitioners, as well as the operations practitioners perform with these texts, are key components of the mathematical culture in the context of which these actors carry out mathematical activity. Like other elements, textual forms are shaped and reshaped over time, illustrating the fluidity of mathematical cultures as I understand them. Texts of procedures contained in Chinese mathematical sources composed between the third century BCE and the seventh century CE provide evidence in support of these two general theses.

As with cuneiform documents, texts of "procedures (*shu* 術)" are a central component of these Chinese mathematical sources. This is notably made clear by the fact that the titles of mathematical works often refer to procedures and computation. Recent research on these texts of procedures allows us to go one step further in the description of how texts reflect the mathematical cultures in the context of which they were composed. Indeed, different Chinese mathematical sources comprise different types of text for procedures. Some of these sources contain only texts referring to operations directly, even though they do so in different ways, which can be associated with different social contexts (Zhu 2022; Zou and Chen 2022). Other sources have in addition texts of procedures prescribing operations indirectly, by stating the reason why these operations should be used: They are in this sense comparable to the cuneiform texts described by Høyrup, even though the ways in which reasons are formulated differ from those analyzed by Høyrup.

The Chinese sources that include the latter type of procedure are characterized by the fact that they also attest to another kind of texts: these texts group together in a single procedure the lists of operations necessary to process several cases to be distinguished in a given task (Chemla 2015). Accordingly, deducing from such a text the sequence of actions necessary to solve a given case requires practitioners to navigate within the text in a non-linear fashion, thus bringing into play an expert use of the list of operations. The fact that actors are interested in and shape texts with properties of this type echoes the emphasis they more broadly place on the epistemological value of generality. As a rule, the textual forms used by practitioners are not isolated facets of the mathematical culture in which they work, but present correlations with other facets of these cultures.

The two highlighted features of the text of a procedure (referring to operations indirectly and requiring a non-linear reading) signal sources communicating actions to be carried out to actors who have been acculturated in a certain way. These textual differences between texts of procedures – in other words, the different ways in which texts communicate about the actions to be performed – thus allow us to distinguish among Chinese mathematical sources different types, which reflect different mathematical cultures.

What is more, the mathematical sources in which we find texts characterized by the two features indicated above reflect transformations in the way of realizing these two features. Both the manuscript *Writings on Mathematical Procedures* (*Suanshu shu* 筭數書), written before about 186 BCE, and the classic *The Nine Chapters* (completed in the first century CE) along with its two layers of commentaries contain texts presenting these two features. However, the modes of indirect reference to operations and the nature of the navigation through texts of procedures evidenced by both documents attest to transformations that seem to illustrate actors' reshaping of the texts with which they worked on lists of operations (Chemla 2015, 2020). These observations illustrate how the elements with which actors carried out mathematical activity are shaped and reshaped in ways that are correlated with other features of the mathematical cultures in the context of which these elements were brought into play. We further see how taking this shaping and reshaping into account is essential for interpreting our written evidence.

So far, we have considered our written evidence from the viewpoint that its aim was to communicate knowledge to readers. As was emphasized above, the fact that historians and philosophers of mathematics commonly assume that this was the purpose of a text can be seen in the criticisms that these modern observers regularly formulate towards these writings. Indeed, the complaint that a text is unclear, or ambiguous, or that it has gaps, derives from the expectation that the text ought to carry out communication, and from the observation that it fails to do so adequately from an observer's viewpoint.

We have argued that not all texts communicate in the same way and that textual communication ought to be contextualized and historicized. When actors make a judgment of this kind, it gives us valuable insights to work on communication as a situated practice. However, is the assumption that the purpose of a text is to communicate legitimate for every written source? Considering this question leads us to wholly different functions of text and to the issue of the relationship between mathematical texts and practice.

3 Texts Written to Work With

Gottfried Wilhelm Leibniz's manuscripts are an excellent starting point to highlight other functions of texts for the practicing mathematician. This fact is brilliantly put forward in Knobloch's inquiry into the scholar Leibniz's "intellectual workshop."

Indeed, Knobloch (2004) explains how for Leibniz, "thinking was thinking in writing" and "text was his instrument of thinking" (p. 51). This is the second function fulfilled by texts for the actors whom we observe. Let us focus on it.

3.1 How Various Uses of the Page Highlight Different Dimensions of the Practice

Leibniz published very little in comparison with the work – notably the mathematical work – he carried out. His manuscripts attest to this fact. However, these documents do not only testify to results Leibniz had obtained without publishing them: they also shed light on his working practices. In particular, Leibniz's manuscripts contain many drafts that illustrate two main facets of his work.

First – and most importantly for us here – some of these drafts were written during the exploratory phases of his reflections on given questions. Drafts of this type attest to how Leibniz used the surface of the page – or of several pages – to work with, developing mathematical ideas and obtaining results through an interaction with the written traces of his exploration. The drafts show, for instance, how Leibniz took advantage of the space of a page to shape tables and other types of spatial layouts in order to investigate problems and look for regularities. In contrast with this "dynamic" use of tables, which exemplifies the processual use of the page (hence the subtitle of Knobloch's chapter: "text as process"), Knobloch identifies a "static" use, in which tables were for Leibniz a reference tool to "avoid repeated computations" (p. 64). The drafts also show how Leibniz drew on specific examples and how he prepared visualizations to form ideas and to think with these "paper tools" – to use Ursula Klein's (2003) apt expression.

The texts borne by these documents did not have the aim of communicating information to a reader, unless one considers the self as a reader. Perhaps in some contexts, drafts were shared with others. However, one might assume that this was not the prime objective of writings of this kind. If historians and philosophers of science were to read such drafts as giving an exposition on a given topic, we can imagine the resulting distortion in the interpretation. Instead, written evidence like this illuminates a research process in mathematics and hence notably mathematicians' practices when they are in a moment of relatively more open exploration. In particular, this type of written evidence sheds light on how in this context mathematicians use texts to think with.

Second, some of Leibniz's drafts were preliminary versions of a text intended for publication and sometimes successive versions of this text. In a sense, in contrast with the open-endedness of the previous ones, in which Leibniz appears to explore freely some mathematical issues, these drafts derive from the pursuit of a specific goal. Drafts of this kind thus shed light on the work process followed to reach this goal. They also draw attention to how the research process continues while preparing a text for communication.

The latter remark suggests that the two functions of texts that we have separated for analytical purposes interact in fact with each other. To put it differently, the different kinds of documents that we have considered in the previous section and now in this section specifically highlight different types of activity actors carry out using the space of the page. However, this contrast does not constitute a binary opposition: it simply sheds light on distinct dimensions of actors' practice that texts reflect. We return below to this issue.

3.2 The Private and the Public

Emmylou Haffner's work on Richard Dedekind's drafts highlights facets of the research process of the practicing mathematician that are strikingly similar to those brought to light by Knobloch in his study of Leibniz's drafts (Haffner 2019). She describes how Dedekind explores ideas in the context of examples, relying on the dynamic making of tables and other spatial devices that the surface of the page allows him to deploy. Haffner further shows how Dedekind shapes static tables and diagrammatic devices to perform computations and avoid repeating them. Interestingly, Leibniz and Dedekind both seek general laws and concepts, and the drafts show how they organize their search for the greatest generality possible. The parallels between these two mathematicians' working practices that their respective drafts allow us to draw suggest that Haffner's project to consider mathematical drafts in general might yield interesting results about regularities in practicing mathematicians' use of writing material in their research. This project might also help us move beyond apparent similarities to highlight personal idiosyncrasies as well as changing collective practices in the engagement with the page.

The key point for us here is that both cases show the central role played by texts that these mathematicians design at the same time as they struggle with the mathematical matter. A case in point clearly supports this conclusion. Haffner examines a case in which Dedekind identifies a symmetry in a corpus of identities and proofs that he develops (we now understand this symmetry as caused by duality). The drafts show how Dedekind shapes specific layouts and notations to observe and study this symmetry and how this work would lead him to thematize – to use Jean Cavaillès' notion – the results of his explorations with the introduction of the concept of *Dualgruppe*. In the previous section of this chapter, I had introduced the opposition between the mathematical knowledge communicated by a local, and that communicated by a global, reading of a text like Euclid's *Elements* or *The Nine Chapters*. Haffner's description of Dedekind's work on this symmetry suggests that this opposition is also valid for the texts with which mathematicians work.

However, Haffner organizes her approach to Dedekind's drafts in a way that also differs from Knobloch's, and this leads her to highlight another important issue with respect to our written evidence. Indeed, in contrast with Knobloch's global approach to Leibniz's drafts, Haffner (2019: 4–6) focuses on drafts relating to the eventual

publication of two research articles in which Dedekind introduces the notion of *Dualgruppe*. Her crucial observation is that in Dedekind's case, the mathematician's procedures in his drafts differ strikingly from the way he presents his results in his publications. Interestingly, Haffner (2021) nevertheless suggests that the same understanding of the epistemic and epistemological value of rigor can be perceived in both types of documents. A striking difference between them is that the exploratory drafts proceed by computation, while the publications put forward mainly a conceptual approach. This conclusion has important implications with respect to our written evidence, which is the focus of this chapter.

In Dedekind's case, Haffner argues, publications allow historians and philosophers of mathematics to observe only part of the mathematician's activity. The other side of Dedekind, which is mostly hidden in his drafts, somewhat invites a more nuanced understanding of his mathematical practice. From the viewpoint of the written evidence, we further see, first, that Dedekind's mathematical work brings into play two separate types of text – one of which is foreign to the purpose of communicating knowledge – and, second, that in his case the distance between the two is radical. The same phenomenon holds true more generally. This can be illustrated from another perspective with the example of Claude Chevalley's lectures to his university students, which proceeded in a highly formal way. Once, Dominique Dumont recalls, Chevalley could not remember a proof. He then moved to a corner of the blackboard and drew a diagram for himself, hiding it from the audience; indeed, a diagram of this kind did not fit the rules according to which he thought he should make a presentation. The diagram – clearly a part of his private mathematical practice – helped him recover the proof. He thus erased it and went on with the formal presentation.[4] Whether the publication of a mathematical exposition is written – as in Dedekind's articles and books – or (semi-)oral, as in Chevalley's lectures, in both cases, the private and the public practices obey different norms, and the resulting writings fall under distinct categories. Is this distance the same for all practicing mathematicians? Or, to put it differently, does the boundary between drafts and publications vary according to the authors and the collectives? Two examples suggest that this distance varies and hence that in other cases, publications reveal sometimes more of the research process than in Dedekind's case.

3.3 Intertwinement Between Exploring and Communicating Knowledge

Unsurprisingly, the first example can be found in Euler's work. I will illustrate this statement using the texts on spherical trigonometry he published between the 1750s and the 1780s (Chemla 2004). In this context, Euler, like Dedekind, encounters a symmetry between statements – and also between proofs of these statements – that we also understand today as an effect of duality.

[4] Dominique Dumont (1947–2016), ULP, Strasbourg, personal communication.

The first substantial memoire that Euler devoted to this topic, in 1753, aims to derive all the formulas of spherical trigonometry on the sole basis of the newly developed calculus of variations. What is striking is that Euler first introduces notations with which he will perform his analytical work. After pages of computations, he has obtained enough results to achieve his goal. He thus shifts to another part of the memoire, in which he will establish that any problem of spherical trigonometry can be solved using the formulas found in the first part. To this end, at the juncture between the first and the second parts – that is, in the middle of the memoire – Euler carries out a complete change of notations (Chemla 2004: CLIX). None of the previous notations is kept. However, the new notations are correlated with a textual phenomenon that affects the second part of the memoire.

In brief, these new notations highlight the symmetry between sides and angles that derives from duality. The dual formula of a given formula – as well as the dual proof – can thus be obtained by a simple and systematic translation of the text of the original formula, and of that of its proof. Accordingly, from this moment of the memoire on, Euler adopts a presentation that proceeds in a way that highlights the symmetry between dual problems, dual formulas, and dual proofs. As above, only a reading of the structure – that is, a reading of a global type – conveys the meanings expressed in this way. Interestingly, in previous articles in which Euler had dealt with spherical trigonometry, he had used other notations.

Consequently, the introduction of these new notations correlates with the fact that Euler deploys a project on spherical trigonometry as a whole: this gives him the opportunity to think about properties that are apparent at a global level. The shaping of notations and of symmetrical parts of text are resources that he creates to this effect. At the end of the eighteenth century, this very symmetry is highlighted in many works on this field, using the same type of textual resources. In other words, readers grasped the point of the new notations, and they carried out the global reading whose meaning I have just emphasized. Accordingly, Euler's notations and type of text would impose themselves in spherical trigonometry.

The fact that Euler introduces the new notations in the middle of the memoire suggests that he is struck by the symmetry, once he has obtained a collection of formulas, and that he changes notations to focus on it and work with it. This remark seems to indicate that the mathematical research presented in the memoire unfolds as Euler writes. In this case, the distance between the draft and the publication appears to be minimal. Noteworthy is the fact that Dedekind carries out a similar change of notation, in the context of his reflection about the generality of the theory he develops. However, he carries out this change in a draft, not in a publication (Haffner 2019: 14).

Euler's last contribution to spherical trigonometry, in 1781, supports the conclusion that he carried out research while writing the text to be published and that the text was a support for his research (Chemla 2004: CLXV *ff.*). The point of this 1781 article is to establish spherical trigonometry on the basis of elementary geometry and to show that all of its formulas derive from three and only three of them. We will limit ourselves to a single remark here. At the outset, Euler introduces the notations that highlight the symmetry. In the article, he tacitly draws on these notations to

derive dual formulas on the basis of dual proofs. The key remark is that, in one of these proofs, he makes a mistake, which is also to be found in the dual proof (Chemla 2004: CLXXII). This implies that Euler actually produced the dual proof and the dual formula by a step-by-step translation of the original proof and the original formula, hence relying on his knowledge of the symmetry and on his notations to produce the text of his article. In other words, Euler is concretely working with the text he writes to generate mathematical knowledge, and this is the very text that would eventually become the publication. Working with a text and communicating results are in this case intimately connected.

3.4 Another Case of Intertwinement Between Exploring and Communicating Knowledge

As shown by Anne Robadey (2006, chap. 2; 2015), exactly the same conclusions seem to hold true for at least some of Henri Poincaré's publications. Robadey obtains these conclusions on the basis – notably but not only – of her study of the way in which Poincaré proceeds to the enumeration of cases in one of his earliest writings on differential equations from 1881.

She highlights that, when Poincaré distinguishes between types of differential equations, his publication presents cases according to a decreasing level of generality. Poincaré's assessment of this degree of generality remains tacit. It can be recovered precisely from the structure of the publication and especially from how Poincaré conducts the enumeration and deals with the successive cases thereby distinguished. Robadey is thus able to make clear the mathematical knowledge that is involved in the way Poincaré proceeds. This knowledge is not the object of a discursive exposition, but the historian can grasp it using clues – which, in this context, requires notably a reading of the publication of a global type.

What is striking is that the way of enumerating illustrated by these writings recurs in other publications by Poincaré that are devoted to different topics. It thus seems to reflect a working procedure that Poincaré regularly put into play in his mathematical work, his exploration of a situation under study repeatedly proceeding from the most general cases to the "special" or "exceptional" ones. Poincaré thus works mathematically with enumerations, and he does so in a specific way. His texts keep a trace of his working process. This remark suggests that the distance between the draft and the public text is, as it was for Euler, minimal. In other words, in these cases at least, his texts were not only meant to present knowledge to readers: they also constituted a key resource with which to work. In particular, they allowed him to carry out these enumerations and deploy their different branches. This hypothesis explains why publications contain clues of the mathematician's work: this work was precisely carried out through an engagement with the text with which he develops his reflection.

Two remarks show that in these situations the structure of the text reflects something fundamental about Poincaré's mathematical work. First of all, the

cases Poincaré considers exceptional – those that are dealt with last in the enumerations – do not receive the same full treatment as the others. In other words, Poincaré first addresses the cases whose complete consideration is in his view essential for a satisfactory resolution of the question under study. In terms of mathematical practice, the type of solution he aims for is correlated with the way he proceeds with the enumerations.

Secondly, the recurring emphasis on those cases that are general enough to suffice for the treatment of a problem and thus those that require the mathematician's attention can also be correlated with mathematical innovations introduced later by Poincaré. Indeed, Robadey (2006, chap. 3; 2016) brings to light how Poincaré was able to correct his winning but erroneous submission for the prize offered in 1888 by King Oskar II of Sweden on the three-body problem, by introducing a new type of mathematical statement. Instead of claiming that his original assertion was universally true – as he incorrectly did in the submission – in his corrected version Poincaré introduced a different statement, which he maintained was true for all trajectories except a set of probability zero.

Statements asserting something to be true for "almost all" objects under consideration have become widespread in mathematics. Robadey argues that these were probably the circumstances under which they were introduced. She also describes the work on probabilities that Poincaré developed in order to put into play this new way of asserting – and also proving – the generality with which a statement holds. The type of statement and the knowledge developed to make its utterance meaningful are clearly correlated with a working procedure assessing systematically the degrees of generality of different cases that might occur: it seems that here, we witness how Poincaré's working procedure, to which the enumerations in his publications testify as early as 1881, was eventually thematized in the form of the possibility offered to the mathematician to consider new types of statements.[5]

To put it differently, the mathematician's results and in particular his conceptual innovations are intimately correlated with his working practice. Reading in texts how they reflect the mathematician's work with these texts highlights facets of the mathematical work that illuminate the production of mathematical concepts, statements, and results. In brief, unlike Dedekind, for whom only the drafts shed light on the work engaged in the observation of the symmetry permeating his object of study, Poincaré's publications reflect part of the work process, which can be recovered using clues. In the terms introduced above, his publications mirror the different dimensions of his mathematical practice.

To conclude, we thus see the diversity of documents that make up our written evidence and also the key reasons why historians and philosophers of mathematics must deploy various types of approach to this evidence. As I have emphasized above, it is not enough to read what we perceive that texts present: we must also learn to grasp clues in texts and documents and develop methods to interpret them. But there is more.

[5] Robadey (2006: 216–220) also discusses the differences between these various works.

4 Contexts and Environments

So far, we have considered actors' work with texts in the context of practitioners' explorations of new territories. However, part of our written evidence reflects other types of activity for which actors engage with textual artefacts in their mathematical work.

4.1 Learning Mathematics with Texts

Let us think, for instance, of the mathematical documents produced in the context of scribal schools from Nippur, during the Old-Babylonian period – early second millennium BCE. They exemplify the type of documents produced in all kinds of learning contexts. Christine Proust (2007) has shown how these Nippur documents – clay tablets inscribed with a stylus – attested to different moments and different facets of the learning process and how they reflected the part devoted to mathematics in the global curriculum.

Some of these documents testify to the learning of the lists of fundamental quantities, into which any given quantity could be decomposed. Proust refers to these lists as "metrological lists." The tablets in question attest to how scribes first reproduced a master's copy written to the left of the tablet and also to how, once said scribes had memorized the lists, they then wrote them on their own.

Other documents reflect the same process of textual learning with the texts to which Proust refers as "metrological tables." These tables contain in the left column lists of the type described above and in the right column the number written in a floating sexagesimal place-value notation associated with the respective quantity written to its left. Copying the master's copy and then writing the memorized table inculcated the difference between two concepts of number, respectively written in the two columns of the layout: one of them (to the left) was used in the expression of quantities, whereas the other (to the right) was used in computations – mainly those of a multiplicative type (Proust 2008). This was the first step in the acquisition of a system of computation to which tablets produced later in the learning process attest.

Particularly significant in this respect is a group of square-shaped tablets, each containing an exercise asking to determine the area of a square that the apprentice scribes had to solve (Proust 2007: 190–196). In the same way as the metrological tables oppose two types of numbers, the tablets under consideration attest to the inculcation of the separation between two kinds of moments, each linked to a type of number, in the resolution of the exercise.

The exercise gave the apprentice scribe the length of the side of the square, asking to determine its area. This length, expressed as a quantity of the type found in the left column of a metrological table for lengths, was written in the lower right corner of the tablet, together with the question and eventually the answer expressed in the same way. The numbers inscribed in the top left corner of the tablet attest to how the actual operation was carried out.

The length was transformed into a number written using the sexagesimal place-value system by means of the metrological table for lengths. The result of this conversion was inscribed twice into the top left corner of the table, and the multiplication was then performed using the numerical table for multiplication with sexagesimal place-value numbers. The result of the operation – a sexagesimal place-value number inscribed in the top left corner of the tablet – was then transformed, using a metrological table for surfaces, to find the area, expressed as a quantity (of the type occurring in a metrological list), which had then to be written in the lower right corner. In other words, a layout was shaped for apprentice scribes to work with, and this layout is correlated with the type of mathematical knowledge inculcated.

These documents highlight many general points about written evidence that are worth discussing.

First of all, clearly scribes learnt the system of computation through the inscription of tables, and then that of the solution of exercises, which they had to write using fixed layouts. The system was inculcated into them using a specific engagement with the surface of the tablet, probably under the supervision of a master. These texts were therefore in no way written with the intention of presenting knowledge to readers: some were produced for personal training, while others were presumably submitted to someone more authoritative, as proof of a successful appropriation of knowledge. These sources illustrate from a new angle the fact that practices of sharing written productions change according to context. Whoever read the document, the work to be performed with the text aimed to impart a specific body of knowledge through what we might call an "embodiment." We have encountered the use of all kinds of layout in drafts in the context of the exploration of new questions. We see again here specific layouts being used for other purposes. In both cases, these layouts give clues on actors' practices and concepts. They also shed light on the cognitive work carried out when engaging with mathematical writings and textual forms.

Secondly, as Proust (2007) emphasizes, what is at stake in the study of this written evidence is that these tablets shed light on practices that were taught to apprentice scribes: These practices were later taken for granted and hence no longer explicitly mentioned in the more advanced texts on which historians of mathematics have tended to concentrate.[6] Studying the former tablets hence gives us tools to better interpret the latter. Seen from another angle, with Dedekind's case, we had emphasized above how different contexts for mathematical activity and different types of moment are reflected by the fact that they leave different types of written different types of written evidence. The same conclusion recurs here, even though it takes on a different meaning, with the opposition between the activity of learning at school and that of producing more advanced mathematics. This remark again highlights what is at stake in the development of a historicized and contextualized approach to our mathematical documents.

[6] Ehrhardt (2008) makes use of a similar approach to situate Galois's research publications against the background of the training that he received.

4.2 How Texts Reflect Material and Social Environments

The remark applies notably to learning environments, which, in different contexts, have left us documents of different types. The scribes who learnt in the Old-Babylonian schools from Nippur and the masters who taught there do not seem to have produced lecture notes – at least not of the type that are specific to later contexts. The lecture notes surviving from other educational contexts thus seem to attest to different practices of teaching. These are only some examples of a more general phenomenon that would require that more research be devoted to it: texts and documents shed light on the social environments and more broadly the kinds of contexts in which they were produced.

To return to lecture notes, they also testify to how speakers as well as listeners engaged with the page, for the production of oral performances as well as for their assimilation of the knowledge thereby dispensed.[7] Some of these documents were meant for a quite limited circulation, others not. Actors considered some of these documents as worth keeping – for further personal use or to be passed on to future users. By contrast, others that actors employed were meant to be ephemeral, and so it is somehow miraculous that some of these ephemera survived. For drafts like Dedekind's, this is in particular related to the establishment of institutions like archives. For school documents like the Nippur texts discussed above, this is due to the fortuitous contiguity between the material used for writing and that used for construction works. When ephemera were produced with calculating rods – as in mathematical practices in ancient China[8] – in dust, as in the Arabic world,[9] or on a blackboard, they were destined by nature to disappear.[10] We return to these ephemera below.

Whatever the case might be, the various types of documents produced in related contexts and environments that we have distinguished present several forms of contiguity with each other. Our last general remark inspired by the school tablets from Nippur will bear precisely on this contiguity. It is not surprising that drafts and publications might display continuities with each other and that school documents might present correlations with mathematical writings produced outside a teaching context. However, the case of the square-shaped tablets from Nippur highlights another form of contiguity that will prove to have a certain generality. As I have explained, the solution to the exercise displayed on each of these tablets requires the bringing into play of other texts, which are not explicitly mentioned on these tablets, notably metrological tables. At the moment of their training when scribes dealt with

[7] Bruter (2008, 2012), Bustamante de la Ossa (2019).

[8] See, for instance, Chemla (2010, 2015).

[9] See, for instance (Saidan 1978).

[10] The introduction of the blackboard in teaching practices is dealt with in (Belhoste 2003: 322–323; d'Enfert 2014). As for the blackboard in research practices in mathematics, see Greiffenhagen (2014).

the computation of areas of this kind, these texts had been inculcated into them. Accordingly, they put these tables into play in the solution of the exercise.

Put differently, the production of a tablet involves the use of other texts – some in written form, others known by heart – which, in the case of our exercise, can be detected only through clues. This is a general phenomenon: the mathematical work in relation to which a given document is produced brings into play other texts. Some of them pre-exist to the making of said document. Others are designed while producing the document. This textual environment is reflected in the document produced. Sometimes – especially in modern times – reference to these other texts is made explicit, and even follows fixed norms, as in bibliographies and footnotes.

The type of contiguity between the document produced and the textual environment to which it adheres remains an important yet understudied research topic. However, the documents and inscriptions present in the background and put into play to produce a given text are not always the object of specific reference, as is clear from the square-shaped tablets from Nippur. Again, in such cases, historians can rely only on any clues they may find, to explore the nature and extent of the textual environment. This raises important methodological issues: which methods are available to historians and philosophers of mathematics to recover this environment and understand how actors worked with it? And also: what kind of contiguities does this environment present with the text written in relation to it? The case of mathematical texts from Early imperial China is quite illuminating on these issues. We will now focus on it, which will require that we return to ephemera of a certain type.

4.3 Ephemera

Most Chinese mathematical texts written between the third century BCE and the seventh century CE are composed of problems and of procedures solving them. The key point for us is that texts of procedures refer to inscriptions formed with calculating rods, which are in general neither explicitly described nor represented in the texts. In particular, as a rule, the procedures tacitly imply the use of calculating rods to write numbers, and sometimes they prescribe placing those numerals in specific locations, without making clear the nature of the space used for that purpose. These inscriptions were, at the time, by nature ephemeral, and yet the texts of procedures were written in contiguity with them. What can we say about this contiguity and why does it matter for historians and philosophers of mathematics? Clues that the mathematical works contain enable us to restore, at least partly, some of the features of these ephemeral inscriptions and thus to address these questions.

To begin with, at a very elementary level, procedures prescribe computations to accomplish a task illustrated by a mathematical problem. The execution of these prescriptions seems to have been carried out by means of inscriptions written with calculating rods, since the texts of procedures regularly refer to "placing" the representation of the numbers before operating on them. The complete procedure is therefore made up of the text recorded in the mathematical document and of the actual computations that put this text into play outside the written document. I will

refer to a configuration of this kind by suggesting that the text has an "aura," in which we find in particular inscriptions with calculating rods. As a result, in line with the fact that texts of procedures shed light on the inscriptions that accompany them, the restoring of inscriptions regularly enables us to interpret these texts of procedures. This would suffice to justify the importance of considering the material environment of a text, and its contiguity with the written evidence, even when this environment consists of ephemera. But there is more.

The inscriptions that the texts of procedures allow us to restore attest to facets of actors' knowledge that the works do not explicitly address. For instance, the clues that *The Nine Chapters* contains about how, at the time, calculating rods were used to represent numbers and execute operations indicate that actors employed a decimal place-value system comparable to the one we use today, when we write 123 to represent one hundred and twenty three. However, at the time, this system was only written on ephemera. We thus see the importance of recovering these inscriptions.

This is not an isolated phenomenon: Exactly the same conclusion can be drawn about the earliest extant Sanskrit and Arabic texts that deal with a decimal place-value system.[11] Incidentally, this remark illustrates a key feature of these inscriptions: they were not written in an informal way, but followed strict norms. It might be tempting to assimilate drafts and ephemera: This would be misleading and obscure our reflection on the written evidence available to us. In particular, drafts can be quite casual. This does not seem to have been the case at all for the inscriptions formed with calculating rods. In line with the use of a decimal place-value system, positions on the calculating surface were put into play for mathematical work according to strict rules. From the viewpoint of the constraints ruling over their arrangement and transformation, these inscriptions should rather be compared with the formulas on which modern mathematical practice relies.[12]

Further, restoring not only the inscriptions that are drawn in relation to a text but also how they change throughout the computations prescribed by the text of the related procedure allows us to discover and describe actors' mathematical practices with these inscriptions, on which the works cast light (Chemla 1994, 2020). The practices in turn reflect actors' mathematical work. This is all the more crucial as, in some cases, knowledge of these practices enables us to perceive historical changes and circulations that the texts alone would not make clear (Chemla 1994). This fact has a key implication. Our first remark about these inscriptions tended to suggest that, from a temporal viewpoint, the inscriptions with calculating rods were subsequent to and derived from the composition of the text: the prescription would come first, and the execution only second. However, if, as it seems, actors carried out a mathematical work with the inscriptions that is reflected in the text, presumably the write-up of the text and the inscription of the calculating rods were more intertwined. This feature also distinguishes the draft – which constitutes a pre-text – from these inscriptions.

[11] Chemla, in dialogue with Keller and Proust (2022).

[12] Chemla, in dialogue with Keller and Proust (2022).

4.4 Inscriptions as Spatial and Dynamic Entities

Finally, the case of the inscriptions with calculating rods to which Chinese mathematical documents refer is most interesting insofar as, after centuries during which these inscriptions remained purely ephemeral, from the tenth century at the latest, representations of these inscriptions were included in the pages of works (Chemla 2010). In other words, what was once contiguity between two types of textual work now took the shape of the embedding of the two practices into different zones of the same page. In fact, probably, actors continued to use calculating rods in addition to the page. However, now, writings displayed part of the work carried out on ephemera. This key shift in the nature of our written evidence enables us to perceive that the norms and practices that governed ephemera show strong continuity with those to which later mathematical writings attest. In addition, the concepts elaborated and the mathematical work carried out with ephemeral inscriptions continued to be further developed when inscriptions were included in the pages of the writings.

These remarks highlight why a long-term historical approach to mathematical activity in ancient China requires that historians and philosophers of mathematics strive to restore the ephemera used by practitioners of mathematics for their mathematical work – as can be done using clues in the mathematical writings – and even that they include these inscriptions in the corpus of written evidence. Similar conclusions can be drawn for cases in which mathematical activity is carried out only with ephemera. I argue that the mathematical explorations performed by actors in the Trobriand Islands with string figures are of this type. Eric Vandendriessche (2015) has shown why we might consider that the corpus of string figures, and of procedures to generate them, attests to an activity at least partly of a mathematical type. In this case, finger operations with strings and the resulting figures are the only material ephemera testifying to this activity, if we set aside the works of anthropologists who attempted to keep traces of these practices.

Why, one may ask, were some inscriptions that appear crucial for practitioners of mathematics apparently kept purely ephemeral for so long, and even sometimes until now? It is impossible to answer a question of this kind with certainty. However, one may note that what was inscribed on and with ephemera such as those we have described was spatial and dynamic: actors represented in this way processes – in these cases, processes of computation or of drawing. Dynamic inscriptions seem to play a specific role in mathematical activity. This is also the conclusion that Knobloch (2004) draws from the study of Leibniz's manuscripts, which he eloquently titles "text as process." Perhaps the same conclusion could be drawn to explain the persistence of the use of the blackboard among practitioners of mathematics – working mathematicians, lecturers, as well as teachers – at a time when the practitioners in other disciplines are mostly using projections of material prepared beforehand (Greiffenhagen 2014).

More generally, the Chinese writings on which the previous developments rely highlight another issue: mathematical work is carried out, more broadly, in a given material environment, which does not always consist only of documents and whose components are not always illustrated in the pieces of evidence written in this

context. For instance, Chinese works written before the eighth century refer to diagrams and blocks. Both seem to have been material objects, at least before the tenth century, and they were not represented in the texts composed before the eighth century that were handed down. However, restoring them is nevertheless essential for the interpretation of the texts written in contiguity with them. To use the expression introduced above, these material objects belong to the "aura" of the text written in relation to them. Moreover, after the tenth century, exactly as was the case of the inscriptions with calculating rods, diagrams began being included in the pages of writings, and they display striking continuities with the diagrams to which the earlier texts refer (Chemla 2010). In other words, contiguity needs not be considered only with respect to a textual environment – as I have argued above – but more broadly with a material environment.

We had evoked circulation of types of inscriptions and practices between drafts and published texts and between school texts and more advanced texts. We now see that the synchronic and diachronic circulation between different types of support needs to be considered more broadly to understand, in an appropriate textual and material context, the production of the texts that form our written evidence. How the elements of the textual and material environments are put to work to produce another document is a complex question, which would require a historical study of the reading of mathematical documents and inscriptions. We will leave it here.

5 Conclusion

Clearly, to treat the written evidence available to historians and philosophers of science as solely having the aim of communicating concepts, ideas, and results makes us run the risk of a rather narrow approach to these texts, even for published works. Indeed, texts like documents contain clues about the mathematical work carried out in producing them as well as about the environment in which this work was undertaken. In other words, seen from this perspective, our written evidence gives us insights into actors' mathematical practice in context. However, it does so according to modalities that depend in part on the context and the practices of the actor observed. Indeed, as we have seen, publications by Euler, Dedekind, and Poincaré do not reflect the work from which they derive in the same way. What is more, documents produced in different contexts and at different moments of mathematical work yield clues of different kinds. Whatever the case might be, I suggest that, in order to deploy a full approach to mathematical practices, historians and philosophers of mathematics might benefit from a reflection about the nature of the written evidence on which they rely.

In this chapter, I have foregrounded a key type of mathematical practice: actors' engagement with texts and documents. These reflections lead me to a more general remark about how writing is often discussed. In *The Domestication of the Savage Mind* (whose purpose seems to be better rendered using the title of the 1979 French translation, *La raison graphique*), Jack Goody (1977) put forward an argument that has had a tremendous impact. He argued that societies in which writing was

available were able to shape and work with new cognitive technologies, like lists and tables, in such a way that their members were able to develop bodies of knowledge more elaborate than those in other societies. In this proposition, writing referred to the writing down of speech, the existence of which, in Goody's view, brought about a major shift for the societies in which this technology was developed and adopted.

However, what we have seen above suggests an inflection of this argument. Indeed, if we take the example of the ephemera to which ancient Chinese mathematical texts refer, we see that their inscription was independent from the writing of the Chinese language and that mathematical bodies of knowledge were developed with inscriptions of this kind – notably a decimal place-value notation.[13] More generally, tables and other types of layout that abound in drafts do not all seem to depend on the writing of utterances. The study of the written evidence that attests to mathematical activity thus invites us to rethink how we approach writing. In particular, it questions the assumption that the writing of speech would be the basis for all the other uses of writing. The latter assumption might well be another way of asserting that communication was the primary function of writing in general. However, I have argued how limited this view is.

By contrast, a reflection about the written evidence available to historians and philosophers of mathematics cogently suggests that the writing of utterances is only one of many writing practices and that other writing practices have been essential to the development of mathematics. Writing may have played a central role in mathematics primarily insofar as practitioners shaped the types of text with which they could work for the issues they were addressing. Such an assumption would easily account for a patent phenomenon: mathematical texts have changed constantly in the course of history, and they have done so much more substantially than the technologies of writing speech might explain.

References

Austin JL (1962) How to do things with words. The William James lectures delivered at Harvard University in 1955. Clarendon Press, Oxford

Belhoste B (2003) La Formation d'une technocratie. L'Ecole polytechnique et ses élèves de la Révolution au Second Empire. Belin, Paris

Bruter A (ed) (2008) Le cours magistral, XVe-XXe siècles. 1. Publics et savoirs. Special issue of Histoire de l'éducation 120. Online at: https://journals.openedition.org/histoire-education/1035

Bruter A (ed) (2012) Le cours magistral, XVe-XXe siècles. 2. Le cadre institutionnel et matériel, Special issue of Histoire de l'éducation 130. Online at: https://journals.openedition.org/histoire-education/2324

Bustamante de la Ossa, MC (2019) À l'aube de la théorie des quanta. Notes inédites d'Émile Borel sur un cours de Paul Langevin au Collège de France (1912–1913). Brepols, Turnhout

Chemla K (1994) Similarities between Chinese and Arabic mathematical writings. I. Root extraction. Arabic Sciences and Philosophy A Historical Journal 4:207–266

[13] Chemla, in dialogue with Keller and Proust (2022).

Chemla K (1999) Compte rendu de Un théorème de Fermat et ses lecteurs. Gazette des Mathématiciens 79:121–122. https://smf.emath.fr/system/files/filepdf/Gaz-79.pdf
Chemla K (2004) Euler's work in spherical trigonometry: contributions and applications. In: Radelet-de Grave P, Speiser D (eds) Euler. Opera Omnia. Commentationes physicae ad theoriam caloris, electricitatis et magnetismi pertinentes. Appendicem addidit Karine Chemla. Birkhäuser Verlag, Basel, pp CXXV–CLXXXVII
Chemla K (2008) Classic and commentary: an outlook based on mathematical sources, Preprint/Max-Planck-Institut für Wissenschaftsgeschichte, vol 344. Max-Planck-Institut für Wissenschaftsgeschichte, Berlin. http://www.mpiwg-berlin.mpg.de/Preprints/P344.PDF
Chemla K (2010) Changes and continuities in the use of diagrams tu in Chinese mathematical writings (3rd century-14th century) [I]. EASTS East Asian Sci Technol Soc Int J 4:303–326
Chemla K (2012) Prologue: historiography and history of mathematical proof: a research program. In: Chemla K (ed) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge, pp 1–68
Chemla K (2015) Describing texts for algorithms: how they prescribe operations and integrate cases. Reflections based on ancient Chinese mathematical sources. In: Chemla K, Virbel J (eds) Texts, textual acts and the history of science. Springer, Dordrecht, pp 317–384
Chemla K (2020) Comment l'activité mathématique façonne sa langue et ses formes textuelles. In: Robert J-N (ed) Langue et science, langage et pensée. Odile Jacob, Paris, pp 283–311
Chemla K (in dialogue with Agathe Keller and Christine Proust). 2022. Cultures of computation and quantification in the ancient world: an introduction. In: Chemla K, Keller A, Proust C (eds) Cultures of computation and quantification in the ancient world. Numbers, measurements, and operations in documents from Mesopotamia, China and South. Dordrecht: Springer, pp 1–140
Chemla K, Fox Keller E (eds) (2017) Cultures without culturalism: the making of scientific knowledge. Duke University Press, Durham/London
Chemla K, Virbel J (eds) (2015) Texts, textual acts and the history of science, Archimedes. Springer, Dordrecht
Chemla 林, Karine 力娜, Zou 鄒 Dahai 大海 (2018) Parts in Chinese mathematical texts. Interpreting the chapter form of the nine chapters on mathematical procedures. In: Bretelle-Establet F, Schmitt S (eds) Pieces and parts in scientific writings. Springer, Dordrecht, pp 91–133
d'Enfert R (2014) Les objets de l'école, XIXe-XXe siècles. Une approche matérielle de la culture scolaire. In: Condette J-F, Figeac-Monthus M (eds) Sur les traces du passé de l'éducation… Patrimoines et territoires de la recherche en éducation dans l'espace français. MSHA Maison des Sciences de l'Homme d'Aquitaine, Pessac, pp 149–162
Ehrhardt C (2008) Evariste Galois, un candidat à l'école préparatoire en 1829. Revue d'histoire des mathématiques 14:289–328
Ehrhardt C (2017) E Uno Plures? Unity and diversity in Galois theory, 1832–1900. In: Chemla K, Fox Keller, E (eds) Cultures without culturalism in the making of scientific knowledge. Duke University Press, Durham, pp 327–351
Goldstein C (1995) Un théorème de Fermat et ses lecteurs. Presses Universitaires de Vincennes, Saint-Denis
Goody J (1977) The domestication of the savage mind. Cambridge University Press, Cambridge
Greiffenhagen C (2014) The materiality of mathematics: presenting mathematics at the blackboard. Br J Sociol 65:502–528
Haffner E (2019) From modules to lattices: insight into the genesis of Dedekind's Dualgruppen. Br J Hist Math 34:23–42. https://doi.org/10.1080/17498430.2018.1555928.hal-02426462
Haffner E (2021) The shaping of Dedekind's rigorous mathematics: what do Dedekind's drafts tell us about his ideal of rigor? Notre Dame J Formal Logic 62:5–31. https://doi.org/10.1215/00294527-2021-0001
Høyrup J (1990) Algebra and naive geometry: an investigation of some basic aspects of old Babylonian mathematical thought. Altorientalische Forschungen 17:27–69, 262–324
Høyrup J (2002) Lengths, widths, surfaces. A portrait of old Babylonian algebra and its kin. Springer, Berlin/London

Keller A (2006) Expounding the mathematical seed. A translation of Bhaskara I on the mathematical chapter of the Aryabhatiya. Birkhaeuser, Basel
Keller A (2010) On Sanskrit commentaries dealing with mathematics (fifth–twelfth century). In: Bretelle-Establet F (ed) Looking at it from Asia: the processes that shaped the sources of history of science. Springer, Dordrecht, pp 211–244
Keller A (2015) Ordering operations in square root extractions. Analyzing some early medieval Sanskrit mathematical texts with the help of speech act theory. In: Chemla K, Virbel J (eds) Texts, textual acts and the history of science. Springer, Cham, pp 183–218
Klein U (2003) Experiments, models, paper tools: cultures of organic chemistry in the nineteenth century. Stanford University Press, Stanford
Knobloch E (2004) Leibniz and the use of manuscripts: text as process. In: Chemla K (ed) History of science, history of text. Springer, Dordrecht, pp 51–79
Lloyd GER (1990) Demystifying mentalities, Themes in the social sciences. Cambridge University Press, Cambridge/New York
Netz R (1999) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press, West Nyack
Plofker K (2009) Mathematics in India. Princeton University Press, Princeton
Proust C (2007) Tablettes mathématiques de Nippur: reconstitution du cursus scolaire. Edition des tablettes conservées à Istanbul. Translittération des textes lexicaux et littéraires par Antoine Cavigneaux. Préface de Christian Houzel, Varia Anatolica XVIII. IFEA, De Boccard, Istanbul
Proust C (2008) Quantifier et calculer à Nippur. Rev Histoire Math 14:5–72
Robadey A (2006) Différentes modalités de travail sur le général dans les recherches de Poincaré sur les systèmes dynamiques. PhD University Paris 7 Paris Diderot. http://tel.archives-ouvertes.fr/tel-00011380/, http://tel.archives-ouvertes.fr/tel-00011380/
Robadey A (2015) A work on the degree of generality revealed in the organization of lists: Poincaré's classification of singular points of differential equations. In: Chemla K, Virbel J (eds) Texts, textual acts and the history of science. Springer, Dordrecht, pp 385–419
Robadey A (2016) Elaboration of a statement of the degree of generality of a property: Poincaré's work on the recurrence theorem. In: Chemla K, Chorlay R, Rabouin D (eds) The Oxford handbook of generality in mathematics and the sciences. Oxford University Press, Oxford, pp 169–222
Saidan AS (1978) The Arithmetic of al-Uqlidisi. The Story of Hindu-Arabic Arithmetic as told in Kitab al-Fusul fi al-Hisab al-Hindi, by Abu al-Hasan Ahmad ibn Ibrahim al-Uqlidisi, written in Damascus in the year 341 (A.D. 952/3). D. Reidel, Dordrecht
Vandendriessche E (2015) String figures as mathematics ? An anthropological approach to string figure-making in Oral tradition societies, Studies in history and philosophy of science 36. Springer, Dordrecht
Zhu Yiwen 朱一文 (2022) Another culture of computation from 7th century China. In: Chemla K, Keller A, Proust C (eds) Cultures of computation and quantification in the ancient world. Numbers, measurements, and operations in documents from Mesopotamia, China and South Asia. Springer, Dordrecht, pp 555–602
Zou Dahai 鄒大海, Chen Wei 陳巍 (2022) The characteristics of mathematical methods in the Wu cao suanjing and its social background. In: Chemla K, Keller A, Proust C (eds) Cultures of computation and quantification in the ancient world. Numbers, measurements, and operations in documents from Mesopotamia, China and South Asia. Springer, Dordrecht, pp 603–645

What Happens, from a Historical Point of View, When We Read a Mathematical Text?

Lucien Vinciguerra

Contents

1 Introduction 3074
2 The Reading Process 3076
3 The Partition of Truth and Falsehood 3079
4 The Partition of Language and Image 3082
5 The Partition of the Mathematical and the Non-mathematical 3084
6 The Case of the Equation in Descartes' *Geometry* 3086
7 Conclusion 3097
References 3098

Abstract

The history of mathematics can be read in two ways. On the one hand, unlike the history of physics, it does not proceed by conjectures and refutations. New theories rarely refute old theories, but give them new foundations, generalize them, and reinterpret them through new concepts. This reading is unifying, highlighting the unity of the history of mathematics from its origins, through the permanence of its truths. On the other hand, many contemporary historians of mathematics have insisted on the diversity of methods, objects, and practices throughout history. They have shown that the objects of the past mathematics are not the same as those of today. That second reading shatters the illusion of unity. It disjoins what the first reading unifies.

Rather than deciding between these two readings, this chapter examines what makes them both possible and legitimate, based on an analysis of the acts of reading that we have to perform when confronted with the pages of a mathematical text, its signs, diagrams and images, and the different language modalities it employs. Starting with a discussion of Husserl's theses in The Origin of Geometry on the role of writing in the history of mathematics, and using a number of

L. Vinciguerra (✉)
Université de Lille UMR 8163 – STL – Savoirs Textes Langage, Lille, France
e-mail: lucien.vinciguerra@univ-lille.fr

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_80

historical examples, the chapter shows that any reading of a mathematical text involves at least three divisions, three delimitations made by the reader's gaze: divisions between the true and the false, between language and image, and between what is inside mathematics and what is outside. For the same text, these divisions are traced by the gaze in different ways throughout history.

The chapter develops an example in greater detail, concerning the status of the equation in Descartes' Geometry. It discusses the contrasting recent readings of the equation by two historians of science, Henk Bos and Enrico Giusti. It shows the concepts of The Geometry and the ideas the philosopher develops on reading and the nature of the sign. It shows how these ideas can be used to reconstruct how Descartes might have read a mathematical text, and it explains how successive readings of The Geometry can at the same time insert it into the mathematics of later centuries.

Keywords

Husserl · Peletier du Mans · Descartes · Pascal · Bachelard · Quine · Davidson · Mathematical sign · Henk Bos · Enrico Giusti

1 Introduction

Mathematics has undergone some upheavals and overhauls. It no longer conceives its rationality as Euclid's contemporaries did. It has never ceased discussing and reconstructing its foundations and what defines the nature and the rationality of its objects; but neither has it ever experienced the great collapses endured by physics: indeed, while Galilean physics might have owed some of its concepts and problems to medieval science, it simply wiped out the earlier scientific concepts of cosmos and motion; after Einstein, Newton's physics can only be described as having, at best, approximate validity. The history of experimental science is replete with rectifications, refutations, and renunciations, the analysis of which is at the heart of the philosophy of physical and biological sciences. How can a theory be held true for centuries, defended by arguments to which all contemporaries give their unreserved assent, confirmed by numerous experiments – and then simply discarded, and refuted? By contrast, much of ancient mathematics still holds true for us. We still teach Pythagoras' and Thales's theorems as part and parcel of our own knowledge. In other words, mathematical theories endure. To read a mathematical text from the past means, first and foremost, to experience permanence: its theories are still entitled to a place in our knowledge. Mathematics is reread, reinterpreted, and extended, new foundations are discovered – but its theories are not forsaken.

Of course, locally, sequences of past knowledge have been shown to be erroneous or meaningless. Mathematicians' papers often contain mistakes, corrected through posterior research. Poorly justified methods or principles sometimes lead to system-forming errors: thus, in the mid-eighteenth century, Leonhard Euler wrongly asserted that a principle he thought to be fundamental, the "law of generality of algebra"

(discarded some decades later), would lead us to admit, against all evidence, that $-1 = 1 + 2 + 4 + 8 + 16 + \ldots$[1] However, such errors exist only against a background of truths, truths that were born someday and have remained valid for the most part until now. Mathematical theories last, and they last until us.

In *The Origin of Geometry*, written in his later life, the philosopher Edmund Husserl marvels at such a permanence:

> Our interest shall be the inquiry back into the most original sense in which geometry once arose, was present as a tradition of millennia, is still present for us, and is still being worked on in a lively forward development.[2]

The mythical act of the first geometer who, out of mere land surveying, invented the ideality of the figure, endures to this day in geometry; and geometry still has it embedded within its body of knowledge.

And yet, such permanence contradicts the science historian's experience in her concrete practice of reading texts – an experience of strangeness and difference. When reading ancient texts, the historian almost inevitably finds out that, even if they still hold some truth today, their objects are not identical to those we manipulate, taking this idea in its most immediate and naive sense. Even if the discussions since the beginning of the twentieth century on the foundations of mathematics may lead us to think that mathematics is not really a discourse about objects, it may nonetheless be said that Euclid's *Elements* is about points, straight lines, triangles, and magnitudes, or that Euler's *introductio in analisin infinitorum* speaks of functions, series, logarithms, and roots of equations: both texts use terms referring to such objects, and assert their existence. The historian may well find contemporary truths in ancient texts – but the objects will not be identical. While most statements of the ancient theory of magnitudes may still be considered true, the very notion of magnitude was abandoned with the arithmetization of Analysis in the nineteenth century; similarly, although much of the Italian scholar Bonaventura Cavalieri's geometry of indivisibles could be retained as part of integral calculus, the geometric concept of indivisible is no longer relevant in today's mathematics. The theorems of Thales and Pythagoras have lost their generality in contemporary geometry, and retain only a limited value in particular axiomatics. The straight lines and triangles involved in both theorems no longer have the "thickness" and intuitive evidence they had in ancient science.

[1] Leonhard Euler (1751) On the controversy between Messrs. Leibniz and Bernoulli concerning the logarithms of negative and imaginary numbers. Berlin Acad Sci 5:139–179, 1749. The text is reproduced in the *Opera omnia*, XVII, series 1, pp 195–232. Euler obtained this result by giving x the value -2 in the series $\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + \ldots$ Whose validity, formally demonstrated, is for him necessarily universal (Euler 1996).

[2] Jacques Derrida, *Edmund Husserl*'s *Origin of Geometry: An Introduction*, translated with a preface and afterword by John P. Leavey, University of Nebraska Press, Lincoln/London (Derrida 1962, p. 158).

Let us take one last example, on which we will dwell a little later. The historian who examines Descartes' *Geometry* will discover familiar equations, such as, in Book II, the equation of the hyperbola, the first analyzed by Descartes: $yy = cy - (cx/b)y + ay - ac$. For us, such an expression represents a curve only insofar as the variables and constants represent real numbers; we establish a correspondence between ordered pairs of real numbers and points on the coordinate plane. However, in the Cartesian text, the letters of constants and variables do not refer to numbers, but to lines of finite length. Even operations such as addition and multiplication are based on lines, not on numbers. So again, his equation is not ours, and does not refer to the same objects.

All the above-mentioned examples point to one conclusion: the historicity of mathematics seems markedly different from that of experimental sciences. When attentive to the details of texts and arguments, the historian of mathematics will not stumble upon a field of ruins made of discarded theories and misconceptions, so typical for the historian of experimental sciences: She discovers familiar propositions, which he can identify with those she learned in her own classes of mathematics. But these rediscovered truths are not based on genuine identities. The objects are no longer the same and we erroneously posit an identity between what are, in fact, non-identical objects. Such a discrepancy, while here mentioned in the context of the present reading of past texts, may, however, also be found within the history of mathematics itself: the protagonists in the history of science do not refer to the same objects as the very predecessors they are commenting – or sometimes, even their quasi-contemporaries. Bernoulli and Leibniz might debate about logarithm, and even agree on the essentials – yet is it really the same object? And is it the same as that of Descartes, who constructs the curve of a logarithm without naming it, all the while not even considering it as a genuine mathematical object?[3]

How is this possible? How can the history of mathematics present us both with a science that maintains the truth of past statements as it expands and transforms, and, at the very same time, a body of knowledge in which these very same statements continually refer to different objects and concepts? How can there be a durability of truths, in the sense that we continue to assent to the statements of past texts, while their terms have change meaning?

2 The Reading Process

To try to understand this, we should return to what is first given to the historian. In truth, past mathematical knowledge is not handed to us in the form of concepts, truths, and theories from the outset. We read texts, and it is only through the signs and traces marked on their pages, which somehow act upon us, that we have access to the truths and errors of the past. The mediation between us and past texts is therefore an act of reading, and this act allows us to enter the world of a particular

[3] In his letter to Florimond de Beaune dated 20 February 1639 (AT, III, 409).

theory. What is the nature of this singular reading-act that it allows for both identity and difference? What property of its gaze enables us to recapture both the Pythagorean Theorem and the difference that separates us from what it was at the time of its discovery? What is the relationship between our own reading of past texts and the reading that past scholars made of each other's works?

Husserl, in the above-mentioned text from *The Origin of Geometry*, precisely addresses this act of reading. According to him, geometry's permanence from its origins to the present day, the fact that its truth has been maintained until now, may be clarified by interrogating the meaning of the act of reading and the status of written signs. The knowledge of first geometers was recorded in written signs, and this is what allows us to revive it today and acknowledge it as our own. A modern reader of Euclid's *Elements* can find truths in it – but only because reading a mathematical text differs from the mere reception of a message, as happens in ordinary reading and its process of passive elucidation:

> Consider, for example, the way in which we understand when superficially reading the newspaper, and simply receive the "news"; here there is a passive taking-over of ontic validity such that what is read straightway becomes our opinion.[4]

Reading as a way to convey information consists merely in the reception of a meaning received passively by the reader. It presupposes knowledge of the language, but does not guarantee the truth of what is meant, nor does it rely on such a truth. On the contrary, according to Husserl, reading a mathematical text involves a requirement that of its own substantiation as a mathematical text by demonstration:

> In accord with the essence of science, then, its functionaries maintain the constant claim, the personal certainty, that everything they put into scientific assertions has been said 'once and for all,' that it 'stands fast,' forever identically repeatable, usable in evidence and for further theoretical or practical ends-as indubitably able to be reactivated with the identity of its genuine sense.[5]

To identify a text as mathematical when reading it is to conceive of it as written in the past as a sequence of universally valid deductions. Thus, for Husserl, we inevitably insert this text into our currently accepted science, which may be seen both as the text's historical aftermath (with the successive extensions, amplifications and justifications which make up the history of mathematics) and as its clarification and elucidation. If we accept Husserl's position, then twentieth century geometry, with its new foundations, would actually be the appropriate way to tackle the authentic content of an ancient theorem like that of Thales, and grasp the grounds for its truth and demonstrative force. For the German philosopher, the relationship between past and present mathematics is thus rational in two ways: the past is the reason for the present, in the sense that the latter is the consequence of the former,

[4] Husserl, *op. cit.*, p. 165.

[5] Ibid., pp. 169–170.

because the history of mathematics consists of a series of rationally solved problems; and the present is the reason for the past, in the sense that our contemporary mathematics, with its founding principles and rules, clarify the true meaning and nature of ancient objects and theorems. Zermelo-Frankel's axiomatics, for example, can thus shed light on what a number is and always has been, as well as on what founds its properties.

There is therefore a kind of reciprocal entanglement of the present in the past and of the past in the present, which for Husserl is at the heart of the mechanism of reading the signs written on the pages of ancient texts, and which guarantees the permanence of geometric truths.

Husserl's interpretation may undoubtedly seem naive to a historian of mathematics. It assumes that our contemporary but retrospective readings of ancient texts are based on a core of identity and permanence whereas, as we have seen, what is experienced is rather the dispersion and heterogeneity of objects in the texts. Historians of science have long denounced the illusion of retrospective readings that attempt to clarify the past by simply projecting ideas of our present onto it. They insist that mathematics should be considered as a practice rather than as knowledge, and have emphasized the diversity of practices and ways of doing things throughout history. Hilbertian axiomatics used games and their rules as template to think about the nature of mathematics. They turned this metaphor against the axiomatic conception and, following Wittgenstein, thought of mathematics as a plurality of language games with diverse and changing rules, identifiable by a "family resemblance" rather than by lists of strict rules. Just like Husserl, they emphasized writing, the role of signs and their different uses; but they refused to think of the sign merely as a means for placing and registering a meaning, and to examine it solely from a semantic point of view. They proposed looking at the signs of mathematical texts in relation to the images, tables, and graphs also present on the page, in a semiotics for which they sought to create tools.[6] This fertile new approach has revealed a diversity and variety in the history of mathematics that had been erased by a retrospective reading of past texts.

And yet, Husserl's reflections emphasize that such a retrospective reading is not only possible, but is also at the heart of what makes a text mathematical; and that this needs to be explained: mathematics is the practice and the science in which the present can find its own truths in the past, and thus embed the past as a part of itself. Husserl believed that it was possible because of permanence and identity, and on this

[6] In his *Mathematics as Sign* (Stanford University Press, 2000), Brian Rotman proposes just such a semiotic, drawing in part on Peirce's theory, and expressly criticizing Husserl's theory in *The Origin of Geometry* (cf. p. 47). Brian Rotman is undoubtedly unfair to Husserl in considering that the latter merely grants the sign the role of recording an ideality born originally in the solitary consciousness of a mathematician. Actually, the German philosopher defends the idea that the object or the mathematical truth only becomes truly objective, and thus mathematical, properly speaking, through the mediation of writing, its "linguistic flesh." The sign thus has a constitutive role for him. Nevertheless, Brian Rotman is right to point out that a reflection on the signs' nature and form, and on how they influence and transform the nature and form of mathematical objects, is utterly foreign to Husserl's thinking.

issue, he was wrong. But when historians saw mathematical texts as a locus of multiple rules and sign-games, and invented tools to describe the diversity of functions these signs held, they also lost the possibility to explain such a retrospective view. In *Mathematics as Sign*, Brian Rotman points out that when we assimilate Euclid's proposition "given any prime number, it is possible to find a larger one" to the modern proposition of the existence of an infinite number of primes, we do not discover a timeless truth, but we produce a "historically imposed continuity."[7] But nowhere does he explain how imposing such a continuity in mathematics is possible.

Actually, such an "imposed continuity" is linked to the mathematician's own view of her history. When the mathematician is interested in the past texts, she needs to find in them the questions that she himself asks. His gaze reads his own truths in the ancient texts when she seeks in them the original form of the questions she asks and, sometimes, the means to retrieve the answers he gives. In contradistinction, the historian views history as a kind of divergence and disagreement. There are thus two poles in any reading of a mathematical text which, to simplify matters, we will call the historian's pole and the mathematician's pole. But what matters here is not to take sides, but to understand the nature of both a mathematical text and our gaze on it for such an equivocality to be possible. How can the very same text disclose ever-differing objects yet mirror today's knowledge and objects?

Perhaps we should conceive of the text as a divisible space, the partitioning of which would be done by reading and looking. Reading is not about deciphering signs, or looking at diagrams, graphs, and figures. A mathematical text should not be understood as a collection of signs with their semiotic, diagrams with their specific classification, and images of mathematical objects like figures that could be analyzed with the tools available to the historian of science. Instead, we should think of reading as the act of drawing lines of division in the text, an act of the gaze that is the result of an interaction between the reader and the text. And among all these divisions, we note that which separates signs, diagrams, and figures, which, as we have seen, do not exist as such in the text, but only as the effect of said interaction. To study the different forms that the present reading of the past may take, therefore means to first and foremost analyze the various ways of drawing these dividing lines.

The mathematician's backward glance rediscovers her own knowledge because she draws lines of division in the text – lines which are not the historian's. We may thus in turn distinguish at least three major partitions possible, which are the conditions for reading a mathematical text:

3 The Partition of Truth and Falsehood

The French tradition of history of science from the 1930s to the 1970s has taught us that no reading of a scientific text can happen without an act of judgment and a value position: in other words, without an act of division between truth and falsehood. In

[7] *Op. cit.,* p. 40.

the tradition of philosophers of science such as Gaston Bachelard in physics, Georges Canguilhem in biology, and Jean Cavaillès in mathematics, highlighting the meaning and problems of a scientific text first requires us to distinguish what Bachelard called the "sanctioned past," which has received the seal of approval of the present and is still accepted as true, and the "outdated past," abandoned to the world of error and scientific nomadism.[8] This division is useful to understand the difficulties faced by past mathematicians when establishing and constructing knowledge, the epistemological obstacles unknowingly encountered, and the detours and side roads taken – often indispensable errors in order to reach a truth. As Jean Cavaillès said, "the after explains the before":[9] for instance, Galois' group theory explains the difficulties encountered by earlier mathematicians in their attempts to find a solution in radicals of a quintic equation. The historian of mathematics must therefore rely on present knowledge in order to shed light on the progression of past knowledge; but the difficulty is to maintain this requirement while avoiding the pitfall of a reading that would only project today's knowledge onto the past, and thereby lose the past's own rationality: in other words, the challenge is to avoid the naive view of the historian looking for "discoveries" and "first times."

Thus understood, the French tradition has put great emphasis on the retrospective point of view, which places it at the opposing end of sociology of science, or of a conception that would make science an anthropological practice of the same type as other cultural artifacts. If science is not merely a cultural phenomenon, and if it is not to be considered only as such by the historian, it is thanks to this division, which it continuously draws in its past, and which the historian must take into account. The historian's view of the text is based on this division, and can therefore never completely disregard what her present holds to be true; however, she must be aware that such a division does not exist *as such* in the text from the past, but is merely the inevitable outcome of the past and the present's encounter. It is the effect of an interaction. Thus, the history of science tags along today's science, and will also need to be rewritten, at least in parts, with each transformation of science. Such a process allows for all sorts of reversals. In the eighteenth century, the Marquis de l'Hôpital believed he had discovered a new world of numbers, the world of the infinitely small, in differential calculus. He asserted the existence, in addition to ordinary quantities, of infinitesimally small quantities. In the nineteenth century, Cauchy and Weierstrass brutally erased this delusional belief, by rigorously constructing the concept of convergence of a function. The derivatives of functions do not imply the existence of infinitesimal numbers. They are defined as limits of ratios, without assuming the existence of vanishing quantities: the Marquis de l'Hôpital's statements then appeared as mere errors and naiveties. But in the mid-twentieth century, following work on Tarski's model theory, Abraham Robinson invented a new discipline, non-standard analysis, which restored the possibility of

[8] Gaston Bachelard (1965) *L'activité rationaliste dans la physique contemporaine*. PUF, p. 25.

[9] Jean Cavaillès (1962) *Remarques sur la formation de la théorie abstraite des ensembles. Philosophie Mathématique*, éd Hermann, p. 27.

infinitesimal numbers.[10] This in turn led the historians to modify the division line between truth and falsehood in eighteenth century texts. It goes without saying that it would be naive to try to find non-standard numbers in the Marquis de l'Hôpital's text: the recurrence of the mathematical gaze is never simply the discovery of present objects in a past text, nor was its author a mere forerunner – we all know the weakness of such a concept. But Robinson's invention forces us to re-read the past and to analyze differently the obstacles to, and the conditions of the emergence of truth.

Now this new division of truth and falsity should not be understood as the reattribution of a truth value to statements whose meaning would remain unchanged. By redistributing these values, the historian also transforms the meanings she uncovers in the texts. The philosophy of language of the second half of the twentieth century, and that of Quine and Davidson in particular, has taught us that it is impossible to separate the question of belief from that of meaning in discourse. We assign a meaning to an interlocutor's words and sentences that is always determined by what we hold to be true in her discourse. In a sentence, we have the choice of deciding that it is true and refers to a given object, or that it is false and refers to another one. There is thus an irreducible indeterminacy, which will be made manifest in the experience of radical translation, but which we encounter in fact in the most mundane conversations. In radical translation, I am faced with a discourse in a language of which I know nothing. In order to interpret it, I have to decide what is to be held true, in order to correlate the words to the surrounding context. But such a decision is not unequivocal. I have to make choices, from which the meanings I attribute to those words will derive. Imagine two explorers-translators separately trying to understand the language of an isolated tribe: they may come up with two manuals with radically different meanings to the words of the tribal language, without it ever being possible to decide which one is correct.[11]

Such an indeterminacy, here illustrated by the fiction of the radical translation of an absolutely unknown language, is at the heart of our own language. When someone speaks, we too have to decide whether the words uttered are true about one thing or false about another:

> Let someone say (and now discourse is direct), "There is a hippopotamus in the refrigerator"; am I necessarily right in reporting him as having said there is a hippopotamus in the refrigerator? Perhaps; but under questioning he goes on: "It's roundish, has a wrinkled skin, does not mind being touched. It has a pleasant taste, at least the juice, and it costs a dime. I squeeze two or three for breakfast". After some finite amounts of such talk, we slip over the lines where it is plausible or even possible to say correctly that there was a hippopotamus in the refrigerator, for it becomes clear he means something else by at least some of his word than I do. The simplest hypothesis so far is that my word 'hippopotamus' no longer translates his word 'hippopotamus': My word 'orange' would do better.[12]

[10] On this question, see Jean-Michel Salanskis, *Le Constructivisme non standard*, Septentrion, 1999, which presents Robinson's conceptions in "Robinson" while discussing the question of historical recurrence in the introduction (Salanskis 1999).

[11] On this issue, see Quine (1999). *Word and object*. MIT Press.

[12] Donald Davidson (1984) *Inquiries into truth and interpretation*. Oxford University Press, "On saying that," pp. 100–1.

And thus, adds Davidson, "beyond a point, there is no deciding, even in principle, between the view that the Other has used words as we do, but has more or less weird beliefs, and the view that we have translated him wrong."[13]

This experience of the orange-hippopotamus is at the heart of our reading of past science. The historian's gaze on a past text leads her to constantly make choices. If such and such term, such and such expression, means this particular concept, then we must accept that this particular sentence in the text is false, whereas that other one is true. But if it means another concept, then truth and falsehood are reversed. These choices are not arbitrary. If I want to understand this past text, I have to hold it largely as true. According to the Davidsonian formulation, I must maximize agreement.[14] But there are many ways to do this. Hence the equivocality of the ancient text, depending on whether I maximize its truths in relation to those of today's science, conveying the partition between truth and falsehood in some way, and thereby disclosing certain meanings rather than others; or whether I choose to correlate its sentences to those of its contemporaries, or to other sentences by its author. In the latter case, I interpret the words I read, as a historian does, on the basis of what I know of what the author and his contemporaries held to be true. And this single assumption of truth reorients and reorganizes the meanings of the text. There is nothing here but the normal game in which all the meanings of language are given. In one case, I discover the unity of science by separating it from its surrounding culture, and, thereby decipher my own meanings in the text; in the other, I choose to see its dispersion and heterogeneity throughout history. These are two ways of conveying the division of truth and falsehood.

4 The Partition of Language and Image

But the act of reading past texts cannot be understood merely as an operation of interpretation, which concerns both oral and written discourse, and is not specific to scientific discourse. The written nature of mathematical knowledge introduces new uncertainties. When facing the text, the eye does not merely have before it a set of signs given as positive objects it could simply record. As mentioned above, we must draw the division (and this is the second division) between language and image, between sign and fragment of a sign, between a significant difference in a sign and a negligible variation, between what we see and what we read. But here again, all sorts of indeterminacies remain. Let us, e.g., take a look at Pascal's short 1654 treatise on the arithmetic triangle.[15] He describes the construction of a set of lines intersecting at right angles, delimiting squares called cells. In each cell, there is a number.

[13] Ibid., p. 101. In "Radical interpretation," Davidson discusses "the trade-off between the beliefs we attribute to a speaker and the interpretations we give his words" (Ibid., p. 207).

[14] On the relation between these general analyses of beliefs and meanings and some problems in history of science, see "On the Very Idea of a Conceptual Schema," Ibid., pp. 183–98.

[15] Blaise Pascal (1970) *Œuvres complètes*, eds Jean Ménard, Desclée de Brouwer, pp. 1176–1332. On this treatise, see Lucien Vinciguerra (1999) *Langage, visibilité, différence. Histoire du discours mathématique de l'âge classique au XIXe siècle*. Vrin, Chapter 1.

What kind of object is here open to our reading? On the one hand, we see the construction of a figure, which is the object studied by the mathematician. Pascal describes it to us as Euclid would, in the form of a progressive construction. The numbers in the cells, identified by Pascal in his text to the cells themselves,[16] thus appear as parts of the object triangle. Such a geometrical reading is unquestionable, and is buttressed in the second print of the pamphlet by a reinforced use of a geometrical vocabulary. But on the other hand, one may also see in it a multiplicity of statements summarized by their graphic representation, since each line of the triangle may be read as an arithmetical statement. The figure is merely a way of synoptically formulating a list of arithmetical propositions concerning binomial coefficients. Now if we turn to the discursive organization of Pascal's discourse, we find good reason to favor the first hypothesis. Pascal first describes the triangle as a genuine figure, whose construction he goes on to produce in a separate treatise; further treatises then make suggestions as to how to use this arithmetic-geometric construction. But what are we to do with such an object if we want to situate it in the history of mathematics, and make some sense of it in today's mathematics, where it appears as a nothing but a chimera? Our decisions as readers will embed Pascal's treatises either in the history of mathematics, or in the works of the philosopher.

Let us consider another example in the history of algebraic notations. Jerome Cardan's *Ars Magna*, published in 1543,[17] developed problem-solving procedures without using symbolic notations. In the French publications that followed his work, at the time of the Renaissance and the early classical age, we witness an increasing use of symbols. Jacques Peletier du Mans, in 1554, introduces the letter *R* to designate the unknown, the *res* of the *Ars Magna*, but all the other quantities in a geometric or arithmetic problem were still named discursively.[18] Shortly after, François Viète, in the *Isagoge*, used different letters to designate the unknowns and the data in problems, and arrived at true symbolic equations, but at the cost of massive preparatory work made necessary by the difficulty he faced in conceiving the multiplication of quantities. Descartes used similar symbolism for unknowns to what we use today, and simplified the notations. But for Jacques Peletier, was the letter R really designating an object? He writes:

> So for the unknown number, I put R, that is to say I put R to be equal to the unknown number. Then I reason as follows: The third part of R is 1/3R, and the fourth part of R is 1/4R; these parts substracted from R leave 5/12R. But since R is equal to the entire unknown number, 1/3R is equal to 1/3 of it, and 1/4R to 1/4. Thus, after substracting 1/3R and 1/4R from R, and similarly 1/3 and 1/4 from the number to be found, as the question requires, the two that remain are equal. Now 1/4R and 1/3R substracted from R leave 5/12R, and 1/4 and

[16] For example, we read: "In any arithmetic triangle, the sum of the cells of each base is the double those of the preceding base" (Pascal, *op. cit*., p. 1292).

[17] Gerolamo Cardano (1974) *The great art of the rules of algebra*, trad: Richard Witmer T. MIT Press.

[18] Jacques Peletier (1554) *L'Algèbre. . . departie en deus livres*. J. de Tournes, Lion.

> 1/3 substracted from the number I am looking for leave 10. So 10 must be equal to 5/12R. Here is our equation.[19]

The calculation is oddly duplicated. Although Peletier uses R, each statement involving this R is then repeated in another statement about the unknown number, called "the number to be found" or "the number I am looking for." If you think about it, R is not a term for Jacques Peletier, or else his duplication would not make sense. When he writes: "For the unknown number, I put R," he perhaps means that R replaces the unknown number in his calculation: R is thus another object, substituted for the unknown object in order to have something more familiar to compute. Peletier's algebra would thus maintain an essential proximity to the old geometry, which requires its object to stand before the reader, and cannot make sense of a calculation on an absent and unreachable object. There is no doubt that the decision to make R a letter, a term of a language to be interpreted, re-embeds Peletier's work in the universal history of algebra. But to make it an object in the stead of another, something closer to a figure than to a word, opens up a very different story.

5 The Partition of the Mathematical and the Non-mathematical

A third division remains. A mathematical text is never mathematical through and through. The reader must know how to draw what, in the text, is truly mathematical knowledge and theory, apart from what is external, and has to be rejected. Text and theory are always in reciprocal excess. A theory is never wholly contained in a given text; it always extends toward other real or possible texts, toward statements that may be deduced from the text before my eyes, toward constructions that may be added to those represented in it, toward objects other than those in the text. Conversely, a text is never purely theoretical through and through. Alongside the process of the proof, the concepts and objects that define what one would might call the theoretical content, the text presents us with remarks to accompany the reader: preliminaries, digressions, possibly historical or biographical annotations. But above all, there are statements, terms, and objects that may not evidently be elements of the theory constructed in the text.

Let us take another example from Cartesian mathematics. Descartes' *Geometry* may be read as a theory of equations, a new way of solving problems of planar or solid geometry. It mentions curves and lines. Depending on how we look at it, we can emphasize its place either in the history of algebra or in the geometric tradition. But there are several instances in the book where one also discovers the description of mechanisms used by the philosopher to build geometrical curves, such as, in Book II, small set squares sliding or pivoting in relation to each other.[20]

[19] Ibid., p. 24.

[20] This instrument generalizes another one, sometimes called "mesolabum," known before Descartes. It is used by Descartes for the first time in the *Cogitationes privates* of 1619.

The curves AH, AF, AD, are described by the points H, F, D as the compass XYZ is progressively opened. But what is the use of such a mechanism in geometry? Are these set squares also mathematical objects in his discourse, and do they belong to Cartesian theory? Are they necessary to understand the status of curved lines of geometry, by defining the "regular motion" that allows them to be "received in geometry?"[21] Are they themselves theoretical objects? Or are they merely auxiliary procedures and practical illustrations which may be erased? Today's science no longer grants them any importance, and, reasonably enough, is led to neglect them. But is this legitimate, and if so, in what sense?

This leads us to a third division to be made: we must, in the text, distinguish between an interior and an exterior of the text, a theoretical core and an impure husk. In a way, mathematics has never ceased to separate the pure and the impure throughout its history, sever what is truly mathematical from what is not. But such a division is not fixed, and mathematicians constantly trace its line in various ways and places in past texts. Descartes excluded the so-called mechanical curves from geometry, those produced by movements incommensurable with each other, such as logarithmic curves or spirals. According to him, what defines a curve as mathematical is the possibility of generating it from a finite sequence of ratios of proportions; the device of set squares gives us an image of this finite constructability of truly geometrical curves. The mechanism with set squares, which, for us, is deeply non-mathematical, is on the contrary for him what establishes the mathematicality of a curve. It plays more or less the same role as set theory or axiomatic rules in the foundation of twentieth century mathematics: a hitherto foreign element one day introduced into mathematics, of which it is claimed that it contains the very essence of mathematics. Thus, what defines mathematical purity for Descartes is no longer, as in the Platonic conception of geometry, the figure as object of contemplation, an ideal figure separate from the concrete and impure figures in the sand or on the page, but a finite succession of mind-acts establishing clear chains of relations. His geometry thus establishes a new division between what, in it, is mathematical and what is not.

This partition of the mathematical and the non-mathematical in mathematical texts runs through the history of mathematics. In ancient Platonism, it took the form of a rejection, outside the ambit of mathematics, of the concretely drawn figures. In *Geometry*, Descartes places the line of division differently – not only in his own text, but, just as much, in all the ancient texts he reads. And it can be said that, by rejecting the idea of mathematics as a discourse about objects, i.e., by refusing the very notion of mathematical objects, Hilbert's axiomatics once again drew a boundary between the texts' interior and exterior. According to Hilbert, there are no mathematical objects, but only mathematical structures that can be applied to any class of objects, as long as these verify the structure's axioms. This means in turn that, in the mass of historical mathematical texts, all the terms which are used to refer to these objects no

[21] Descartes (1954) *The Geometry*, trad: Smith DE, Latham M. Dover Publications, p. 47 (AT, VI, 442).

longer designate anything determined; they are no longer part of the truth of mathematical knowledge. Instead of discussing points, lines, and planes, we might as well be talking about chairs, tables, and beer mugs, the Göttingen mathematician once joked. The purity of mathematics is in the axioms, not in the definitions.[22] What the sentences in mathematical literature refer to must therefore be excluded from mathematics. While constantly reactivating its own past, mathematics continually re-draws this third line of partition.

These three dividing lines, which do not exclude others,[23] are not independent of each other. To move one of them in a text is necessarily to move all the others. We should not therefore think of them as elements of a combinatorial system with several entries leading to a classification. They are really the result of a historical shift of the gaze. The mathematician who reads the past texts is in reality in front of a moving landscape, implying zones of indeterminacy. In fact, he has a choice. He can choose to read the arithmetic triangle as a figure or as an abbreviated set of arithmetic statements. He can choose to read a triangle geometry proposition as dealing with an ideal object, or as a set of symbols subject to rules. He can choose to admit the Euler's law of generality of algebra, or to reject it. Some choices made in the course of history have allowed us to overcome obstacles that appeared insurmountable to the old mathematicians. They thus give an impression of necessity, and the new mathematics seems to emerge inexorably from the problems that were posed to those that preceded it. But this necessity is probably illusory. It is the effect of our difficulty, once these limits have been moved all together, to conceive another set of divisions.

6 The Case of the Equation in Descartes' *Geometry*

We have on several occasions drawn on Cartesian geometry, and while we have discovered an easy familiarity with it, we have also found objects we are not familiar with, another way of applying arithmetic operations to geometry, as well as another way of providing a foundation for mathematics. Let us take a look again at the 1637 text of *The Geometry*. It contains equations. What does it means? What happens when we read an equation?

[22] "Mathematics goes towards its purity as much and more than it comes from it," Michel Serres writes in "La querelle des anciens et des modernes," *Critique*, 1963, reprinted in *La communication*. Seuil, 1969, p. 73 (translation ours). Suzanne Bachelard comments: "Mathematics never stops moving toward purity. The principle of purity remains a regulating principle. The principle of purity is a principle native to mathematics. Recognizing it does not prevent us from recognizing that impure conceptions, when they are born with the great mathematicians, can be catalysts" (*La Représentation géométrique des quantités imaginaires au début du XIXe siècle*, Palais de la Découverte, D 113, translation ours) (Bachelard 1966).

[23] In a recent book, I tried to describe what we can call a fourth limit, the limit between the subjective position of the mathematician in the text and the space of its objects. On this question, see Lucien Vinciguerra, *Celui qui parle. Science et roman* (Hermann 2019).

Since the end of the nineteenth century, an algebraic equation with two unknowns x and y is an expression, a set of symbols combined according to given rules, and whose two unknowns will take their values in the set of real numbers. Depending on the values of x and y, the expression will be true or false. It can thus be considered as the representation of the calculation procedure for a function, whose starting set is $R \times R$ (set of ordered pairs of real numbers) and the ending set is the pair {true, false}. Solving the equation then means finding the elements of the starting set that are mapped to "true." An equation may in turn be represented as a curve by associating each couple of numbers verifying the equation with a point, whose coordinates would be that very couple, in a coordinate plane. The equation $x^2 + y^2 = a^2$ is verified by the couples of numbers (x,y) such that M(x,y) is on the circle whose center is O and radius a. It is thus the equation of this circle.

In this conception, the equation is fundamentally related to the set concept of function, and to the construction of real numbers as developed in the nineteenth century. The equation is like a sentence, but about numbers, to which it refers by the number names it contains, and the letters a, b, c, etc., which are parameters. But it is an incomplete sentence, with some empty spaces (x and y); it is this incomplete or, in Frege's vocabulary,[24] unsaturated character that makes the sentence the expression of a function.

Now, if we conceive of the equation in this way, the mathematician's view of it is in fact twofold, as the purpose of the calculation is its resolution (i.e., finding the objects that make it true). As an incomplete sentence, it has a meaning, and gives me the stage of a calculation: an expression like $2x = y$, for instance, tells us that "y is double x"; it has the same linguistic status as the latter sentence in quotation marks. When I read it, I access a certain content. But the equation is also an object, the object of a discourse which is in a position of a metalanguage vis-à-vis it; such discourse argues and questions the equation's truth, by linking it to a pre-existing set of numbers that must be constructed and posed independently of it: e.g., it would allow me to state "$(2x = y)$ is true for (1,2) and false for (5,7)." This is why the algebraic equation presupposes the prior construction of its world of objects (the set of couples of real numbers), and a gaze that, in the text of the equation's resolution, must go back and forth from the expression as object to the world of objects that make it true or false. The equation thus does not have the same mode of existence as the language of the mathematical text that surrounds it, refers to it, and tells us how to transform it; a language which, with respect to the equation, is in a position of metalanguage.

Now, let us open Descartes' *Geometry.* The issue of the nature and the place of equations in Descartes' book has, in recent decades, given rise to a lively debate among science historians. Enrico Giusti emphasizes the importance of the relationship between equation and geometric curve in the treatise. *The Geometry* was originally conceived as a general theory of curves as geometric knowledge, for

[24] See Gottlob Frege (1960) Function and concept. In: Geach P, Black M (eds). *Translations from the philosophical writings of Gottlob Frege*. Basil Blackwell.

which the theory of equations would act simultaneously as criterion of identity, of admissibility, and of classification.[25]

The thesis that curves are the objects of geometry was quite radical in the seventeenth century, as it departed from the idea that geometry's true objects were figures – entities, originally introduced together with boundaries in Book I of Euclid's *Elements*,[26] which the mathematician has to imagine and contemplate, and which he can divide as he wishes by drawing new lines. Now, stating that there was a correlation between equations and curves also amounted to demote an operation which, until now, had been at the heart of geometry: the typical "geometric construction" (whereby successive acts of drawing lines are undertaken on given figures in order to reach the points solutions of a problem) ceased to be the fundamental mode of access to a geometric object and to a proper resolution of a problem. Two instruments – the ruler and the compass – had been considered by Greek geometers as the privileged means of geometric construction; when faced with their limits, mathematicians had then attempted to invent new ones, which would enable the resolution of problems unsolvable with the mere use of compass and ruler. The above-mentioned *Mesolabum* played in the seventeenth century a role in keeping with just such an instrumental and constructivist tradition. Nonetheless, these methods (which we do find in *The Geometry*) merely led to disparate techniques. As Giusti notes, Descartes' universal method in science could not be founded on them:

> There is no universal compass. Any new machine would merely generate a new curve, without affecting the overall situation. And to such an array of instruments, there can be no corresponding array of methods and research approaches.[27]

The theory of equations unifies what, in construction methods, remains scattered, and thus makes Cartesian science coherent with the principles of its own method. The degree of the equation offers a universal criterion to classify all geometric curves. In the eyes of a long tradition of science historians, of which Giusti is one of the most rigorous representatives, *The Geometry* therefore undoubtedly represented a theory of equations.

But other readings of *The Geometry* are possible. If we look more closely at the internal structure of the book, which until now seemed to lack coherence,[28] if we take into account its position in the slow development of Cartesian science starting from his first works in 1619, as well as its situation in mathematical debates in the

[25] See Enrico Giusti (1990) Numeri, grandezze e *Géométrie.* In: *Descartes, il methodo e i saggi*. Armando Paoletti, pp. 419–431.

[26] "A figure is that which is contained by some boundary or boundaries." Euclid, *Elements*, Book I, Def. 14 (Euclid 1956).

[27] Enrico Giusti (1998) La Révolution cartésienne en géométrie. In: *Descartes et son œuvre aujourd'hui.* Pierre Mardaga, pp. 47–62

[28] See Henk JM Bos (1990) The structure of Descartes' *Géométrie*. In: *Descartes, il methodo e i saggi*. Armando Paoletti, pp. 349–69.

early seventeenth century, then a whole new project emerges, as shown by Henk Bos.[29] According to Bos, *The Geometry* is not a theory of curves, nor even a theory of figures. Its goal was rather the attempt to uncover a general method of resolution of geometric problems. It aimed at classifying and solving problems, and the starting point to understand the meaning of his work should precisely be this notion of problem. If that is the case, the formulaic rendering of an equation could not offer a satisfactory resolution of problems; indeed, finding the solution to an equation does not necessarily allow us to construct its solution geometrically. This is why, according to Bos, Descartes insists on retaining construction devices. The *Mesolabum*'s concatenated set squares thus remained, in his eyes, the device *par excellence* to highlight a problem's complexity. The number of necessary set squares, i.e., the number of basic moves to complete in order to reach a solution (moves ordered by the progressive opening of the compass), represented the consecutive stages of a deduction which could then be grasped in one single intuition, as per the *Rules for the Direction of the Mind*.[30] The set square device thus plays a major role in the implementation of Descartes' method (although, as Bos notes, this is probably truer of the method exposed in the *Rules* than that in the 1637 *Discourse*). Contrary to the first uses of the *Mesolabum* in his works dating from 1619, Descartes could now relate his instrument to the theory of equations. This new relation led him to progressively opt for the uniformity and the universality of algebra over geometry's instrumental processes.

> Descartes finally adopted algebraic criteria, abandoning earlier attempts to classify according to methods of tracing. The process may be described as an infiltration of algebraic thinking into the interpretation of geometrical exactness. It is indeed a natural process.[31]

[29] Henk JM Bos (2001) *redefining geometrical exactness. Descartes' transformation of the early modern concept of construction*. Springer.

[30] See rule 7: "Si igitur, ex.gr., per diversas operationes cognoverim primo, qualis sit habitudo inter magnitudines A and B, deinde inter B and C, tum inter C and D: non idcirco video qualis sit inter A and E, nec possum intelligere praecise ex jam cognitis nisi omnium recorder. Quamobrem illas continuo quodam imaginationis motu singula intuentis simul and ad allia transeuntis aliquoties percurram, donec a prima ad ultimam tam celeriter transire didicerim, ut fere nullas memoriae partes relinquendo, rem totam simul videar intueri" (AT, X, 388). "If, for example, by way of separate operations, I have come to know first what the relation between the magnitudes A and B is, and then between B and C, and between C and D, and finally between D and E, that does not entail my seeing what the relation is between A and E; and I cannot grasp what the relation is just from those I already know, unless I recall all of them. So I shall run through them several times in a continuous movement of the imagination, simultaneously intuiting one relation and passing on to the next, until I have learnt to pass from the first to the last so swiftly that memory is left with practically no role to play, and I seem to intuit the whole thing at once" (*Philosophical writings*, vol. 1, p. 42). On the relationship between the compass device and the *Rules*, see Vuillemin J (1960) *Mathématiques et métaphysique chez Descartes*. PUF and Serfati M (1993) Les compas cartésiens. *Arch Philos* 56: 197–230.

[31] Henk JM Bos, *op. cit.*, p. 361.

Let's now take a sidestep, and look at the position of *The Geometry*'s reader, and at the body of knowledge accessible to Descartes about the act of reading and the type of gaze it enables. *The Geometry* can certainly be understood from the standpoint of all these mathematical "objects" bequeathed by his predecessors, and debated by his contemporaries: curves, figures, equations, constructions, techniques, and instruments. But it also involves the act of reading, the mode of existence of signs on a written page, as well as the action of light rays travelling from the page to the geometrician's eye – all elements which intertwine Descartes' optics, mathematics and philosophy in a most profound way.

Let us examine the first pages of *The Geometry.* How does the text begin*?* Not with the definition of objects, numbers, lines, curves or figures, nor with the construction of a new language, nor with axioms or principles; it opens with the definition of operations on straight lines: addition, subtraction, multiplication and root, geometric and non-arithmetical operations, named using the traditional symbols of corresponding arithmetic operations. Once these operations are constructed, Descartes names the lines using letters, whether the lines are given and known, or still unknown; he reserves the last letters of the alphabet for the latter:

> If, then, we wish to solve any problem, we first suppose the solution already affected, and give names to all the lines that seem needful for its construction, to those that are unknown as well as to thoses that are known.[32]

From there onward, the equation never appears in Cartesian discourse as a language-object, whose truth conditions must be found. Its resolution is not the search for a world of objects under a set of written signs and numbers, itself another object. It is conceived as an operation of ordering, the ordering of a set of relations between lines. These relations are expressed in the equation, but in an initially confused form, which will need to be disentangled and ordered:

> Then, making no distinction between known and unknown lines, we must unravel the difficulty according to the order which shows most naturally of all in what way they mutually depend on each other, until we find it possible to express a single quantity in two ways. This will constitute an equation.[33]

A bit further down, Descartes characterizes a division operation as "démêlant ces équations" ("disentangling these equations").[34]

If we take all this vocabulary of confusion and entanglement seriously, we discover that the Cartesian equation does not exist on the page of *The Geometry* as it does in our algebra textbooks. According to Descartes, the operation of resolution

[32] Descartes (1954), *The Geometry*, op.cit., p. 8 (AT, VI, 372).

[33] Ibid. (translation modified).

[34] Ibid., 374. This metaphor is not translated in the English version. And at the next paragraph, "lorsque la dernière équation aura été entièrement demêlée" is translated by "when the last equation shall have been entierely solved" (*op. cit.*, p. 13).

is not to be understood as the mapping of an expression (a linguistic object) onto another type of object, which would be, for example, numbers. The equation is fully a discourse. The letters, a, b, c, x, y, are the names of lines or line segments, and the calculation symbols refer to operations on these lines. When I write $z^2 = -az + bb$, I am stating the relation between a line z to two other known lines a and b. The line is either in whole or in part determined by these relations. And by ordering these relations between lines, I find the line z, and I know it distinctly: I can now determine it by distinguishing it from the others, in accordance with the Cartesian theory of the distinct idea, "so precise and different from all the others that it includes in itself only what appears manifestly to the one who considers it as it was meant to be."[35] To solve then, does not mean finding the objects that make an expression true: it, very simply, means to order. Such an ordering process does not refer to the signs of the equation conceived as material objects. The equation is a discourse on the pages of *The Geometry* at exact par with the other surrounding sentences, which explain how to draw lines, join points, etc.: a discourse about the world of geometric objects, without any attention ever being paid to the materiality of its signs. The only peculiarity of the equation is to refer to these objects with letters and symbols rather than words. Thus, in the text of *The Geometry*, we do not witness the double modality of discourse, which would see one language-object on the one hand, and a metalanguage, on the other. In the Cartesian text, the equation is nothing but discourse: it is never an *object* for a discourse. And its resolution is the ordering of this initially confused set of relations between known and unknown things.

When we pay attention to the detail of the Cartesian text without immediately dividing it into language and object, we may find the traces of a specific mode of existence of what is, for him and for us, an equation. Certes, this equation is not the same for Descartes as it is for our gaze today; but we might still be able to reconstitute the Cartesian gaze to a certain extent. Descartes was not only a mathematician: he was also a philosopher, who reflected on the nature of language, of writing and signs, of looking and seeing. Unlike many other mathematical texts for which the act of reading can only be inferred from the organization and structure of the text, Descartes conducted analyses in various non-mathematical writings that shed light on this issue.

In a famous passage in the beginning of the *The World*, the dissimilarity between the linguistic sign and the object it signifies appears to be the template of that between the sensation – or the idea received in the mind – and the object that causes it:

> For although everyone is commonly convinced that the ideas we have in our mind are wholly similar to the objects from which they proceed, nevertheless I cannot see any reason which assures us that this is so. On the contrary, I note many observations which should make us doubt it [4]. Words, as you well know, bear no resemblance to the things they signify, and yet they make us think of these things, frequently even without your paying attention to the sound of the words or to their syllables. Thus it may happen that we hear an utterance whose

[35] Descartes, *Principles*, I, 45.

> meaning we understand perfectly well, but afterwards we cannot say in what language it was spoken. Now if words, which signify nothing except by human convention, suffice to make us think of things to which they bear no resemblance, then why could nature not also have established some sign which would make us have the sensation of light, even if the sign contained nothing in itself which is similar to this sensation?[36]

Ideas of sensations are not similar to the material things that gave rise to them. The content of a sensation such as light does not resemble the material sign that caused the sensation; the experience of hearing someone speak is a better template of their relation than the resemblance of images. In the text, Descartes uses an analogy the order of which must not be reversed: what the material sign that causes sensation is to the idea or sensation, speech is to the thing it signifies, that is, to the meaning of that speech. But for Descartes, this experience of listening is, surprisingly enough, the event of a discourse whose audible materiality has disappeared for the person at the receiving end: the word is not heard as an object, a sound which would simultaneously be associated to a meaning; rather, it opens up a meaning at the exact moment it is erased as such from my representation, so that I do not even know in what language they were uttered. There is no doubt that Descartes' mastery of foreign languages led him to experience such a situation on numerous occasions: linguistic communication is an exchange in which one's mind's thought can be conveyed to another through the medium of discourse, itself a material object which does not require to become an object of representation and thought in order to perform its function. I hear the meaning without thinking about words or language, and this is valid for both oral or written speech. Years later, in a letter to Clerselier dated April 23, 1649, Descartes writes:

> Forgive me for having interlarded this letter with Latin; the brief time I have had to write it has not allowed me to think of words.[37]

To pay attention not to ideas but to words requires a special effort of the mind. If we are distracted, preoccupied, or tired, we become oblivious to them, and yet, we still speak, and understand what is said. We may therefore conclude that it is not because we perceive the sound of words that we grasp their meaning: the two processes are autonomous.

In the fourth part of the *Principles*, Descartes again refers to this experience; it concerns writing too, and he is again seeking a model of the action of the body on the mind:

> For we see that spoken or written words excite all sorts of thoughts and emotions in our minds. With the same paper, pen and ink, if the tip of the pen is pushed across the paper in a certain way it will form letters which excite in the mind of the reader thoughts of battles, storms and violence, and emotions of indignation and sorrow; but if the movements of the

[36] Descartes (2009), *The Philosophical Writings*, vol. 1, p. 107 (AT, XI, 4).

[37] Descartes (1991), *The Philosophical Writings*, p. 378 (AT, V, 357).

> pen are just slightly different they will produce quite different thoughts of tranquillity, peace and pleasure, and quite opposite emotions of love and joy.[38]

The linguistic sign is a material thing. When we think of its effects, we are in the field of causality and action: the sign "excites," "produces," "occasions acts." What the sign-thing affects is the mind, which, modified by this action, then has a thought. At that moment, the mind gets a representation: it imagines battles or peace, it feels the passions of sadness or love, and imagines or perceives the objects of these passions. Writing is thought of within the framework of the union of soul and body, of which it is also the template.[39] Writing is not intelligible without that union. But this in turn means that the sign is not an object of interpretation or decryption: this would imply that it becomes, at some stage, an object of representation, to which the mind would associate another object – another thought. The relation of the sign to what it represents is not an association in the mind of one representation with another. The meaning is entirely clarified by the action of a body on a mind.

How and why can the sign operate in such a way? Actually, the sign belongs to the bodies explored by Descartes in his physiology and his optics, of which he says they act on the soul because they belong to the mechanical process that makes us perceive. This process must be conceived as a material movement: the traces and impressions in the brain, the nerve fibers, the image on the retina, the vibrations of light. A sign is all these things for Descartes. Now, such signs cause the soul to perceive and to have a representation, as per the institution of nature, only insofar as they are not seen.[40] The cerebral impression and the physical substratum of light lead to a perception only for those who do not see them. These signs are invisible for those who perceive, and visible only to a third party, such as a physicist depicting light in the form of straight lines in an optical treatise, or an anatomist describing someone else's eye and brain. If we were to see them, as Malebranche noted, it would then be impossible to see the representation they excite.

The linguistic sign, though a human institution, draws its efficacy from a similar process. Descartes conceives the linguistic sign on the general model of the sign as unperceived trace; this is why, when he analyzes language, he refers to the experience of a discourse whose meaning we grasp without perceiving its material reality.

But is there truly nothing in common between the sign and the representation it generates? Let us dwell on the above-mentioned text in the fourth part of the *Principles.* It refers to a quill that moves in different ways. And the differences in its movement arouse "quite different thoughts." Similarly, in the following article,

---

[38] Descartes, *Principes*, IV, 197, in *The Philosophical Writings*, vol. 1, p. 329 (AT, IX, 316).

[39] On this issue, we may of course also refer to the fourth discourse of *Dioptrics*, op. cit., p. 202 (AT, VI, 112).

[40] In the fourth discourse of *Dioptrics,* it is indeed what is sent "inside our brain," such as the vibration of the nerve fiber or the image on the back of the eye, which receives the qualification of sign likely to excite thoughts in the soul, "as for example signs and words, which do not resemble in any way the things they signify" (Ibid.).

Descartes evokes "les divers mouvements de quelques corps [qui] suffisent pour lui faire avoir tous les divers sentiments qu'elle a."[41] There is no resemblance between the sign and what it means for the soul, but there is a corresponding difference. In the same French edition of the *Principles* dated 1647, article 71, Part I, Descartes mentions the relation between color and its signs, the bodily movements which he describes as the "sentiments tels que sont ceux qu'on nomme goût, odeur, son, chaleur, froid, lumière, couleur et autres semblables, qui véritablement ne nous représentent rien qui existe hors de notre pensée, mais qui sont divers selon les diversités qui se rencontrent dans les mouvements qui passent de tous les endroits de notre corps jusqu'à l'endroit du cerveau auquel elle est étroitement jointe et unie."[42] And in his *Dioptrics*, Descartes, in a new attempt to explain both the dissimilarity between sensations and their material causes and the possibility for one to represent the other, brings in the character of a blind man touching different things with a stick. He writes:

> They certainly do not transmit anything to him except in so far as they cause his stick to move in different ways according to the different qualities in them, thus likewise setting in motion the nerves in his hand, and then the regions of his brain where the nerves originate. This is what occasion his soul to have sensory perception of just as many different qualities in these bodies as there are differences in the movements caused by them in his brain.[43]

The bodies' relations to each other are similar to those of the stick's movements, and to those of the qualities experienced by the soul. Yet again, the same differences and the same relationships between the bodily sign and its content may be found. While there is no resemblance between the sign and its content, something does remain from one to the other insofar as one is erased to make the other visible: the same differences and the same movement.

We can now return to the Cartesian theory of the equation in *The Geometry*, and to the historians' debate concerning the respective place of equations and geometric problems within it. In a way, Giusti's and Bos' interpretations of Descartes' work reflect different views of its position in the history of mathematics. Giusti, a mathematician first and foremost, legitimately reads it from the starting point of what happened later. Bos, on the other hand, emphasizes the concepts common to Descartes and his contemporaries. However, both agree that *The Geometry* contains something like a theory of equations, these peculiar objects progressively and collectively constructed by algebraists in the sixteenth and seventeenth centuries; they further agree that other shared objects exist, such as curves, as do geometric problems. But what is the relation between these two classes of objects?

[41] Descartes, *Principes*, IV, 198 (French translation of Claude Picot revised by Descartes).

[42] AT, IX, 58. In this two passages, the french text dated 1647 is very different from the latin text dated 1644. It was probably modified by Descartes himself.

[43] *Dioptric,* discourse four, in Descartes (2009), *The Philosophical Writings*, vol. 1, p. 202 (AT, VI, 114).

Descartes could not have made his equation the true object of the reader's gaze. Its signs must disappear under the gaze of the geometer reading the treatise on *Geometry*, as must all the traces that made it visible.[44] Descartes notes in his definition of an equation that what is given to the reader consists indeed of relations – readily known relations, between known and unknown lines, but which require some ordering in order for the reader to know what exactly they relate. It is immediately clear that, in Cartesian science, it is impossible to interpret the algebraic expression as a set of material signs on a page, which we would have to associate to a field of objects. Strictly speaking, there is no set of objects to be associated with the equation. The equation is not a function with values in a set, be it a set of numbers or of anything else. It presents me with a diversity, which, if the solution is not readily apparent, may be ill-ordered and confused. The resolution of the geometric equation is therefore first and foremost the act of a gaze which orders and re-arranges things. It is this very reading device, which is intimately linked to Descartes' philosophy, which allows us to find the objects of his *Geometry.*

The relationship between the geometric curve, the problem, and the construction that will generate its solution must thus be conceived in light of the Cartesian dualism of body and mind. For the philosopher, a problem is a mental entity (which it was not for the Greeks who, like Menaechmus, conceived of geometry as a theory of problems). As such, its reality is that of a meaning, something I have "in mind," endowed with both clarity and confusion, disclosing some of its objects and obfuscating others. The equation is nothing but a problem, when my gaze pierces through the material symbols on the page and reaches the entangled relations that require ordering. Ancient geometricians conceived the resolution through a construction, a process whose starting point were specific objects, from which one would then amble to reach new objects, through the drawing of lines and determination of properties; this whole process was thought of as wholly undertaken in the world of mathematical objects. Descartes, on the other hand, conceptualizes a process which arises in the mind to reach the world of geometric objects that stems from a question in the geometer's mind to reach an answer whose job is first and foremost to order that which, in the question, had remained muddled, thus eventually uncovering its objects clearly and distinctly. This, in turn, also alters the very notion of construction. Insofar as it now pertains to analysis and no longer necessarily to synthesis, a geometric construction may now be conceived as a

[44] It might be argued that *The Geometry* never presents the problem in this light, and that it never evokes the theory of the sign of *The World* or of the *Dioptrics*. However, it should be remembered that *The Geometry* is the third essay of a series of works which began with the *Discourse of Method*, *Meteors* and the *Dioptrics*; and that Descartes prefaces his table of contents at the end of the work with the following warning, which remains relevant to this day: "Those who visit the tables of books only in order to choose the subjects they wish to see, and to exempt themselves from the trouble of reading the rest, will derive no satisfaction from this table: for the explanation of the questions mentioned therein almost always depends so expressly on what precedes them, and often also on what follows them, that it cannot be understood perfectly unless the whole book is read with attention" (AT, VI, 486; this passage is not translated in the English version).

succession of operations whereby the mind grasps the object it seeks thanks to the analysis of given relations in the question it itself had asked.[45]

A geometric construction is a set of acts – the act of drawing lines. To draw circles or lines, whether straight or not, means to initiate a movement. We know, though, how much the ancient and the modern conceptions of movement differ. For the Ancients, movement is a whole, defined by its starting point and its end. The successive acts undertaken in the construction of a figure is thus a succession of indivisible steps. When a point is constructed as the intersection of two semi-circles, the outline of the two semi-circles is not envisaged as the succession of points passed through by the compass's mobile tip, but rather, as the encounter of two singular acts, which each find their fulfilment in that one remarkable point of intersection. In Descartes' thinking, on the contrary, the outline is the path delineated by the succession of indifferent points on lines and circles, a succession that may be measured by the y-intercept along a line of reference.[46] Descartes thus opposes the curves in which only a few remarkable points may be traced to the "good" geometric curves, in which any point whatsoever may be reached. He correlates this possibility (intrinsic to the form of the equation) to the notion of regular and continuous movement.

> On the other hand, there is no point on these curves which supplies a solution for the proposed problem that cannot be determined by the method I have given. But the fact that this method of tracing a curve by determining a number of its points taken at random applies only to curves that can be generated by a regular and continuous motion does not justify its exclusion from geometry.[47]

Descartes emphasizes the idea that the delineation of a curve necessarily implies a device which allows for a regular movement passing by every single point of the curve; this explains why, in the opening lines of *The Geometry*'s Book II, he expresses reservations about the term "mechanical," in reference to the Ancient geometers' opposition between geometric and mechanical curves (an opposition which he incidentally retains):

> Nor do I see why they called the latter mechanical, rather than geometrical. If we say that they are called mechanical because some sort of instruments has to be used to describe them, then we must, to be consistent, reject circles and straight lines, since these cannot be described on paper without the use of compasses and a ruler, which may also be termed instruments.[48]

[45] This idea is at the heart of the fifth and sixth *Rules for the direction of the mind.*

[46] Descartes evokes "the fact that all points of those curves which we may call 'geometric', that is those which admit of precise and exact mesurement, must bear a definite relation to all points of a straight line, and that this relation must be expressed by means of a single equation, *The Geometry, op. cit.*, p. 48, livre second (AT VI, 392)."

[47] Ibid., p. 91 (AT VI, 411).

[48] Ibid., p. 43 (AT VI, 389).

Let's take a new glance at Descartes' set square device. As the compass is progressively opened, the points D, F, and H move away from point A. When the compass is fully closed (an impossible position, practically, but theoretically and mathematically rigorous), all these points merge into one. All along their path, ratios are preserved throughout the continuous and ordered movement undertaken: YA/YC = YC/YD = YD/YE = YE/YF = YF/YG. They remain identical, but progressively more explicit as they progress, and it is always possible to stop the compass on any given point of the curves' trajectory to find random proportional averages. We may thus see this ratio-preserving movement in light of another movement: that which, in the theory of perception and of reading, correlates meaning, sign, and thing: from start to end, the differences remain identical. The mathematician who reads an equation in Descartes has a muddled meaning in mind; in a way, she is akin to the closed compass where all the ratios are present, but reduced to one single point, with no extension, all jumbled. In order to solve the problem, she needs to open the compass: the same ratios will emerge, but ordered thanks to the movement of the device. The construction's movement is similar to that of the mind progressing methodically until it finds the solution of the problem, from the unextended thing to the extended thing.[49]

There is a deep solidarity between the opening pages of *The Geometry,* in which Descartes defines the equation on the basis of ratios between lines, as expressed by arithmetic operations on the one hand, and the mechanisms whereby Descartes represents these very ratios, linked together so as to disclose the very object they purport to be the ratio of, on the other. What we may derive thereof is that the issue is not to reconcile, in Descartes' work, a theory of equations and a search for geometric constructions, as much as it is to understand how the partitions drawn by Descartes' make the equations and figures objects which, while embedded in a tradition, are related to his philosophy, and thus differ from his predecessors' notions.

7 Conclusion

Tracing back the mechanism of reading at play in Cartesian thought, we may now be able to read *The Geometry* in a manner closer to how Descartes himself read the very mathematical text he was writing, and probably the other texts he read. The case of Descartes is naturally special: as a philosopher, he was able to explicitly examine the nature of the sign and of reading. And yet, it reveals a more general reality, present in all texts in the history of mathematics. The way in which we make objects and meanings appear in them depends intimately and irreducibly on the form of the reading that we choose to implement. Such an interaction in reading always consists

[49] Michel Fichant, in "La géométrisation du regard. Réflexion sur la *Dioptrique* de Descartes" (re-edited in *Science et métaphysique dans Descartes et Leibniz*, PUF, 1998), has emphasized the role of movement in the Cartesian analysis of vision: "It is the ordered set of all these movements (and not the more or less resembling traces they leave in their wake) which signifies to the soul what it interprets as a thought of seeing" (p. 48; translation ours) (Fichant 1998).

in linking the current text to other texts, which will help us delineate the above-mentioned divisions between truth and falsehood, between the visual and the linguistic, between the interior and the exterior of mathematics, in our current reading. Thanks to these divisions, objects, concepts, and identities are disclosed. By linking *The Geometry* to other texts written by Descartes that do not belong to mathematics, we make these objects appear, and discover that they are not the same as ours; we also find out that Descartes' text followed a reasoning foreign to us. Does it allow us to conclude that we have found the "true" objects and the "true" principles of *The Geometry*? Would it mean that science is not universal in any way, but consists only in a multiplicity of private knowledge systems, each with its own objects? In which case, how can there be a scientific community? In truth, by linking the text of *The Geometry* to the texts of *The World* and the *Dioptrics*, by trying to identify a properly Cartesian way of reading, we have gained something: it becomes clearer how the problems and questions presented themselves to the mathematician himself. But this is essentially only another partition line, and, with it, we lose an essential element: we are no longer able to explain how *The Geometry* exists as a milestone in the history of mathematics. We have wrested Cartesian science from its future by linking sentences in *The Geometry* to other contemporary texts, which we can no longer say belong to mathematics. As such, this reading is no more true than a retrospective reading, which is just as indispensable. At every step of the historical process of science, it is indeed this backward glance that restores the unity of science. Such a possibility was and remains crucial.

References

Bachelard G (1965) L'activité rationaliste dans la physique contemporaine. Presses Universitaires de France (PUF), Paris

Bachelard S (1966) La représentation géométrique des quantités imaginaires au début du XIXe siècle. In: Les Conférences du Palais de la découverte, Série D, Paris, p 113

Bos HJM (1990) The structure of Descartes' Géométrie. In: Belgioioso G (ed) Descartes, il methodo e i saggi. Armando Paoletti, Florence

Bos HJM (2001) Redefining geometrical exactness. Descartes' transformation of the early modern concept of construction. Springer, New York

Cardano G (1974) The great art of the rules of algebra (trad: Witmer TR). MIT Press, Cambridge, MA

Cavaillès J (1962) Philosophie mathématique. Hermann, Paris

Davidson D (1984) Inquiries into truth and interpretation. Oxford University Press, Oxford, UK

Derrida J (1962) Edmund Husserl's origin of geometry: an introduction (translated with a preface and afterword by Leavey JP). University of Nebraska Press, Lincoln/London

Descartes R Œ (1954) The geometry (trad: Smith DE, Latham M). Dover Publications, New York

Descartes R Œ (1991) The philosophical writings, vol 3 (trad: Cottingham J, Stoothoff R, Murdoch D). Cambridge University Press, Cambridge, UK

Descartes R Œ (1996) Adam C, Tannery P (eds). Vrin, Paris (for all Descartes' citations, the reference to this edition is given in brackets)

Descartes R Œ (2005) The philosophical writings, vol 2 (trad: Cottingham J, Stoothoff R, Murdoch D). Cambridge University Press, Cambridge, UK

Descartes R Œ (2009) The philosophical writings, vol 1 (trad: Cottingham J, Stoothoff R, Murdoch D). Cambridge University Press, Cambridge, UK
Euclid (1956) The elements (trad: Heath T). Dover Publications, New York
Euler L (1996) Opera omnia, series 1, vol XVII. Springer, New York
Fichant M (1998) Science et métaphysique dans Descartes et Leibniz. Presses Universitaires de France (PUF), Paris
Geach P, Black M (eds) (1960) Translations from the philosophical writings of Gottlob Frege. Basil Blackwell, Oxford, UK
Giusti E (1990) Numeri, grandezze e Géométrie. In: Belgioioso G (ed) Descartes, il metodo e i saggi. Armando Paoletti, Florence
Giusti E (1998) La révolution cartésienne en géométrie. In: Descartes et son œuvre aujourd'hui. Pierre Mardaga, Enmore
Pascal B (1970) In: Ménard J (ed) Œuvres complètes de Blaise Pascal. Desclée de Brouwer, Paris
Peletier du Mans J (1554) L'Algèbre. . . departie en deus livres. J. de Tournes, Lion
Rotman B (2000) Mathematics as sign. Stanford University Press, Stanford
Salanskis J-M (1999) Le constructivisme non standard. Septentrion, Paris
Serfati M (1993) Les compas cartésiens. Archives de philosophie 56:197–230. Beauchesne éditions
Serres M (1969) La communication. Seuil, Paris
van Orman Quine W (1999) Word and object. MIT Press, Cambridge, MA
Vinciguerra L (1999) Langage, visibilité, différence. Histoire du discours mathématique de l'âge classique au XIXe siècle. Vrin, Paris
Vinciguerra L (2019) Celui qui parle. Science et roman, Hermann, Paris
Vuillemin J (1960) Mathématiques et métaphysique chez Descartes. Presses Universitaires de France (PUF), Paris

Symbolic Algebra as a Semiotic System

Ladislav Kvasz

Contents

1 Introduction 3102
2 The Compositional Component of the Language of Mathematics 3104
2.1 The Logical Power of the Language of Algebra 3106
2.2 The Expressive Power of the Language of Algebra 3109
2.3 The Methodological Power of the Language of Algebra 3111
2.4 The Integrative Power of the Language of Algebra 3112
2.5 The Explanatory Power of the Language of Algebra 3113
2.6 The Constitutive Power of the Language of Algebra 3114
2.7 Summary 3116
3 The Referential Component of the Language of Mathematics 3116
3.1 Pictorial Form as a Tool of Epistemological Reconstruction 3119
3.2 Aspects of the Pictorial Form of the Language of Geometry 3121
3.3 The Development of the Pictorial Form of the Language of Geometry 3123
3.4 Transferring the Concept of Pictorial Form to Algebra 3126
3.5 Aspects of the Pictorial Form of the Language of Algebra 3127
3.6 Objectification as a Fundamental Semiotic Process 3130
4 Conclusion 3131
References 3132

Abstract

The invention of symbolic algebra in the sixteenth and seventeenth centuries fundamentally changed the way we do mathematics. If we want to understand this change and appreciate its importance, we must analyze it on two levels. One concerns the *compositional function* of algebraic symbols as tools for representing complexity; the other concerns the *referential function* of algebraic symbols, which enables their use as tools for describing objects (such as

L. Kvasz (✉)
Institute of Philosophy, Czech Academy of Sciences, Prague, Czech Republic

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_65

polynomials), properties (such as irreducibility), relations (such as divisibility), and operations (such as factorization). The reconstruction of both the compositional function and the referential function of algebraic symbols requires the use of different analytic tools and the taking of different temporal perspectives. In this chapter, we offer both: a reconstruction of the compositional function of algebraic symbols, and a reconstruction of their referential function.

Keywords

Logical power · Expressive power · Integrative power · Pictorial form · Epistemic subject · Symbolic algebra

1 Introduction

There is no lack of literature on the various aspects of the language of mathematics.[1] The different approaches describe important aspects of the language of mathematics, but they usually do not provide an overall picture. The purpose of this chapter is to *propose an analytic framework* based on the insights of Frege and the early Wittgenstein *for the study of mathematical language* in its complexity, and illustrate this framework through the specific case of the symbolic language of algebra.

To achieve this goal, we propose a specific understanding of the symbolic language of algebra. On the one hand, we suggest understanding it in an inclusive sense, thereby including *all symbolic tools* used in books, papers, and manuscripts devoted to algebra from the fifteenth century onwards, when the use of algebraic symbolism became sufficiently complex.[2] Thus, we include the language of polynomial, linear, and multilinear algebra just like that of group theory, ring theory, or lattice theory – so practically everything that algebraists use to communicate their ideas and results.[3] However, on the other hand, we propose to understand the language of algebra in a purist sense and exclude all symbols imported into algebra from differential and integral calculus (e.g., the functional notation), predicate calculus, and set theory. More radically, we exclude all commentaries in ordinary

[1] See (Rotman 2000, Lakoff and Núñez 2000, O'Halloran 2005, Barton 2008, Avigad 2015).

[2] To a historian, the choice of the fifteenth century as a starting point may appear arbitrary, as all the symbolic tools of fifteenth-century European algebra had already been available in Latin, Arabic, Sanskrit, and other cultures. Nevertheless, as our aim is not a historical analysis of the birth of algebra but an *epistemological* reconstruction of the functioning of its language, we prefer taking the fifteenth century as the starting point because the language of algebra at that particular time was already sufficiently complex for epistemological reconstruction.

[3] We do not differentiate between the use of words, letters, and special symbols – they are all included in the symbolic language. The key distinctive feature is whether a particular symbol obeys some rules specific for algebra, such as substitution, raising to a power, extracting roots, and permutation. Thus, if in a text a word appears by which the mathematician enacts such an operation, it belongs to the language of algebra. If not, then not.

language.[4] Thus, in short, we will be interested in changes of the language of algebra over long periods of time while excluding all external influences.[5]

We shall use material from the history of mathematics because taking a historical perspective allows us to see the different aspects of the language of algebra in the process of change, thereby making it possible to discriminate these aspects more easily. Nevertheless, we have no intention of offering any contribution to the discipline of the history of mathematics. As far as is possible, we will stick to the standard interpretation of the historical material and simply use it for making epistemological claims. Thus, in order to understand the way symbols are used in algebra, the language of algebra will be approached from a diachronic point of view.

The following pages focus on two developmental processes taking place simultaneously, but on *two different time scales*. This will help to discriminate *two semiotic components* of the language of mathematics, the compositional and the referential. The compositional component of the language of algebra concerns how individual signs are put together to form composite expressions, such as those used for representing polynomials, matrices, and tensors. This component is closely connected to the syntax of algebraic language. However, syntax is a well-defined notion in logic, and it automatically uses functional notation, so we prefer using the expression *compositional component*, to which we can attach our own meaning and strip the symbols of any functional connotations. Similarly, the *referential component* concerns how the composite expressions represent aspects of reality and how they acquire meaning from this representation. Just as in the previous case, the referential component is closely connected to the semantics of the algebraic language. However, semantics is a well-defined notion, and we prefer speaking about the referential component in order not to give the symbols unnecessary logical connotations.

Having distinguished the compositional and referential components, we will leave the history of mathematics and shall examine the individual semiotic

[4] These commentaries in ordinary language, of course, are part of the *practice of algebra*, yet we suggest *not* counting them as part of *algebraic language proper*. It is as if algebraic symbols – and these symbols can have the surface form of words, for what turns them into symbols are the rules, which are applied to them – were read in English and the commentary, say, in Hungarian. This difference is even more radical in the case of geometry, where the iconic language of geometry belongs only to that which occurs on paper due to the rules of construction and letters, as (Netz 1999) convincingly showed. All commentaries in ordinary language are excluded. We all are, of course, bilingual; we understand both English and Hungarian, but if we wish to reconstruct the functioning of the language of algebra in this restricted sense, we must pretend that we cannot speak Hungarian. We can do whatever the instructions in Hungarian require, but they are not part of the English language. By a *rational reconstruction* of the symbolic language of algebra (or the iconic language of geometry), we mean the reconstruction of the rules that govern this English part of our common (bilingual) linguistic practice.

[5] This second move is, of course, a methodological decision. If we sufficiently well understand both the dynamics of the development of the language of algebra and the dynamics of the development of the language of differential and integral calculus, it will be possible to analyze their mutual influences.

components separately. We understand semiotics in the most neutral and uncontroversial way as a theory that describes how signs acquire meaning and reference. Applied to the language of algebra, the role of semiotics is to elucidate how particular signs, such as the sign for the cube root, are used, what their meaning is and to what, if anything, they refer, how they join other signs to form more complex expressions, etc.[6] These semiotic components of the language of mathematics have already been studied, first by Frege when he developed his system of logical notation (Frege 1879) and then by Wittgenstein in his picture theory of meaning (Wittgenstein 1921). In a sense, the following pages are an attempt to apply Frege's and Wittgenstein's theories to the understanding of the epistemological structure of the language of algebra.[7]

In the first part of this discussion, algebraic symbolism is distinguished from other similar systems, such as arithmetical symbolism, the symbolism of the differential and integral calculus, and the symbolism of the predicate calculus; and epistemological characteristics such as the logical, expressive, methodological, integrative, and explanatory power of algebraic symbolism are characterized. In the second part, the picture theory of meaning is introduced, first as a tool for analyzing the development of the iconic language of geometry; then it is transferred to algebra to be used in the reconstruction of the referential component of algebraic language.

2 The Compositional Component of the Language of Mathematics

Frege called his variant of logical calculus *Begriffsschrift* – concept-script (Frege 1879). In the subtitle he characterized it as "*eine der aritmetischen nachgebildete Formelsprache des reinen Denkens*" (a formula language of pure thought created according to the model of the formula language of arithmetic). Later, in *Funktion und Begriff*, Frege outlined the *history of mathematics* as a sequence of such formula languages (Frege 1891, p. 30; English translation p. 40):

> If we look back from here over the development of arithmetic, we discern an advance from level to level. At first people did calculations with individual numbers, 1, 3, etc.

[6] Signs are human artifact, and so they are rooted in social practices of communication, problem-solving, knowledge preservation, and the like. This chapter will not deal with these questions, which are well described in the literature by scholars such as Karine Chemla, Albrecht Heeffer, Jens Hoyrup, Raviel Netz, Vincenzo de Risi, and Roy Wagner in the works listed in the references.

[7] The reader may argue that we also apply general theories to the language of mathematics. However, compared with the approaches mentioned in footnote 1, the difference is that both Frege's concept-script and Wittgenstein's picture theory of meaning are theories that generalize some principles of the language of mathematics. In Frege's case, this is clear from the subtitle of his book. In contrast to the books mentioned in footnote 1, which apply notions from cognitive psychology, general semiotics, or linguistics to the language of mathematics, here concepts that were created on the basis of the analysis of the language of mathematics itself shall be used.

$$2 + 3 = 5 \qquad 2.3 = 6$$

are theorems of this sort. *Then they went on* to more general laws that hold good for all numbers. What corresponds to this in symbolism is the transition to the literal notation. A theorem of this sort is

$$(\mathrm{a} + \mathrm{b}).\mathrm{c} = \mathrm{a.c} + \mathrm{b.c}.$$

At this stage they had got to the point of dealing with individual functions; but were not yet using the word, in its mathematical sense, and had not yet formed the conception of what it now stands for. *The next higher level* was the recognition of general laws about functions, accompanied by the coinage of the technical term "function." What corresponds to this in symbolism is the introduction of letters like f, F, to indicate functions indefinitely. A theorem of this sort is

$$\frac{dF(x).f(x)}{dx} = F(x).\frac{df(x)}{dx} + f(x).\frac{dF(x)}{dx}.$$

Now at this point people had particular second-level functions, but lacked the conception of what we have called second-level functions. By forming that, we make *the next step forward*.

In this passage, Frege distinguishes four stages in the development of mathematics.[8] Importantly, he distinguishes these stages based on the way they use symbols to express *generality* and, in particular, he characterized algebra on the basis of its use of *letters* to express generality. We now turn to the language of algebra, i.e., the second stage in the development of mathematical symbolism, and thereby complement Frege's description of it as a system of "*laws that hold good for all numbers*" with more detail in order to obtain a more comprehensive set of semiotic characteristics.

The point of introducing the long quote from Frege is that Frege found *a level of abstraction* (and the corresponding temporal scale) at which the development of mathematics may be analyzed if we wish to understand the main innovations that shape the language of mathematics. Of course, generality is not the only aspect that is changed in this development. By supplementing generality with other aspects,[9] which might be called the *potentialities of the language of mathematics*, we can offer a comprehensive view of the first component of the semiotic system of mathematics:

[8] We start our analysis of the compositional component of the language of algebra with Frege. Rotman criticized Frege's views: "*Frege gives no idea, explanation, or even hint as to what this 'something' might be which allows the subjective, temporally located bearer* [of thought] *to 'aim' at an objective, changeless thought*" (Rotman 2000, p. 33). Below we shall answer Rotman's criticism by complementing Frege's analysis with the analysis of the pictorial form of the language of mathematics.

[9] These aspects are *analogous to generality* in the sense that they change on the same scale as generality and characterize the language of mathematics at the same level of abstraction as the quote from Frege. Potentialities 1, 2, 4, and 5 are presented in more detail in Kvasz (2008, pp. 11–84) while 3 and 6 are in Kvasz (2012).

1. *The logical power*, which shows how general formulas can be proven in the language
2. *The expressive power*, which shows what new facts and relations, which in the previous stages defied expression, can now be expressed in the language
3. *The methodological power*, which shows what new methods can be introduced into the language where at previous stages there was only a tangle of unrelated tricks
4. *The integrative power*, which shows how language discloses unity and order where, in the previous stages, only unrelated cases were seen
5. *The explanatory power*, which shows how the new language makes it possible to explain the failures of language that were incomprehensible at the previous stage
6. *The constitutive power*, which shows how language helps us to transcend the limits of the previous language and to establish radically new kinds of objects

When it comes to mathematical symbolism, linguists are usually attracted by its effectiveness, accuracy, and brevity, and interpret it as a tool for the effective coding of meaning that already existed before it had been coded (e.g., O'Halloran 2005, pp. 110 and 126). However, symbolism can do more. Accuracy, brevity, and efficiency can appear in some contexts as secondary to the logical, expressive, and methodological power of the language. Thanks to these potentialities, mathematical symbolism is able to radically change the scope, reach, and depth of our thinking. Symbolism can give our thinking new qualities, such as generality.

The main feature of the symbolic language of algebra used during the fifteenth and sixteenth centuries was the introduction of a symbol representing the unknown quantity (see, e.g., Scholz 1990, Alten et al. 2014, and Hoyrup 2019). We all learned a version of this symbolic language in high school. Let us consider its potentialities.

2.1 The Logical Power of the Language of Algebra

The language of algebra contains a fundamental innovation – the symbol for the unknown. In this context, Frege spoke about counting with letters and the ability of the language to express generality in an explicit way. The introduction of letters for the unknown has two sides. The symbol x stands for *an arbitrary object* to which the language can refer and for something with which algebraic operations can be done, i.e., it stands for *an arbitrary argument*. It seems that the unknown originally represented not an arbitrary object but an arbitrary argument, something that could be manipulated. Algebra was born as a language for general operations – as a language for *al-jabr* (if on one side of the equation there are terms that need to be subtracted, then the corresponding value is added to both sides), *al-mukabala* (if the same powers appear on both sides, subtract the smaller term on one side from the larger one on the other), and *al-rad* (if the coefficient of the highest power of the unknown is different from one, then it divides the entire equation), and one of its earliest documents is (Al-Chorezmi, 850).

A letter, for instance x, represents an arbitrary number, generality thereby becoming a fully explicit aspect of the formula language.[10] Nevertheless, the symbolic language of algebra contained a further aspect that the language of synthetic geometry lacked: functionality. Let us take the formula expressing the solution of the equation $ax^2 + bx + c = 0$:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{1}$$

The parameters a, b, and c give this formula generality. However, the formula expresses in a symbolic way not only the solution of *any quadratic equation*, but it also expresses the *steps* of its calculation.

A diagram of synthetic geometry contains all the elements of a particular construction. Each line or point used during the construction process remains as part of the resulting diagram. What is lost, however, is the sequence of steps. Therefore, in geometry textbooks, the diagram is supplemented by a commentary in ordinary language which indicates the order of steps of the construction.[11] The symbolic language of algebra does not need such a commentary because it expresses the order of the steps of the calculation to a sufficient degree.[12] The formula (1) contains a set

[10] Of course, if we were to ask from whence comes this generality, the answer would be, from ordinary language. In ordinary language, we have words such as every, all, any, etc. which allow us to express generality. However, this is not the point. In a sense, all aspects of the formula languages of mathematics are derived from the corresponding aspects of ordinary language. The difference is, however, that a formula language is governed by explicit rules, and so we can build and manipulate complex representations which would cause comprehension problems in ordinary language. The quantifiers, as used in ordinary language, allow us to go rather far in developing calculus. If we use a notion with three quantifiers, like that of convergence, it is not easy to maintain control of its use, as Cauchy discovered, when he believed he had proven the false theorem that the sum of a convergent series of continuous functions is continuous. The theorem is valid if the convergence is uniform, but the difference between the pointwise and the uniform convergence is the order of quantifiers, which is difficult to handle in ordinary language. Even more delicate is the case of four quantifiers, like in the notion of Hausdorff dimension. The same thing applies to generality, where algebra is the symbolic tool that allows us to build and manipulate complex representations of general properties and relations in the form of polynomials of higher degrees. Linear and quadratic equations can be handled in ordinary language, as the Babylonian texts sufficiently manifest. However, to solve a cubic equation, to understand this solution, and first of all to be able to explain this solution to others, requires at least a rudimentary formal language, that is, a language in which you can handle substitutions, raise composite terms to powers, transform equations, and the like.

[11] This does not influence the rigor of proofs, as shown in (Netz 1999 or Manders 2008a, b). However, from the point of view of our reconstruction, these commentaries are not part of the language of synthetic geometry.

[12] To express the order of steps in full detail, some conventions concerning the priority of arithmetic operations and the scope of the root must be met. However, the letter labels in geometry, as shown by Netz, can tell us a lot about the order of operations, although it is still true that an algebraic formula indicates to a sufficient degree the order of steps of its execution, while in a geometric diagram this is mostly lacking. If we take, for instance, formula (1), it is true that it uses some conventions, as for instance the priority of multiplication over addition. In our view, this means only that the formula is not indicating the steps of the calculation unequivocally, thus we could produce

of dependencies (we must first evaluate the difference of b^2 and $4ac$, before we can take the square root of it; but before we can evaluate the difference, the two terms, the b^2 and $4ac$ must be themselves evaluated), which introduce some order into the steps of the calculation of its value.

An algebraic formula thus denotes not only *the quantities* that appear as the result of say, a geometric construction, but it also indicates *the order* of the steps. Thanks to this, it is possible to represent the *process* of a geometric construction (a sequence of steps) as an algebraic *object* (an algebraic formula, the different elements of which correspond to the steps of the construction). It turns out that in a geometric construction, only quantities expressible by the four operations and a sequence of quadratic root extractions can occur. Thus, it is possible to describe the system of all constructible quantities as a number field, that is, the union of all possible sequences of quadratic extension. To prove the non-constructability of a particular geometrical object, it is sufficient to prove that it contains a line segment, the length of which does not belong to that field (see, e.g., Stewart 1989).[13]

The language of algebra can prove theorems of this kind because the process of solution can be represented in the symbolic language. The question of solvability thus becomes a question of whether a certain linguistic expression belongs or does not belong to a certain set of expressions. In the iconic language of synthetic geometry, we cannot *represent* the set of all constructible objects in the form of a mathematical object (such as, for instance, a number field), because this iconic language cannot represent a procedure (the process of construction) but represents only its outcome (the constructed object). In contrast to this, the language of algebra can express a procedure (in the form of an algebraic formula). Thus, in algebra, it is possible to reconstruct the steps of a geometric construction. As the construction goes on and on, the diagram becomes more and more complex, and the accompa-

two or three alternative readings, the choice made being based on conventions. The point is that in a geometrical figure, we have no alternative readings of the process of construction, as we have no reading of the order of steps at all. It is not that the determination of the order of the steps of construction in a diagram is not perfect, there is simply none. (We have, of course, to keep in mind footnote 4 and the distinction of the "Hungarian" and "English" part of the text. When I say that in a geometrical diagram there is no indication of the order of steps of construction, I mean of course the "English," that is, the iconic part of the diagram. I admit that in "Hungarian" everything is present, but we are reconstructing the "grammar" of "English," that is, of the compositional component of the language of mathematics).

[13] It is important to realize that in geometry, we can perform a construction, but we cannot turn it into an object. Only the outcome of a construction, but not the process itself, is iconically represented, i.e., represented as a geometrical object. Algebra changes this situation as it represents the steps by means of operations, and thus a formula indicates not only the result but also how it was obtained. So, the question of the constructability of an object is, from the algebraic point of view, identical with the question about whether the features of the object can be expressed by algebraic means that correspond to the different construction steps.

nying formulas will become more and more involute. Nevertheless, thanks to the functionality of the algebraic language, the order of steps is represented.[14]

At the end, we thus have a set of algebraic formulas (and the corresponding set of quantities) that represent the process of construction. The point is that we can characterize by algebraic means the sets of quantities that correspond to geometric constructions. They belong to a finite sequence of quadratic extensions of the field of rational numbers. So, it is possible to form the union of all such finite sequences, which is the set K of all constructible quantities. This set has several algebraic properties. To prove that a quantity, such as the length of the side of a regular heptagon inscribed into the unit circle, *cannot be constructed* by ruler and compass, it is sufficient to show that it *does not belong* to the set K. So, the nonexistence of a construction was turned into the relation between a set and an element, details of which can be found in Courant and Robbins (1941). The proofs of insolvability are thus an illustration of the logical power of the language of algebra.

2.2 The Expressive Power of the Language of Algebra

The expressive power of the language of algebra is based on the operation of exponentiation, i.e., the creation of an indefinite sequence of powers of the variable: $x, x^2, x^3, x^4, x^5, x^6, x^7, x^8, \ldots$. The first three members of this sequence, x, x^2, and x^3, are anchored in geometry: x is length, x^2 is area, and x^3 is volume.

Nevertheless, algebraists have overcome the limitations of three-dimensional space by means of a syntactic operation that could be iterated indefinitely. It is this ability to transgress the boundary of space that establishes the expressive power of the language of algebra. In geometry, an unknown quantity is expressed by a line segment of indefinite length, the square of the unknown by a square constructed above this line segment, and the cube of the unknown by the corresponding cube above the line segment.[15] The language of algebra is able to create powers of any degree. Italian mathematicians of the sixteenth century, such as Girolamo Cardano,

[14] Symbols of algebraic operations are *functional symbols*. In expressions like $\sqrt{5+4}$, 5 and 4 are arguments of the (binary) *function* of addition, while 9 is its value. This value is the argument of the (unary) *function* of square root, and the number 3 is its value. So, in this formula, the addition must be executed prior to taking the square root. If we have a more complex *formula*, composed of, say, 250 elementary symbols, it is possible to form a tree indicating the dependence of different steps. In the case of a geometric diagram that consists of 250 points and lines, something of this sort is not indicated by the diagram itself, but only by a commentary attached to it. This commentary is in ordinary language, and so is not part of the language of geometry.

[15] Many Greek and Arab mathematicians discussed sequences of proportional lines, which allowed them to speak about quantities, which algebra would represent as monomials and polynomials of high degrees. Nevertheless, these discussions occurred in ordinary language, and thus, due to the rule expressed in footnote 4, they cannot be counted as belonging to the language of geometry, understood as a representational tool. This only shows that linguistic practice is always much richer than the rules explicitly allow.

called the unknown *res* and symbolized it as *r*, the square of the unknown *zensus* and symbolized it as *z*, so they wrote the fourth power as *zz* (*zenso de zensu*), the fifth as *rzz*, the sixth as *zzz*, etc. In this way, the language of algebra surpassed the boundaries placed on the language of geometry by three-dimensional space.

In *Ars Magna Sive de Regulis Algebraicis* (Cardano 1545), Cardano formulated the equation of the third degree as "*De cubo & rebus aequalibus numero*" and stated its solution in the form of a rule. "*Cube one third of the number of things; add to it the square of one-half of the number; and take the square root of the whole. You will duplicate this and to one of the two you add one-half the number you have already squared and from the other you subtract one-half of the same. You will then have a binomium and its apotome. Then subtracting the cube root of the apotome from the cube root of the binomium, the remainder [or] that which is left is the thing*" (Cardano 1545, p. 98).[16] To understand what Cardano is doing, we can write the equation in modern notation as $x^3 + bx = c$ and its solution as:

$$x = \sqrt[3]{\frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3}} - \sqrt[3]{-\frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3}}. \qquad (2)$$

Cardano never wrote such a formula. In his time, algebra was a set of verbal rules.[17] The language of algebra contained names for powers and roots as well as the terms *binomium* and *apotome*. Cardano called $\frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3}$ a *binomium* and $-\frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3}$ an *apotome.* These seem to be some of the first technical notions denoting *terms*.[18] The solution of the cubic equation thus illustrates the expressive power of the language of algebra.

[16] Cardano expressed the rule for the solution of the *general* cubic equation of the form as $x^3 = bx + c$ only in words, as he did not have *symbolic means* for expressing coefficients of the general case. Nevertheless, his language contained *technical terms* representing the coefficients of the equation, and it described the steps of the construction of the solution in a precise technical way.

[17] It seems that Cardano's book has so impressed some philosophers and linguists that they have claimed that mathematical expressions are rules and not descriptions (see, e.g., Barton 2008, p. 127). What they claim is partly true. The case of Cardano shows that there was a time when algebra was a set of rules (*regula della cosa*). However, to interpret all of mathematics on the basis of this limited fragment of algebra is difficult to justify.

[18] These terms are, of course from geometry, from Book X of Euclid. However, in the iconic language of geometry, they are present only as line segments, i.e., without an indication of the steps of their construction; such steps in Euclid are given in ordinary language, and so in our account (expressed in footnote 4), they do not belong to the language of geometry proper. Algebra turned these quantities into terms. Thus, despite the fact that Cardano's proof of the "formula" was geometric – a long time passed before mathematicians realized that algebra also needs axiomatic foundations – the formula itself is expressed in the language of algebra, and thus it illustrates the expressive power of the language of algebra. Ancient geometers were not able to solve cubic equations – the expressive power of the language of geometry did not allow this.

2.3 The Methodological Power of the Language of Algebra

Euclidean geometry is a set of construction tricks.[19] Each problem is solved in a special way that needs to be remembered. Algebra replaced these tricks with the analytic method presented by François Viète in *In Artem Analyticam Isagoge* (Viète 1591).[20] Before Viète, algebraists used different letters (*r*, *z*, *c*, *zz*, *rzz*, etc.) to denote *different* powers of the *same* unknown, so normally they used equations with only one unknown. This does not mean that mathematicians could not introduce more than one unknown. Nevertheless, their handling required special attention and was not done automatically by the syntax of the symbolic language. Viète marked different powers of the same unknown with the *same letter*, with its power indicated by a word. He wrote the first three powers of the unknown *A* as *A-latus*, *A-quadratum*, and *A-cubus*. The capital letter *A* thus indicated the identity of the unknown, and the word indicated its power. This made it possible to work simultaneously with several unknowns.

Another important linguistic innovation introduced by Viète was to mark the coefficients of the equation with letters and thus to introduce a distinction between a parameter and an unknown. Viète's analytic art, as he called his method, was based on *expressing* all aspects of a problem, the known as well as the unknowns, by symbols, "*lifting*" the problem in its generality into the symbolic language, *turning* the problem into an equation. This gives rise to the analytic method which is *a general procedure for solving problems of a certain kind, which is explicitly expressed in language*.

The methodological power of the language of algebra is thus based on a linguistic innovation – the introduction of a *conventional distinction* between the symbols for the unknowns, Viète using the vowels *A*, *E*, *I*, etc., Descartes the letters from the end of the alphabet *x*, *y*, *z*, etc., and the symbols for the parameters, Viète using consonants *B*, *C*, *D*, Descartes the letters from the beginning of the alphabet *a*, *b*, *c*, etc. Nevertheless, it is a strange distinction – it exists only in the language and does not correspond to any real difference in the world described by the language. It corresponds to the difference in the degree of our knowledge: the unknown is unknown to us, while the values of the parameters are known to us. Thanks to this distinction, a problem can be expressed in its generality, using

[19] This view of geometry is due to Descartes. Later, Poncelet and Chasles were able to change the practice of geometry, and under the heading of projective geometry, they also introduced into geometry aspects of the methodological power of the language of algebra. However, even if one accepts that projective geometry has some esthetic appeal, Klein's *Erlanger Program* shows the methodological superiority of algebraic methods also in this area.

[20] Pappus described the analytic method temporarily: one should imagine the problem already solved and work backwards in time along the possible steps of its construction or proof until one comes to a contradiction or to something familiar. Thus, Pappus' analysis operated outside the iconic language of geometry proper. Viète's main achievement was that he incorporated analysis into the language of mathematics.

equations. The distinction between the unknowns and the parameters forms the basis of the analytic method and thus establishes the methodological power of the language of algebra.

2.4 The Integrative Power of the Language of Algebra

Cardano considered the equations $x^3 + bx = c$; $x^3 + c = bx$; and $x^3 = bx + c$ to be different and dealt with each of them separately. It was natural to deal with these types of equations separately, because European mathematicians of his time usually accepted only positive numbers both for the coefficients and for the solutions. Michael Stifel found a way to simplify this situation. In *Arithmetica integra* (Stifel 1544), he introduced rules for counting with negative numbers. He also started to use negative numbers as the coefficients of equations. Thus, it became possible to unite the three kinds of equations of the third degree into one form, which was an important step on the way to the notion of a polynomial. Polynomials, just like matrices, are formal objects. By their means, algebra achieved a deeper understanding of the unity of algebraic expressions. Where Cardano saw three different problems, namely, $x^3 + bx = c$; $x^3 + c = bx$; and $x^3 = bx + c$, Descartes presented only one, the general equation of the third degree. This unity illustrates the integrative power of the language of algebra.

By iterating an operation with the variables beyond any boundary, and combining the thus created expressions into *terms* and forms, the language of algebra has acquired a wealth of expressive possibilities. The language of algebra makes it possible to represent a situation by means of several syntactically different but semantically equivalent expressions. The process of solving an equation is based upon passing from one such representation to other semantically equivalent ones. If we take Cardano's equation: "*Cubus et rebus equalibus numero*," i.e., $x^3 + bx = c$, (in our symbolism), there are four different but semantically equivalent ways of writing it:

$$x^3 + bx = c, \quad x^3 = -bx + c, \quad x^3 - c = -bx, \quad x^3 + bx - c = 0.$$

The equation "*Cubus equalis res et numero*," i.e., $x^3 = bx + c$ can be similarly written as:

$$x^3 = bx + c, \quad x^3 - bx = c, \quad x^3 - c = bx, \quad x^3 - bx - c = 0.$$

And the last type, "*Cubus et numero equalibus rem*," i.e., $x^3 + c = bx$ can also be written in four forms:

$$x^3 + c = bx, \quad x^3 - bx = -c, \quad x^3 = bx - c, \quad x^3 - bx + c = 0.$$

Descartes unified the equations $x^3 + bx - c = 0$, $x^3 - bx - c = 0$ and $x^3 - bx + c = 0$ into a single form from which the notion of a polynomial could be created.[21] Thus, the notion of a polynomial was formed by uniting *formally similar* representations of *semantically different* situations.[22]

The four equations in each line are *semantically equivalent*. They are syntactically distinct representations of semantically identical content. However, Descartes did not use the semantic affinity of the expressions in a particular line, but *united syntactically similar but semantically unrelated expressions*.[23] He did not unite the expressions which say something similar, but he united those that are syntactically alike, even though each of them says something completely different. The unity he introduced into the language is not semantic, but formal. What unites these expressions is the common process of solution. The integrative power of the language of algebra consists in its ability to introduce a unity for which there is no support in reality that is represented by the language. Polynomials are new objects created "inside" the language. They were introduced formally, and *this* allows the language of algebra to find connections between situations where there is no relation at the semantic level. The aspect that establishes the integrative power of language is thus the *union of terms into polynomial forms*.

2.5 The Explanatory Power of the Language of Algebra

The language of algebra makes it possible to understand why some geometric problems, such as the trisection of an angle or the construction of a regular heptagon, are insolvable with ruler and compass. In problems that can be solved with ruler and compass, there occur only line segments whose lengths can be expressed using the four arithmetical operations and square roots. The reason for the insolvability of the two aforementioned problems is that their solution requires the construction of line segments, the lengths of which are solutions of irreducible cubic polynomials. Such

[21] Descartes came up with the idea of moving all terms to the left-hand side. The concept of a polynomial was thus created in two steps. The first was the unification of different types of equations into a single form by accepting both the positive and negative values of coefficients. This step was made by several mathematicians in the second half of the sixteenth century, among others by Stifel. The second step consisted in a specific choice of this form by moving all members to the left-hand side and putting zero to the right.

[22] A polynomial is an objectification of the equation, replacing the equation (which is a relation) with an object. A polynomial arises when the left-hand side of the equation is extracted from the equation and manipulated separately from the equation. Nevertheless, first all terms of the equation should be put on the left-hand side, otherwise the polynomial would not correspond to the equation, would not represent it.

[23] The unification is, of course, due to suspending the positive/negative difference. Nevertheless, the unification as such, that is as a change in the writing of expressions (the introduction of a new practice of writing one and the same expression where formerly one would write three different ones), is a change of the language of algebra.

lengths cannot be expressed using only square roots (see Courant and Robbins 1941, pp. 134–139).

Algebra can explain the failure of geometry to construct a regular heptagon using ruler and compass because the polynomial forms that establish its integrative power enable us to introduce the predicate "*to be the root of an irreducible third-degree polynomial*." This predicate is a tool for explaining why the side of a regular heptagon is unconstructible. No line whose length is the root of a third-degree irreducible polynomial is constructible, and the side of a regular heptagon inscribed in the unit circle is just that. In this explanation, we classify quantities that describe real phenomena on the basis of which polynomial they suit. We thus characterize a geometric object by means of whether it suits a certain artificially created linguistic expression. A polynomial (at least at the first sight) has nothing to do with geometry. Geometric lines seem to have nothing to do with polynomials. Surprisingly, if we project polynomials onto the world of geometry, they make it possible to understand the difference that is crucial for their construction. The boundary that separates problems that can be solved with ruler and compass from those that are not solvable in this way is determined by the predicate "to be the root of an irreducible third-degree polynomial." We can understand insolvability with ruler and compass because the language of algebra allows us to make certain distinctions. The explanation of the insolvability of the above problems illustrates the explanatory power of the language of algebra.

2.6 The Constitutive Power of the Language of Algebra

In addition to the equation $x^3 + bx = c$, the solution of which was presented in Sect. 2.2, Cardano also found a rule for the equation $x^3 = bx + c$, which can be written as:

$$x = \sqrt[3]{\frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 - \left(\frac{b}{3}\right)^3}} + \sqrt[3]{-\frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 - \left(\frac{b}{3}\right)^3}}.$$

When he tried to solve the equation $x^3 = 7x + 6$, he obtained a solution, which we would write as:

$$x = \sqrt[3]{3 + \sqrt{-\frac{100}{27}}} + \sqrt[3]{3 - \sqrt{-\frac{100}{27}}}.$$

A negative number appeared below the square root sign, $\sqrt{-\frac{100}{27}}$ which does not make sense.

When we liberate algebraic expressions from their close connection to reality and start to consider them as separate objects, the square roots of negative numbers can

be seen as a special kind of number. Even if we do not know what they represent, we can calculate with them, as Euler put it:

> However, notwithstanding this, these numbers present themselves to the mind; they exist in our imagination, and we still have a sufficient idea of them; since we know that by $\sqrt{-4}$ is meant a number which, multiplied by itself, produces -4; for this reason also nothing prevents us from making use of these imaginary numbers and employing them in calculations. (Euler 1770, § 145)

The introduction of complex numbers illustrates the constitutive power of the language of algebra.[24]

The constitutive power of the language of algebra is based on the introduction of new objects by means of definite descriptions. An example of a definite description is "the only x for which $p(x) = 0$."[25] For predicates which establish the explanatory power of the language, usually there exist objects that satisfy them. These objects are the ground upon which the corresponding explanation is based. However, there are predicates, such as $P(x)$: $x^2 + 1 = 0$, for which we cannot find any object that satisfies them. There is no real number whose square root is equal to minus one. When mathematicians postulate complex numbers, *they seem to adjust reality to fit language*.

In most languages, operations occur in pairs. In arithmetic, we have addition and subtraction, in algebra, raising to power and root extraction, in mathematical analysis, derivation, and integration. It is these inverse operations, i.e., subtraction, root extraction, and integration, that force us to expand the field of objects so that the inverse operation can always be executed. Thus, the operation of subtraction forces the introduction of negative numbers, root extraction yields complex numbers, and integration forces us to introduce transcendental functions.

Therefore, definite descriptions make it possible to define new objects even in situations where these objects cannot be expressed by means of inverse operations (for example, in the case of fifth-degree polynomials). Thus, the development of the formula language of algebra is completed only when mathematicians realize the limitations of the inverse operation.

[24] There are, of course, several ways the complex numbers might be introduced, some algebraic, some geometrical. The positing of the square root of -1 as an imaginary unit is just one among them. However, to fully appreciate this problem, one must understand why the first complex number was discovered in connection with the cubic equation $x^3 = 7x + 6$ and not much earlier in connection with some quadratic equation, which also has complex roots. The reason is that if a quadratic equation with real coefficients has a complex root, it has no real roots, so it is possible to claim that if the discriminant is negative, no root exists. In the case of the cubic equation, Cardano knew that it had a real root, and so there was a chance (which turned out to be the case) that the strange expression with square roots of negative numbers represents the real root $x = 3$. So, we have a context in which the introduction of complex numbers is not only possible, but also meaningful.

[25] When x is understood as a number, it is not unique because a polynomial of n-th degree has n roots. However, this is a small complication – we can talk about field extensions, and these are, up to isomorphism, unique.

2.7 Summary

For the language of algebra, we obtained a list of potentialities, the growth of which characterizes the development of the language of mathematics at the level of abstraction adopted by Frege. The language of mathematics is a human creation and so the increase in its logical, expressive, methodological, integrative, explanatory, and constitutive power lies in specific semiotic innovations. Introducing objects using definite descriptions is an operation that closes the circle of the construction of a new formula language. In the first step of this construction, a new kind of variable was introduced to enable general reference to *objects* of the traditional kinds (numbers or geometrical quantities). In the sixth and final step, by means of definite descriptions, new objects are introduced. These new objects belong to the domain of the same variable whose introduction started the whole process.[26] The circle of construction of the new formula language ends, returning to individuals. However, while at the beginning these were individuals represented by means of ordinary language (as well as arithmetic and synthetic geometry), and the new language was used only as a *means for their more efficient description* (as many linguists understand the role of the language of mathematics to be), in this last step language becomes a *tool for the introduction of new objects*. The complex numbers could not be coherently thematized before the algebraic operations underwent a series of semiotic innovations.

The process of constructing a new symbolic language has been shown to consist in six steps: *first* – the introduction of a *new kind of symbol*; *second* – the introduction of a *new operation* with these symbols that can be iterated indefinitely; *third* – the introduction of an *epistemological distinction* between symbols; *fourth* – the unification of various expressions representing semantically different situations into a *new form* on the basis of their syntactic similarity; *fifth* – the introduction of *new predicates* using the forms; and *sixth* – the postulating of *new objects*. We see that for each of the six *potentialities* of the language of algebra, a corresponding *syntactic aspect* could be found.

3 The Referential Component of the Language of Mathematics

The most suitable tool for the analysis of the *second, referential component of the semiotic system of the language of algebra* seems to be Wittgenstein's picture theory of meaning. We will start by showing how the picture theory of meaning can be used to analyze texts on synthetic geometry. In Sect. 3, we will transfer the picture theory of meaning into algebra, and finally, in Sect. 4, we will take the example of one

[26] In algebra, complex *numbers* have been introduced, and the construction of the language of algebra started with the introduction of letters referring to arbitrary *numbers*.

specific algebraic symbol and seek to elucidate how the *referential component of the semiotic system of the language of algebra* could be established.

In analyzing the development of the referential aspect of the language of geometry in the present section, the basic idea is to apply the picture theory of meaning to pictures contained in mathematical texts. It seems that it is possible to base the analysis of the development of the referential aspect of the language of geometry upon the picture theory of meaning. We borrow some ideas from Wittgenstein's theory but, from the very beginning, we reject the idea of the existence of a single, universal, unchanging pictorial form common to all languages. Instead, the assumption that the language of geometry passes through developmental stages, which differ from each other in their pictorial form, makes it possible to use Wittgenstein's picture theory of meaning and, above all, the notion of the pictorial form in the identification of stages in the history of geometry. The pictorial form allows a conceptual grasp of the unity of a stage in the development of geometry as a set of propositions together with their representative and argumentative practices, all of which share a common pictorial form.

Nevertheless, Wittgenstein's notion of pictorial form is complex and comprises several core ideas, only some of which are useful for the reconstruction of the referential aspect of algebraic symbolism. We can discriminate six core ideas.

1. The idea that *there exists a form that is shared* by both, i.e., the language and the reality that this language represents. This idea allows Wittgenstein to explain how it is possible that language represents reality.
2. The idea that this shared *form is compositional* so that the parts of language into which the form decomposes a proposition, and the parts into which it decomposes the reality represented by the proposition, correspond to each other. Thus, these parts can be projected onto each other because the corresponding parts are in the same relations to each other on both sides.
3. The idea that this *form is logical* in the sense that it represents the logical relations between propositions and the corresponding states of affairs, which make these propositions true or false.
4. The idea that any linguistic representation comprises *an implicit form* that cannot be expressed by the language but is displayed by it.
5. The idea that the implicit form *is identical* with the compositional form of the language.
6. The idea that this implicit, logical, compositional form of the language is the *one, unique, universal form* common to all languages.

Wittgenstein gave this unique, implicit, logical, compositional form of the language the name *pictorial form* and made it the basis of his picture theory of meaning. Although Wittgenstein later abandoned his picture theory of meaning, it seems that it is still possible to use it for the restricted task of elucidating the referential aspect of geometric pictures or of algebraic symbolism. However, to turn the picture theory of meaning into a working interpretative tool, from the above six core ideas, we will *retain* three (points 1, 2, and 4) and modify the rest (points 3, 5, and 6).

First (contra point 3), it seems that the compositional form of any language must be identical with the (full) logical form. At the time Wittgenstein wrote his *Tractatus*, Frege's logic was the latest tool of symbolic representation that mathematics had developed, and so it was rather natural to identify the compositional form by its means. However, after Gödel, Turing, Post, Church, and others developed the new representational tool in the form of the *theory of algorithms* (or recursive functions, Turing machines, or some other equivalent form) and, on this basis, the theory of generative grammar was introduced, it was natural to drop the identification of the compositional form with logic. Instead of taking one particular tool and identifying the compositional form of language with it, it seems more natural to leave this question open and decide how the compositional form actually works in each particular case. That is why in Sect. 2, we tried to reconstruct the actual logical, expressive, integrative, etc. power of the language of algebra and did not substitute for them the logical, expressive, integrative, etc. power of predicate logic (as Wittgenstein did in the *Tractatus*).

Second (contra point 5), it seems unnecessary to identify the compositional form of a language with its implicit component. The reason is that, even if the compositional structure remains the same, the border between the implicit and the explicit part of the language can move and, as it will appear in this section, it actually did move in history. We will shortly describe three different kinds of implicit form – the perspectivist, the projective, and the coordinative, all three of which have the same compositional structure, thus project the same parts on reality. This indicates that the form of what is implicit has a much finer structure than the compositional form. We must fix the compositional form to be able to study the (changes in) the implicit one.

Third (contra point 6), it seems that there exists a plurality of forms – both compositional (which were described in Sect. 2) and implicit. The development of mathematics can be interpreted as the development of the compositional and the implicit form of its language.

Yet despite these differences, if we *narrow down Wittgenstein's notion of pictorial form* (i.e., retain points 1, 2, and 4) while allowing a plurality of pictorial forms (instead of point 6) and distinguish the pictorial form from the compositional (instead of identifying them), it seems that we can obtain an important tool that allows us to understand the referential aspect of such languages as symbolic algebra and classical mechanics. Because the compositional form of formal languages was already well represented by Frege's notion of a formula language, and Frege understood the compositional form as evolving (as the quote from *Funktion und Begriff* testifies), we suggest retaining Wittgenstein's term "pictorial form" in this narrower sense for the implicit aspect of language which Wittgenstein introduced into the philosophical discourse. This text is thus not an exercise in Wittgenstein exegesis, but rather an attempt to mobilize the insights from the Tractatus to understand the referential aspect of formal languages such as symbolic algebra.

3.1 Pictorial Form as a Tool of Epistemological Reconstruction

As mentioned above, Wittgenstein's picture theory of meaning, understood as a tool for the epistemological analysis of the development of the language of geometry, is based on three theses.

The first thesis is the thesis of the *existence of the pictorial form*. It can be formulated as the belief that, in addition to grammatical structure, language has a further, independent structure, which Wittgenstein called the "pictorial form" TLP 2.2:[27] "*A picture has logico-pictorial form in common with what it depicts*." An important feature of this form is that it cannot be expressed in language. TLP 2.172: "*A picture cannot, however, depict its pictorial form: it displays it*." Wittgenstein also described two elements of this form. The first is the limits of language. TLP 5.6: "*The limits of my language mean the limits of my world*." These limits are, in the light of TLP 2.172, inexpressible in the language. It is not difficult to transfer the notion of the limits of language into the context of geometric representation. For example, in a Renaissance painting, these limits are present as the horizon. However, if we went into the landscape depicted in the painting to the place where, according to the painting, the horizon was located, we would not find anything special. The horizon that appears in the painting is an element to which nothing corresponds in reality. It was not created by strokes of the brush. It only appeared when the painting was nearing completion. The horizon is an illustration of Wittgenstein's distinction between what the picture "depicts" and what it only "displays." Although the horizon is present in the painting and we see it clearly, it is not there because it represents some element of reality. The horizon does not belong to the subject matter but to the form.[28] According to Wittgenstein, the second aspect of the pictorial form is the subject of language. TLP 5.632: "*The subject does not belong to the world*." This means that it is an element of the pictorial form. TLP 5.641: "*The philosophical self is not the human being, not the human body, or the human soul, with which psychology deals, but rather the metaphysical subject, the limit of the world – not a part of it*." Like the horizon, the subject is present in every representation. Each representation is constructed from a particular point of view, but this point of view is not depicted.

The early Wittgenstein believed that the pictorial form is unique, and thus, if a particular representational system claims to be a language, it must have the same pictorial form as the one described in the *Tractatus*. Later, Wittgenstein abandoned the idea of the existence of a single pictorial form common to all languages. He

[27] The sentences of Wittgenstein's *Tractatus Logico-Philosophicus* (TLP) are numbered by a system of decimal classification, whereby sentence TLP 2.2 comments on sentence TPL 2, just as sentence TLP 2.172 comments on sentence TLP 2.17. Therefore, it is customary to refer to the sentences of the *Tractatus* by their number.

[28] The horizon (i.e., the point at infinity) has played a crucial role throughout the entire history of geometry. The systems of Euclidean, projective, and non-Euclidean geometry differ in the nature of the horizon. This indicates that by analyzing the changes of the pictorial form of the language of geometry, we can really characterize from an epistemological point of view the development of this language.

introduced the concept of a language game to capture the diversity and richness of the possible roles of language. He had reasons for abandoning the position of the *Tractatus*, which we shall not question. Nevertheless, the description of the pictorial form contains valuable insights into the functioning of the second component of the semiotic system of mathematics. Therefore, instead of abandoning the picture theory of meaning for the sake of pluralism, we can try to introduce plurality into the picture theory itself. As already mentioned, the aim is not to interpret the views of the historical Wittgenstein, but to use his picture theory of meaning as a tool in a study of the development of geometry. It turns out that many aspects in the development of geometry can be understood if we interpret the development of geometry as a sequence of pictorial forms.

The second thesis is the thesis of the plurality of pictorial forms. This thesis can be formulated as the belief that there are several different pictorial forms and several different ways in which language can be related to the world. The turning points in the development of geometry, such as the discovery of non-Euclidean geometries, are closely tied to a transition from one pictorial form to another. In order not to cause confusion, a few adjectives can be added to the term "pictorial form." We will thus talk about the *perspectivistic form of the language of geometry.* The word *geometry* refers to the discipline whose language we are analyzing and indicates the compositional structure of language, in the framework of which the pictorial form operates. We assume that the pictorial form of the language of algebra will consist of different elements from those that form the pictorial form of the language of geometry. The word *perspectivistic* refers to a specific period in the development of this discipline (i.e., geometry) during which linguistic representations were understood to be pictures of reality, constructed from a fixed point of view. In other developmental periods of the discipline, the pictorial form had a different nature.[29]

The third thesis is the thesis of the *evolutionary nature of pictorial form*. As mentioned above, Wittgenstein identified the pictorial form with a particular compositional structure of language, namely, its logical form. As a consequence of his rejection of the metalanguage in logic, Wittgenstein was committed to rejecting any kind of external position concerning the pictorial form. As soon as we abandon the identification of the compositional and the logical form, we can accept the existence of an evolutionary process by means of which the pictorial form can evolve. It is the process of *incorporating the pictorial form of language into the language* itself. This does not contradict our first thesis because in this way a new language with a new pictorial form emerges, whose pictorial form is again only implicit. The history of mathematics exhibits several examples of such a process.

[29] The phrase *perspectivistic form of the language of geometry* can be abbreviated by omitting some of the words that are obvious from the context. Thus, use will be made of abbreviated terms such as the *perspectivistic form* (to emphasize the particular period), the *form of the language* (to emphasize its conventional character), and the *pictorial form of the language of geometry* (to emphasize the particular discipline). However, all abbreviated forms should be understood as abbreviations of the full name: *perspectivistic form of the language of geometry.*

To sum up, the notion of pictorial form as a tool of the epistemological reconstruction of the development of mathematical disciplines such as geometry or algebra is based on three theses:

The thesis of the existence of pictorial form as a linguistic structure that includes everything that cannot be explicitly represented in the language but is only displayed. The elements of the pictorial form are not accidental defects of language manifesting its failure to represent but form a system establishing the relationship between language and reality.

The thesis of the plurality of pictorial forms, according to which the language of a particular discipline is at each moment in history based on a single pictorial form, but this form can change over time. Changes in the pictorial form are tied to the turning points in the development of the discipline.

The thesis of the development of pictorial form. The pictorial form of the language J_1, which cannot be expressed in J_1 but is only displayed, can be incorporated into the language. For example, in projective geometry, the point of view of the previous pictorial form is represented as the center of projection. This incorporation creates a new language, J_2, with a new pictorial form that cannot be expressed in J_2. However, J_2 will contain the means to express the pictorial form of J_1.[30]

3.2 Aspects of the Pictorial Form of the Language of Geometry

As a result of the above three theses, an epistemological reconstruction of the development of geometry has become a well-defined problem – to find elements of the pictorial form of the geometric language, for its developmental stages as normally distinguished by historians of geometry. To cut to the chase:[31]

1. *The point of view* is the point from which the correspondence between the picture and the world is constructed. The use of a single stationary point in the construction of perspective was invented by Brunelleschi. This point belongs to the pictorial form because it is not depicted in the picture, but is only displayed by it. The position of the viewpoint affects the perception of the picture, and its position can often be determined.

[30] The mechanism of incorporating the pictorial form into the language is an optimal tool for describing the development of a theory. The conceptual structure, present only implicitly, i.e., only displayed but unable to be depicted by the language, will become explicit. This will enrich the conceptual systems of the theory with new concepts, but without any contradiction between the conceptual system of the old and new theories.

[31] The possibility of expressing the pictorial form of language J_1 in language J_2 opens the possibility of the emergence of a *sequence of interconnected languages*. In this sequence, the stage J_n contains the pictorial form of the language of the previous stage J_{n-1} and its own pictorial form becomes expressed in the next stage, J_{n+1}.

2. *The horizon,* which, as already explained, is an aspect of the pictorial form (i.e., nothing corresponds to it in the world, nor does it represent any real object), displays the altitude of the viewer's eyes, and so indicates the position of the point of view. The horizon represents the limits of the world depicted by the picture.[32]
3. *The relation of identity*, i.e., the rules allowing us to identify different representations as representations of the same object, can be interpreted as a further aspect of the pictorial form. Painters discovered these rules when studying the changes in the size of the simplest object – a square – forming part of a regular square grid, called *pavimento*. To represent more complex situations, it is necessary to know the exact construction of the *pavimento*, which helps to properly place the individual figures and elements of architecture. The exact construction of the *pavimento* comes from Leon Battista Alberti. The rules of the reduction in the sizes of tiles forming such a *pavimento* are an aspect of the pictorial form. Nothing corresponds to these rules in reality: the tiles do not get smaller with increasing distance. The reduction of the sizes of the tiles displays the *relation of identity*. If we ask an observer about the tiles, she must answer that they are the same, despite the fact that she is looking at quadrilaterals on the picture that are of different sizes.
4. *The relation of synonymy* is defined as the relation holding between expressions that are in principle different but within the given pictorial form are represented by the same individuals. This means that in the universe of discourse, we are unable to illustrate their differences. Synonymy understood in this way is an aspect of the pictorial form, because when the pictorial form changes, the individuals that became newly expressible by the language may illustrate the difference. Expressions that were synonymous within one pictorial form can cease to be synonymous after the transition to a new form. For example, the expressions "equidistant lines" and "parallel lines" are synonymous within Euclidean geometry: equidistant lines are parallel, and parallel lines are equidistant. However, in non-Euclidean geometry, which is built upon the basis of a new pictorial form, they lose synonymy. In non-Euclidean geometry, equidistants to a straight line are curves, while parallels to a straight line are not equidistant.

   By moving to more complex pictorial forms, *the relation of identity expands* – the language becomes able to represent wider classes of things, which we formerly perceived as different objects, as only different positions of the same object. On the other hand, by moving to more complex pictorial forms, *the*

[32] One could object to our including the horizon among the aspects of the pictorial form. The objection might go as follows: "The objective of a perspectivist representation is to represent a landscape (for example) as one would see it from a certain point of view. However, a landscape seen from a certain point has a horizon, which is an intrinsic property of this view, and not just of the painting. The horizon is part of what I see and therefore also of what must be represented in the painting. I cannot go to this point because it is at infinity. The objective of the painting is to represent the landscape as I see it. And I see the horizon when I am in front of the landscape." This objection is interesting. To answer it requires more technical tools, so we postpone the answer to the next section.

relationship of synonymy narrows down – properties that the language was unable to distinguish, because it lacked suitable objects, become distinguishable.

5. *The background* against which we recognize the identity of objects in geometry is the space. A geometrical object retains its identity when we move it through space. Space is an aspect of the pictorial form because it cannot be depicted in the picture: it is only displayed. If, on a picture, two lines converging toward the horizon are drawn, adopting the perspectivistic form of language means acquiring the ability to see in them two parallels running into the depth of space. To the question "What do you see in the picture?" we must answer not according to what we are looking at (convergent lines), but according to what, strictly speaking, is not present in the picture (parallels).
6. *Ideal elements* are the last aspects of the pictorial form. An example of ideal elements in geometry are points at infinity. The purpose of their introduction is to unify the schemes, so that, for instance, any two straight lines in the plane intersect, which simplifies geometric constructions.

The viewpoint, the horizon, the relation of identity, the relation of synonymy, the background, and *ideal elements* compose the pictorial form of the language of geometry. These aspects are not factual; nothing corresponds to them in reality. Wittgenstein's words are thus given a precise technical content. The pictorial form includes everything that the language displays but cannot depict.

3.3 The Development of the Pictorial Form of the Language of Geometry

We can apply Wittgenstein's notion of the pictorial form to analyze pictures contained in classical texts from the history of geometry and try to identify in these pictures the various elements of the picture theory of meaning. The development of geometry will become a sequence of pictorial forms so that each stage contains the pictorial form of the previous stage incorporated into the language. The pictorial forms differ from each other in the way they relate representation to reality. We shall start with the perspectivist form of geometry but will make a short detour into painting because the concepts of space and of projective transformations were not introduced by mathematicians but by painters. In this way, the history of Renaissance painting, at least indirectly, is part of the history of geometry.

In the *perspectivist form*, a representation is understood as a picture of a real situation. The paintings of Masaccio, van Eyck, and Leonardo, as well as those of many other Renaissance painters, can be characterized by means of the above aspects of the pictorial form, which means that these aspects were introduced as they are present in the perspectivist form. They depict the world from one fixed point chosen as the point of view. The landscape leads in a regular way to the horizon, to which the parallels smoothly converge, letting the objects in a systematic way diminish with distance. The previous section sufficiently explained why we should interpret perspective as a pictorial form in this sense, that is, as a system of conventions regarding how to "see" what is displayed by a painting. It seems that

the Renaissance can be characterized as that period in the history of art which was based upon *perspective as a pictorial form*. Let us turn to the ensuing period, Mannerism.

In the *projective form* that is typical for projective geometry, a representation of representation is created, and so the concept of a geometric transformation becomes explicitly expressible in the language. The projective form was introduced in Mannerism, but we can introduce it by a short detour into astronomy. It is well known that in 1543, Nicolaus Copernicus (1473–1543) published *De Revolutionibus Orbium Coelestium*, in which he proposed a heliocentric astronomy. With this work, he brought a radical shift in pictorial form, that is, in the way of organizing a pictorial representation. Copernicus does not depict the universe from the fixed point from which we observe it. Naturally, from our terrestrial point of view, the Sun and the stars appear to move while the Earth stands still. Nevertheless, Copernicus insisted that the motions of the Sun and stars are only apparent. To be able to say this, he must have been imagining himself, and any other observer of the skies, from an external point of view far apart from the Earth. In his mind's eye, he abandoned his geocentric perspective and, from somewhere "up there," observed some astronomer seated on a rotating, orbiting globe. From this external position, he could see the motionless Sun and understand that the motion of the Sun, which the terrestrial observer perceives, is only an apparent motion.

We can take the viewpoint of the external observer as the point of view (that is, as an aspect of a new pictorial form) and put it on a par with normal perspective, which was the pictorial form of the Renaissance period. To derive a theoretical analysis of the projective form, we can turn to an etching by Albrecht Dürer. Just like Copernicus, who was born only 2 years after him, Dürer observes somebody at work. In his imagination, Copernicus was observing an astronomer observing the skies, whereas Dürer was observing a painter constructing a perspectivist painting. In a sense, the representation in Fig. 1 has two viewpoints. One is external: it is the point from which the space of the room is constructed. The other is internal, the eye of the painter, or more precisely, the point where the rope with the tube is fastened to the

Fig. 1 Albrecht Dürer (1471–1528), *Underweysung der messung mit dem zirckel und richtscheyt* (1525)

wall. This split (or duplicated) viewpoint is one aspect of the projective pictorial form.

If we realize that Copernicus observed an astronomer (and he himself was an astronomer), just like Dürer was observing a painter (and he was a painter), we see that they are incorporating into representation a reflection of the self. We can term this pictorial form *projective form*, because what Dürer depicted is the projection itself; he is representing the process of representation.[33]

The next pictorial form, that is the *coordinative form*, can be introduced with a similar detour to that of the projective form, but this time not into astronomy but cartography. If we take the map of Africa constructed by the Portuguese cartographers in the fourteenth century, it is a representation that is radically different from perspectivist paintings in the sense that before the astronauts flew above Africa, no one could perceive the shape of Africa. Nevertheless, we know that what the cartographers constructed is a sufficiently accurate picture of Africa. However, it was constructed not as perspectivist pictures are, by someone looking at an external object from a fixed viewpoint; neither was it constructed as projective representations used to be, by someone looking from an external viewpoint at how someone else perceives the sky or a vase. A coordinative representation occurs in a different way. There were dozens, maybe hundreds of sailors who sailed around Africa. None of them saw anything like the shape of Africa; they only collected data which they passed to the cartographers and the cartographers coordinated these different data and created the resulting representation. In a sense, the cartographer resembles the external observer of the projective form. The only difference is that he is not observing, but rather coordinating the data supplied by the sailors. Similarly, the sailors resemble the internal observer in Dürer's etching. Yet again, the parallel is not precise, because the sailor, in contrast to the painter depicted by Dürer, does not see the shape that is represented. He has access only to a rather limited amount of information, which is not "visual" but rather "positional." The cartographer is not so much interested in what the sailor saw but where he was positioned.

[33] Now we are in the position to answer the objection from footnote 33. We must realize that in the objection we have a *landscape*, then the *view* of the landscape as seen by the observer, and finally the *painting* that represents this view. Thus, we are in the situation depicted in Dürer's etching, i.e., in the projective form of language. The thesis that the horizon is an aspect of the pictorial form, and thus is displayed by the picture but cannot be depicted, concerns the *perspectivist form*. The projective form is created precisely by incorporating the form of the perspectivist picture into the language, thus turning it into an explicitly represented system. So yes, the horizon that belongs to the internal point of view (in Dürer's etching, the eye of the painter is represented in the etching) is explicitly present in the representation. If we imagined that the painter painted not a vase but a landscape, and so a horizon would appear in his painting, this horizon would not move when Dürer decided to lower his viewpoint and construct the entire etching from a much lower viewpoint. We would see the entire scene from a lower point, but the horizon on the painting would not move. So, in the projective form, the pictorial form of the previous stage is made explicit. The objection is thus based on a good observation, but it does not take into account that, if we consider the painting as a representation of the inner view, we change the entire situation and thus join Copernicus and Dürer in using the projective form.

We could go on and describe the principles of the five further pictorial forms as they occurred in the development of geometry from projective geometry through non-Euclidean geometry to topology. However, we will return instead to algebra and try to transfer the notion of a pictorial form, as well as the aforementioned three particular pictorial forms, from geometry to algebra.[34]

3.4 Transferring the Concept of Pictorial Form to Algebra

The history of algebra is the subject of intense historical research.[35] In 1842, Georg Nesselmann, in *Versuch einer kritischen Geschichte der Algebra*, introduced the distinction between *rhetorical*, *syncopated*, and *symbolic* algebra. This distinction became popular, and several historians of mathematics used it.[36] In *rhetorical algebra*, the solutions of algebraic problems are given in the form of verbal rules, like Cardano's rule given in Sect. 2.2. In *syncopated algebra*, the technical terms are stepwise replaced by their first letters, and for these abbreviations, a kind of calculus is created. In *symbolic* algebra, illustrated by formula (2), the abbreviations are replaced by genuine symbols. After symbolic algebra, we can distinguish *abstract algebra* that studies structures such as groups, rings, and fields, and *categorical* algebra, which studies morphisms between structures (van der Waerden 1985, Boyer and Merzbach 1989, Scholz 1990, Corry 2004, and Gray 2018).

The description of the pictorial forms in geometry in the previous section is sufficiently detailed to enable the identification of the corresponding sequence of pictorial forms in the development of algebra.[37] This is important because we are

[34] The interested reader can find more details concerning the notion of a pictorial form and its development in geometry in (Kvasz 2008). Nevertheless, these details are not important for an understanding of this chapter.

[35] See the works of Chemla (2003, 2005, 2012), Heeffer (2009, 2010), Hoyrup (1994, 2007, 2019), Oaks (2006, 2007, 2009, 2012), and Wagner (2010, 2017) in the list of references.

[36] In recent years, criticism of this periodization has appeared (Heeffer 2009). Nevertheless, as some linguists still use these terms (O'Halloran 2005, p. 57), we will stick to them as well. The focus of the criticism is *syncopated algebra*, which cannot be taken as a stage in the history of algebra because it overlaps to a significant degree with *rhetorical algebra* and is dependent on it. Nevertheless, from a semiotic point of view, syncopated signs can be the subject of study, under the proviso that they are not considered to form a historical period.

[37] One could object that the form of an algebraic expression is not an image in the Wittgensteinian sense of the word, i.e., the relations of the parts of an algebraic expression are not the same as those of its content. The relation of the sign $\surd$ and 8 are not analogous to what the parts of what $\sqrt{8}$ represents, just like the parts of a polynomial expression are not in the same relationship as the parts of what it represents. It is in a way the ability to express complex objects and relations without having to be the image of them that gives algebra its expressive power and its capacity to transform itself in the course of history. Nevertheless, this objection does not take into account the separation of the compositional and pictorial forms in our approach. Thus, in contrast with point 3 (that the compositional form is identical with the logical form), I consider algebraic symbolism to be a formula language with its own kind of compositional form. The compositional form of the language of algebra represents the functional structure of language, thus it represents $\sqrt{8}$ as the application of

unable to directly transfer to algebra the aspects of the pictorial form described in Sect. 3.2. Nevertheless, it is possible to transfer the pictorial forms from geometry to algebra. For example, the *perspectivist form* of the language of algebra seems to correspond to syncopated algebra. In syncopated algebra, only a number that satisfies the conditions of the problem was accepted as a solution to an equation. Thus, an algebraic equation was understood as a *picture* of a real situation. The *projective form* of language seems to correspond to the first stage of symbolic algebra (represented by Viète and Descartes) when, due to the introduction of substitutions, all numbers satisfying an equation were accepted as solutions of that equation, because they could be connected with the real situation by means of some substitution. The equation thus ceases to be seen as a picture of reality. Between the symbolism and the real situation, a distance was created that is similar to the difference between the external and internal points of view. Only some of the solutions of an equation have meaning in the real situation represented by the equation. This is analogous to the fact that only some aspects of reality as represented from an internal viewpoint correspond to some aspects of the same reality represented from an external viewpoint.

Thus, the identification of the historical sequence of pictorial forms in the development of synthetic geometry enabled the discovery of an analogous sequence in the development of algebra and thereby the transfer, to algebra, of the method of reconstruction of the development of the pictorial form.[38]

3.5 Aspects of the Pictorial Form of the Language of Algebra

Confronting our approach with the inspiring text of Brian Rotman (Rotman 1987) facilitated the transfer of some aspects of the pictorial form from geometry to equivalent aspects of the language of algebra. By an algebraic equivalent, we mean an aspect of the pictorial form of the language of algebra that has in that language a role that is equivalent to the role played by the corresponding aspect of the pictorial form of the language of geometry. The idea is that the pictorial form of a language is composed of the above six aspects. So, there is a *structural isomorphism*

the function of $\sqrt{}$ at the argument 8. Similarly, in the case of a polynomial, it represents the result of the particular operations as applied to any number. Furthermore, it is thanks to the fact that the algebraic magnitudes are in the same relations as the relations of the parts of the expression that the algebraic symbols have meaning. Thus, for instance, the quantity represented by the symbol $\sqrt{8}$ is the outcome of the operation $\sqrt{}$ applied at the argument 8, and thanks to that the algebraic symbol represents this quantity. So, if we understand the compositional form of the language of algebras to be a structure on its own, we can interpret its referential aspect in the Wittgensteinian sense (in agreement with points 1 and 2) and turn to the interesting question regarding which aspects of this compositional form remained implicit during a particular historical period. Thus, thanks to the separation of the compositional form from the logical form, we can transfer the notion of a pictorial form to the case of algebra.

[38] A more detailed reconstruction of the history of algebra as the development of the pictorial form of its language can be found in (Kvasz 2008).

between *identical* pictorial forms of *different* languages (i.e., for instance, the perspectivist form of the language of geometry and the perspectivist form of the language of algebra), which maps the epistemic subject of one language onto the epistemic subject of the other, just like it maps the relation of identity of one language onto the relation of identity of the other. Similarly, there is a *structural embedding* between *consecutive* pictorial forms of the *same* language (i.e., for instance, the perspectivist form and the projective form of the language of geometry), which maps the aspects of the earlier form (which are in that form only implicit) onto their explicit expressions in the later pictorial form.

It seems that the algebraic equivalent of the point of view can be identified (as Rotman convincingly showed) with *zero*, which is found in the standard form of algebraic equations on their right-hand side.[39] The algebraic equivalent of the horizon can be seen in the *domain of the coefficients*, i.e., the set of quantities that mathematicians accept as coefficients which can appear in equations. As mentioned, Stifel was one of the first to allow the coefficients in the equations to take on negative values (so he passed from natural numbers to integers as the domain of the coefficient). However, that was not the end of the story. Later, Lagrange realized that the addition of further quantities to the domain of the coefficients results in the decomposition of an irreducible polynomial into factors. The clarification of this problem was brought about later on by the notion of a Galois connection. The richness of the domain of the coefficients determines the scale of relationships that can be expressed in the language of algebra and so determines the limits of the world, in agreement with Wittgenstein's thesis, according to which the limits of language mean the limits of the world.

The algebraic equivalent of the relation of identity, which, in the language of geometry, was illustrated by the rules for diminishing the tiles forming a *pavimento*, could be the *rules of the equivalent transformations of equations*, such as Al Khwarizmi's *al jabr* (after which algebra was named), *al mukábala*, and *al rad*. The algebraic equivalent of the synonymy relations could be the *canonical forms and invariants*, such as the three forms of the cubic equations introduced by Cardano. With an expansion of the domain of the coefficients (for example, by accepting negative numbers), the rules of equivalent transformations are also enriched, which makes it possible to reduce the number of canonical forms, in the case of cubic equations, to one. The algebraic equivalent of the background could perhaps be interpreted as the *system of solutions*, i.e., what is considered to be a solution to an algebraic problem. Finally, the algebraic equivalent of ideal elements could be *complex numbers* and, in general, various algebraic extensions of the domain of numbers.

To indicate the complexity of the development of the pictorial form of the language of algebra and thus to allow an appreciation of the complexity of the task

[39] This supports the interpretation of the early stages of symbolic algebra (first of all the symbolism introduced by Descartes) as projective form, because in the projective form of language, the point of view of the previous pictorial form (the perspectivist one) becomes explicitly represented in language (as the zero on the right-hand side of an algebraic equation).

of giving a possibly complete sequence of pictorial forms in the development of algebra, we will list only the stages in the development of the fifth aspect, i.e., how the system of solutions of an algebraic equation was understood within the particular pictorial forms of the language of algebra.

To solve the equation $g(x) = 0$ means:

1. *To find a regula*, i.e., a rule written in ordinary language, which makes it possible to *calculate* the "thing," that is, the root of the equation.
2. *To find a formula*, i.e., an expression of the symbolic language, which makes it possible to *express* the root of the equation in terms of its coefficients, the four arithmetical operations, and root extraction. The symbols in the formula correspond to steps in the calculation, and so a formula is a representation of the *regula*.
3. *To find a factorization of the polynomial form*, i.e., to represent the polynomial form as a product of linear factors. Each factor represents one root of the equation, and so the number of factors is equal to the degree of the equation.
4. *To find a resolvent*, i.e., to reduce the given problem, by means of a suitable substitution, to an auxiliary problem of a lesser degree. A solution to the auxiliary equation can be transformed into a solution of the original problem. However, here, besides the n roots of the n^{th} degree equation, we also obtain all the associated quantities, in general, $n!$ numbers. By means of n of them, we can split the polynomial form into linear factors.
5. *To find a splitting field*, i.e., to construct the field $Q(\alpha_1, \ldots, \alpha_n)$ that contains all the roots of the equation. This field automatically contains all the associated quantities of the equation, and thus also all the roots of its resolvent.
6. *To find a factorization of the Galois group of the splitting field* $Q(\alpha_1, \ldots, \alpha_n)$, i.e., to decompose the group of symmetries of this field into blocks. Steps in the factorization of the group correspond to particular extensions of the field. Hence, from knowledge concerning the factorization of the group, we can draw conclusions about the field extensions.
7. *To construct a factorization of the ring of polynomials* $Q[x]$ *by the ideal* $(g(x))$, i.e., to find the residual classes of the ring of polynomials after factorization by the ideal that correspond to the equation we want to solve. One of these classes is the root of the equation, and so we have an abstract method for the solution of any algebraic equation.

This list illustrates the levels of abstraction that have been introduced in the history of algebra.[40]

[40] A similar development concerning the different interpretations of the notion of Dirichlet character is presented in Jeremy Avigad's paper *Mathematics and Language* (Avigad 2015, pp. 240–242). Dedekind and Weber played key roles in stages 6 and 7 of the development represented in (Fig. 2), so it is likely that the development of the notion of Dirichlet character may be parallel to some of the stages captured in (Fig. 2).

3.6 Objectification as a Fundamental Semiotic Process

Thanks to the analysis of the incorporation of the pictorial form into language, we can explain why the interpretation of mathematical expressions as rules is inadequate. Even if Cardano did understand algebra as a collection of rules, his syncopated algebra was only the first of the eight pictorial forms of the language of algebra indicated in (Fig. 2). Thus, the case of Cardano shows that the understanding of mathematical expressions as rules is in some cases adequate. Maybe this understanding is based on an illusion created by the fact that the pictorial form of language J_1 is incorporated into language J_2, and so on. Thanks to this incorporation, in any language based on practically any later pictorial form, we can find a fragment which can be translated back to J_1, in which algebra *was* a collection of rules – although for most of the new results obtained thanks to the special properties of the later pictorial forms, such a translation is not possible.

To solve the equation $g(x) = 0$ means:

1. *To find a regula*, i.e., a rule written in ordinary language, which makes it possible to *calculate* the "thing", that is, the root of the equation.
2. *To find a formula*, i.e., an expression of the symbolic language, which makes it possible to *express* the root of the equation in terms of its coefficients, the four arithmetical operations, and root extraction. The symbols in the formula correspond to steps in the calculation, and so a formula is a representation of the *regula*.
3. *To find a factorization of the polynomial form*, i.e., to represent the polynomial form as a product of linear factors. Each factor represents one root of the equation, and so the number of the factors is equal to the degree of the equation.
4. *To find a resolvent*, i.e., to reduce the given problem, by means of a suitable substitution, to an auxiliary problem of a lesser degree. A solution to the auxiliary equation can be transformed into a solution of the original problem. However, here, besides the n roots of the n^{th} degree equation we also obtain all the associated quantities, in general, $n!$ numbers. By means of n of them we can split the polynomial form into linear factors.
5. *To find a splitting field*, i.e., to construct the field $Q(\alpha_1, ..., \alpha_n)$ that contains all the roots of the equation. This field automatically contains all the associated quantities of the equation, and thus also all the roots of its resolvent.
6. *To find a factorization of the Galois group of the splitting field* $Q(\alpha_1, ..., \alpha_n)$, i.e., to decompose the group of symmetries of this field into blocks. Steps in the factorization of the group correspond to particular extensions of the field. Hence, from knowledge concerning the factorization of the group we can draw conclusions about the field extensions.
7. *To construct a factorization of the ring of polynomials* $Q[x]$ *by the ideal* $(g(x))$, i.e., to find the residual classes of the ring of polynomials after factorization by the ideal that corresponds to the equation we want to solve. One of these classes is the root of the equation, and so we have an abstract method for the solution of any algebraic equation.

Fig. 2 An overview of the possible interpretations of the notion of solution of an algebraic equation. (A description of the different pictorial forms of the language of algebra can be found in *Patterns of Change* (Kvasz 2008, pp. 165–200). The summarizing overview of the different interpretations of the solution of an algebraic equation presented in (Fig. 2) can be found on page 199 of that book)

Objectification may be seen as a mechanism of conceptual change in mathematics. In the perspectivist form (i.e., in its syncopated phase), algebra was based on operations, such as raising to the third power or extracting the third root. *The first objectification* consisted in transforming these operations into objects, into powers and roots, and representing them through special signs as done, among others, by Regiomontanus around 1460. Thanks to this first layer of objectification, more complex expressions can be created and, for them, operations of a new kind can be introduced: operations of the transformation of algebraic expressions. *The second objectification* consisted in the replacement of the numerical values of the coefficients by a new kind of symbol – parameters introduced, for instance, by Viète in 1591. Thus, a new algebraic object was born, a general algebraic expression, and with it also a new kind of operation, consisting of transformations into different canonical forms. One of these canonical forms was the one in which all terms were transferred to one side of the equation. However, this presupposed the objectification of negative numbers, i.e., their acceptance as real objects on a par with positive numbers and zero. This was *the third objectification* introduced, in one form, by Stifel in 1544. It took place independently of the second and again introduced new operations. The combination of the last two objectifications led to the formation of polynomials as a new kind of algebraic object. This step can be seen as *the fourth objectification*.

The formation of polynomials gave birth to new operations, the decomposition of a polynomial into factors. It gradually became clear that what enables the decomposition of the polynomial is the presence of specific quantities. *The fifth objectification* brings the objectification of the set of all quantities, which allows the decomposition of a polynomial into a product of linear terms. The set of these quantities is called the splitting field of the polynomial and was introduced, in an implicit form, by Lagrange. The objectification of the splitting field enables the introduction of new operations, such as permutations of the roots of the polynomial, which generate automorphisms of the splitting field. By objectifying these automorphisms (which is *the sixth objectification*), the Galois group of a polynomial is formed. Formation of the Galois groups makes it possible to introduce a new operation, the operation of their factorization. It is only at this sixth level of objectification that a remarkable property of fifth-degree polynomials is revealed: their group of automorphisms contains only one normal subgroup, the alternating group of five elements, which is simple and cannot be further factorized. This fact lies at the bottom of the insolvability of fifth-degree equations. Without the cascade of six objectifications, this fact is not easily discovered. Objectification is thus a *fundamental concept-forming mechanism in mathematics*.

4 Conclusion

In this chapter, we offered a model of semiotic analysis of the language of mathematics based on two components and illustrated this model on the language of algebra. The first component was termed the *compositional component* of

the language, and it was reconstructed by means of the logical, the expressive, the methodological, the integrative, the explanatory, and the constitutive power of the language. We believe that these six aspects of the language of a particular mathematical or scientific discipline are its objective features, which can be further analyzed and perhaps even quantified. Their growth may be responsible for the impression of progress, that many mathematicians and scientists perceive when they view the historical development of their discipline. One cannot deny that, from the point of view of the compositional component, the language of mathematics, say, makes an impression of continual growth and improvement.

The second component was called the *referential component* and was reconstructed by means of the picture theory of meaning. This component was reconstructed using notions like the point of view, the horizon, the relation of identity, the relation of synonymy, the background, and the ideal elements. It turns out that the gains at the compositional level are, in a sense, compensated by making these six aspects of the referential component more indirect, complicated, and difficult to comprehend. So, the development of the compositional component makes it more and more difficult to relate the expressions of the language to a real situation and understand their meaning.

References

Alten H-W, Djafari Naini A, Eick B, Folkerts M, Schlosser H, Schlote K-H, Wesemüller-Kock H, Wussing H (2014) 4000 Jahre Algebra, 2nd edn. Springer, Berlin

Avigad J (2015) Mathematics and language. In: Davis E, Davis PJ (eds) Mathematics, substance and surmise. Springer, Cham, pp 235–256

Barton B (2008) The language of mathematics. Telling mathematical tales. Springer, New York

Boyer C, Merzbach U (1989) A history of mathematics. John Wiley, New York

Cardano G (1545) Ars Magna, or the rules of algebra. MIT Press, Cambridge, MA. 1968

Chemla K (2003) Generality above abstraction: the general expressed in terms of the paradigmatic in mathematics in ancient China. Sci Context 16(3):413–458

Chemla K (2005) Geometrical figures and generality in ancient China and beyond: Liu hui and Zhao Shuang, Plato and Thabit ibn Qurra. Sci Context 18(1):123–166

Chemla K (2012) The history of mathematical proof in ancient traditions. Cambridge University Press, New York

Corry L (2004) Modern algebra and the rise of mathematical structures. Birkhäuser, Boston

Courant R, Robbins H (1941) What is mathematics? Oxford University Press, New York. 1978

De Risi V (2016) Francesco Patrizi and the new geometry of space. In: Vernier, Regier (eds) Boundaries, extents and circulations, vol 2016. Springer Verlag, Cham, pp 55–106

Euler L (1770) Vollständige Anleitung zur Algebra. Reclam, Leipzig 1911. English translation by J. Hewlett: Elements of Algebra. Longman, London 1822

Frege G (1879) Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Georg Olms, Hildesheim 1993. English translation in van Heijenoort J. (ed) 1967, From Frege to Gödel, A source book in mathematical logic 1879–1931. Harvard University Press, Cambridge, MA., pp 1–82

Frege G (1891) Funktion und Begriff. In: Frege G (ed) Funktion, Begriff, Bedeutung. Vandenhoec & Ruprecht, Göttingen 1986, pp 17–39. English translation in: Geach P and Black M (eds): Translations from the philosophical writings of Gottlob Frege, pp 21–41

Gray J (2018) A history of abstract algebra. From algebraic equations to modern algebra. Springer Nature, Cham
Heeffer A (2009) On the nature and origin of algebraic symbolism. In: Van Kerkhove B (ed) New perspectives on mathematical practices. World Scientific, Singapore, pp 1–27
Heeffer A (2010) Algebraic partitioning problems from Luca Pacioli's Perugia manuscript. SCIAMVS 11:3–51
Hoyrup J (1994) In measure, number, and weight. State University of New York, Albany
Hoyrup J (2007) Jacopo da Firenze's Tractatus algorismi and early Italian Abbacus culture. Birkhäuser Verlag, Basel
Hoyrup J (2019) Selected essays on pre- and early modern mathematical practice. Springer Nature, Cham
Kvasz L (2008) Patterns of change, linguistic innovations in the development of classical mathematics. Birkhäuser, Basel
Kvasz L (2012) Language in change. Fernando Gil International prize 2010. Fundacao Calouste Gulbenkian, Lisbon
Lakoff G, Núñez RE (2000) Where mathematics comes from. How the embodied mind brings mathematics into being. Basic Books, New York
Manders K (2008a) Diagram-based geometric practice. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 65–79
Manders K (2008b) The Euclidean diagram (1995). In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 80–133
Netz R (1999) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press, Cambridge
O'Halloran KL (2005) Mathematical discourse. Language, symbolism and visual images. Continuum, London
Oaks JA (2006) Simplifying equations in Arabic algebra. Hist Math 34:45–61
Oaks JA (2007) Medieval Arabic algebra as an artificial language. J Indian Philos 35:543–575
Oaks JA (2009) Polynomials and equations in Arabic algebra. Arch Hist Exact Sci 63:169–203
Oaks JA (2012) Algebraic symbolism in medieval Arabic algebra. Philosophica 87:27–83
Rotman B (1987) Signifying nothing. The semiotics of zero. Macmillan Press, London
Rotman B (2000) Mathematics as a sign. Stanford University Press, Stanford CA
Scholz E (ed) (1990) Geschichte der Algebra. Wissenschaftsverlag, Mannheim
Stewart I (1989) Galois theory. Chapman and Hall, London
Stifel M (1544) Vollständiger Lehrgang der Arithmetik. Konigshausen & Neumann, Würzburg. 2007
van der Waerden BL (1985) A history of algebra, from al-Khwarizmi to Emmy Noether. Springer, Berlin
Vermeir K, Regier J (eds) (2016) Boundaries, extents and circulations. Springer Nature, Cham
Viète F (1591) Introduction to the analytical art. The Kent State University Press, Kent. 1983
Wagner R (2010) The natures of numbers in and around Bombelli's *L'algebra*. Arch Hist Exact Sci 64:485–523
Wagner R (2017) Making and breaking mathematical sense. Princeton University Press, Princeton
Wittgenstein L (1921) Tractatus Logico-Philosophicus. Suhrkamp, Frankfurt am Main. 1989

Reading Mathematical Texts with Structuralist Semiotics

Anna Kiel Steensen

Contents

1 Introduction 3135
2 Framing the Application 3137
3 First Example 3140
4 Second Example 3144
5 Conclusion 3149
6 Cross-References 3149
References 3149

Abstract

In this chapter, we use structuralist semiotics as a framework for analyzing mathematical textual practices. We begin with a short introduction, in which we operationalize certain key aspects of structuralist semiotics. This gives us a tool for semiotic analysis of real mathematical texts. We then illustrate this type of analysis with a concrete text, namely, *Über die Composition der binären quadratische Formen* by Richard Dedekind.

Keywords

Philosophy of mathematical practice · Mathematical textuality · Structuralist semiotics · Dedekind

1 Introduction

In this chapter, we assume as framework Wagner's discussion of structuralist and poststructuralist semiotics (▶ Structural Semiotics as an Ontology of Mathematics). To summarize, in this framework, *signs* emerge from an interaction of two (or more)

A. Kiel Steensen (✉)
ETH, Zurich, Switzerland

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_115

distinct levels or "streams," e.g., sound, visual perception, thought, or emotion. The interaction results in a simultaneous division or "articulation" of each level into units. One level becomes signifying units, called *signifiers*, and the other becomes signified units, called *signifieds*. A *sign* is defined as a connection between a signifier and a signified.

As an example, colors emerge from a division of the visual perception of a spectrum of electromagnetic radiation into subspectra. But we only form signs once we relate this division to another. If the other division happens in the aural stream of speech, the signs are the (sometimes quite subjective) connections of spoken words "red," "blue," "green," etc. to the visual perceptions. The words would be the signifiers and the perceptions the signifieds. However, if instead we correlate the subspectra with our experience of different types of weather, we may form signs where colors are signifiers of (changing) weather (a clear blue sky in winter may signify a freezing cold day ahead, a sudden darkening may signify thunder).

Colors are likely much more semiotically complex than this. But the example illustrates the central point that, as signs, colors do not exist in themselves but only in their differentiation from other colors. This is true for signs in general: From a structuralist perspective, signs only exist in systems of differences. The *meaning* of a sign is its value relative to other signs around it.

Here, our aim is to illustrate how a structuralist semiotic approach can be used to analyze an actual mathematical textual practice in the form of a "real" mathematical text. The essential method of structural analysis is the *commutation tests*, in which we exchange a segment of a stream and observe how a correlated stream changes in response.

In our case, the signifiers are the concrete textual expressions and the signifieds are the content of the text. We define content as an analytical category in the next section.[1] For now, as an example, consider the expression "0 is in {0, 1}." This can be read as a statement about signifiers, which says that "0" occurs in the concrete string "{0, 1}" (like "t" occurs in the string "the"). Alternatively, it can be read as saying that 0 is an element of the set {0, 1}, in which case it is a statement about signifieds.

To study the signs of actual texts, we perform commutations of textual expressions. We use two kinds of commutation tests. In the first, we commute expressions that actually occur in the text. For example, say that both of the following occur:

$$\text{“}0 \text{ is in } \{0,1\}\text{”}$$

$$\text{“}0 \text{ is in } 01110\text{”}$$

Then we might ask if the change between "{0, 1}" and "01110" at the level of signifiers makes a difference at the level of signifieds.

[1] The terms "content" and "expression" have a technical meaning in structuralist linguistics. We use the terms in a sense inspired by it but much simplified (cf. Hjelmslev 1953).

In the second kind of commutation test, we compare an expression that occurs in the text with one that we write ourselves. Suppose we are interested in the mathematical textual practice of mixing natural language with non-phonetic (e.g., symbolic) writing (like "is in" is mixed with numerals and braces in the expressions above). An approach could be to rewrite samples of actual text but using only non-phonetic writing and then note the differences, if any, at the level of signifieds.[2]

To illustrate what a structuralist semiotic analysis may look like, we apply a series of commutation tests to an actual mathematical text, *Über die Composition der binären quadratische Formen* by the mathematician Richard Dedekind. Our commutation tests are chosen with a more general question in mind: Can structuralist semiotics propose a notion of *mathematical textual practice* grounded in the study of actual textual material? It is beyond the scope of this chapter to address this question directly, but our reading of Dedekind's text will illustrate the idea. Specifically, we recapitulate a particular historical-philosophical debate regarding Dedekind's text. Based on this debate, we introduce a distinction between two different kinds of contents, namely, *arithmetical content* and *set content*. We conceive these contents as certain classes of signifieds. We apply commutations to Dedekind's text, which are designed to capture interactions between arithmetical and set contents and indicate how to analyze a textual practice in terms of such interactions.

2 Framing the Application

Dedekind's text is a famous piece of nineteenth-century algebra. It appeared as the tenth appendix or "supplement" to the second edition of Dirichlet's *Vorlesungen über Zahlentheorie* (Lejeune-Dirichlet 1871). For brevity, we refer to Dedekind's text as Dedekind (1964). In Dedekind (1964), Dedekind presents a theory of uniqueness of prime factorization for algebraic integers. We focus on the sections §§159–163 of Dedekind (1964), in which Dedekind introduces the notion of *an ideal*. In Dedekind's definition, an ideal is a special kind of "system"[3] of numbers that is closed under the algebraic operations of addition, subtraction, and multiplication (Dedekind 1964, 251).[4,5] Dedekind then

[2] Wagner (2009) gives examples of both kinds of commutation tests for analyzing the semiotics of "0."

[3] The German term is "System," but there are different translations in the secondary literature. Haffner (2017) translates "System" to "system," whereas, for example, Ferreirós (2007) translates it to "set."

[4] The numbers are *algebraic integers*, i.e., the elements of the ring of integers o of a general algebraic number field. In particular, o is an ideal consisting of all the algebraic integers under consideration – the ambient space, so to speak. In the special case of the field Q of rational numbers, the "algebraic" integers are simply the usual integers and o is Z.

[5] As an example, consider Z, the ring of integers of the field of rational numbers Q. The set of numbers of the form $3n$, where n runs through the integers, i.e., the set $\{\ldots, -6, -3, 0, 3, 6, \ldots\}$ is an ideal. Similarly, any set of numbers of the form αn, where α is a fixed integer, is an ideal.

gives a prime factorization for ideals and proves that the prime factorization is unique.[6]

Dedekind (1964) was the first version of the theory. Dedekind later reworked and republished it three times.[7] There is a consensus that Dedekind's successive reworking of ideal theory almost seems to encapsulate an anticipation of certain developments in mathematics during the twentieth century. Summarized by Avigad:

> Whereas Dedekind's first version remained fairly close to Kummer's computational style of presentation, the later versions became increasingly abstract and algebraic. As a result, the development is an early and salient example of a transition to types of reasoning (set-theoretic, algebraic, structural, infinitary, nonconstructive, and so on) that are characteristic of modern mathematical thought. (Avigad 2004, 2)

The cut is not clear, though. The precise relation in Dedekind's work between, on the one hand, the "new" ways of doing mathematics (set-theoretic, algebraic, structural, infinitary, nonconstructive…) and, on the other hand, the "old" ways remains a contested issue.[8] For our purpose, the important fault line in this debate is the interaction between arithmetic and set theory in Dedekind's writing.

According to Mehrtens (1979), the concepts that Dedekind introduces in Dedekind (1964) are abstract and their relationship to arithmetic is merely accidental (Mehrtens 1979, 76–77). The concepts may seem to be concretely arithmetical, but this is because they were "tailored to a specific frame of reference," not because they were not abstractly conceived. Haffner (2014, 2017, 2019) instead suggests that Dedekind uses "arithmetical rewriting" (2017, 32), a set of writing practices deliberately invoking arithmetic at several levels, including notation, conceptualization, and proof techniques. In this view, Dedekind establishes analogies to number theoretical concepts and methods, which would have been familiar to readers of the time.[9]

How novel was it to work with sets? Edwards writes:

> Indeed, despite the fact that Dedekind defines an ideal as a "system" of algebraic integers rather than a "set," the set-theoretic point of view is unmistakable and appears here well before the work of Dedekind's colleague Cantor on this subject […] [T]his theory of ideals (in 1871) can surely be regarded as the first appearance of the set-theoretic approach that was later to become so influential. (Edwards 1980, 346)

[6] For a full account (besides Dedekind's text itself), see (Edwards 1980, 1990) or Stillwell's introduction (Dedekind and Stillwell 1996).

[7] Dedekind published a second version of the theory in *Bulletin des Sciences Mathèmatiques et Astronomiques* 1876–1877, which already differs substantially from Dedekind (1964). The third and fourth versions came in the later editions of *Vorlesungen*, in 1879 and 1894, respectively.

[8] Loosely, the "old ways" would be constructive, computational, object-oriented, or focused on specific (geometric or arithmetical) problems.

[9] See also Corry (2004, esp. Chaps 2, 3, 5), Awodey and Reck (2002), Ferreirós (2007), Sieg and Schlimm (2005, 2017), Reck (2003, 2009), and Avigad (2006).

This is the debate that motivates our analysis. Dedekind's writing, and Dedekind (1964) in particular, lends itself to be read as two seemingly conflicting ways of doing mathematics. We suggest that structuralist semiotics can be used to grasp the relationship between them as a concrete mathematical textual practice.

In Dedekind (1964), if α is a number contained in an ideal a, then a is said to *divide* α (Dedekind 1964, 251).[10] An ideal b is said to *divide* an ideal a if b contains each of the numbers in a:

> If all the numbers of an ideal a are also contained in an ideal b, [...] we will say that a is a multiple of b or divisible by b, b is a divisor of a or divides a. (Dedekind 1964, 252)[11]

All of our examples are based on the following simple observations. Dedekind continues to use "contain" and "in" when writing about numbers in ideals. That is, sometimes he writes that an ideal contains a number and sometimes that the ideal divides the number. Given the definition above, we might expect the expressions "divide" and "contain" to be completely interchangeable, i.e., that commuting between signifiers "divide" and "contain" has no effect at the level of signifieds. However, Dedekind never writes that an ideal "contains" or "is contained in" another, only that it "divides" or is "divisible by" another.[12] This asymmetry in the actual use of expressions indicates that there is some difference at the level of signifieds.

Expressions such as "+," "−" frequently occur together with "divide" in arithmetical texts. We use such expressions to characterize what we call *arithmetical content*. Specifically, we call the following the *arithmetical expressions*: "addition," "subtraction," "division," "multiplication," "sum," "difference," etc. and verbs such as "to add," "to divide," etc. We include also "+," "−," the use of concatenation instead of "·", and fractions.[13] If an expression in Dedekind (1964) can be commuted with a construction (sentence or formula) of arithmetical expressions, we say that it has *arithmetical content*.

Similarly, we use the expressions "contain" and "in" that are used frequently in writing about sets or collections and their elements but less frequently in arithmetical writing to characterize what we will call *set content*. If an expression in Dedekind (1964) commutes with an expression of the form "X in Y" or "X contains Y," then we say that it has *set content*. We include expressions that can be "built" from the

[10] In what follows, we adopt Dedekind's use of lowercase Latin letters to denote ideals and lowercase Greek letters to denote numbers.

[11] All translations from the German original are mine.

[12] Consider again the example of the ideal of integer multiples $3n$, i.e., $\{\ldots, -6, -3, 0, 3, 6, \ldots\}$. This ideal divides (or contains) for example -3, and it also divides the ideal $\{\ldots -12, -6, 0, 6, 12, \ldots\}$ of integer multiples $6n$, but in Dedekind (1964), we do not find expressions such as "$\{\ldots, -6, -3, 0, 3, 6, \ldots\}$ contains $\{\ldots -12, -6, 0, 6, 12, \ldots\}$."

[13] For example, xy and $\frac{x}{y}$.

simple form, for example, "let k be any number contained in S." We call "in" and "contain" the *set expressions*.[14]

The starting point of our analysis is thus the concrete textual expressions rather than Dedekind's supposed intentions with the text. For simplicity, we sometimes speak about "Dedekind," but we do not mean the historical person – we only speak about the expressions and signs of the actual text.

The setup operationalizes the structuralist assumption that linguistic signs and their meaning can be studied – traced – by applying commutation test to actual use of expressions. To this we add the further assumption that written mathematics, including all its non-linguistic expressions, behaves like a language in this sense. As can be seen, our approach involves a sort of handheld implementation of the distributional hypothesis assumed in current NLP research (Sahlgren 2008). Such manual computational analysis is sometimes called "reading at a middle distance" to distinguish it from both close reading and the "long-distance reading" of automated corpus analysis (Moretti 2000, 2013).[15] The relationship between the distributional hypothesis and structuralism is discussed in Gastaldi (2021) and Gastaldi and Pellissier (2023). There has been very little work on mathematics in the structuralist semiotic tradition with Herreman (2000), Vinciguerra (1999), and Wagner (2009) as important exceptions, all of which show the different layers of meaning at work within the seemingly regimented semiotic systems of mathematical writing. We are essentially doing the same here with a distinction between two kinds of content as analytical tool.

3 First Example

Consider the following passage:

> The integer $\lambda = x\mu : \nu = \sqrt[e]{\mu x^{e-1}}$ is divisible by p but not by p^2; therefore λ^r is divisible by p^r but not by p^{r+1}, from which it also follows that the ideals p^r and p^{r+1} really are different. (Dedekind 1964, 257)

Here, Dedekind asserts that some specific (algebraic) integer λ is divisible by (the prime ideal) p but not by (the ideal) p^2 and, as a direct consequence, that (the algebraic integer) λ^r is divisible by (the ideal) p^r but not by (the ideal) p^{r+1}. Dedekind

[14] Dedekind does not use notation such as "$\subset$" or "$\in$," which today are standard ways of expressing containment, in Dedekind (1964).

[15] Middle-distance reading is nothing new in the humanities. It goes back at least to the work in the 1950s and 1960s of literary theorist Josephine Miles, e.g., Miles (1967). Miles observes that certain literary styles can be distinguished by certain proportions of word classes. This allows Miles to consider literary style in terms of (quantitative) relations between the words that occur in actual texts, instead of considering style as a reference to the nature of the subject matter of the text or as a product of the individual psychology of the author. See also Buurma and Heffernan (2018). Arguably, our analysis here is more subjective and idiosyncratic than Miles'.

does not explicate this inference further. Such gaps are ubiquitous in written mathematical proofs. Here, we focus on how, from a semiotic perspective, the proof may be sufficiently convincing despite its gaps.

We begin with some definitions from Dedekind (1964) and an important observation about the organization of ideal theory in that particular text. An ideal p is called a *prime ideal* if the only divisors of p are o and p itself (Dedekind 1964, 253).[16] The *simple ideals* are defined as follows:

> If μ is a number in o that is different from zero and not a unit, then [. . .] there always exists a number ν such that all the roots π of the congruence $\nu\pi \equiv 0\ (mod\ \mu)$ constitute a prime ideal. prime ideals that consist in all the roots of such a congruence we will call in what follows simple ideals. (Dedekind 1964, 255–256)[17]

Intuitively, the way Dedekind uses powers of λ and p in the presentation of the proof indicates that we are doing something like taking the rth power and that this operation preserves the relationship between λ, p, and p^2 (i.e., such that the relationship between λ^r, p^r, and p^{r+1} is the same as that between λ, p, and p^2).

Prior to the passage in question, Dedekind has defined the *rth power of p*, for p any simple ideal (i.e., a p given by a congruence as above):

> Now, if r is any whole, rational, non-negative exponent, then all the roots ρ of the congruence $\rho\nu^r \equiv 0\ (mod\ \mu^r)$ form an ideal, which is called the rth power of p and will be denoted by p^r. (Dedekind 1964, 256)

Simple ideals are prime by definition and Dedekind shows some pages later that every prime ideal is simple (Dedekind 1964, 257–258). So, from then on, this definition of power applies to prime ideals. For ideals in general, there will not be a notion of power until later in Dedekind (1964).

The passage in question involves a prime ideal p and the powers p^2, p^r, and p^{r+1}. But what is really the connection between a prime ideal p and the powers of p? We recall that Dedekind has already defined the multiples of an ideal: If p contains every number contained in some other ideal b, then b is a multiple of p. In rational arithmetic, we may consider say 5^r, the rth power of 5, to be obtained by iterated multiplication by 5, but it is not given that p relates to the powers of p by means of some "multiplication operation" in the way that 5 relates to the powers of 5 by arithmetical multiplication.

The problem is that rth powers of prime ideals precede the product of prime ideals in the concrete textual organization of ideal theory in Dedekind (1964). So what in

[16] The ideal $\{\ldots, -6, -3, 0, 3, 6, \ldots\}$ is prime; the ideal $\{\ldots -12, -6, 0, 6, 12, \ldots\}$ is not. In general, all and only the ideals $\{\ldots, -2p, -p, 0, p, 2p, \ldots\}$, where p is a prime number, are prime ideals in Z. However, this simple characterization does not generalize to other rings.

[17] The unit (German "Einheit") mentioned in the definition is any number in o with the property that it divides every other number in o (like 1 divides every integer). Equivalently, ε is a unit if there exists an ε' such that $\varepsilon\varepsilon' = 1$ (Dedekind 1964, 239).

the proof seems to be manipulation of powers of p comes before the establishment of an operation on ideals that would justify such manipulations. Indeed, at this point in Dedekind (1964), Dedekind has introduced "multiple" (and "divisible," "divisor," and "divides") for ideals but not multiplication or division in the sense of operations on ideals. At this point in the text, we thus cannot obtain in a rigorous way the powers of p by iterated multiplication by p, simply because Dedekind has not given a multiplication of ideals yet.[18] Rather, p^r is obtained from p by manipulating the congruence that determines p. When p is given by the roots π of $\nu\pi \equiv (mod\ \mu)$, then p^r is given by the roots ρ of $\rho\nu^r \equiv 0\ (mod\ \mu^r)$. In terms of rth powers, ν and μ are replaced by ν^r and μ^r, respectively.

Only toward the end of the pages devoted to ideals do we get the product of two ideals:

> If one multiplies all the numbers of an ideal a by all the numbers of an ideal b, then these products and their sums form an ideal divisible by a and b, which is called the product of the factors a and b and will be denoted by ab. From this statement it is immediately clear that $ao = a$, $ab = ba$; furthermore $(ab)c = a(bc)$. (Dedekind 1964, 259)

This construction defines a multiplication on ideals and allows us to take the r-fold product $ppp\cdots$ (r times), but what is not immediately clear is whether this $ppp\cdots$ (r times) is the same ideal as p^r. Because Dedekind has defined p^r independently of a notion of a product of ideals, the fact that the two ideals $ppp\cdots$ (r times) and p^r coincide is something that has to be shown.[19]

With this in mind, the way Dedekind uses powers of λ and p in the presentation of the proof might aim to make a reader anticipate an operational relationship between p and rth powers of p similar to that between λ and rth powers of λ. We now explore the idea of such "anticipation" using a commutation test. We have the following two expressions, in which powers of λ and p occur:

1. "$\lambda = x\mu : \nu = \sqrt[e]{\mu x^{e-1}}$ is divisible by p but not by p^2"
2. "λ^r is divisible by p^r but not by p^{r+1}"

[18] Since set-theoretic approaches were not standard when Dedekind wrote Dedekind (1964), we should not assume that Dedekind expects the reader to come equipped with an idea of a multiplication of ideals or other "systems" of numbers.

[19] To be precise, what Dedekind shows is that $p^a p^b = p^{a+b}$. The identity $p^r = pppp\cdots$ (r times) follows from $p^a p^b = p^{a+b}$ and associativity of the multiplication of ideals (Dedekind 1964, 259). Just in terms of sheer textual volume, the connection between powers and product suggested by the power notation is the fruit of a solid amount of work. In *Vorlesungen*, Sup X spans the pages 380–497. "Ideal" occurs regularly up to page 495 (included) but only the pages 452–462 are the introduction of ideal theory. The definition of ideals is at page 452; the definition of the product of ideals is not until page 459. After stating it, Dedekind relates products of ideals to powers of prime ideals (Dedekind 1964, 259–260) culminating with a theorem. Its proof relies on results established over the seven intermediate pages. The conclusion that the two ideals p^r and $pppp\cdots$ (r times) coincide for p prime (Dedekind 1964, 260) takes three pages to reach. In the end, this leads to the possibility of defining a notion of power for any ideal, not just the simple ones, which agrees with the initial congruence definition of p^r (Dedekind 1964, 260).

(1) concerns how p and p^2 relate to λ as divisors. (2) resembles (1), but λ is replaced by λ^r, p by p^r, and p^2 by p^{r+1}. Intuitively, Dedekind's presentation of (2) as derived from (1) may make it seem as if going from p to p^r and from p^2 to p^{r+1} is similar in some way to going from λ to λ^r.

In this sense, the multiplication of an ideal with itself as an operation – which satisfies the same rules of divisibility and manipulation of exponents as with arithmetical powers – exists as signified before it exists as a mathematical object. This is what we take away from this example: *Something may exist as signified before it exists as a rigorous mathematical object*.

We can view the change from expression (1) to (2) in terms of a commutation test, which replaces λ by λ^r, p by p^r, and p^2 by p^{r+1}. In the first commutation, λ^r is interchangeable with $\lambda \cdot \lambda \cdot \lambda \cdots \lambda$, the usual arithmetical product of λ with itself r times. But a similar commutation of the power of p, i.e., the replacement of p^2 by $p \cdot p$, p^r by $p \cdot p \cdot p \cdots p$ (r times) and p^{r+1} by $p \cdot p \cdot p \cdots p$ ($r + 1$ times), has a different effect at the level of signifieds. Changing p to p^r and p^2 to p^{r+1} is correlated with the change of λ to λ^r. Commuting λ with λ^r has arithmetical content so passing between the powers of p inherits arithmetical content. One way of expressing these different effects is by saying that the arithmetical content attached to the relation between λ and λ^r is a rigorous mathematical operation (taking the rth power of an algebraic integer), but going from p to p^r is not.[20]

There is still some work to be done before Dedekind can introduce a multiplication for ideals and show that the power of an ideal corresponds to iterations of that multiplication. This is the concrete organization of ideal theory in this specific text. The question then arises: Does the arithmetical content in the passage correspond to operations on ideals when multiplication of ideals will only be established later in the text?[21] If so, the operation seems to exist as signified before it exists as a mathematical object.

[20] Without the parallel occurrence of "λ" and "λ^r," perhaps we would have been more likely to consider the various powers of p in terms of replacing ν and μ by appropriate powers in the congruence that defines p. For the sake of contrast, the superscript r of p^r could also just be considered as an index, which indicates both a position in the sequence of ideals $p, p^2, \ldots, p^r, \ldots$, and the powers of the numbers μ and ν used in the congruence $\rho\nu^r \equiv 0 \ (mod\ \mu^r)$ to define p^r. The superscript r of p^r creates an index sign in the Peircean sense, since the superscript r of p points directly to the superscript r of μ^r and ν^r in the congruence $\rho\nu^r \equiv 0 \ (mod\ \mu^r)$.

[21] In a certain sense, it can only be established later in the text. In this example, changing the order of presentation is more complicated than simply giving some definitions in a different order. The different version of ideal theory testifies to this: The three later versions of ideal theory, Dedekind not only presents the theory differently but develops it in a different order. In Dedekind (1964), the notion of the product of two ideals comes at the end. Edwards notes: "In the second version – and all subsequent versions – Dedekind proceeded very differently. He defined products of ideals early in the development of the theory and elevated to the level of "the [principal] difficulty of the theory" the proof that if A and B are ideals and if $A \subset B$ then there is an ideal C such that $A = B \cdot C$. [...] In the original version, this theorem did not even appear, but it is a simple consequence of the theorems which did appear; [...] in connection with this, the *multiplicities* with which integers are divisible by a prime ideal are not defined in terms of powers of the ideal. (Edwards 1980, 350–351. Emphases in the original.). For more details, see Edwards (1980, 349–352).

The example suggests that it is worth taking the order of presentation into account as part of mathematical textuality. One could object that ideal theory has an intrinsic order and is interconnected in determinate but perhaps unknown ways. In this view, the different presentations of ideal theory would merely instantiate this order in seemingly different but ultimately accidental ways. The differences between presentations would be cognitive, aesthetic, cultural, historical, or other "auxiliary" factors, on which the intrinsic order of the theory would not depend. Even from that position, however, the recovery of a theory from its different instantiations can hardly be claimed to be trivial. Here, structural semiotics becomes relevant as tool to articulate differences.

4 Second Example

Dedekind defines two numbers ω, ω' to be *congruent modulo the ideal a* if a contains the difference $\omega - \omega'$. This is written $\omega \equiv \omega' \ (mod\ a)$. If a contains the number α, then α is said to be *congruent to zero modulo a*, which is written $\alpha \equiv 0 \ (mod\ a)$ (Dedekind 1964, 251). The notation is the same when the modulus is a number.[22]

We now have three expressions that would seem to be interchangeable: "a divides α," "a contains α," and "$\alpha \equiv 0 \ (mod\ a)$." However, ideal congruences can be manipulated: "[S]uch congruences can (according to I.) be added, subtracted, and (according to II.) multiplied, like equations" (Dedekind 1964, 251).[23] They can be manipulated not just like the congruences of rational arithmetic (when the moduli would be integers) but "like *equations*." At this point of Dedekind (1964), writing actual equations involving sums, differences, or products of ideals would be something of a leap since addition, subtraction, and multiplication have not been introduced for ideals yet. But "added, subtracted, and multiplied like equations" purports that the congruence notation can be used for arithmetical style computations with ideals. It is this "ideal arithmetic" made possible by the ideal theoretic congruence notation that we now explore.

In the following excerpt, Dedekind proves that the roots π of the congruence $\nu\pi \equiv 0 \ (mod\ a)$ constitute a prime ideal under a certain condition. We focus on three points, which we have marked **(1)**, **(2),** and **(3a–b)**:

> If the number η is not divisible by the ideal a, then there always exists a number ν not divisible by η such that all roots π of the congruence $\nu\pi \equiv 0 \ (mod\ a)$ constitute a prime ideal. **(1)** All the roots β of the congruence $\eta\beta \equiv 0 \ (mod\ a)$ constitute an ideal b dividing a. If b is not a prime ideal, that is, if there exist two numbers η', ρ' not divisible by b, for which their product $\eta'\rho' \equiv 0 \ (mod\ b)$, **(2)** then all the roots γ of the congruence $\eta'\gamma \equiv 0 \ (mod\ b)$, i.e., the congruence $\eta\eta'\gamma \equiv 0 \ (mod\ a)$, constitute an ideal c dividing b, [. . .]; if c is a prime ideal, then

[22] Also when the modulus is a so-called module, which (like ideals) are a specific kind of systems of numbers and which Dedekind introduces and discusses in §161–162 in Dedekind (1964).

[23] I and II refer to the two conditions in the definition of an ideal. I: An ideal contains the sum and difference of any two numbers that it contains. II: Any product of a number in the ideal with a number in o (the universe) is again a number in the ideal (Dedekind 1964, 251).

> the theorem is proven. If however c is not a prime ideal, then one can continue in the same way; eventually **(3a)** in the sequence of ideals b, c, d, ... [...] a prime ideal p must occur, which consists in all the roots π of the congruence $\nu\pi \equiv 0 \ (mod\ a)$, where **(3b)** $\nu = \eta\eta'\eta'' \ldots$ is divisible by η. (Dedekind 1964, 255)

The proof begins with the construction of an ideal b by means of a congruence $\eta\beta \equiv 0 \ (mod\ a)$. The first of our focus points **(1)** is that this ideal b divides the ideal a. The point **(2)** is that the roots γ of the congruence $\eta'\gamma \equiv 0 \ (mod\ b)$ are the same as the roots γ of $\eta\eta'\gamma \equiv 0 \ (mod\ a)$. That is, the ideal b divides $\eta'\gamma$ if and only if the ideal a divides $\eta\eta'\gamma$. We note that Dedekind does not explicate these inferences beyond what is contained in the excerpt above.

In **(2)**, Dedekind moves directly from $\eta'\gamma \equiv 0 \ (mod\ b)$ to $\eta\eta'\gamma \equiv 0 \ (mod\ a)$,[24] which might be an actual example of what Dedekind means by treating ideal congruences "like equations": Multiplying both sides with the same factor preserves equivalence. The move presents the change of modulus from ideal b to ideal a as equivalent to multiplication by η, as if the difference between the ideals a and b amounts to a factor η. This is at a point in the text where something like an operation of the form ηb has not been introduced. A simple commutation test confirms this impression. If we view the expressions "$\eta'\gamma \equiv 0 \ (mod\ b)$" and "$\eta\eta'\gamma \equiv 0 \ (mod\ a)$" as a commutation, then the change from $\eta'\gamma$ to $\eta\eta'\gamma$ is correlated with the change from $(mod\ b)$ to $(mod\ a)$. Commuting $\eta'\gamma$ and $\eta\eta'\gamma$ has arithmetical content (multiplication by η). By correlation, commuting b with a would seem to inherit arithmetical content. The equation-like appearance of moving directly from one congruence to the other is thus important for the textuality of this concrete proof.

In **(3a)**, ideals b, c, d, ... are generated until the output is a prime ideal (the procedure for generating c from a and b is explained in the proof; d is generated by applying the same procedure to b and c and so on).[25] A commutation test suggests that the sequence b, c, d, ... has arithmetical content: First, generating the next ideal of the sequence can be viewed as the commutation of a term in the sequence b, c, d, ... with its successor. Second, this commutation is correlated with the exchange of the expression "$\nu = \eta\eta'\eta'' \ldots$" in **(3b)** with a similar expression having one more term "$\eta^{(i)}$" (for appropriate i) at the rightmost end. The latter commutation has arithmetical content (multiplication by $\eta^{(i)}$). In this sense, we correlate the sequence of ideals with arithmetical multiplication of numbers.[26]

[24] a contains all numbers of the form $\eta\beta$ for β in b by construction of b, so taking a number in b and multiplying it with η is sure to bring us inside a, but a might also contain numbers that are not of this form.

[25] Dedekind uses the so-called norm, which associates a unique positive integer to any ideal, in order to associate the sequence of ideals b, c, d, ... with a strictly decreasing sequence of positive integers, all of which are strictly greater than 1. Such sequence of integers necessarily stops and thus the sequence of ideals also stops. At the same time, the procedure that produces the sequence of ideals stops if and only if the output is a prime ideal, so this proves the existence of the prime ideal asserted by the theorem.

[26] Note also that the order of the sequence and the order of factors together may make it seem as if we "move into" the next ideal when we multiply by $\eta^{(i)}$, while in fact we multiply by $\eta^{(i)}$ to make sure that the product remains an element in a.

Is the ideal congruence purely arithmetical? From a formal point of view, there is no difference between "divide," "contain," and "congruent with 0." But Dedekind uses all three expressions – if anything, their formal identity obscures this practice. What difference does using the congruence notation make when there is no formal difference? For this, we are looking for a difference between the signifieds of "divide" and "contain." Our strategy to approach this question is to write possible verifications of **(1)** and **(2)** beyond what the excerpt contains, i.e., to fill out some of the gaps in Dedekind's proof. We then apply commutation tests to the text of the verifications. Specifically, we write one version using the arithmetical expression "divide" and another using the set expressions "contain" and "in." We avoid using congruences altogether. This strategy allows us to describe the congruence notation as signs in terms of differences between the arithmetical version and the set version.

Possible verification of **(1)** using "contain" and "in":

> Let b be the ideal containing precisely those numbers β for which $\eta\beta$ is in a. We show that b contains every number contained in a. Let α be in a. Then a contains $\eta\alpha$. Thus, by definition of b, b contains α. Thus, b contains every number that a contains. So b divides a.

Possible verification of **(1)** using "divide":

> Let b be the ideal dividing precisely those numbers β for which $\eta\beta$ is divisible by a. We show that b divides every number divisible by a. Let α be divisible by a. Then a divides $\eta\alpha$. Thus b divides α. Thus, b divides every number that a divides. So b divides a.

In the version where we used set expressions, we (implicitly) appeal to one of the defining characteristics of ideals and to Dedekind's definition of division as containment. Even when we use "divide," we appeal (implicitly) to the definition of an ideal and of division of ideals. This introduces set content. The question is now, what happens if we "bracket" this set content and consider only the arithmetical content of "divide." Concretely, we supress the set content by applying one more commutation to the "divide" justification. We replace "(the ideal) a" by "(the integer) α'" and "(the ideal) b" by "(the integer) β'" in accordance with Dedekind's use of Latin letters to denote ideals and Greek letters to denote numbers. In this sense, we "pretend" that the passage is about ordinary integers rather than ideals. If we can do this without changing anything else in our text, it might suggest that, in the original excerpt, the argument requires the plane of set content in an irreducible way to make sense.

> Let β' be the number dividing precisely those numbers β for which α' divides $\eta\beta$. We show that β' divides every number that α' divides. Let α be divisible by α'. Then α' divides $\eta\alpha$. Thus β' divides α. Thus, β' divides every number that α' divides. So β' divides α'.

In this arithmetical version, what happens with the appeals to the ideal theoretic definitions? Concretely, if we had made these appeals explicit, we would have the expression "Then by the definition of an ideal, a contains $\eta\alpha$" instead of just "Then a contains $\eta\alpha$." But it is not clear what we should write instead of "by the definition of an ideal." Two immediate candidates are "by the definition of an integer" and "by

the definition of division." The first emphasizes the relational aspect ("by the definition of division"), and the second emphasizes the object aspect ("by the definition of an integer"). If we allow ourselves some speculation, perhaps this unclarity can be seen in relation to the tension between object-oriented mathematics on the one hand and relational or structural mathematics on the other. Gathering more examples like this one might be used as textual support for discussing the idea that in ideal theory, objects are defined relationally. The "vanishing" of objects, suggested by some scholars, from Dedekind's mathematics might here be expressed as the doubt, or the choice, between two clearly separate possible substitutions for "by the definition of an ideal," one that emphasizes the relational aspect ("by the definition of division") and one that emphasizes the object aspect ("by the definition of an integer"). In ideal theory, these choices might collapse into one, whereas they might be relatively separate in traditional arithmetic.

In the arithmetical version, we have instead the inference that if β' divides every number that α' divides then β' divides α'. This seems almost trivial, since α' divides itself (we consider integers to divide themselves). So if β' divides every number that α' divides, then α' is one among the numbers that β' divides. From this perspective, β' divides α' follows from β' divides every number that α' divides as a particular case from the general case.

The set version seems to work differently. For ideals a and b, "b divides every number that a divides" does not include "b divides a" as a special case; it is the definition of ideal divisibility. From this perspective, "if b divides every number that a divides then b divides a" is not an inference but the repetition of a definition given earlier in Dedekind (1964).

This suggests that the commutation between "(ideals) a and b" and "(integers) α' and β'" has an effect at the level of signifieds, expressed here as a difference between inferential structures**.** Where the proof using "(ideals) a and b" and set expressions appeals to a definition, the proof using "(integers) α' and β'" and the arithmetical expression "divide" infers a special case from the general. This difference can be taken to suggest that the passage in its original form does not rely only on arithmetical content. Rather, the passage in its original form would seem to indicate a set content that is not directly translatable to arithmetical content at this point in the text (but which may be put in rigorous correspondence later in the text or by other means). We do not exclude that the passage can be read purely arithmetically, i.e., a reading treating the ideals as a kind of arithmetical objects (as numbers, basically). Neither do we exclude a purely set-theoretic reading of the passage. But we may say that the space of possible readings includes at least one more option, which combines set theory and arithmetic, despite the fact that the two perspectives are not aligned in terms of explicit axioms or definitions at this point in the Dedekind (1964).

We explore this third possible reading by commuting possible verifications of **(2)**.

A possible proof of **(2)**:

> If $\eta'\gamma \equiv 0 \ (mod\ b)$ then b contains $\eta'\gamma$ so, by definition of b, $\eta'\gamma$ is a root β such that $\eta\beta \equiv 0 \ (mod\ a)$, i.e., $\eta\eta'\gamma \equiv 0 \ (mod\ a)$. Conversely, if $\eta\eta'\gamma \equiv 0 \ (mod\ a)$, then $\eta'\gamma$ is a root β of $\eta\beta \equiv 0 \ (mod\ a)$ so, again by definition of b, b contains $\eta'\gamma$, i.e., $\eta'\gamma \equiv 0 \ (mod\ b)$.

Under the same commutation as in we applied to the proof of **(1)**, we get:

> If $\eta'\gamma \equiv 0 \ (mod\ \beta')$, then β' divides $\eta'\gamma$ so, by definition of β', $\eta'\gamma$ is a root β such that $\eta\beta \equiv 0 \ (mod\ \alpha')$, i.e., $\eta\eta'\gamma \equiv 0 \ (mod\ \alpha')$. Conversely, if $\eta\eta'\gamma \equiv 0 \ (mod\ \alpha')$, then $\eta'\gamma$ is a root β of $\eta\beta \equiv 0 \ (mod\ \alpha')$ so, again by definition of β', β' divides $\eta'\gamma$, i.e., $\eta'\gamma \equiv 0 \ (mod\ \beta')$.

This proof relies on the definition of β' by the congruence $\eta\beta \equiv 0 \ (mod\ \alpha')$, just as the argument for ideals relies on the definition of the ideal b by the congruence $\eta\beta \equiv 0 \ (mod\ a)$. Neither argument appeals to definitions of divisibility. It was precisely around the use of this definition in our proof of **(1)** that we found the discrepancy between the arithmetical and the set version. When we commute "divide" with "contain" and Greek letters with Latin, we would expect to get a change at the level of content. But compared to the change we observed when we did the same commutation in the verification of (1), here something seems to stay the same under the commutation. This indicates that another part of what we called the inferential structure above is not sensitive to variation between "contain" and "divide." This suggests that part of the inferential structure does not depend on the specific kind of content (set or arithmetical).

This lack of effect leaves us in the open with regard to a specification of the signified of the congruences. We could not find it above, since inserting "divide" and "contain" for "$\equiv 0$" did not give enough of a difference at the level of content. Perhaps our analytical framework is inadequate to describe the congruence as sign. But another way to view the invariance would be that the signified of the congruence notation is intentionally underspecified so as to create a sign that mediates between set theory and arithmetic. Perhaps Dedekind's point is that the two kinds of content must interact even at a point when they are not compatible. We test this idea by a commutation designed to test what happens if we bracket one kind of content completely and consider only the other. In other words, we consider two "extreme" readers: an extremely arithmetical reader and an extremely set-oriented reader.

(i) The arithmetical reader may interpret the arithmetical content of **(2)** as a direct multiplication of an ideal b by a number η, ignoring the set content of b. This reader interprets the arithmetical content of concatenation by η as the "quotient" of a by b, that is, effectively treats a and b as integers.
(ii) The set-oriented reader translates the arithmetical content of concatenation by η into arithmetical multiplication of η with the elements of b.

Option (i) is in conflict with the inferential structure of **(1)**, which changes if we interpret the ideals as integers. Option (ii) is a total transition to set content, effectively ignoring the presence of arithmetical content, including the "equation-likeness" of the manipulations of the congruences.

Based on this, Dedekind's choice of notation thus seems to enable a mediation between the two kinds of content, rather than to encourage a total transition to any one of them. This may be possible precisely because of the congruence notation as sign remains underspecified, which in turn may be because Dedekind does not

justify the statements **(1)** and **(2)** further. Our counter-textual arguments for **(1)** and **(2)** brought to the surface some tensions between set content and arithmetical content that are not explicit at the level of expression of Dedekind's text. But with the level of detail of the original text, the reader does not need to articulate a signified that captures how set content and arithmetical content interact in the passage.

Dedekind's notational choices thus seem to give rise to a new signified mediating between arithmetical and set content. If we require mathematical signifieds to be rigorous, this mediating signified is not quite mathematical. If we consider the signified to be mathematical, this mathematical object is not rigorous in the way rigor is usually conceived. By the end of Dedekind (1964), Dedekind will have provided the reader with the means to a rigorous reconstruction of this mediating signified; our examples illustrate the functioning of a mathematical textual practice, which may or may not adhere to such reconstruction.

5 Conclusion

At the heart of this chapter is a single analytical move: The semiotic analysis produced different possible readings of the same passage of text. These readings were set up against each other. By opposing and mediating between the different readings, certain semiotic features of the text emerged.

Perhaps the most important feature was the underspecified mediating signified: a possible gap between, on the one side, the signifieds of the text and, on the other side, rigorous mathematical reconstructions. By pointing out this gap, we do not question the possibility of rigorous reconstruction. But we emphasize how Dedekind's writing seems to initiate, while at the same time depend upon, interpretative interaction with a reader and the reader's intervention is not necessarily a rigorous reconstruction (or rigorously reconstructable). From the perspective of our structuralist semiotic approach, the underspecification illustrates (dare we say explains) how mathematical text can be read before the mathematical content suggested along the way is rigorously established.

6 Cross-References

► Structural Semiotics as an Ontology of Mathematics

References

Avigad J (2004) Dedekind's 1871 version of the theory of ideals. Unpublished. http://www.contrib.andrew.cmu.edu/~avigad/Papers/ideals71.pdf. Accessed 13 June 2021

Avigad J (2006) Methodology and metaphysics in the development of Dedekind's theory of ideals. In: Ferreirós J, Gray J (eds) The architecture of modern mathematics: essays in history and philosophy. Oxford University Press, Oxford, pp 159–186

Awodey S, Reck H (2002) Completeness and categoricity. Part I: nineteenth-century axiomatics to twentieth-century metalogic. Hist Philos Logic 23(1):1–30. https://doi.org/10.1080/01445340210146889

Buurma RS, Heffernan L (2018) Search and replace: Josephine Miles and the origins of distant Reading. Modernism/modernity 3, cycle 4

Corry L (2004) Modern algebra and the rise of mathematical structures, 2nd edn. Birkhäuser, Basel. https://doi.org/10.1007/978-3-0348-7917-0

Dedekind R (1964) (Sup X). Über die Komposition der binären quadratischen Formen. In: Über die Theorie der ganzen algebraischen Zahlen. Vieweg+Teubner Verlag, Wiesbaden, pp 223–261

Dedekind R, Stillwell J (1996) Translator's introduction. In: Stillwell J (trans) Theory of algebraic integers. Cambridge University Press, Cambridge

Lejeune-Dirichlet JPG (1871) Vorlesungen über Zahlentheorie (ed: Dedekind R). Vieweg und Sohn, Braunschweig

Edwards HM (1980) The genesis of ideal theory. Arch Hist Exact Sci 23:321–378

Edwards HM (1990) Divisor theory. Birkhäuser Boston, Boston

Ferreirós J (2007) Labyrinth of thought: a history of set theory and its role in modern mathematics, 2nd edn. Birkhäuser, Basel. https://doi.org/10.1007/978-3-7643-8350-3

Gastaldi JL (2021) Why can computers understand natural language? Philos Technol 34(1):149–214. https://doi.org/10.1007/s13347-020-00393-9

Gastaldi JL, Pellissier L (2023) The logic of language: from the distributional to the structuralist hypothesis through types and interaction. hal-03064480v2

Haffner E (2014) The "science of numbers" in action in Richard Dedekind's works: between mathematical explorations and foundational investigations. History, philosophy and sociology of sciences. Doctoral dissertation, Université Paris Diderot, Paris 7. English. https://hal.archives-ouvertes.fr/tel-01144626

Haffner E (2017) Strategical use(s) of arithmetic in Richard Dedekind and Heinrich Weber's Theorie der algebraischen Funktionen einer Veränderlichen. Hist Math 44(1):31–69. https://doi.org/10.1016/j.hm.2016.09.002

Haffner E (2019) From modules to lattices: insight into the genesis of Dedekind's Dualgruppen. Br J Hist Math 34(1):23–42. https://doi.org/10.1080/17498430.2018.1555928

Herreman A (2000) La topologie et ses signes: éléments pour une histoire sémiotique des mathématiques. L'Harmattan

Hjelmslev L (1953) Prolegomena to a theory of language. Wawerly Press, Baltimore

Mehrtens H (1979) Die Entstehung der Verbandstheorie. Hildesheim, Gerstenberg

Miles J (1967) Style and proportion: the language of prose and poetry. Brown & Little, Boston

Moretti F (2000) Conjectures on world literature. New Left Rev 1:54–68

Moretti F (2013) Distant reading. Verso, London

Reck EH (2003) Dedekind's structuralism: an interpretation and partial defense. Synthese 137:369–419. https://doi.org/10.1023/B:SYNT.0000004903.11236.91

Reck EH (2009) Dedekind, structural reasoning, and mathematical understanding. In: Van Kerkhove B (ed) New perspectives on mathematical practices: essays in philosophy and history of mathematics. World Scientific, Singapore, pp 150–173

Sahlgren M (2008) The distributional hypothesis. Ital J Linguist 20:33–53

Sieg W, Schlimm D (2005) Dedekind's analysis of number: systems and axioms. Synthese 147:121–170. https://doi.org/10.1007/s11229-004-6300-9

Sieg W, Schlimm D (2017) Dedekind's abstract concepts: models and mappings. Philos Math 25(3):292–317. https://doi.org/10.1093/philmat/nku021

Vinciguerra L (1999) Langage, visibilité, différence: histoire du discours mathématique de l'âge classique. Vrin, Paris

Wagner R (2009) S(Zp, Zp): post-structural readings of Gödel's proof. Polimetrica, Milan

The Design of Mathematical Language

Jeremy Avigad

Contents

1 Introduction 3152
2 Perspectives 3154
2.1 Philosophical Orientation 3154
2.2 Methodological Orientation 3155
3 Fundamentals 3156
3.1 Sorts 3156
3.2 Operations 3158
3.3 Predicates and Relations 3159
3.4 Definedness and Partiality 3159
4 Abstract Objects 3161
4.1 Numbers 3161
4.2 Dedekind Abstraction 3162
4.3 Equality 3164
4.4 Identifying Objects 3165
4.5 Identifying Structures 3166
4.6 Managing Abstraction 3167
5 Mathematical Theories 3169
5.1 Definitions 3170
5.2 Theorems 3171
5.3 Proofs 3171
5.4 Notation 3174
5.5 Calculation 3175
6 Set-Theoretic Abstraction 3175
6.1 Functions 3175
6.2 Sets 3177
6.3 Quotients 3178
7 Algebraic Abstraction 3179
7.1 Implicit Structure 3179
7.2 Algebraic Hierarchies 3180
7.3 Structures as Objects 3182

J. Avigad (✉)
Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: avigad@cmu.edu

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_64

8 Other Aspects of Mathematical Language 3183
8.1 Diagrams 3183
8.2 Algorithms 3183
8.3 Plans 3184
8.4 Intuitions and Heuristics 3185
9 Conclusions 3186
References 3187

Abstract

As idealized descriptions of mathematical language, there is a sense in which formal systems specify too little, and there is a sense in which they specify too much. On the one hand, formal languages fail to account for a number of features of informal mathematical language that are essential to the communicative and inferential goals of the subject. On the other hand, many of these features are independent of the choice of a formal foundation, so grounding their analysis on a particular choice of a formal system introduces unnecessary specificity. This chapter begins to map out the design features of mathematical language without descending to the level of formal implementation, drawing on examples from the mathematical literature and insights from the design of computational proof assistants and their libraries.

Keywords

Mathematical language · Formalization · Proof assistants

1 Introduction

Mathematics is governed by a network of norms that determines what we can say and how we are supposed to say it. Learning how to speak about mathematics properly is an important part of learning to do mathematics. But, when pressed, it turns out to be surprisingly hard to say what it is we are talking *about*; mathematics seems to be, as the title of a book by John Burgess and Gideon Rosen proclaims, a subject with no object. And without making sense of what we are talking about, it is hard to see how we can make sense of what we are thinking about. The outward manifestations of mathematical thought are found in the way we communicate mathematics to others.

Here we adopt the view that an important way to make sense of what we are talking about is to develop a better understanding of mathematical language itself. From a grammatical point of view, the language of mathematics is rather simple. There are no subtle variations of tense, modality, or aspect, and the subject is generally devoid of subjunctives and counterfactuals. Mathematical statements make claims as to what is true, always has been true, and always will be true, and mathematical proofs back these claims up with more statements of the same sort. The subtleties of mathematical language stem rather from the features that render it suitable for reasoning about abstract objects and properties.

Formal languages provide informative models of mathematical discourse. It is by now well understood that ordinary mathematics can be formalized in various ways in first-order logic and the language of set theory, in variants of simple type theory or dependent type theory, or in category-theoretic foundations. These variations provide complementary perspectives and serve as bases for formalization in proof assistants like Mizar (Grabowski et al. 2010), Isabelle (Nipkow et al. 2002), Coq (Bertot and Castéran 2004), HOL Light (Harrison 2009), Metamath (Megill 2006), and Lean (de Moura et al. 2015).

As valuable as they are, there is a sense in which formal languages specify too much, and there is a sense in which they specify too little. The fact that an ordinary mathematical text can be represented equally well in any of the foundations above shows that these foundations should be viewed as alternative *implementations* of mathematics. It is therefore reasonable to look for descriptions of mathematical language and patterns of inference that abstract away the superficial differences and clarify the specifications that the implementations are designed to meet.

Formal languages specify too little in the sense that many essential features of ordinary mathematical language are not addressed by a formal specification. Instantiating the formal foundation is only the first step in implementing a mathematical proof assistant, and the bulk of the work then goes into supporting the interactions that make them usable in practice. Such systems have to parse user input, disambiguate notation, manage libraries of theorems and definitions, and keep track of algebraic structures and relationships between them. They need to provide convenient manners of expression and support efficient inference. The design of a proof assistant requires countless engineering decisions that bear on the system's usability, and these decisions can, in turn, be viewed as attempts to capture the functionality of informal mathematical language and inference. We might optimistically seek a better philosophical understanding of this functionality, one that can help us make sense of the implementation goals.

Another approach to thinking about mathematical language is to view it as a part of natural language, albeit a part of natural language with its own characteristic features. Ganesalingam (2013) leans in this direction, which allows him to bring the methods of generative linguistics to bear on the analysis of ordinary mathematical texts. Various systems of *controlled natural language* offer a complementary approach, presenting structured languages with enough flexibility to incorporate a range of natural language constructs, so that users can write texts that read like natural language but can be translated to the language of an underlying formal system (Paskevich 2007; Cramer et al. 2009).

Here we will not try to account for the range of grammatical and stylistic variation that one finds in natural language. We will rather seek idealizations that illuminate the features of mathematical language that are specifically adapted to supporting mathematical activity. One way of describing what we are after is to say that we are looking for a description of mathematics as a *semiformal language*. We want a description that renders it regimented and precise, like a formal language, designed to support certain types of abstraction and inference. At the same time, we want a model that is more informative than formal logic, one that tells a story of how

mathematical language serves to support the processes that are needed to track mathematical objects and the relationships between them. Our overarching goal will be to develop such a design specification for mathematical language without descending to the level of a fully formal implementation.

This chapter falls short of that goal, but it aims to make progress by cataloging some of the features of mathematical language that are essential to its communicative and inferential aims. It also describes some of the mechanisms that contemporary proof assistants use to model those features and support those aims. The interaction between informal and formal mathematics gives rise to a fruitful dialectic: the engineering efforts clarify the nature of mathematical language and inference, and, conversely, philosophical understanding guides the implementation.

2 Perspectives

2.1 Philosophical Orientation

Philosophy of mathematics has traditionally been concerned with the nature of mathematical objects, knowledge, and thought. Talking about these in a rigorous way requires some sort of conception of mathematics itself, and some sort of understanding of the features of mathematical practice that fall under the scope of the analysis. Toward forming such a conception, what we have the most direct access to is the mathematical literature: the historical record of statements, questions, arguments, definitions, and other textual artifacts that are constitutive of the subject. These artifacts, rooted in language, form the starting point for philosophical study.

No matter how we ultimately try to characterize the goals and methods of mathematics; we have to start with language. Whether we view mathematics as the practice of solving problems, abstracting from experience, or getting at a certain type of truth, what we say about that practice has to fit with what we see in the mathematical literature. We need to understand how mathematical language enables us to carry out those activities and how those activities are manifested in language.

One central thesis of this chapter is that, when we study mathematical language, it is important to understand not just what is *allowed*, but also what is *desired*. Formal systems specify rules that tell us when a formula is well-formed and when an inference is justified, but it doesn't tell us which definitions are good definitions, or which among the myriad inferential steps that can be taken at any given point are most worthy of our attention. Mathematics calls upon its practitioners to carry out complex tasks and to do so creatively, efficiently, and reliably. Reflection on mathematical practice should help explain how it helps us manage complexity and carry out fruitful exploration.

Another central thesis of this chapter is that it is helpful to view mathematical language as the object of *design*. Mathematical language and methods have evolved over the centuries, presumably for good reasons. Some features of mathematical language and method have remained remarkably stable, again, we may presume, for good reasons. Mathematics provides powerful means for abstraction and for

managing information, making data salient when it is needed and suppressing it when it is a distraction. Recognizing that mathematical language has evolved to serve such purposes should encourage us to try to understand the general principles by which it succeeds, and, conversely, understanding how it functions can help clarify the nature of mathematics itself. Thinking of mathematical language in terms of design is not at odds with viewing it as the outcome of an organic evolutionary process; some have even found it profitable to think about natural language in design terms (Chomsky 2005; Hockett 1960a, b).

2.2 Methodological Orientation

With respect to the goal of developing a semiformal description of mathematical language, how can we tell whether we are on the right track? This chapter uses two sources of data as touchstones.

The first is the mathematical literature itself. In the sections that follow, I will frequently resort to examples pulled from a random sample of textbooks. There is nothing exceptional about the examples I chose; the point is that they represent patterns that one can find in virtually any mathematical text. A theory of mathematical language should help us make sense of the instances of mathematics that we find all around us.

The second source of data comes from contemporary proof assistants, and from the formal methods that have been developed to support mathematical communication and inference in that setting. Processing a mathematical text requires tokenizing, parsing, and elaboration, and a variety of methods are used to infer information that is generally left implicit, including class inference, canonical instances, and unification hints (Wenzel 1997; Sozeau and Oury 2008; Asperti et al. 2009; Garillot et al. 2009). Theories and libraries are organized using environments, modules, and namespaces. Decision procedures support domain-specific reasoning, and indexing techniques provide fast ways of locating relevant information for automated search (Robinson and Voronkov 2001). Ongoing research is bringing the methods of machine learning to bear on mathematical reasoning (Gauthier et al. 2021; Lee et al. 2020).

Describing this technology in detail would take us too far afield. (For overviews, see Hales (2008); Avigad and Harrison (2014); Avigad (2018); Blanchette and Mahboubi (n.d.).) The technological developments are informative, however, because they make explicit the challenges that need to be met in order to bridge the gap between informal mathematics and its formal representation. Highlighting these challenges helps clarify the features of informal mathematical language that are essential to mathematical practice. A second measure of success for the analysis here, then, is the extent to which it helps us understand these features and meet the technical challenges.

We will focus on contemporary mathematical language. It is also worthwhile to study the way mathematical language has changed over time and to try to understand the reasons behind those changes, but I will make no attempt to do that here. When I refer to historical sources in this chapter, it is either to highlight features of

mathematical language that have remained fairly stable over time or to highlight features of mathematical language that are relatively new. In both cases, the aim is to characterize mathematical language as it is used today.

3 Fundamentals

This section deals with features of mathematical language that are close to the basic grammar of the subject. It considers the way that mathematical objects fall into categories that shape the way we talk about them, the expressions that we use to name objects, and the ways that we talk about the relationships between such objects.

3.1 Sorts

We start with the simple observation that mathematical objects fall into categories. Consider the first proposition of Euclid's *Elements*:

> Proposition 1. *On a given straight line to construct an equilateral triangle.*
> Let *AB* be the given finite straight line.
> Thus it is required to construct an equilateral triangle on the straight line *AB*.
> With centre *A* and distance *AB* let the circle *BCD* be described;
> again, with centre *B* and distance *BA* let the circle *ACE* be described;
> and from the point *C*, in which the circles cut one another, to the points *A*, *B* let the straight lines *CA*, *CB* be joined. (Euclid 2002, p. 3)

The proof refers to a number of different kinds of objects: points A, B, and C, a finite straight line AB, a circle BCD, and the equilateral triangle that is under construction. Some objects arise in the course of the proof: given the point A and the finite straight line AB (now construed as a distance), a circle is described, and given the points C and A, the straight line CA is joined. Objects bear relationships to one another: two circles can cut one another at a point C, and an equilateral triangle can lie on a straight line.

It is informative to consider the way that objects are introduced into contemporary mathematical discourse.

> Let X be a locally convex topological vector space and F a closed convex subset. (Royden 1988, p. 241)

> Let p and q be odd primes. (Ireland and Rosen 1990, p. 53)

> Let k be a field, and let $k[x] = k[x_1, \ldots, x_n]$ be a finitely generated ring over k. Let $\varphi : k \to L$ be an embedding of k into an algebraically closed field L. (Lang 2002, p. 378)

> Let $\mathfrak{M}$ be a maximal ideal of $k[x]$. Let σ be the canonical homomorphism $\sigma : k[x] \to k[x]/\mathfrak{M}$. (Lang 2002, p. 378)

> Let $\tau_2 = n$; let t_1 be the smallest k such that $X_k \geq \alpha$, if there is one, and n otherwise. (Billingsley 1995, p. 5)

The first three examples appear in the statements of theorems, whereas the next two are taken from the beginnings of proofs. The point is simply that we would not expect to see a theorem or proof begin with the words "let p and q be objects" or "let p and q be things." Mathematical variables always range over a *sort*.

It is not always clear what we should take the fundamental sort of an object to be and what we should view as ancillary attributes. It seems sensible to say that an equilateral triangle is fundamentally a triangle that happens to have the property of being equilateral, and it seems sensible to say that a prime number is a number that is prime. But is a nonnegative integer an integer that has the property of being nonnegative, or a number that has the property of being nonnegative and integral? Is a finite straight line in the *Elements* a straight line that happens to be finite? A maximal ideal of $k[x]$ can be construed as an ideal of $k[x]$ that has the property of being maximal, or a subset of $k[x]$ that has the property of being a maximal ideal, or simply a set that has the property of being a maximal ideal of $k[x]$. In axiomatic set theory, everything is a set, but that stance is not more informative than considering everything an object.

The fact that objects belong to sorts is not specific to mathematics. Aristotle distinguished between an object's *essential* and *accidental* properties. Socrates was both human and short, and we can imagine what ancient Greece might have been like had Socrates been a few inches taller, but when we try to imagine Socrates as anything other than a human being, it's hard to argue that we are still thinking about Socrates. The ascription of sorts to mathematical objects, however, has specifically mathematical uses.

- It can be used to disambiguate notation. The expression x^y might denote exponentiation, but it might also denote conjugation in a group. The expression $x \cdot y$ can denote multiplication in any structure, so its meaning depends on the sort of objects that x and y are.
- It can be used to determine meaning. An element x can be *maximal* only if it is viewed as an element of a structure with an associated ordering. A function f is *surjective* only with respect to a specification of its codomain, *continuous* only with respect to a topology on its domain and codomain, a *homomorphism* only with respect to algebraic structures on those, and *essentially bounded* only with respect to a measure on the domain and an ordering of the codomain.
- It can be used to justify an expression as meaningful; for example, $\sqrt{x}$ makes sense if x is a nonnegative real number, and $1/x$ is meaningful if x is a nonzero real.
- It can be used to support immediate inference. If $f : G \to H$ and $g : H \to K$ are group homomorphisms, their composition is again a group homomorphism. If we view them fundamentally as functions, then their composition is merely another function, and we need an additional inference to conclude that they are again homomorphisms.
- It can be used to support higher-level inferences and heuristics. Knowing that we are dealing with integers or real numbers or triangles or elements of a group cues us to a body of methods and background knowledge that are relevant to the inferential tasks at hand.

We can designate the sorts of objects that occur in the passage from Euclid with tokens like *Point*, *Line*, and *Circle*. In the theory of programming languages and in a number of computational proof languages, it is common to view these as *data types*, or, more simply, *types*. We can write $A : Point$ to express that a variable or expression A is a point, and write $L, M : Line$ to express that L and M are lines.

In formal axiomatic foundations like *simple type theory* and *dependent type theory*, every expression denotes an object of some type. Proof assistants based on those foundations therefore have built-in mechanisms for *type inference*, which is the task of determining the type of a given expression, and *type checking*, which is the task of checking whether the type of an expression matches an expected type. The latter can be used to report errors when users apply operations to the wrong sorts of arguments, or when the system cannot infer the information it needs to make sense of an expression. Even systems based on set theory sometimes incorporate notions of type. *Soft typing* refers to the use of a typing discipline separate from the axiomatic foundation to infer meaning and to report errors.

3.2 Operations

Mathematical operations provide constructions or descriptions of objects. Two points A and B determine a line segment AB and, in the excerpt from the *Elements*, ACE is specified as the circle with center B and radius equal to the distance AB. The first operation may be modeled with a constant $segment\text{-}of : Point \to Point \to Segment$, where the type indicates that *segment-of* can be applied to two points to yield a segment. The convention in proof assistants is to take the arrows to associate to the right, so that the type above is read $Point \to (Point \to Segment)$. This means that when *segment-of* is applied to a point A, the result, *segment-of A*, has type $Point \to Segment$. When this is applied to another point B, the resulting expression, *segment-of A B*, has type *Segment*. The constant *segment-of* in *segment-of A* is said to be *partially applied*, and considering such expressions is natural in some mathematical contexts but not others.

The second operation may be modeled by a constant $circle\text{-}of : Point \to Segment \to Circle$, or perhaps $circle\text{-}of : Point \to Point \to Circle$ if it turns out that the distance in question is always measured from the center. Arithmetic offers other prototypical examples of mathematical operations: we can refer to the sum of two numbers, $x + y$, the square x^2 of x, or the greatest common divisor *gcd x y* of x and y.

In most formal foundations, a function f has an intended domain A and an intended codomain B, which is what is expressed by writing $f : A \to B$. In ordinary mathematics, it is well understood that the interpretation of the codomain is fluid: the specification $f(x) = x^2$ can be viewed equally well as defining a function $f : \mathbb{R} \to \mathbb{R}$ and as defining a function $f : \mathbb{R} \to \mathbb{R}^{\geq 0}$, where $\mathbb{R}^{\geq 0}$ denotes the nonnegative real numbers. More confusingly, the interpretation of the domain is also fluid: in a mathematical context, $f(x) = x^2$ may be viewed as a function defined on the complex numbers, real numbers, rationals, integers, or natural numbers, or any other domain with a multiplication. These considerations bear on whether we take the squaring

function to be injective, surjective, both, or neither. In mathematical contexts there is generally a broadest construal of the domain that makes sense, and circumstances dictate when we need to pass to a restricted interpretation.

The fact that the dot in $x \cdot y$ can denote multiplication in any of a number of structures and that an expression *intersection X Y* may denote the intersection of two lines, two circles, or a line and a circle shows that mathematical notation and terminology are *overloaded*. In these cases, information about the sorts of the arguments can be used to disambiguate meaning. In some cases, it is natural to view information like the structure in which multiplication occurs as an *implicit argument* to the notation. Proof assistants invariably have mechanisms to infer such information. A semiformal specification of mathematical language should say something general about how ambiguities are resolved and how and when implicit information is inferred.

3.3 Predicates and Relations

If we use *center-of* to denote the relation of being the center of a circle, computer scientists are apt to write *center-of A* γ to express that A is the center of γ. In a proof assistant, we might write

$$\textit{center-of} : \textit{Point} \to \textit{Circle} \to \textit{Prop}$$

to specify the type of the relation. We continue to interpret the arrows as associating to the right, so that *center-of A* γ denotes a proposition, namely, the proposition that A is the center of γ. Modeled in this way, relations can take any number of arguments. A relation that takes only one argument is usually called a *predicate* or *property*. For example, *equilateral T* expresses that a triangle T is equilateral, and *even n* expresses that an integer n is even.

Notation and terminology for relations, like notation and terminology for operations, can also be overloaded. We can use a symbol $\textit{on} : \textit{Point} \to \textit{Line} \to \textit{Prop}$ to express that a point A lies on a line L, and a symbol $\textit{on} : \textit{Point} \to \textit{Circle} \to \textit{Prop}$ to express that a point A lies on a circle γ. Similarly, $x \leq y$ can be used to denote an order comparison in any structure that has such an order. And we have already seen that predicates can rely on implicit information. A predicate *maximal x* depends on the order structure with respect to which x is judged to be maximal, and *continuous f* depends on the topologies with respect to which f is judged to be continuous.

3.4 Definedness and Partiality

Another complication that arises in the formalization of mathematics is that mathematical language often uses operations that can only meaningfully be applied to some of the objects in the natural domain of interpretation. Euclid can only speak of the intersection of two circles when the two circles intersect, and the phrase "the

greatest common divisor of x and y" only makes sense when x and y have a greatest common divisor, which is not the case when x and y are zero.

Set theory allows us to view each of these as a partial function, whose value may or may not be defined at a given input. Formally, a partial function from A to a B is a function from some subset of A to B. Proof assistants sometimes provide mechanisms for defining *subtypes* or adding *preconditions*, namely, proof obligations that must be fulfilled every time such a function is applied. For example, one can insist that every inscription a/b is accompanied by a proof that b is nonzero.

In ordinary mathematics, however, the implicit preconditions can be subtle. Consider a statement of the prime number theorem:

> Let $\pi(x)$ be the number of primes between 1 and x...
>
> $$\lim_{x \to 1} \frac{\pi(x)}{x/ln(x)} = 1.$$
>
> (Ireland and Rosen 1990, p. 2)

The expression in the limit is undefined when x is negative, equal to 0, or equal to 1. None of this matters because to determine the limit, we can restrict attention to arbitrarily large values of x, but in a formal foundation, the range of relevant values and the interpretation of the expression at those values have to be made explicit. Similarly, consider the statement of Bézout's theorem in an introductory number theory textbook:

> ... Thus the greatest common divisor (b, c) is defined for every pair of integers b, c except for $b = 0$, $c = 0$, and we note that $(b, c) \geq 1$.
>
> **Theorem 1.3**. If g is the greatest common divisor of b and c, then there exist integers x_0 and y_0 such that $g = (b, c) = bx_0 + cy_0$. (Niven and Zuckerman 1980, p. 7)

Here, the phrase "the greatest common divisor" is rendered meaningful by an implicit assumption that b and c are not both zero.

One always has the option of modeling a partial operation as an operation on a smaller sort. For example, one can take the variable x in the statement of the prime number theorem as ranging over positive real numbers. This often leads to problems; we should be able to supply a positive real number in any context where a real number is expected. (See the discussion of *identification* in Sect. 4.) But in other contexts, it is more natural. The product $M \cdot N$ of two matrices makes sense only if the number of columns of M is equal to the number of rows of N; in other words, if M is an $m \times k$ matrix and N is a $k \times n$ matrix for some m, k, and n. One can interpret multiplication as a partial operation on matrices, but formalization is often smoother if we interpret it as a family of total operations indexed by m, k, and n, and take these values to be implicit in the expression. This requires that a variable like M can range over a sort *Matrix m k* of $m \times k$ matrices, or even *Matrix R m k*, where R specifies a ring of underlying

elements. Such a sort is said to be a *dependent type*. The fact that *Matrix R n n* can be viewed as a ring for every $n > 0$ provides additional support for viewing *Matrix R m k* as an independent sort of objects rather than a predicate on a larger one.

4 Abstract Objects

Mathematics is the art of rigorous abstraction. Mathematization amounts to identifying the features that are essential to the inferential and calculational steps that are needed in specific reasoning contexts and making them explicit. This results in *modularization* and *encapsulation*: the objects and assumptions on which an argument depends are presented as part of the interface, and the argument can then be applied to any data that instantiates the relevant hypotheses.

Abstracting and modularizing in this way serves two purposes (Avigad 2020). The first is to manage complexity. If an argument about a geometric object only depends on the fact that it is a triangle, presenting it as such means we can temporarily ignore other information we have about it, such as its particular dimensions or its relationship to other objects in a diagram. If an argument about the roots of an equation only depends on the fact that the set of permutations of those roots is a group, presenting the argument in group-theoretic terms allows us to set aside everything else we know about equations and roots. If an argument about a complicated space of functions depends only on the fact that it is a convex Banach space, presenting it as such allows us to suppress specific details of the space that are irrelevant to the argument.

The second purpose is, of course, generality. A theorem about triangles holds of all triangles, not just the one in a particular diagrammatic configuration. A group-theoretic result holds of all groups, of which there are many, and a theorem about convex Banach spaces holds of every convex Banach space. An abstract argument can be applied not only in settings we are currently interested in, but also in settings we cannot yet even imagine.

In Sects. 6 and 7, we will consider the ways that set-theoretic language and the use of algebraic structures, respectively, support these goals. But first, it is helpful to think about the way we talk about abstract objects in general, and the role that abstract objects play in mathematical thought.

4.1 Numbers

For centuries, mathematics was understood to be the general science of *quantity*, with continuous quantities, or *magnitudes*, as the subject matter of geometry, and discrete quantities, or *numbers*, as the subject matter of arithmetic. The first mention of magnitudes occurs in Book V of Euclid's *Elements*, which introduces the theory of ratios:

> A magnitude is *part* of a magnitude, the less of the greater, when it measures the greater. The greater is a *multiple* of the less when it is measured by the less. (Euclid 2002)

In the *Elements*, two magnitudes may be equal, or one may be less than another. A magnitude a is a *part* of a magnitude b if b can be *measured* as some whole number multiple of a. Book VII extends the terminology to *numbers*:

> An *unit* is that by virtue of which each of the things that exist is called one.
> A *number* is a multitude composed of units.
> A number is *a part* of a number, the less of the greater, when it measures the greater. (Euclid 2002)

Today, we are apt to interpret Euclid's magnitudes as nonnegative real numbers and his numbers as natural numbers. But Euclid's presentation and the early history of mathematics bring some interesting differences to light. One curious feature of the *Elements* is that the smallest number is 2; one sheep does not make a multitude. This fact is manifest in the first two propositions of Book VII, which describe the Euclidean algorithm for calculating the greatest common divisor of two numbers. The first proposition deals with the case where the result is a unit, whereas the second deals with the case where the two numbers have a common measure.

A more significant difference is that, in Euclid, a magnitude is always a magnitude of some type. It makes sense to add and compare lengths, areas, and angles, but not to add a length to an area or compare a length to an area. Even toward the end of the sixteenth century, we find François Viète, in the *New Algebra*, insisting that only homogenous quantities can be compared or added. So if x is a length, instead of $x^3 + x$, we must consider $x^3 + ax$ for some area a. Descartes' landmark *Geometry* of 1637 resolved the problem of comparing lengths and areas by providing a geometric construction that reduces products to lengths, modulo a choice of a unit.

Although today we consider numbers to be relatively concrete mathematical objects, we should keep in mind that they nonetheless represent a substantial abstraction. Mathematics allows us to add the average temperature in Pittsburgh on April 19, 2020 to the number of moons of Jupiter, and then multiply the result by the length of the Nile river, whether or not the resulting number is useful or interesting. The fact that we consider such a quantity to be a legitimate mathematical object says something significant about the nature of mathematical abstraction.

4.2 Dedekind Abstraction

A parable by Paul Benacerraf imagines two children who, thanks to a solid logical upbringing, each possess a set-theoretic definition of the natural numbers (Benacerraf 1965). The definitions serve each of the children well until they come to realize that although the definitions denote isomorphic set-theoretic structures, they are not identical. Since every object of set theory is, fundamentally, a set, mathematical objects have ancillary properties qua sets. So we can imagine that according to one definition, 2 is an element of 3, and according to the other, it is not. Since the structures are isomorphic, the two students are in a position to come to agreement regarding all substantive claims about the set of natural numbers. But the

identification of abstract numbers with concrete set representations forces us to the unsettling conclusion that, if the natural numbers really are sets, then either 2 is an element of 3 or it isn't, and any definition we give has to adjudicate the matter.

The title of Benacerraf's paper, "What numbers could not be," is a nod to an essay by Richard Dedekind from 1888, titled "Was sind und was sollen die Zahlen?" (roughly, "What are numbers and what should they be?"). That essay, and Dedekind's 1871 construction of the real numbers, foreshadow the problem. Dedekind explicitly addressed the issue in a letter to Heinrich Weber in 1888, in which he resists identification of the real numbers with his particular construction as a system of *cuts*:

> . . . I would advise that by [real] number . . . one understand not the class itself . . . but something new (corresponding to this class) which the mind creates. We are a divine race and undoubtedly possess creative power, not merely in material things (railways, telegraphs) but especially in things of the mind. This is precisely the same question that you raise at the end of your letter in connection with my theory of irrationals, where you say that the irrational number is nothing other than the cut itself, while I prefer to create something new (different from the cut) that corresponds to the cut and of which I say that it brings forth, creates the cut. . . . The rational numbers also produce cuts, but I would certainly not call the rational number identical with the cut it produces; and after the introduction of the irrational numbers one will often speak of cut-phenomena with such expressions, and ascribe to them such attributes, as would sound in the highest degree peculiar were they to be applied to the numbers themselves. Something quite similar holds for the definition of cardinal number . . . as a class; one will say many things about the class (e.g. that it is a system of infinitely many elements, namely, of all similar systems) that one would apply to the number only with the greatest reluctance; does anybody think, or won't he gladly forget, that the number four is a system of infinitely many elements? (Dedekind 1888)

These circumstances have led a number of philosophers to adopt a *structural* view of mathematics (Resnik 1997; Shapiro 1997; Reck 2003; Parsons 2004), which holds that when we say something about the number two, we are never talking about a unique object, but, rather, an element of an axiomatically characterized natural-number structure. Some go so far as to give the notion of a place in a structure a metaphysical standing of its own. The more sober mathematical practice, inaugurated by Dedekind, is to characterize mathematical structures uniquely up to isomorphism, provide a specific mathematical construction that shows that such a structure exists, and then refer to the structure only in terms of its axiomatic characterization, so that anything one says about the structure holds equally well of any isomorphic copy.

Perhaps there is no harm in identifying mathematical objects like numbers with particular set-theoretic objects if one then subsequently restricts oneself to the proper mathematical statements about them, but we should recognize that mathematical language obeys this restriction. Proof assistants implement various mechanisms to support this. In a number of proof assistants, once one has a construction of the real numbers as Dedekind cuts (or equivalence classes of sequences, or whatever), one can introduce a new type, $\mathbb{R}$, with an abstraction function *abs* from the type of representatives to $\mathbb{R}$, and a representation function *repr* in the other direction. These

are assumed to satisfy *abs* (*repr y*) $= y$. The type $\mathbb{R}$ provides the "new things" suggested by Dedekind. The function *abs x* maps every representation to the abstract thing it represents, and the axiom says that every real number has a representation, not necessarily unique.

Inductive types in a proof assistant allow one to declare the natural numbers axiomatically as a type that is freely generated by zero and a successor function. Such declarations are underwritten by a set-theoretic interpretation, but the form of the declaration in the system helpfully abstracts away the specifics. Alternatively, some systems, like Agda and Coq, offer systems of *modules*, which allow users to declare constants and properties axiomatically, reason about them, and instantiate them later on.

4.3 Equality

We will see in the next two subsections that mathematics allows us to substitute one expression for another in contexts that respect the sense in which we can treat them as the same thing. The equality relation in mathematics is generally reserved for substitutions that are allowed in all mathematical contexts. Saying $a = b$ means that we should be able to substitute a for b in any meaningful mathematical statement *salva veritate*, that is, preserving the truth of that statement. Frege's "On sense and reference" (Frege 1892) famously provides examples of linguistic contexts, like ascriptions of knowledge or belief, where identity and substitutability come apart. One of the hallmarks of mathematical language is that it is devoid of such modalities.

If we assume that any meaningful mathematical statement about an object x describes a property of x, the requirement above follows from the principle that if two arguments are equal, they have all the same properties. Conversely, if two objects have all the same properties in common, then they must be equal: being equal to x is a property of x, so if y has all the same properties that x does, it is equal to x as well. This characterization of equality is known as the *Leibniz principle* or the *principle of identity of indiscernibles*.

Leibniz' principle says something about equality in general, but nothing about what it means for two objects to be equal in particular. It is perfectly reasonable to ask when two natural numbers are equal to each other; the inductive characterization of the natural numbers provides an answer to that, namely, that they are equal if they are both zero or have equal predecessors. Similarly, if we define the real numbers as equivalence classes of Cauchy sequences, asserting the equality of two real numbers is tantamount to saying that the Cauchy sequences representing them are equivalent. Questions of equality involve representations of objects, and settling such questions can require substantial mathematical knowledge.

We have seen that foundational approaches often overspecify mathematical objects. In set theory, we can ask whether $1 + 1$ is equal to the set $\{\emptyset, \{\emptyset\}\}$, a mathematical variant of Frege's Julius Caesar problem. Dedekind's abstraction helps; in most implementations of type theory, a proposition $s = t$ is grammatically acceptable only when s and t are objects of the same type.

One thing that plagues proof assistants is that, in addition to substantial judgments of equality, there are also implicit equality judgments that are so fundamental that they have a silent bearing on the very grammar of the language. For example, if x and y are column vectors of length 3, then concatenating them results in a vector of length 6. In a mathematical text, one might use

$$M \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

to denote the product of an $n \times 6$ array with that vector. Strictly speaking, the concatenation of x and y yields a vector of length $3 + 3$, and a tiny bit of mental arithmetic is needed to see that the expression is well-formed. Accepting this fact opens the door to allowing any background mathematical knowledge to bear on such determinations, and it isn't clear where to draw the line. Systems of type theory that allow ordinary mathematical expressions to occur in an expression's datatype (like $\mathbb{R}^{3+3}$ and $\mathbb{R}^6$) distinguish between *definitional equality*, which is the kind that the system needs to recognize in order to judge whether an expression is well formed, and *propositional equality*, which is the more substantive kind. Informal mathematics does not make the distinction, and we need to better understand how to model the identifications that mathematicians are not even conscious of performing.

4.4 Identifying Objects

If there is any such thing as a concrete mathematical object, the number two should fit the bill. But we have seen that even the number two is a rather abstract thing, and its nature is slippery. There are a 2 in the integers and a 2 in the rationals, and also 2s in the real numbers and complex numbers. Set theory tells us that we have $\mathbb{Z} \subseteq \mathbb{R} \subseteq \mathbb{C}$, so we can view these all as the same 2. But such a view is hard to maintain as the 2s proliferate. There are also 2s in the quaternions and octonians, and there is a 2 in the ring of p-adic integers for every prime number p. In fact, there is a 2 in every ring. In $\mathbb{Z}/2\mathbb{Z}$, we have $2 = 0$, so that particular 2 probably isn't the same as the integer 2. Even if we restrict attention to rings of characteristic zero, which contain an isomorphic copy of $\mathbb{Z}$, it is hard to maintain that all the 2 s are the same. The 2 in the ring $M_2(\mathbb{R})$ of 2×2 matrices over $\mathbb{R}$ is

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

and it is hard to see how that can be equal to the integer 2.

The general mathematical stance is that what we are really doing is *identifying* all the isomorphic copies of the integers, allowing us to transfer properties between them. In particular, identifying the integers with a subset of the real numbers means that we can consider an object n to be an integer and a real number at the same time. So, in an expression like $3x^2 + 2x + 7$ we can view the coefficients as integers, we

can view x as a real number, and we can view the multiplication as the usual multiplication operation on the real numbers. But each number domain has its own special character. We can prove a general expression involving a natural number n by induction on n only if we view n as a natural number and not a real number, and we can use the implication from $n < m + 1$ to $n \leq m$ when we know m and n are integers. When we write the binomial theorem,

$$(x + y)^n = \sum_{i \leq n} \binom{n}{i} x^i y^{n-i},$$

we understand that i ranges over natural numbers, or, equivalently, nonnegative integers, x and y are elements of any commutative ring, and the apparent multiplication by $\binom{n}{i}$ is interpreted as an action of integers on ring elements. In fact, in the identity

$$\binom{n}{i} = \frac{n!}{i!(n-i)!}$$

we recognize that the division symbol is justified a priori because any integer can be viewed as a rational number, though at the same time we are aware of the fact that in this case, the operation keeps us within the integers.

The fluidity with which we pass between views of mathematical objects poses challenges for formalization. We will consider some of the practical mechanisms used to cope with this in Sect. 4.6.

4.5 Identifying Structures

We also identify structures. Any mathematician will tell you that the set $\{0, 1\}$ with addition modulo two is the two-element group, and that the set $\{1, -1\}$ with multiplication is also the two-element group. The reflexivity of equality tells us that the identity element of the two-element group is equal to the identity element of the two-element group. By substitution, we can conclude that 0 is equal to 1.

As a mathematical joke, this isn't a very good one. There is nothing mysterious going on here: mathematicians are well aware of the fact that we can treat distinct objects as being the same only insofar as we limit ourselves to operations and properties that respect the senses in which they are the same and avoid talking about the differences. The two structures in the last paragraph are the same only insofar as we talk about them as groups. If we allow ourselves to talk about the elements of the carrier set, they are patently different.

In logical terms, it is permissible to identify isomorphic structures in the sense that any statement that can be expressed in the language of the structure's signature cannot distinguish between them. But then it is important to recognize that the

relevant signature depends on the reasoning context at hand. Keränen (2001) challenged the philosophical view that mathematical objects are merely places in structures with the observation that some structures have nontrivial automorphisms. If the complex number i is just a place in a structure, which one is it? Anything that can be said about i can be said equally well about $-i$, so there is nothing to distinguish the two.

Keränen used this observation to reject a metaphysical stance, *realist structuralism*, which holds that references to elements of a mathematical structure should be interpreted at face value, in terms of an ontology of *places* in abstract structures. He takes the problem to rule this out in favor of a more mathematically familiar stance, *nominalist structuralism*, in which terms refer to elements in any system of elements that satisfies the relevant structural axioms. But the question remains as to how an expression like i can refer to any particular structure in which its denotation is ambiguous, and Keränen does not address that problem.

Mathematical practice itself suggests a straightforward solution: when we talk about the complex numbers, we know full well that there are two square roots of -1, and that i can be designated to be either one. So the best way to construe talk about the complex numbers while viewing them as an algebraically characterized structure is to view them as a structure with a fixed choice of i. With that choice, the structure is rigid, and the problem is solved. But this calls for a more robust account of how and when such choices should be viewed as having been fixed, implicitly, by the mathematical context, and what ensures that the various implicit choices are coherent with one another. We will see a variation on this problem in Sect. 7.2.

4.6 Managing Abstraction

The conclusion of the previous subsection is that it is safe to identify structures as long as suitable linguistic hygiene is maintained. But mathematicians don't go around declaring languages and signatures; in any given context, they simply know what is allowed. When we formalize mathematics with a proof assistant, we have to operationalize that. It's not enough to make vague gestures at reasoning contexts; the rules need to be spelled out in enough detail that a proof assistant can supply the relevant axiomatic justification. We also need to make sure the different contexts interact, so that we can establish mathematical claims once and for all and then apply them wherever it is appropriate to do so.

In this subsection, I will describe four general strategies that are used in formal libraries to support the application of abstract theorems in the various contexts in which they can be applied. The four strategies are as follows:

1. Generalize.
2. Use equality.
3. Use algebra.
4. Use coercions.

The first strategy can also be phrased in negative terms, in which case it sounds like cynical advice to a young politician: avoid saying anything specific. For instance, proof assistants usually establish $x + y = y + x$ as a general property of a suitable class of algebraic structures. That means that when we need to use the fact that $2 + 3$ is equal to $3 + 2$, we don't have to worry about whether this is a fact about integers or real numbers, or even whether they are the same fact or not. Nor do we have to worry about transferring the result from one domain to another. We just need to know that the relevant algebraic hypotheses are satisfied.

For that reason, mathematical libraries tend to favor theorems about commutative rings and ordered fields over theorems about specific number structures. Other opportunities for generality often crop up in unexpected ways. Theorems of multivariate analysis hold of the structure $\mathbb{R}^n$ of n-tuples of natural numbers, which we can view as maps from the set $\{0, 1, 2, \ldots, n-1\}$ to $\mathbb{R}$. But this can lead to the annoying need to reindex when we try to view $\mathbb{R}^n$ as a subspace of $\mathbb{R}^{m+n}$. In fact, most theorems one might want to prove about $\mathbb{R}^n$ remain true if we interpret n as an arbitrary finite set, and we can interpret the $+$ in $\mathbb{R}^{m+n}$ as disjoint union. This avoids the need to treat any particular n-element set as canonical. If we incorporate this design decision into a formal library, we don't need to identify one n-element set with another because our theorems hold for all of them. In a similar way, Buzzard et al. (2022) found it advantageous to reason about arbitrary localizations of a ring R at a submonoid S rather than any particular representation.

Nonetheless, there are invariably times when we want to consider a mathematical structure as both an entity in its own right and a substructure of a larger one. It is often convenient to have the integers be a subset of the reals. For a more striking example, consider the number field obtained by adjoining both $\sqrt{2}$ and $\sqrt{3}$ to the rationals. We can describe this structure as $\mathbb{Q}[\sqrt{2}][\sqrt{3}], \mathbb{Q}[\sqrt{3}][\sqrt{2}]$, or $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$. These describe distinct mathematical constructions, any of which may come up in practice, leaving us struggling to identify them when the contexts clash. But viewed as subsets of the real numbers, they are the same subset, and any element of one is an element of the other.

A variation on this second strategy played a fundamental role in the celebrated formalization of the Feit–Thompson theorem (Gonthier et al. 2013). The formalization of group theory used for that project might more accurately be called a formalization of *subgroup* theory because theorems in the underlying library were phrased as theorems about subgroups of a larger ambient group. Every group is a subgroup of itself, so, mathematically, there is no difference. But an element of a group can generally be viewed as an element of multiple subgroups of interest, and centering the library on the concept of a subgroup avoids the need to mediate between the different views.

The third strategy in the list is really a special case of the first: sometimes the right way to generalize a theorem involves adopting a proper algebraic perspective. Mathematicians are well aware of the fact that any abelian group A can be viewed, equivalently, as a $\mathbb{Z}$-module, which means that it comes equipped with an action $n \cdot x$ of $\mathbb{Z}$ on A. On the additive part of any ring R, this action coincides with multiplication by image of n in R. Reasoning in terms of scalar multiplication means that we don't have to worry about this identification; often the properties of the action

are all we need. Similarly, instead of reasoning about a subring R of a ring S, it is more convenient to view S as an R-algebra. This is equivalent to saying that there is a homomorphism from R to S, namely, the one that maps each element r in R to $r \cdot 1$. Any subring is a homomorphic image of the identity map, so, once again, there is no substantial mathematical difference. But building a library around images of morphisms rather than substructures avoids the need to identify one structure with another; the identification is built into the theorems. (See Baanen et al. (2021) for a nice instance of this.)

In fact, reasoning about structures and mappings between them from a structural algebraic perspective often makes it unnecessary to reason about elements at all. Every subspace of a vector space is the image of a linear map, and it is often natural to build a library around properties of morphisms and maps. There is no need to worry about identifying elements if the theorems are about the identifications themselves.

Finally, there are times when it is impossible to avoid identifying two distinct structures and transferring results between them. Sometimes we really need to identify the real numbers with the one-dimensional vector space over the reals, and sometimes we really need to identify the elements of an algebraic number field with their representations in the complex numbers. In that case, proof assistants often use *casts* or *coercions* from one structure to another. There are various mechanisms that can be used to infer the need for such coercions and insert them automatically, as is done in most programming languages. There are also mechanisms that can be used to translate statements across an embedding or isomorphism (Barthe et al. 2003; Huffman and Kuncar 2013; Lewis and Madelaine 2020). Coercions are generally not required to be injective; so in a sense, elements of the source are identified with one another as well as with their images. (See also the discussion of quotients in Sect. 6.3.) Homotopy type theory offers another solution, based on an axiom known as *univalence* (Univalent Foundations Program 2013).

It may seem disappointing that I have offered a grab-bag of ad hoc approaches to managing identification, especially when mathematicians seem to do it so effortlessly. Surely, one might think, there must be a simpler explanation as to how mathematicians manage to transfer results from domain to domain without violating foundational norms. But it is a mistake to look for an easy explanation. Mathematics requires extensive training, and learning to do mathematics requires learning how to manage abstraction appropriately. The strategies I have described are all natural, though partial, explanations of what is going on under the hood. The fact that mathematicians have internalized the allowable moves does not mean that we cannot make them explicit. In that respect, formalization can help us understand the inner workings of the practice.

5 Mathematical Theories

So far, we have focused on some of the basic components of mathematical statements. These, in turn, are used to formulate definitions, state theorems, and prove them. Collections of definitions and theorems are then structured into mathematical theories. This section considers some of the syntactic mechanisms that the subject uses to structure knowledge in this way.

5.1 Definitions

Mathematical definitions are often set apart typographically, but they are also often introduced on the fly, in the flow of a mathematical text.

> **Definition 2**. R is said to be a *principal ideal domain (PID)* if every ideal of R is principal. (Ireland and Rosen 1990, p. 9)

> **Definition**. Given a point x of a set X and an open set U of the space Y, let
>
> $$S(x,\ U) = \{f | f \in Y^X \text{ and } f(x) \in U\}.$$
>
> The sets $S(x, U)$ are a subbasis for a topology on Y^X, which is called the *topology of pointwise convergence* (or the *point open topology*). (Munkres 2000, p. 281)

> By a *measure algebra*, we mean a Boolean σ-algebra $\mathcal{A}$ together with a nonnegative real-valued function μ defined on $\mathcal{A}$ such that $\mu(A) = 0$ if and only if $A = 0$ and
>
> $$\mu\left(\bigvee_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu A_i$$
>
> if $A_i \wedge A_j = 0$ for $i \neq j$. (Royden 1988, p. 398)

In either case, the text serves to associate the body of the definition, the *definiens*, to the *definiendum*, which is the word or phrase being defined. In the first and third examples, the body of the definition associates a predicate on the type of objects it describes, in the first case, a ring R, and in the second, a pair $(\mu,\ \mathcal{A})$ of the kind indicated. In other cases, a definition describes a new object or structure.

Making sense of a definition by description requires recognizing that there is exactly one object meeting the description. When a function is defined in a programming language, the definitional apparatus itself guarantees the existence of a function, or at least a partial function, meeting the specification. Modern mathematics, however, allows for more liberal descriptions, and it is often not obvious that the object described has been well-defined. Indeed, it is a common procedure to justify such a definition explicitly by first proving existence and uniqueness. (And sometimes uniqueness means uniqueness up to isomorphism, along the lines described in Sect. 4.5.)

Sometimes a definition is a construction – we speak of the construction of the Haar measure or the Radon–Nikodym derivative or a direct limit. The definition of a function sometimes comes with an implicit algorithm (see Sect. 8.2), but mathematics tends to favor definitions that characterize the concept's role in a theory over the means of computation. In other words, the definition of the greatest common divisor of two integers or the Legendre symbol may or may not be accompanied by an algorithm to calculate it.

5.2 Theorems

Theorems are the building blocks of theories. They are noteworthy accomplishments in their own right and stepping stones to even greater ones. The word "proposition" is generally reserved for smaller theorems, facts that are generally unsurprising but which form the basis for more substantial results. A "lemma" is a fact established expressly to support the proof of a theorem, and the label "corollary" is used to denote a straightforward consequence.

Setting out the inferential structure of a mathematical theory requires having ways of referring to theorems that have previously been established. In a journal article or textbook, theorems generally have unique identifiers like "Theorem 5.3" or "Proposition 4.2.3," and these can be used across publications, as in "Theorem 4.3 of [5]." Some fundamental results have names, like the prime number theorem, the ergodic theorem, the dominated convergence theorem, and the fundamental theorem of algebra. Some names celebrate their authors, like Birkhoff's theorem, Picard's theorem, Szemerédi's theorem, Whitehead's theorem, Minkowski's inequality, and the Hilbert basis theorem. Some, like the Yoneda lemma, König's lemma, and Zorn's lemma, show that apparently minor results can take on added significance after the fact.

Sometimes a theorem is referred to descriptively, like the compactness of S^n, the uniqueness of Haar measure, the existence of an orthogonal complement, the monotonicity of the exponential function, and the fact that a positive semidefinite matrix has real eigenvalues. Using such a description suggests that supporting references are not hard to come by, and that the results are assumed to be part of a common background knowledge.

5.3 Proofs

In mathematical texts, proofs are often introduced with the label "*Proof*" and closed with a box at the end, a notational device introduced by Paul Halmos. The flow of information is generally linear. The primary justification for a claim typically comes from the statement or two preceding it, though additional background knowledge is often needed, and occasionally facts established earlier in a proof. Statements of fact are generally favored over the reasons that they hold, though a number of devices are used to indicate the latter.

> Hence the closure of S in E is $\overline{E} \cap S$. (Royden 1988, p. 151)

> By our previous results the first case obtains when $p \equiv 1 \bmod 4$ and $p \equiv \pm 1 \bmod 12$. (Ireland and Rosen 1990, p. 54)

> Because H is connected, the fiber F over $\gamma(0) = \gamma(1)$ is connected. (Duistermaat and Kolk 2000, p. 175)

> Now we can see that all the maps f_* and $\overline{f}_*$ in the commutative diagram of transfer sequences are isomorphisms by induction on dimension, using the evident fact that if three maps in a commutative square are isomorphisms, so is the fourth. (Hatcher 2002, p. 175)

> From $P[Y \in D_f] = \mu(D_f) = 0$ it follows that $f(Y_n) \to f(Y)$ with probability 1, and so by change of variable (see (21.1)) and the bounded convergence theorem, $\int f\, d\mu_n = E[f(Y_n)] \to E[f(Y)] \to \int f\, d\mu$. (Billingsley 1995, p. 335)

> Using Hilbert's theorem 90 again, together with the fact that $|\sigma\alpha| = |\alpha|$ for all $\alpha \in K^*$ we see that $H = K^{*1-\sigma}$. (Lang 1994, p. 189)

Sometimes the justification for a claim is an explicitly named theorem, like the bounded convergence theorem or Hilbert's theorem 90. Sometimes it is a named property, like the commutativity of addition or the compactness of a space. Sometimes it involves repeating a fact previously established. Sometimes it requires a few words of explanation, and sometimes it involves reminding the reader of something they already know.

The biggest difference between informal proofs and the proofs that are parsed by proof assistants is the amount of justificatory information that is left implicit in the informal ones. Every fact in a formal library, no matter how mundane or trivial, has a unique identifier, and one of the hardest parts about using a proof assistant is that one generally has to learn the names of facts in the library and refer to them explicitly when needed. Justificatory steps are also spelled out in far greater detail in a proof assistant, whereas ordinary mathematics has fewer resources for naming facts and composing them. Informal proofs rely instead on contextual hints and cues that enable a reader with sufficient expertise to spell out the details. This helps explain why expertise is so important; checking an informal proof is far from a mechanical process.

Sometimes the justification for a later inference is flagged in advance.

> Note that adjoining a copy of $\mathbb{Z}$ induces an injection on π_1 since the induced homomorphism is the free product of the injection $\pi_1(A) \to \pi_1(\mathbb{Z})$ with the identity map on the complementary free factor. (Hatcher 2002, p. 172)

> Note that each set U_n' is open, being the difference of an open set U_n and a closed set $\cup_{i=1}^n \overline{V}_i$. (Munkres 2000, p. 201)

Consider the significance of *noting* as a cognitive act; the proof asks the reader to keep the fact close at hand so that it can be deployed when needed, often without further comment. Mathematical texts call on us to perform other cognitive acts as well:

> Expressing each generator as a sum of homogeneous elements, ... (Lang 2002, p. 428)

> We can choose the indices i in such a way that $\sigma_i A_1 = A_i$. (Lang 2002, p. 183)

> Now identify $\mathfrak{g}$ with $\mathbb{R}^3$ by means of an orthonormal basis with respect to the Ad G-invariant inner product on $\mathfrak{g}$. (Duistermaat and Kolk 2000, p. 169)

> . . . consider the n simplex $[\nu_0, \ldots, \nu_n]$ with ν_0 at the origin and ν_k the unit vector along the kth coordinate axis for $k > 0$. (Hatcher 2002, p. 270)

> We now apply the bilinear form of Parseval's identity (Lemma 1.5, Chapter 3) to the integral defining $\mathcal{A}$. (Stein and Shakarchi 2003, p. 104)

To make our way through a proof, we have to express things, choose things, identify things, consider things, and apply things. (See Avigad (2006), especially Sect. 3.2, as well as Tanswell (n.d.) and Weber and Tanswell (2022).) Traditional logic doesn't account for these acts, only the result of performing them, which should be a network of inferences that takes us from a theorem's premises to its conclusion.

Logicians are fond of representing proofs diagrammatically as trees that branch at inference rules that have more than one premise. But the flow through an ordinary mathematical proof is remarkably linear. Sometimes the application of a theorem requires checking side hypotheses that are dispensed with quickly, and the base case in a proof by induction is usually dismissed easily. Occasionally, one comes across a case split where the cases have equal footing, and in those situations, the branching nature of the proof is made clear.

> If $d \equiv 2$ or 3 (4) then. . .
> If $d \equiv 1(4)$ . . .. (Ireland and Rosen 1990, p. 191)

> There are three cases. If $m < n$ then $f^{-1}(C)$ closed implies $C \cap X^{n-1}$ closed, hence $C \cap \overline{e}^m_\beta$ is closed since $-e^m_\beta \subset X^{n-1}$. If $m = n$ then e^m_β is one of the cells e^n_α, so $f^{-1}(C)$ closed implies $f^{-1}(C) \cap D_\alpha$ is closed, hence compact, hence its image $C \cap -e^n_\alpha$ under f is compact and therefore closed. Finally there is the case $m > n$. (Hatcher 2002, p. 521)

Such case splits, however, are few and far between.

Although a proof is generally presented as a sequence of factual assertions, the narrative has to pull off the difficult task of enabling us to determine how each successive statement follows from what has come before. Mathematical language supplies an array of devices that can be used to suggest inferential patterns, flag relevance, and indicate the necessary connections. Some of these devices are implemented in proof assistants. The Mizar proof language (Grabowski et al. 2010) was a pioneering effort in this sense, providing the first formal representation of informal proof vernacular. (See also de Bruijn (1987).) Most proof languages today contain similar means for structuring formal proofs. Those who design the languages have to balance competing goals, such as making proofs easy for users to write, making proofs easy for users to read, and making proofs easy for the system to process. But there is still more to be learned, and paying closer attention to the literature should provide better insight as to how well-chosen language supports mathematical inference.

5.4 Notation

Mathematical texts often introduce new notation.

> Let V, V' be two vector spaces, and suppose given a mapping
>
> $$V \times V' \to K$$
>
> denoted by
>
> $$(x, x') \mapsto \langle x, x' \rangle$$
>
> for $x \in V$ and $x' \in V'$. We call the mapping *bilinear* if ... (Lang 2002, p. 144)

> We define the *convolution* $f \star g$ on $[-\pi, \pi]$ by ... (Stein and Shakarchi 2003, p. 44)

> We define the *Dirichlet L-function associated to* χ by the formula
>
> $$L(s, \chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s}.$$
>
> (Ireland and Rosen 1990, p. 254)

Exploring what makes for good notation would take us too far afield, but it is worth recognizing that good notation has remarkable staying power (Cajori 1993). Even fonts and letter choices tend to stick around: the use of π goes back to the early eighteenth century, and the use of $f(x)$ and $g(x)$ for functions in analysis, as well as the use of e, i, and φ for the familiar mathematical constants, goes back to Euler. We still speak of Riemann's ζ function and a Dirichlet L series, and we tend to use Fraktur letters $\mathfrak{a}, \mathfrak{b}$, and $\mathfrak{c}$ for ideals, following Dedekind.

One challenge for proof assistants is that notation, like identifiers, can be overloaded. The addition symbol in $x + y$ can refer to addition in the integers or reals, but also to addition in any ring or additive group. The addition symbol can also be used to denote the pointwise sum $f + g$ of two functions, or the sum $\mu + \nu$ of two measures.

When a symbol is used in two unrelated ways, computer scientists refer to it as ad hoc polymorphism. In all the examples above, one can view the domains over which the addition is defined as being instances of an additively written semigroup, which makes them instances of *parametric polymorphism*. We can therefore interpret them as the same addition, where only the semigroup itself is left implicit. We will return to this when we discuss algebraic language in Sect. 7. Computational implementations of mathematical language use various mechanisms to infer the relevant structure. Even ad-hoc uses of an addition symbol can be used as instances of an "addition structure" (Wadler and Blott 1989).

5.5 Calculation

Once seen as the essence of mathematics, calculation is now often associated with applied mathematics, but it plays a role in pure mathematics as well. Here is an excerpt from a proof of the ergodic theorem:

$$\begin{aligned}\int_{[|a_n|>\lambda]} |a_n|\, dP &\le \frac{1}{n}\sum_{k=1}^{n}\int_{G_\lambda} |f(T^{k-1}\omega)|P(d\omega)\\ &\le \frac{1}{n}\sum_{k=1}^{n}\left(\int_{|f(T^{k-1}\omega)|>\alpha} |f(T^{k-1}\omega)|\,P(d\omega) + \alpha P(G_\lambda)\right)\\ &= \int_{|f(\omega)|>\alpha} P(d\omega) + \alpha P(G_\lambda)\\ &\le \int_{|f(\omega)|>\alpha} P(d\omega) + 2\frac{\alpha}{\lambda}E[|f|].\end{aligned}$$

(Billingsley 1995, p. 318)

In this case, the calculation is a chain of equalities and inequalities. Formal systems rarely specify syntax for calculations or distinguish them from other types of arguments, but proof assistants generally provide special mechanisms to support them.

6 Set-Theoretic Abstraction

Sets, functions, and relations are so fundamental to contemporary mathematics that it is easy to forget that they are recent additions to the subject. In the middle of the eighteenth century, Euler grounded calculus on the notion of a function, but the term did not take on anything approximating its contemporary significance until late in the nineteenth century. Similarly, it is hard to find anything like the modern notion of a set before that time.

The goal of this section is to characterize the contemporary notions of set and function and consider their role in mathematical language. In large part, set-theoretic language evolved in the nineteenth century as a means of supporting algebraic abstraction, which is discussed in the next section.

6.1 Functions

In the early nineteenth century, mathematicians dealt with a number of mathematical objects that we view as instances of the general function concept today, but were not then seen as such. These include sequences, series, geometric transformations, and permutations of roots of an equation. The word *function* and its cognates in western

languages were used exclusively for functions that are defined on the real or complex numbers and take values in one of those domains (Monna 1972; Youschkevitch 1976/77; Avigad and Morris 2014). Even number-theoretic functions, like the Euler ϕ function or the totient function, were described as *symbols* or *characters* rather than functions. The general view of a function as a mapping between any two domains first emerged later in the nineteenth century, in Dedekind's supplements to Dirichlet's *Lectures on Number Theory* and Frege's *Begriffsschrift*.

The modern concept of function has these features:

1. A function $f : A \to B$ maps elements from a domain A to a codomain B.
2. Any description that associates a unique element of B to each element of A serves to define a function. There need not be any algorithm or procedure for producing an element of B from an element of A; in fact, there need not be any means of representing particular elements of A in a way that an algorithm can act on them.
3. The axiom of choice allows us to consider functions for which we don't even have a means of description.
4. Two functions from A to B are considered to be the same when they take the same values at every argument. There is nothing more to a function than the association of output values to inputs.

Foundational accounts differ as to whether the specification of the codomain is part of the function, for example, whether the function $f(x) = x^2$ from $\mathbb{R}$ to $\mathbb{R}$ is simultaneously a function from $\mathbb{R}$ to the nonnegative reals. Mathematicians freely identify the two. Logical foundations often differ as to whether applying a function f to an element outside its intended domain is a syntactic error or a semantic one, that is, whether it is grammatically incorrect or only mathematically dubious.

In set theory, a function $f : A \to B$ is usually defined to be a set of ordered pairs, where $f(x) = y$ means that the pair (x, y) is an element of f. A function $f(x, y, z)$ with multiple arguments is generally viewed as a function $f : A \times B \times C \to D$. In type-theoretic foundations, the notion of a function $f : A \to B$ is a primitive notion, and the type of inputs and the type of the outputs are both an inherent part of the syntactic specification. We have seen that in type theory, a function $f(x, y, z)$ is commonly represented as an object of type $A \to (B \to (C \to D))$ and application is commonly expressed by writing $f\,x\,y\,z$, where x, y, and z are of type A, B, and C, respectively. This has the sometimes useful side effect that an expression like $f\,x\,y$ denotes an object of type $C \to D$, in which case, f is said to be *partially applied*. In type theory, it is harder to view a function that takes values in the integers as simultaneously a function that takes values in the real numbers. In compensation, the fact that the type of an expression like $f\,x$ is unambiguous is often useful.

The modern viewpoint allows us to treat functions as objects in their own right, on par with other mathematical objects, like natural numbers. We can quantify over functions the same way we quantify over numbers, we can define structures whose elements are functions just as we can define structures whose arguments are numbers, and we can define functions (also known as *functionals*) whose arguments are functions. This is what distinguishes the discussion of functions here from the discussion of

operations in Sect. 3.2. We can talk about multiplication as an operation on the integers, but when we talk about the multiplication *function* on the natural numbers, we are talking about a first-class object. Any implementation of mathematical language has to mediate between these two views, that is, the treatment of functions as objects and the treatment of functions as operations in the sense of Sect. 3.2.

6.2 Sets

In his book, *Labyrinths of Thought*, José Ferreiros observes that in the nineteenth century, there was no standard terminology to denote the set concept (Ferreirós 1999, page xx). Although French quickly settled on *ensemble*, Italian on *insieme*, and Spanish on *conjunto*, German terminology was not as uniform. Dedekind used words like *System* and *Gebiet*, while Cantor tended to favor *Mannigfaltigkeit* and *Menge*. It is interesting to read the following snippet from a letter from Dedekind to Cantor.

> (incidentally, I should like to see the shorter and equally Riemannian word "domain" [Gebiet] given clear preference over the clumsy word "manifold" [Mannigfaltigkeit]. . .). (Ewald 1996, p. 870)

The term *manifold* has, of course, survived, though it is used to describe sets with a specific geometric structure. Language and notation for sets in general are found everywhere in mathematics.

> Let $A_0 \subset C_0$ be the set of cubes disjoint from K, and inductively, let $A_k \subset C_k$ be the set of cubes disjoint from K and not contained in cubes of A_j for $j < k$. (Hatcher 2002, pp. 525–526)

> Consider the set of ideals $\{A^i \mid 1 \leq i \leq h_F + 1\}$. (Ireland and Rosen 1990, p. 179)

> W acts transitively on the set of connected components of $G^{\text{reg}} \cap T$. (Duistermaat and Kolk 2000, p. 154)

> Let $\{e_1, \ldots, e_m, \ldots\}$ be a basis of F. . . (Lang 2002, p. 521)

> A set A is a μ-continuity set if it is a Borel set and $\mu(\partial A) = 0$. (Billingsley 1995, p. 335)

Talk of sets is often implicit in the definitions of mathematical spaces and structures because they are often defined by specifying the underlying set.

> The space of bounded linear functionals on a normed space X is called the *dual* (or conjugate) of X and is denoted X^*. (Royden 1988, p. 326)

> Given a map $f: A \to B$, let E_f be the space of pairs (a, γ) where $a \in A$ and $\gamma : I \to B$ is a path in B. (Hatcher 2002, p. 407)

> Let $[M]$ denote the isomorphism class of a finite module M. We define the sum to be the direct sum. Then the isomorphism classes over the ring form a monoid. (Lang 2002, p. 139)

Any implementation of mathematical language has to make it easy to describe sets and reason about them.

In axiomatic set theory, the notion of set is basic; everything is a set, and a function is a set of ordered pairs. In axiomatic type theory, in contrast, the notion of a function between types is basic. A predicate on a data type α is viewed as a function from α to truth values, and sets are in one-to-one correspondence with predicates: given a predicate like *Even* on the natural numbers, we can form the set of even numbers, and, of course, given any set s, we can consider the property of being an element of s.

There is an important difference between set theory and type theory. In set theory, any collection of objects can form a set. We can consider the set consisting of the number π, the multiplication function on the integers, and the Eiffel Tower. In type theory, any set is a set of objects of some fundamental type. We can consider a set of integers or a set of functions, but there is no notion of a set of *arbitrary* mathematical objects. Whereas this is mathematically natural, the distinction between sets and types in type theory forces us to distinguish between the integers as a fundamental type and the integers as a subset of the real numbers. But, as we have seen in Sect. 4, the problem of identifying objects and structures is not limited to type theory. It poses challenges for set-theoretic foundations as well.

6.3 Quotients

In a familiar set-theoretic construction, if A is a set and $\equiv$ is an equivalence relation on A, the quotient $A/\equiv$ is defined to be the set of equivalence classes of A. Functions and relations on A that respect the equivalence relation descend to functions and relations on $A/\equiv$. (Confusingly, computer scientists often use the term "lift" instead of "descend.") This construction provides a useful way to treat equivalent objects as identical.

Quotient constructions come up often.

> We denote by $P_{\mathfrak{c}}$ the subgroup of P consisting of those principal fractional ideals (α) with $\alpha \in k_{\mathfrak{c}}$. Then it is clear that $P_{\mathfrak{c}}$ is a subgroup of $I(\mathfrak{c})$. The factor group $I(\mathfrak{c})/P_{\mathfrak{c}}$ will be called the group of c-*ideal classes*. (Lang 1994, p. 125)

> From the diagram above, $H_n(X)$ can be identified with $H_n(X^n)/\text{Im } \delta_{n+1}$. (Hatcher 2002, p. 140)

> Let X be the mapping torus of f, the quotient space of $\left(S^2_\alpha \vee S^2_\beta\right) \times I$ under the identifications $(x, 0) \sim (f(x), 1)$. (Hatcher 2002, p. 589)

> Let $\mathfrak{M}$ be the class of measurable subsets of [0, 1], $\mathfrak{N}$ the class of subsets of measure zero, and m Lebesgue measure. Then $\langle \mathfrak{M}/\mathfrak{N}, m\rangle$ is a separable measure algebra without atoms. (Royden 1988, p. 399)

> The equivalence classes of fractional ideals form a finite group, . . . which we call the *ideal class group*. (Lang 1994, p. 123)

> Because the canonical projection π: $M \to G\backslash M$ is continuous and maps open subsets of M onto open subsets of $G\backslash M$, the orbit space $G\backslash M$ is locally compact. (Duistermaat and Kolk 2000, p. 104)

Mathematicians often pass silently between talking about elements of a quotient and talking about representatives in the quotiented set. For example, it is common to conflate functions with equivalence classes of functions up to almost everywhere equivalence in a function space.

7 Algebraic Abstraction

Algebra involves calculating with variables that range over elements of some domain, using rules that are valid for that domain. The practice was used in the late sixteenth century by Viète, who saw the algebraic method as the foundation for a universal science. In the seventeenth century, Descartes and Leibniz held similar views.

We obtain modern algebra by making the rules explicit and abstracting the underlying domain, so that variables range over any domain that obeys the general laws, and the theorems we prove can be applied to any such domain. In other words, we get modern algebra (also known as *abstract algebra*) when we treat the objects under consideration as elements of an algebraic structure. This is a powerful means of abstraction that makes the properties that are relevant to a particular reasoning context explicit.

Conventional terminology wavers between using the expression "algebraic structure" to describe a class of mathematical structures and using it to describe any particular instance. In other words, sometimes we call the group concept an algebraic structure and sometimes we call a particular group an algebraic structure. For clarity, here I will reserve the term "structure" for the individual instances, and use the phrase "abstract structure" for the axiomatically characterized concept. With that terminology in place, we can describe the general practice: we define an abstract structure by specifying a linguistic signature and a collection of properties, and this allows us to reason uniformly about all its instances. The process of reasoning about such algebraically characterized classes becomes an art unto itself, and it is worthwhile to consider how mathematical language supports it.

7.1 Implicit Structure

An abstract structure is specified with a signature. We start by specifying that it has one or more carrier sets, and then we specify the functions and relations that are to be found in any instance, together with the axioms they are assumed to satisfy. A group

$(G, \cdot, e, \cdot^{-1})$ is any structure with a carrier set G, an element $e \in G$, a binary operation $\cdot$ on G, and a unary operation $\cdot^{-1}$ on G, all satisfying the group axioms. A partial order (X, R) consists of a transitive, reflexive, and antisymmetric relation R on a set X. An abstract structure can sometimes refer to higher-order objects over the carrier sets. A topological space consists of a pair $(X, \mathcal{T})$ where $\mathcal{T}$ is a collection of subsets of X satisfying the axioms for a topological space, and a hypergraph (X, E) consists of a set X and an arbitrary collection E of nonempty subsets of X.

When a mathematical object is an element of the carrier set of a structure, properties ascribed to the object often implicitly invoke the signature of the ambient structure. An element $a \in X$ of a partial order (X, R) is *minimal* if it is R-related to every other element of X. Similarly, properties of a function $f: X \to Y$ between carrier sets of two structures or a subset $A \subseteq X$ of the carrier set of a structure often make implicit reference to those structures. Saying that a set $A \subseteq X$ is *closed* or *compact* presupposes that there is a topology on X, saying that $f : X \to Y$ is continuous presupposes that there are topological structures on X and Y. This is an important sense in which sorts make a difference: to make sense of a statement like "a is minimal" or "f is continuous," it is not enough to view a as an object and f as a function; we have to view them as objects of appropriate sorts, and we have to view those sorts as coming equipped with ambient structures.

In a mathematical text, the relevant structures are often determined implicitly by the context. The same can be said for operations like the maximum function on an order or the closure operation on sets in a topology, and for notation like $x \cdot y$ and x^y. In proof assistants, various methods are used to infer implicit structure (Wenzel 1997; Asperti et al. 2009; Sozeau and Oury 2008; Garillot et al. 2009; Selsam et al. 2020). A good semiformal description should clarify the means by which the relevant associations are registered by a mathematical text.

7.2 Algebraic Hierarchies

Abstract structures bear relationships to one another. An abstract structure can add axioms to one described previously. A *total order* is a partial order (X, R) such that for any two elements a and b of X, either $R(a, b)$ or $R(b, a)$ holds. A *well-founded partial order* adds a different assumption, and a *well-founded total order* retains both. An *abelian group* is a group in which the group operation is commutative. Adding an axiom to an abstract structure means that the resulting class of instances is contained in the class of instances of the prior one, which is to say, any instance of the class of structure with the additional axiom is an instance of the class without it. This is one way in which structures form a hierarchy.

An abstract structure can also extend another by adding additional data. For example, a *pointed topological space* is just a topological space $(X, \mathcal{T})$ with an additional point $a \in X$. More commonly, an abstract structure will extend another with both additional data and additional axioms. A *group* is a monoid with an inverse; an *ordered group* is a group with a partial order on the carrier that is compatible with the group operation. A *ring* $(R, \cdot, +, 1, 0)$ is best described as the

combination of a monoid $(R, \cdot, 1)$ and an abelian group $(R, +, 0)$, satisfying the distributivity axioms.

Viewing a network of abstract algebraic structures and the relationships between them as a hierarchy supports modularity and abstraction in the sense of Sect. 4. A definition made in the context of one abstract structure instantly carries over to any others that inherit that structure, and a theorem proved about an abstract structure immediately applies to any of its instances. The reason that this is relevant to our discussion of mathematical language is that the transfer is built into the grammar. Any concepts defined in the context of a group can be interpreted in any abelian group or any ordered group, and, in practice, such concepts are deployed without comment or explanation. So a semiformal description of mathematical language has to specify not only how implicit algebraic structure is inferred, but also how algebraic relationships are tracked and used to make sense of mathematical statements.

The combined need to infer structure and to track relationships in the algebraic hierarchy gives rise to an interesting problem, namely, the *diamond problem*, that has been made manifest by the practice of formalization. It arises from the fact that, given a complex hierarchy, structure can be inferred in multiple ways, and the coherence of the practice relies on the fact that the results agree.

Here is an example of the phenomenon. Every metric space is an instance of a topological space; given a metric space, there is a unique topological space that is determined by the metric space. Given topological spaces with carriers X and Y, there is a canonical topology on their product, $X \times Y$. If X and Y are metric spaces, there are various natural ways of imposing a metric on $X \times Y$ that induces the same topology. Suppose we fix one and let f be a function from $X \times Y$ to some other space, and suppose we then say that f is continuous. We have seen that this requires us to view $X \times Y$ as a topological space. The question is: what is the topological space that is implicit in the statement?

The problem is that there are two ways to find such a space. Given that X and Y are metric spaces, we can view them as topological spaces, and then take the canonical topology on $X \times Y$. Alternatively, we can view $X \times Y$ as a metric space, and take the metric-space topology on $X \times Y$. Now, assuming we have chosen the metric on $X \times Y$ to have this property, the two choices coincide. But it may be a nontrivial mathematical *theorem* that they do. Suppose, in proving a theorem, we establish the continuity of f using one understanding of the relevant topological space, and then apply a theorem that presupposes the continuity of f using the other. We can hardly expect our proof assistant to run off and prove a substantial mathematical theorem just to infer that the two statements are the same. This is yet another instance of the identification problem raised in Sect. 7: mathematicians are comfortable shifting between equivalent readings of a statement once they have internalized the equivalence, but explaining what is going on formally requires some effort.

One solution found in the literature (Buzzard et al. 2020; Affeldt et al. 2020) is to take the structure of a metric space to include a specification of the associated topology, together with a proof that the topology is compatible with the metric. That makes it possible to ensure that both paths yield answers that are easily seen to

be the same. But more experimentation is needed to determine whether this is a viable way of solving all the diamond problems that crop up in practice, and whether the method can be made principled and general enough to become part of our semiformal specification.

Diamond problems serve to highlight fundamental questions about meaning and reference. As in Sect. 4.5, in cases where these are underspecified by the surface syntax, we would like to say that the mathematical context supplies the information needed to fix meaning and reference appropriately. But then we need a better story of how that works, and, in particular, a story about what ensures that the "canonical" interpretations that the context supplies are coherent. Implementations of mathematical language in proof assistants clarify the problem and open the door to philosophical analysis.

7.3 Structures as Objects

What gives an entity the status of a mathematical *object*? Section 3 presented some of the hallmarks: variables can range over objects, objects bear properties and relationships to one another, and mathematical operations build new objects from old ones. In Sect. 4, we observed that objects can be identified with other objects in various ways. The current section provides yet another aspect of objecthood: mathematical objects are, or can be seen as, elements of axiomatically characterized structures.

In contemporary mathematics, individual structures themselves are mathematical objects in all these senses. We quantify over groups, rings, and fields the same way we quantify over numbers and points on the Euclidean plane. We talk about products, powers, and limits of sequences of algebraic structures the same way we talk about products, powers, and limits of sequences of numbers. Various operations can be used to build new structures out of old ones; the automorphisms of a group, ring, or field form a group, the square-integrable real-valued functions on any measure space form a Hilbert space, and for every field F and natural number n, the structure F^n is a vector space. We identify structures, though here the relevant notion of "sameness" is isomorphism, which is coarser than the notion of equality that derives from viewing a structure as a tuple consisting of a carrier set and some functions and relations. And algebraic structures can be inhabitants of larger algebraic structures. The category of abelian groups (or abelian groups in a fixed set-theoretic universe) is an algebraic structure whose elements are abelian groups, a sheaf of rings is a collection of rings bearing an algebraic structure, and measure-theoretic probability allows us to consider measures on spaces whose elements are themselves measures.

This means that algebraic structures play dual roles. We can treat a group G as a mathematical object, and at the same time consider elements x and y of G. In the latter case, G specifies the sorts of objects that x and y are, and this is used to make sense of expressions that involve them. This duality is handled in various ways by proof assistants. One approach is to take a group, G, to consist of a structure, one

component of which is the carrier type, *carrier G*. The statement that x and y are elements of G is interpreted as shorthand for saying that x and y are elements of *carrier G* by coercing G to its carrier. An alternative is to use *type classes*, which provide a means of using context to associate information with a given data type. In that case, the statement that G is a group is interpreted as saying that G is a data type with an associated group structure. The statement that x and y are elements of G is interpreted as just that; but an expression like $x \cdot y$ is interpreted as making reference to an ambient structure that is inferred from the context.

8 Other Aspects of Mathematical Language

There are a number of important aspects of mathematical language that we have not even begun to address. We use language to state conjectures, raise questions, launch research programs, describe methods, and assess the value of mathematical contributions and artifacts. In this final section, we briefly consider some other aspects of mathematical communication.

8.1 Diagrams

Diagrams and images are important components of mathematical language. They are used to justify and explain, often more efficiently and effectively than text-based representations. In the last couple of decades, there has been a substantial literature on the use of diagrams in reasoning; see Giaquinto (2007); Manders (2008); De Toffoli and Giardino (2014); De Toffoli (2017); Hamami and Mumma (2013) for a representative sample. Work like this provides a diagrammatic analogue of the type of textual analysis proposed here, by developing a taxonomy of representational devices and their uses.

Given the size of the literature on mathematical diagrams, it is reasonable to ask why the analysis of textual communicative devices is lagging. One possibility is that the differences between diagrammatic representations and formal logic are so dramatic that it is clear that there is something interesting going on, whereas, in contrast, it is easier to convince ourselves that formal logic captures the essence of text-based reasoning. I hope this chapter has made the case that the mechanisms employed in text-based reasoning are just as elaborate, subtle, and mysterious as the ones employed in diagrammatic reasoning, and that they are just as worthy of study, if not more so.

8.2 Algorithms

For most of its history, mathematics was algorithmic in nature. Euclid's *Elements* provided recipes for constructing geometric objects, early algebraists were concerned with recipes for calculating quantities of interest, early probabilists were

concerned with calculating odds in games of chance, and early analysts were concerned with calculating the motions of the planets.

Pure mathematics tends to keep computation at arm's length, relegating questions about how to calculate quantities described or guaranteed to exist by a theory to the realm of application. A contemporary textbook on modern algebra might well prove that every matrix can be put in Jordan canonical form without saying explicitly how to do it (Lang 2002, p. 559). Still, algorithms sometimes creep into textbooks in pure mathematics.

> As a first application of quadratic reciprocity we show how, in conjunction with Proposition 5.1.2, it can be used in numerical computations of the Legendre symbol. (Ireland and Rosen 1990, p. 54)

> *Example 4L.3. Stable Splittings.* The formula ($\star$) tells us how to compute Steenrod squares for $\mathbb{R}\,P^\infty$, hence also for any suspension of $\mathbb{R}\,P^\infty$. (Hatcher 2002, p. 491)

Numerical and symbolic computation are an important part of mathematics, and the language that textbooks and journal articles use to describe such algorithms should be viewed as an important part of mathematical language.

Fortunately, computer science has developed ample means of representing algorithms formally – this is precisely what programming languages do. And proof assistants are commonly used for software verification, so there is a vast literature to draw on when thinking about mathematical algorithms.

8.3 Plans

Our discussion of proof in Sect. 5.3 ignored a number of narrative elements that are found in mathematical texts. A proof serves as a tour guide, drawing our attention along, highlighting key facts, and shaping the experience of understanding. A good proof reminds us where we have been, and it gives us a sense of where we are going and what we need to have with us on the way.

> Now let c be a connected component of t^{reg}. (Duistermaat and Kolk 2000, p. 147)

Think about the significance of the word "now." It is hard to pin down its inferential significance, but it plays an important role. It tells the reader that the part that has come before has been preparatory to that moment; it asks us to check our footing, and prepare ourselves for the next phase of the ascent.

> To show that ψ is onto, it is sufficient to show that ...
> ...
> Having shown that ψ is onto, we now investigate the kernel. Clearly, $\ker \psi = A_1 \cap A_2 \cap \cdots \cap A_g$. We must show that under the hypotheses the intersection is equal to the kernel. (Ireland and Rosen 1990, p. 181)

> Thus it only remains to show that (ii) ⇒ (iii). To do this. . . (Royden 1988, p. 206)

> We are now in a position to state the main theorem of this chapter. The proof will be spread out over the next three sections. (Ireland and Rosen 1990, p. 251)

> There would be no difficulty if we could write $B = A[\alpha]$ for some α. This is true only locally. Hence we shall use the approximation theorem to reduce our problem to the local case. (Lang 1994, p. 63)

Rebecca Morris has observed that we expect a proof to be *motivated*, in the sense that the structure of such a proof should make it clear how each step might have been anticipated, and how each step gets us closer to the goal (Morris 2020). Her work explores some of the devices that contribute to that aim. Morris and Yacin Hamami have argued that good proofs show evidence of a rational *plan* (Hamami and Morris n.d.). We tend to understand other people's behavior by assuming that they have particular goals and by interpreting their actions in relation to plans that achieve those goals. In a similar way, we make sense of proofs by interpreting them in terms of a systematic overall plan. Proof assistants give us languages for expressing definitions, theorems, and proofs. What can we say about the components of mathematical language that are used to put forth a plan?

I am grateful to David Waszek for pointing out that, contrary to a claim I made in the introduction, the last quotation includes an instance of a modality: "There *would be* no difficulty if *we could* write $B = A[\alpha]$ for some α." One also sometimes sees modal language in proofs by contradiction, for example: "We need to show x is even. Suppose x *were* odd. . .." In each case, modal language is used to modify the reasoning context by temporarily introducing a new assumption before drawing out conclusions. Logically and mathematically, such hypothetical reasoning is easy to explain, and it is telling that mathematicians more often avoid the use of subjunctive or counterfactual language in the second case, writing, instead, "suppose x *is* odd. . .." It may well be that the closest thing to the notion of a possible world in mathematical discourse is the notion of a temporary reasoning context, one that is related to a more primary reasoning context in some fashion. It is an interesting question whether conventional philosophical ways of thinking about modal language can shed light on the way we pass between such contexts.

8.4 Intuitions and Heuristics

Seventeenth century mathematicians like Descartes and Leibniz were enamored of the prospect of having methods of solving mathematical problems, like geometric construction problems and optimization problems. One can view this as a search for computational algorithms. But one can view their interest more broadly as having a body of *heuristics* that limit the search space and guide reason to its desired end. The mathematical literature is filled with such offerings.

> The methods to be used are identical with those already developed in the previous sections. (Ireland and Rosen 1990, p. 103)

> Two corollaries, interesting in themselves, will make clearer the structure of the proof of sufficiency given above. (Billingsley 1995, p. 350)

> The next section treats a general scheme for dealing with path-function questions by in effect replacing an uncountable time set by a countable one. (Billingsley 1995, p. 503)

> We now return, for the remainder of the chapter, to the consideration of various methods for imposing topologies on sets. (Munkres 2000, p. 112)

> The notions of a Galois extension and a Galois group are defined completely algebraically. Hence they behave formally under isomorphisms the way one expects from objects in any category. (Lang 2002, p. 264)

> This section contains a general discussion of invariant densities, especially on homogenous spaces. Although in the next section this is only used in a relatively simple case, the general discussion will be useful for future reference. (Duistermaat and Kolk 2000, p. 179)

> Notice that if f already sends some vertices of K to vertices of L then we may choose g to equal f on these vertices, and hence the homotopy from f to g will be stationary on these vertices. This is convenient if one is in a situation where one wants maps and homotopies to preserve basepoints. (Hatcher 2002, p. 179)

Statements like these do not make precise mathematical claims. Instead, they convey information that is meant to shape the way we organize, categorize, and store the knowledge we have, and the way we deploy that knowledge in the future.

Automated reasoning depends crucially on using heuristics to limit the space of options and fruitfully guide a search through a combinatorial explosion of possibilities. Successful machine learning algorithms seem to acquire representations that also structure spaces of options in fruitful ways. If we focus on mathematical statements like the ones above and think about what they are doing, we may be in a better position to design systems that provide better support for mathematical reasoning.

9 Conclusions

Understanding mathematical language can help us formalize mathematics, design better automated reasoning tools, and harvest data mechanically from mathematical texts. But mathematical language is also interesting in its own right, and understanding how it works adds to our appreciation of mathematics. A good piece of mathematics is a work of art, and there is a deep satisfaction to be had in admiring the creativity and insight that went into it. But that is not at odds with studying the mechanics of the underlying language. Whether we view that language as the medium for mathematical thought or merely an outward manifestation, understanding how it works is an important part of understanding the nature of mathematics itself.

References

Affeldt R, Cohen C, Kerjean M, Mahboubi A, Rouhling D, Sakaguchi K (2020) Competing inheritance paths in dependent type theory: a case study in functional analysis. In: Peltier N, Sofronie-Stokkermans V (eds) International joint conference on automatated reasoning (IJCAR) 2020. Springer, Berlin, pp 3–20

Asperti A, Ricciotti W, Coen CS, Tassi E (2009) Hints in unification. In: Berghofer S, Nipkow T, Urban C, Wenzel M (eds) Theorem proving in higher order logics (TPHOLs) 2009. Springer, Berlin, pp 84–98

Avigad J (2006) Mathematical method and proof. Synthese 153(1):105–159

Avigad J (2018) The mechanization of mathematics. Not Am Math Soc 65(6):681–690

Avigad J (2020) Modularity in mathematics. Rev Symb Log 13(1):47–79

Avigad J, Harrison J (2014) Formally verified mathematics. Commun Assoc Comput Mach 57(4): 66–75

Avigad J, Morris R (2014) The concept of "character" in Dirichlet's theorem on primes in an arithmetic progression. Arch Hist Exact Sci 68(3):265–326

Baanen A, Dahmen SR, Narayanan A, Nuccio Mortarino Majno di Capriglio FAE (2021) A formalization of Dedekind domains and class groups of global fields. In: Cohen L, Kaliszyk C (eds) Interactive theorem proving (ITP) 2021, vol 193. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, LIPIcs, pp 5:1–5:19

Barthe G, Capretta V, Pons O (2003) Setoids in type theory. J Funct Program 13(2):261–293

Benacerraf P (1965) What numbers could not be. Philos Rev 74(1):47–73

Bertot Y, Castéran P (2004) Interactive theorem proving and program development – Coq'Art: the calculus of inductive constructions. Springer, Berlin

Billingsley P (1995) Probability and measure, 3rd edn. Wiley, New York

Blanchette J, Mahboubi A (eds) (n.d.) Handbook of proof assistants and their applications in mathematics and computer science. Springer, Berlin

Buzzard K, Commelin J, Massot P (2020) Formalising perfectoid spaces. In: Blanchette J, Hritcu C (eds) Certified programs and proofs (CPP) 2020. ACM, pp 299–312

Buzzard K, Hughes C, Lau K, Livingston A, Mir RF, Morrison S (2022) Schemes in Lean. Exp Math 31(2):355–363

Cajori F (1993) A history of mathematical notations. Dover Publications, New York. Two volumes. Reprint of the 1928 and 1929 originals

Chomsky N (2005) Three factors in language design. Linguist Inquiry 36(1):1–22

Cramer M, Fisseni B, Koepke P, Kühlwein D, Schröder B, Veldman J (2009) The Naproche project: controlled natural language proof checking of mathematical texts. In: Fuchs NE (ed) Workshop on controlled natural language (CNL) 2009. Springer, Berlin, pp 170–186

de Bruijn NG (1987) The mathematical vernacular, a language for mathematics with typed sets. In: Dybjer P et al (eds) Proceedings of the workshop on programming languages, Marstand, Sweden

de Moura LM, Kong S, Avigad J, van Doorn F, von Raumer J (2015) The Lean theorem prover (system description). In: Felty AP, Middeldorp A (eds) Conference on automated deduction (CADE) 2015. Springer, Berlin, pp 378–388

De Toffoli S (2017) 'Chasing' the diagram – the use of visualizations in algebraic reasoning. Rev Symb Log 10(1):158–186

De Toffoli S, Giardino V (2014) Forms and roles of diagrams in knot theory. Erkenntnis 79(4): 829–842

Dedekind R (1888) Letter to Heinrich Weber (24 January 1988), translated in Ewald (1996), vol 2, pp 834–835

Duistermaat JJ, Kolk JAC (2000) Lie groups. Springer, Berlin

Euclid (2002) Euclid's *Elements*. Green Lion Press, Santa Fe. The Thomas L. Heath translation, Edited by Dana Densmore

Ewald W (1996) From Kant to Hilbert: a source book in the foundations of mathematics. Clarendon Press, Oxford. in two volumes
Ferreirós J (1999) Labyrinth of thought: a history of set theory and its role in modern mathematics. Birkhäuser Verlag, Basel
Frege G (1892) Über Sinn und Bedeutung. Zeitschrift für Philosophie und philosophische Kritik 100:25–50
Ganesalingam M (2013) The language of mathematics: a linguistic and philosophical investigation. Springer, Heidelberg
Garillot F, Gonthier G, Mahboubi A, Rideau L (2009) Packaging mathematical structures. In: Berghofer S, Nipkow T, Urban C, Wenzel M (eds) Theorem proving in higher order logics (TPHOLs) 2009. Springer, Berlin, pp 327–342
Gauthier T, Kaliszyk C, Urban J, Kumar R, Norrish M (2021) Tactictoe: learning to prove with tactics. J Autom Reason 65(2):257–286
Giaquinto M (2007) Visual thinking in mathematics: an epistemological study. Oxford University Press, Oxford
Gonthier G, Asperti A, Avigad J, Bertot Y, Cohen C, Garillot F, Roux SL, Mahboubi A, O'Connor R, Biha SO, Pasca I, Rideau L, Solovyev A, Tassi E, Théry L (2013) A machine-checked proof of the odd order theorem. In: Blazy S, Paulin-Mohring C, Pichardie D (eds) Interactive theorem proving (ITP) 2013. Springer, Berlin, pp 163–179
Grabowski A, Kornilowicz A, Naumowicz A (2010) Mizar in a nutshell. J Formal Reason 3(2): 153–245
Hales TC (2008) Formal proof. Not Am Math Soc 55(11):1370–1380
Hamami Y, Morris R (n.d.) Plans and planning in mathematical proofs. Rev Symb Log
Hamami Y, Mumma J (2013) *Prolegomena* to a cognitive investigation of Euclidean diagrammatic reasoning. J Log Lang Inf 22(4):421–448
Harrison J (2009) HOL light: an overview. In: Theorem proving in higher order logics (TPHOLs) 2009. Springer, Berlin, pp 60–66
Hatcher A (2002) Algebraic topology. Cambridge University Press, Cambridge
Hockett C (1960a) Logical considerations in the study of animal communication. American Institute of Biological Sciences, Washington, DC
Hockett C (1960b) The origin of speech. Sci Am 203:88–96
Huffman B, Kuncar O (2013) Lifting and transfer: a modular design for quotients in Isabelle/HOL. In: Gonthier G, Norrish M (eds) Certified programs and proofs (CPP) 2013. Springer, Berlin, pp 131–146
Ireland K, Rosen M (1990) A classical introduction to modern number theory, 2nd edn. Springer, New York
Keränen J (2001) The identity problem for realist structuralism. Philos Math 9(3):308–330
Lang S (1994) Algebraic number theory, 2nd edn. Springer, New York
Lang S (2002) Algebra, 3rd edn. Springer, New York
Lee D, Szegedy C, Rabe MN, Loos SM, Bansal K (2020) Mathematical reasoning in latent space. In: International conference on learning representations (ICLR) 2020, OpenReview.net
Lewis RY, Madelaine P (2020) Simplifying casts and coercions (extended abstract). In: Fontaine P, Korovin K, Kotsireas IS, Rümmer P, Tourret S (eds) Practical aspects of automated reasoning (PAAR) 2020, CEUR-WS.org, CEUR workshop proceedings, vol 2752, pp 53–62. http://ceur-ws.org/Vol-2752/paper4.pdf
Manders K (2008) The Euclidean diagram. In: Mancosu P (ed) The philosophy of mathematical practice. Oxford University Press, Oxford, pp 80–133
Megill ND (2006) Metamath. In: Wiedijk F (ed) The seventeen provers of the world. Springer, Berlin, pp 88–95
Monna AF (1972) The concept of function in the 19th and 20th centuries, in particular with regard to the discussions between Baire, Borel and Lebesgue. Arch Hist Exact Sci 9(1):57–84
Morris RL (2020) Motivated proofs: what they are, why they matter and how to write them. Rev Symb Log 13(1):23–46

Munkres JR (2000) Topology, 2nd edn. Prentice Hall, Upper Saddle River
Nipkow T, Paulson LC, Wenzel M (2002) Isabelle/HOL – a proof assistant for higher-order logic. Springer, Berlin
Niven I, Zuckerman HS (1980) An introduction to the theory of numbers, 4th edn. Wiley, New York/Chichester/Brisbane
Parsons C (2004) Structuralism and metaphysics: the foundations of mathematics and logic. Philos Q 54(214):56–77
Paskevich A (2007) Méthodes de formalisation des connaissances et des raisonnements mathématiques: aspects appliqués et théoriques. PhD thesis, Paris XII University
Reck EH (2003) Dedekind's structuralism: an interpretation and partial defense. Synthese 137(3): 369–419
Resnik MD (1997) Mathematics as a science of patterns. The Clarendon Press/Oxford University Press, New York
Robinson JA, Voronkov A (eds) (2001) Handbook of automated reasoning. Elsevier/MIT Press. in two volumes
Royden HL (1988) Real analysis, 3rd edn. Macmillan Publishing Company, New York
Selsam D, Ullrich S, de Moura L (2020) Tabled typeclass resolution. CoRR abs/2001.04301. https://arxiv.org/abs/2001.04301
Shapiro S (1997) Philosophy of mathematics: structure and ontology. Oxford University Press, New York
Sozeau M, Oury N (2008) First-class type classes. In: Mohamed OA, Muñoz CA, Tahar S (eds) Theorem proving in higher order logics (TPHOLs) 2008. Springer, Berlin, pp 278–293
Stein EM, Shakarchi R (2003) Fourier analysis: an introduction. Princeton University Press, Princeton
Tanswell FS (n.d.) Go forth and multiply: on actions, instructions and imperatives in mathematical proofs. In: Brown J, Bueno O (eds) Essays on the philosophy of Jody Azzouni
Univalent Foundations Program T (2013) Homotopy type theory: univalent foundations of mathematics. Institute for Advanced Study, Princeton
Wadler P, Blott S (1989) How to make ad-hoc polymorphism less ad-hoc. In: Principles of programming languages (POPL) 1989. ACM Press, pp 60–76
Weber K, Tanswell FS (2022) Instructions and recipes in mathematical proofs. Educ Stud Math 111: 73–87
Wenzel M (1997) Type classes and overloading in higher-order logic. In: Gunter EL, Felty AP (eds) Theorem proving in higher order logics (TPHOLs) 1997. Springer, Berlin, pp 307–322
Youschkevitch AP (1976/77) The concept of function up to the middle of the 19th century. Arch Hist Exact Sci 16(1):37–85

How to Do Maths with Words: Neural Machine Learning Applications to Mathematics and Their Philosophical Significance

Juan Luis Gastaldi

Contents

1 Introduction 3192
2 What Are Deep Neural Nets? 3194
3 How to Do Maths with Neural Nets 3196
3.1 Proof-Oriented 3197
3.2 Object-Oriented 3202
3.3 Skill-Oriented 3207
3.4 Heuristic-Oriented 3211
4 Philosophical Significance 3215
4.1 The Return of Language 3216
4.2 New Challenges for a Linguistic Approach to Mathematical Practice 3218
5 Conclusion: But How? 3220
References 3221

Abstract

Recent years have seen a remarkable development of deep neural network techniques for data analysis, along with their increasing application in scientific research across different disciplines. The field of mathematics has not been exempted from this general trend. The present chapter provides a survey of recent applications of neural models to mathematics and assesses their philosophical significance, related to the role of language in mathematics.

Keywords

Deep Neural Networks · Mathematical Practices · Artificial Intelligence · Mathematical Language · Natural Language · Natural Language Processing · Mathematical Language Processing · Linguistics · Computation ·

J. L. Gastaldi (✉)
ETH Zurich, Zurich, Switzerland
e-mail: juan.luis.gastaldi@gess.ethz.ch

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5_142

Large Language Models · Mathematical Proofs · Mathematical Objects · Mathematical Skills · Mathematical Heuristics

1 Introduction

On February 15, 2023, Yuhuai (Tony) Wu, researcher in Artificial Intelligence at Google at that time, gave a talk at the UCLA Institute for Pure & Applied Mathematics (IPAM).[1] In his talk, Wu presented his most recent results in the automatic formalization of proofs through deep neural networks. The workshop was organized by a small and highly select committee, including two Fields Medalists, and the audience was composed of well-established world specialists in the field of automated theorem-proving. While explaining the details of the neural model in question and providing elements for its epistemological guarantees to this intrigued yet somewhat skeptical public, Wu affirmed: "We know for sure that the informal proof, the natural language proof, is correct because it is produced by a human." The statement was followed by a second of sincere bewilderment in the room, only resolved by one glorious collective laugh. Unable to hide his confused embarrassment, Wu tried, however, to reassure the crowd by adding: "Ok, so, we make sure that the human. . . we scrape the solutions from the Internet!" inevitably provoking a new burst of laughter, this time sounding like a redemption.

Wu's statement is not improvised or misguided, nonetheless. Let alone isolated. For a proof, take the official presentation of the second version of Google's Large Language Model PaLM (Ghahramani 2023), released shortly after Wu's talk. In it, one can read the following passage, flagged as "weird" by the renowned researcher at the Santa Fe Institute, Melanie Mitchell:

> PaLM 2's wide-ranging dataset includes scientific papers and web pages that contain mathematical expressions. As a result, it demonstrates improved capabilities in logic, common sense reasoning, and mathematics. (Ghahramani 2023)

The rapid proliferation of this peculiar kind of statement in the current state of research at the crossroads of mathematics and computing is not as much anecdotal as it is *symptomatic*. As is symptomatic the perplexed reaction of well-established specialists in the fields concerned. Both are the symptom of the surge of a perspective on mathematical knowledge and practices radically at odds with the one that governed our image of mathematics for more than a century. This circumstance is the effect of an ongoing revolution in the field of machine learning resulting from the remarkable development of Deep Neural Network models (DNNs) during the last two decades. As it has been widely observed, the fact that tech corporations are the main actors of this technological revolution entails that most research efforts

[1] Wu's talk was part of a Workshop on Machine-Assisted Proofs held between February 13 and 17, 2023. The recording of the talk is available on the website of the IPAM.

associated with this process are inseparable from the development of market products and the increase of profit under the brand of "Artificial Intelligence" (AI). However, the past years have also witnessed the application of the new neural methods to various aspects of scientific research across the disciplinary spectrum, from natural to social sciences. Significantly, mathematics has not been the exception to this general trend. Indeed, since the mid-2010s, many efforts have been dedicated to treating numerous aspects of mathematical knowledge using DNNs of various kinds.

Mathematics has been in the sights of AI champions since the very emergence of the AI program. In his pioneering 1948 paper on "Intelligent Machinery," for instance, Turing (1948/2004) already considered mathematics a privileged field for the application of "thinking machines." And the celebrated 1955 Dartmouth project proposal included a supplement by Newel and Simon with a plan to extend to "learning, self-programming and the self-selection of theorems to be proved (innovation)" their work on "Mathematics Machines" that "discover proofs in the propositional calculus [...] by employing search procedures that appear at least grossly similar to those used by humans in dealing with the same problems" (Newell and Simon 1956). At any rate, considering the conception of intelligence motivating the AI program, one should not be surprised that mathematics falls within its scope.

However, given the nature of recent neural techniques and the unexpected results they can afford, as natural as it may be for the AI ideology to address mathematical tasks, the very possibility of their successful application to mathematical knowledge cannot but come as a surprise from a conceptual and philosophical perspective. Hence, faced with the results of the renewed AI momentum, it seems inevitable for well-established conceptual positions to return to their foundations and reopen old questions: Wasn't the whole idea of proof assistants and automated theorem provers to *prevent* the frailty of human proving practices instead of *reproducing* them and *relying* on them as a gold standard? Wasn't mathematical competence a *condition* to write and read scientific papers and mathematical expressions rather than their miraculous *effect*? Wasn't natural language the *cause of* rather than the *solution to* the multiple problems preventing mathematics from achieving higher degrees of precision? And more generally, didn't the *formal* nature of mathematics, finally conquered by the logical regulation of its symbolic means, make it impassive to the strong *empirical* position assumed by connectionist approaches guiding the application of DNNs?

In the current state of the art, it is not possible to determine who will have the last laugh. Yet, whatever the outcome, the rapid spread of deep learning techniques will likely have, for better or worse, a significant impact on future mathematical practices. For this reason, the current situation represents a most exciting moment in the eyes of the historian and philosopher of sciences – a moment where established certainties can and need to be regarded afresh to assess if they hold still in a context that suddenly is no longer the same. Hence, even as the symptom of a bewildered state, the philosopher can only rejoice at the reopening of those fundamental epistemological questions, hoping that a philosophical treatment of the latter can

contribute to critically assessing the stakes of the new situation and indicating novel orientations.

The present chapter intends to make a first step in this direction. As such, its primary purpose is to offer a survey of the many recent applications of neural machine learning techniques to different aspects of mathematical knowledge, calling attention to what can be of significance for a philosophy of mathematics. It will appear that, disregarding traditional boundaries in the mathematical landscape, the orientations in this emerging research area tend to be distributed following different mathematical *practices*. Yet, despite their apparent dispersion across those practices, all applications share, if only unwittingly, a common philosophical assumption. It is this assumption that the statements opening these pages announced. Namely, the idea that whatever we do when we do mathematics, we do it with *words*. More precisely, the analysis of different mathematical applications of DNNs shows that their potential success should force us to revise our conception of the relation mathematics maintains with language, and understand that and how *natural language, in its most immediate and unaffected expression, constitutes a decisive component of mathematical knowledge*.

Accordingly, the plan of the chapter is as follows. In Sect. 2, we provide a brief presentation of deep neural network models. Section 3 constitutes the principal contribution of this chapter, offering a survey of recent applications of DNN models to mathematical tasks. The section is organized into four parts, featuring four distinct lines of research identifiable in the field, according to the task assumed by researchers to characterize mathematical practice: proving theorems, manipulating objects, acquiring skills, or addressing open problems. Finally, in Sect. 4, we address the philosophical significance of those results, focusing on the relation between mathematics and natural language and the challenges for a theory of language where the semantic aspects of mathematical knowledge can be inferred from pure syntax. We conclude in a programmatic tone, indicating the perspectives open to a language-driven philosophy of mathematical practice.

2 What Are Deep Neural Nets?

A detailed presentation of DNNs falls outside the scope of these pages. Our account will thus remain deliberately simplistic, offering the elementary features needed to grasp the main characteristics of their application to mathematical tasks in the following section. Fortunately, owing to the growing popularity of AI applications in recent years, a wide range of resources are now readily accessible to the interested reader, featuring different levels of granularity and difficulty. In particular, for a systematic and detailed presentation, we refer the reader to Goodfellow et al. (2016) or Brunton and Kutz (2022). Also, Williamson (2023) does an excellent job at providing a rigorous yet brief and accessible introduction to neural nets intended for the uninformed working mathematician.

From a bird's-eye view, artificial neural networks are *parametric functions between two vector spaces*. Therefore, their application to any field relies on the

possibility of encoding some information as a vector (i.e., a list of numbers), feeding the latter into the function as an argument (i.e., the "model" 's input), and retrieving a target vector as output, to be decoded as some other kind of information. The nature of the information is assumed to be of little significance as long as it can be encoded as a vector and a vast set of examples of pairs of input and target vectors is available to train the model. Indeed, neural network architectures have been proven to be universal function approximators (Hornik et al. 1989). Therefore, given enough examples of input and target vectors, a neural net will approach a function transforming one into the other. The hope is that such an approximation will perform in some expected way for arguments outside the training data.[2] To encourage this "out-of-distribution generalization," neural models are typically trained on only a fraction of the available data (typically around 90%) and tested on the remaining held-out part.

In the most elementary cases, the internal structure of the model is composed of a series of transformations, each having a similar configuration: a linear transformation (i.e., a matrix), a bias (an added vector), and an activation (nonlinear) function. More precisely, a DNN can be described as a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ explicitly expressed as a composition:

$$f : \mathbb{R}^n \xrightarrow{f_1} \mathbb{R}^{n_1} \xrightarrow{f_2} \ldots \xrightarrow{f_K} \mathbb{R}^{n_K} \xrightarrow{g} \mathbb{R}^m \tag{1}$$

with

$$f_i(\mathbf{x}_{i-1}) = a(\mathbf{M}_i \mathbf{x}_{i-1} + b_i) \tag{2}$$

where $\mathbf{x}_i = f_i(\mathbf{x}_{i-1})$; $\mathbf{x}_0$ is the input vector; the $\mathbf{M}_i$ are $n_i \times n_{i-1}$ matrices representing the linear functions; the $b_i \in \mathbb{R}^{n_i}$ are the biases; the a is the nonlinear activation function; and g is the output function.

Each one of these complex transformations (i.e., each f_i) constitutes a *layer*. Hence, when a vector is fed into a layer, it is multiplied by the matrix, the bias vector is added, and the nonlinear function is applied element-wise. The *deep* character of neural nets is determined by the composition of many of these layers so that the output of one layer becomes the input of the next one until a final output vector is produced.

Finally, the model is *trained* by iteratively updating the components of the matrices and bias vectors (or *weights*) of the different layers. This is achieved by performing a gradient descent on a *loss function*, that is, on a function that measures the error between the vector output by the model and the target vector provided by

[2] This does not mean that any task is solvable by means of neural models, as a significant part of the AI community tends to think. Many problems, maybe most of them, do not accept being constructed as a predictive task based on a finite number of supposedly similar cases expressible as vectors. Problems in politics, law, art, and many other domains of social life cannot, by their nature, be framed in such terms without transforming, if not wholly obliterating that very nature.

the training set. Through an algorithm known as backpropagation, the weights of different layers are progressively adjusted so that the average error over the set of examples hopefully converges to some minimum value.

There exist many training strategies. In *supervised learning*, the target vectors are handcrafted to encode a property of the data encoded in the input vector. In *self-supervised learning* (sometimes mistaken for unsupervised learning), the target vector is some part of the input vector masked when the latter is fed to the model. Finally, in *reinforcement learning*, each output by the model receives a score through a carefully designed reward function, and the model's objective is to maximize that score.

As we said, the scheme presented here corresponds to the most elementary neural models. In recent years, DNNs have become increasingly complex and diversified, from convolutional neural nets (CNNs) and recurrent neural nets (RNNs, LSTMs) to transformers and diffusion models, to name only the most popular architectures. Those architectures remain particularly sophisticated cases of the basic model presented above. Due to their significance in the current state of the art, the case of *transformers* deserves special mention. Introduced by Vaswani et al. in 2017 and popularized by models like BERT (Devlin et al. 2018) and GPT-3 (Brown et al. 2020), the transformer architecture is characterized by the implementation of an *attention* mechanism. Simply put, the idea is to enhance each layer f_i of the net with a set of vectors capable of determining, for each component of the input vector $\mathbf{x}_{i-1}$, which other components of that vector are relevant to compute the correct output vector (i.e., which other components the component in question "attends" to). Like the linear components of the architecture, attention vectors are randomly initialized and jointly trained with the rest of the model. Among other technical advantages, the attention mechanism allowed the exploitation of parallelism during training, significantly reducing the training time and paving the way for a massive increase in scale, both of model size (parameters) and training data through self-supervised learning. Since their introduction, transformers have exhibited surprising capabilities across the most diverse tasks and domains and have been widely adopted, defining the state of the art at the moment of writing these pages.

3 How to Do Maths with Neural Nets

Applying DNN models to the processing of mathematical knowledge is not a straightforward and unambiguously defined task. As is usually the case with current AI applications across domains, craftsmanship in computational models is sometimes more prominent than proficiency in the objects those models are expected to analyze. Accordingly, the different DNN applications to mathematics are determined less by a refined elaboration on the nature of mathematical objects and knowledge than by the AI researcher's implicit assumptions as to what characterizes a mathematical task. Given the current exploratory state of a field that evolves at a remarkable speed, the answers to this unraised question are multiple, with new results and sometimes original perspectives released on a monthly basis. Any ambition to

provide an exhaustive account of the current state of the art is, therefore, doomed to fail from the start. Hence, we will content ourselves with presenting what we recognize as the main trends underlying these developments. Significantly, those trends are informed by their implicit understanding of what it is that we do when we do mathematics. From this perspective, it is possible to identify four main trends in the field, depending on whether the research is oriented toward finding *proofs*, manipulating *objects*, acquiring *skills*, or supporting *heuristics*. The following sections present each one of these orientations, providing some details of what we identify as their respective seminal papers, and briefly accounting for a few landmark results and developments.[3]

3.1 Proof-Oriented

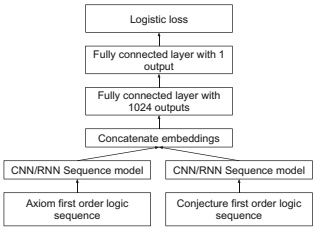
What can certainly be taken as the first significant attempt to apply by then well-established DNN models to the solution of mathematical tasks dates back to 2016, with the *DeepMaths* paper by Alemi et al. (2016). This work pioneered a prolific line of research oriented by the implicit assumption that theorems are the key to mathematical knowledge and theorem proving constitutes, therefore, the heart of mathematical practice. Accordingly, from this perspective, the main goal is to train DNN models to find *proofs*. Ideally, one would train a neural net on an extensive corpus of existing proofs, after which one could feed a yet unproved mathematical statement (a "conjecture") to the model, which would output a proof, if any. Notice that this is not the same as constructing a logical framework for automatic theorem proving because there is no definition of an explicit logical system in this case. One could imagine logical systems to be just a specific kind of function – even if a partial one – mapping the domain of statements onto that of proofs, and we are asking the neural model to approximate such a function.

Researchers within this trend tend to think that solving this task is practically sufficient to automate mathematical practice altogether. Indeed, in a paper significantly entitled *Towards the Automatic Mathematician*, Rabe and Szegedy (2021) affirm: "Ideally, we could simply talk to an automatic mathematician like a colleague, and it would be able to contribute to mathematical research, for example by publishing papers without human support." While the authors vaguely mention other tasks such as "formulate and explore its [the automatic mathematician's] own theories and conjectures," theorem proving is acknowledged as "an important instrument of our plan" and, indeed, the only strictly mathematical task considered in their manifesto.

[3] The reader should keep in mind that, despite our efforts to cover as much of the field as possible, the following constitutes only a partial view on the state of the art, biased toward works exhibiting philosophical interest. A more precise perspective can be obtained by inspecting the "Related Work" section usually included in the papers we will present. Lu et al. (2023) also provide a useful systematic survey of DNN models for mathematical reasoning.

In practice, however, things are more subtle. Actual research in this direction has focused chiefly on specific subtasks identified within the framework of existing interactive or automated theorem provers (ITP/ATP), such as E, HOL Light, Mizar, Metamath, Isabelle, or Lean. The focus is mainly put on those tasks that represent computational bottlenecks. In particular, there has been intensive research on *premise selection*, namely, "the selection of a limited number of most relevant facts for proving a new conjecture" (Alemi et al. 2016, p. 1). This task was the one motivating Alemi et al.'s (2016) pioneering work, where it was implemented as establishing the pairwise relevance between conjecture and axioms by learning to predict the usefulness of given axioms in the proof of given conjectures. The overall strategy consists in feeding both a conjecture and an axiom into the model, processing the input vectors independently through different network architectures producing intermediate vector representations of various kinds, concatenating those vectors, passing them to a fully connected DNN, and outputting a probability measure of the usefulness of the axiom for proving the conjecture (see Fig. 1). The dataset used was the Mizar Mathematical Library, a library of almost 60K theorems proved in Mizar, together with unnamed top-level lemmas. The input axioms and conjectures were then the strings (i.e., sequences of digital characters) representing the first-order logic formalization extracted from the Mizar Library (see Fig. 1 for an example). The authors tested two ways of feeding these texts into the model: character by character and word by word (token by token). In the latter case, tokens recognized as identifiers were mapped onto their already processed definition. Other than that, the expressions were fed sequentially with no parsing information. For each conjecture, the model was trained to produce a probability distribution over the set of possible premises. To evaluate the model, the authors assessed the number of conjectures that could be proved from the top-k premises in the corresponding probability distributions (for different values of k) using the E automated prover. They also evaluated the ranking quality, measuring if relevant premises appeared at the top of the rank. Such measures were then used to compare the different DNN architectures.

The results obtained by the best of those early models exhibited a slight increase in performance compared to existing benchmarks. Although modest, the outcome was far from wholly unreasonable, showing neural models could compete with more




```
:: t99_jordan: Jordan curve theorem in Mizar
for C being Simple_closed_curve holds C is Jordan;

:: Translation to first order logic
fof(t99_jordan, axiom, (![A] : (
(v1_topreal2(A) & m1_subset_1(A,
k1_zfmisc_1(u1_struct_0(k15_euclid(2))))) =>
v1_jordan1(A)) ) ).
```

Fig. 1 Network structure (left) and input example (right) for the premise selection task in *DeepMath* (Alemi et al. 2016)

classically engineered techniques and handle mathematical proofs to some extent. This work sparked then a series of other studies in the same direction, where the integration of neural techniques within ITPs and ATPs can be seen to increase. For example, Loos et al. (2017) provided evidence that integrating neural techniques for premise selection in the proof search procedure of the E ATP results in an increased amount of theorems proved under time and memory constraints. On the training side, Kaliszyk et al. (2017) proposed a new dataset for higher-order theorem proving, comprising over 2 million training examples of proof steps extracted from almost 11,400 proofs formalized by humans in the HOL Light ITP. The authors also provide a series of benchmarks for different machine learning models over that dataset. HOL Light was also used by Bansal et al. (2019) to propose a different training strategy, adding to the usual supervised learning on human proofs a reinforcement learning step. Instead of being trained to predict the correct premise, the model interacts as an "artificial agent" with the ITP, proposing both a tactic to be applied and a rank over the set of possible premises. The success or failure of the ITP to apply the received tactics on the premises is then used as a reinforcement signal to train the model to predict pairs of tactics and premises.

With the evolution of this line of work, researchers started considering a wider range of theorem proving subtasks. Kaliszyk et al. (2017), for example, mention, among others: "Predicting whether a statement is useful in the proof of a given conjecture," "Predicting whether a statement is an important one (human named)," "Predicting the name given to a statement," or "Generating the conjecture the current proof will lead to" (p. 4). Bansal et al. (2019) add some others to that list, such as predicting the same tactics as the human proofs or predicting goals and subgoals. Also, *tactics prediction* has gained increasing attention due to the significance of ITPs and ATPs in the concrete developments of this research orientation (Bansal et al. 2019; Paliwal et al. 2019; Lee et al. 2020). More generally, recent years have seen a proliferation of datasets, models, and benchmarks, which contribute to the expansion of this perspective.

The arrival of the *transformers* architecture (Vaswani et al. 2017) brought about new developments in neural theorem proving. As we saw, transformers are designed to deal with sequential data in a self-supervised and highly scalable way. Their success in the field of AI is inseparable from the surprising capabilities exhibited in the processing of natural language, motivating the emergence of so-called *Large Language Models* (LLMs). We will have the opportunity to say more about LLMs in the *Skill-Oriented* section below. For now, it is sufficient to note that, encouraged by their linguistic capabilities, researchers have explored multiple ways to utilize transformer-based models for neural theorem proving.

A noteworthy case is *GPT-f* (Polu and Sutskever 2020), an automated prover and proof assistant for the Metamath language, which was capable of significantly outperforming the state of the art in neural theorem proving at the moment of its publication. *GPT-f* is based on a model architecture similar to GPT-2 and GPT-3, the celebrated LLMs by OpenAI. It is designed to perform backward proof search, starting from a root goal and exploring tactics possibly leading to such a goal. The process may involve identifying other subgoals for which new tactics are explored

iteratively (with a prespecified bound on the iterations) while keeping track of the tree structure. The neural model is therefore trained to predict a "proofstep" (i.e., a tactic and the corresponding subgoals) for a given goal, both encoded as strings, which are then passed on to the Metamath kernel for formal verification. The use of transformers has the advantage of unifying the neural architecture, as opposed to using complex combinations of different architectures as in previous models. It also allowed for pretraining as a language model over various kinds of corpora (cf. LLMs in the *Skill-Oriented* section), including the mathematical fragments of *arXiv* and *StackExchange* as well as computer code from *GitHub*. Other than achieving a new state of the art for Metmath (56.22% closed proofs in the test set vs. 21.16% for the best existing models at the time), the authors could show that, within this framework, both language modeling pretraining (especially on mathematical corpora) and increased model size result in improved performance. Interestingly, *GPT-f* could generate 23 shortened proofs of existing statements in the Metamath library, reported as enthusiastically received by members of the Metamath community. For the authors, this was "the first effective contribution of a deep learning system to a formal mathematics library" (Polu and Sutskever 2020, p. 14).

One of the drawbacks of transformer architectures is that their high performance is conditioned to training over a massive amount of data, orders of magnitude larger than existing hand-crafted mathematical libraries. A common strategy has been to augment the latter with synthetic data generated in various ways out of known mathematical properties and algorithms (cf., for instance, Wang and Deng 2020). However, Polu and Sutskever (2020) showed that scaling the portion of synthetic data up to 5% of the training data actually hurt the performance of their model. While the authors did not inquire into the reasons for this empirical behavior, they proposed an alternative way to circumvent the scarcity of proof data by training a "value function" helping to guide the proof search more efficiently. Essentially, such a function predicts the probability that a subgoal is actually used in the final proof. The function is trained on a synthetic dataset produced by sampling proofs *from the trained neural model itself*, keeping *only those that the formal verifier can prove*. Their experiments show that the iterative use of this procedure, also known as *expert iteration,* leads to significantly better results, suggesting a strategy of "continuous self improvement." Following this line of research, Lample et al. (2022) showed that replacing this iterative process with an *online training* over proofs generated by the model and verified by an ITP results in substantial performance gains.

In recent years, the success of DNN models – and transformers in particular – in the processing of natural language motivated what can be conceived as a subtle but notable reorientation of the research efforts within proof-oriented approaches toward what came to be known as *autoformalization*. In a highly programmatic paper, Szegedy (2020) defines an autoformalization system as "an automated system that is capable of automatically formalizing significant portions of mathematics from a natural language input and verifying it automatically" (p. 4). From this perspective, formalizing essentially means expressing statements as computer programs susceptible to be verified computationally following transformation rules. Implicit assumptions concerning the nature of mathematics within this body of work become here

explicit: "Mathematical reasoning is just reasoning about anything specified formally" (p. 4). Yet, this recent turn pushes these assumptions further and presents autoformalization as a natural step toward AI models endowed with "general reasoning" capabilities. In this way, mathematical practice, which had already been reduced to formalizing practices, is spontaneously raised to the level of "the discipline of pure reasoning": "Mathematical reasoning is not about mathematics per se, it is about reasoning in general" (p. 3). The distance between thought and mathematical thought becomes thus so misty it is then "natural to ask: Will we ever arrive at the point where an AI agent can learn to do reasoning as well as the best humans in the world in most established domains of mathematics" (p. 4).

However, all these prejudices about mathematics, reasoning in general, and the best humans in the world should not divert our attention away from an essential fact: This somewhat unexpected emergence of informal mathematics through natural language at the heart of the most extreme attempts to fully automate theorem proving looks more like an admission of weakness than a *tour de force*. Indeed, when one examines the recent evolution of DNN models, it is easy to identify a subtle glide away from highly measurable functions toward more informal tasks, such as text and image generation. For the latter, unforeseen capabilities can be achieved across multiple fields, including mathematics and programming. Yet, the counterpart of that originality is a loss of robustness and reliability, all the more fundamental that failure is intrinsically more difficult to assess in informal settings (how can a poem or a picture be "incorrect"?). The very nature of this success, *informal by design*, together with the relative scarcity of formal proof corpora, left AI researchers little choice other than focusing on bridging a gap secretly undesirable for most of them between formal methods on one side and neural linguistic models on the other.

Besides explicit or implicit motivations, recent explorations of natural language expressions of mathematical statements such as Ferreira and Freitas (2020, 2021) and Welleck et al. (2021), and autoformalization tasks in particular, are of interest in their own right. A noteworthy work concerning the latter is that of Wu et al. (2022), exploiting the "few-shot learning capabilities of LLMs" to produce translations into Isabelle code of natural language mathematical statements and problems contained in the MATH (Hendrycks et al. 2021b) and MiniF2F (Zheng et al. 2022) mathematical competition datasets. The relative success of this approach is nonetheless remarkable considering the negligible amount of aligned pairs of informal-formal expressions expected to be present in the vast training corpus of the LLMs. Autoformalization faces, however, the challenge of articulating the high-level structure of informal proofs with the low-level details required for automated proving. To address this difficulty, Jiang et al. (2023) proposed a method where an initial informal statement is associated with an informal proof *draft* (either through available data or through automatic generation by an LLM), which is, in turn, broken down by an LLM into a formal proof *sketch*, i.e., a high-level structure of formal statements, to be then passed on to an ATP. The experiments on the miniF2F dataset suggested that this strategy is indeed capable of successfully guiding the ATP and boosting its proving capabilities, with comparable results for the cases where the informal drafts were produced by humans or LLMs.

Considering the development of this research orientation from its inception up to the current state in 2023, one thing seems clear: Despite a declared inclination toward end-to-end black-boxed models inherited from AI perspectives and coated with multiple ideals of intelligent machines and dispensable humans, DNN applications to theorem proving have actually become increasingly reliant on existing (nonneural) interactive or automated theorem provers. To the extent that, as of today, it is hard to imagine how the latter could be dispensed with. And yet, against all odds, neural techniques have proved to contribute to the performance of formal theorem proving within those frameworks. So much so that a hybrid strategy, integrating neural techniques and I/ATPs, appears today as a promising research program for the years to come. In this sense, Jiang et al. (2022) proposed a framework for neural models to interact with any hammer-enabled ITPs, showing that neither the neural model nor the ITP alone can match the performance achieved by their careful integration. However modest and far from its promises the contribution of neural nets may be, the latter have at least succeeded in making such contributions through means that were entirely out of the radar of traditional formal methods. From a philosophical standpoint, the leveraging of informal expressions resulting from various mathematical practices, and the reliance on natural language in particular, are distinctive features of this approach, to which we will have the chance to return. For now, let us say that, despite the significant challenges and the ideological noise surrounding AI applications to theorem proving, DNNs have shown that they can do *something* with mathematical proofs.

3.2 Object-Oriented

A second research orientation for applying DNN methods to mathematics shares a similar use of the learning models with the first. However, the focus is put on manipulating mathematical *objects* rather than statements. Admittedly, proofs can also be understood as mathematical objects, in which case proof-oriented approaches could also fall under this category. At any rate, as learning models, DNNs are, in principle, indifferent to this distinction. Yet, both from a philosophical and a formal standpoint, there is arguably a significant difference between, say, manipulating numbers to perform arithmetical operations and proving a general arithmetical property for a possibly infinite set of numbers.

In line with the neural models' general setting, the goal in object-oriented mathematical applications is to feed the expression of a mathematical object (e.g., equation, terms of a series, etc.) into the model and expect the latter to output some specific result or property associated with it (e.g., solution, recurrence relation, etc.).

The most elementary case of this strategy is given by arithmetical operations and their corresponding results. Early assessments of deep neural models already showed that recurrent architectures (RNNs) could be trained to perform integer addition and multiplication based on binary representations with strong generalization capacities (Łukasz Kaiser and Sutskever 2016). Those results were extended to include decimal multiplication through a binary encoding of each decimal digit (Freivalds and Liepins

2017). Notice that, in all cases, from the model's perspective, input numbers are simply *tokens*, so there is no reason a priori that their numerical content is preserved when represented as the usual sequence of digits, hence the binary representation preferred by these early approaches. Indeed, Trask et al. (2018) showed that other numerical representations exhibit serious difficulties generalizing outside the range of values seen in training and proposed to enhance neural models with a specific architecture designed to incorporate an inductive bias for numerical computation.

More recently, researchers have explored the capabilities of DNNs to perform numerical calculations on more complex mathematical objects. A landmark in this sense is the work by Charton (2021), training transformers to compute solutions to linear algebra problems, including transposition, addition, multiplication, and inversion of matrices, as well as eigenvalues, eigenvectors, and singular value decomposition (SVD). His results showed the models can indeed compute approximate solutions with high accuracy, generalizing out of their training distribution under certain conditions and even extrapolating to larger numbers. Unlike most of the work presented in the previous section, models were here trained on synthetic corpora specially tailored for the intended tasks. This work is emblematic of an object-oriented approach; therefore, it is worth examining it more closely.

The input and output in Charton (2021) are, then, matrices of real numbers. However, their representation was carefully designed to fit the kind of data on which transformers are known to exhibit high performance, namely, sequential data. Matrices were thus encoded as sequences of coefficients prefixed with the matrix dimensions. Remember that each term in a sequence is simply a token. For this reason, numerical coefficients were also encoded following very precise rules. Based on the general principle that real numbers can be represented by a triplet of a sign, a mantissa, and an exponent, four alternative encodings were proposed for that purpose: base 10 and 1000 encoding of the mantissa, signed base 1000, and floating point (see Table 1 for Charton's (2021) original examples). Datasets were produced by online sampling dense random matrices, from 5×5 up to 15×15 dimensions, with coefficients sampled uniformly within the interval (–10, 10). The data was used to train a transformer for the various tasks studied. For all tasks, an output sequence of tokens was considered a correct solution if it could be decoded as a matrix and represented a correct solution within a predetermined numerical tolerance.

Under those conditions, the study showed that DNNs can perform matrix transposition – which does only involve permutations of tokens in the sequence, without arithmetic operations – with practically perfect accuracy. This is also the case with

Table 1 Examples of different encodings for numerical coefficients in Charton (2021)

Encoding	3.14	$-6.02.10^{21}$	Tokens/ coefficient	Size of vocabulary
P10	`[+, 3, 1, 4, E-2]`	`[-, 6, 0, 2, E21]`	5	210
P1000	`[+, 314, E-2]`	`[-, 602, E21]`	3	1100
B1999	`[314, E-2]`	`[-602, E21]`	2	2000
FP15	`[FP314/-2]`	`[FP-602/21]`	1	30000

matrix addition up to 10 × 10 dimensions and 1% tolerance, with a slight decrease in performance for matrices of larger sizes. The model's accuracy dropped noticeably when trained over matrices of different sizes. Although more complex, matrix multiplication presented very similar results to addition, but only for the first two numerical encodings. The study also exhibited surprisingly high performance in the much more complex task of eigenvalue computation. The correct prediction of eigenvectors remained restricted to matrices of small sizes. However, the accuracy increases with the numerical tolerance, indicating that *errors are not random but somewhat akin to rough approximations*. Matrix inversion proved to be hard for DNNs, with no more than 80% accuracy at 5% tolerance. Finally, while achieving high performance on 4 × 4 matrices, SVD could not be predicted for matrices of 5 × 5 or more dimensions.

The fact that the properties of the objects under analysis are well-known contributed to providing many principles of interpretability for these results, both for the success and the failure cases (Charton 2021, 2022). In this way, Charton could show that, for eigenvectors, even when the predictions were incorrect, the model correctly learned some underlying properties, such as their unit norm and their orthogonality. Also, a fine-grain analysis of the metric used in matrix inversion could show that failures were to be attributed to the computation of matrix inverses rather than the neural model itself. The study also provided evidence against the possibility that the results depend on the model simply memorizing the data and could point to strong predictors of success or failure for some of the most challenging tasks. Overall, this work showed that the success of DNNs in performing numerical calculations on matrix operations is highly dependent on the encoding (the best being P1000 and FP15), with greater difficulty for increasing matrix size and the capacity of generalizing out of the training distribution if the latter is carefully chosen.

The solution of partial differential equations through DNN techniques has also been the object of much research effort. Blechschmidt and Ernst (2021) provide a survey of approaches to numerical solutions. However, Lample and Charton (2020) introduced an original perspective, testing the DNNs' capabilities to solve differential equations and integration *symbolically*. The overall strategy of this seminal paper was to feed a function or a first- or second-order differential equation as input and require the model to output its integral or its solution, respectively. Here, too, we find that the training corpora were carefully designed for the intended tasks. In particular, mathematical expressions of functions and differential equations were represented as binary trees, where the internal nodes represent operators, while the leaves can represent either numbers, constants, or variables. The trees were then sequentialized through prefix (i.e., Polish) notation by enumerating nodes in prefix order and then encoded as a vector to be fed into the model (see Fig. 2 for an example). As the authors observe, using trees to represent expressions also provides a principled way to tackle the otherwise thorny problem of sampling from an expression space. Training and test datasets were carefully constructed using Mathematica to produce pairs of equations and solutions, which were then turned into the chosen representation and carefully cleaned up to eliminate any possible source of noise. The model used was an early version of a transformer, considered by the authors as a sequential model aimed at natural language translation.

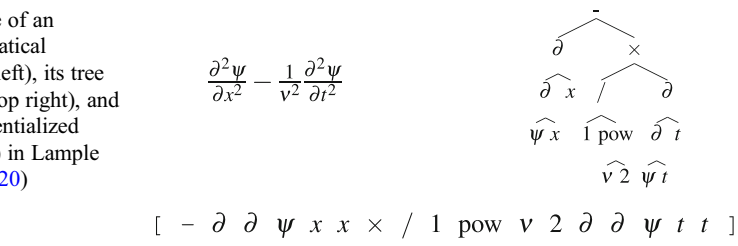


Fig. 2 Example of an original mathematical expression (top left), its tree representation (top right), and the latter's sequentialized version (bottom) in Lample and Charton (2020)

Relying on all this careful handcrafting of the mathematical object's representations, the authors trained models on five tasks, each one relying on a dedicated synthetic dataset: integration of randomly generated functions solvable by a computer algebra system; integration of derivatives of randomly generated functions; integration of the composition of one random function and the derivative of another random function (i.e., integration by parts); first-order differential equations; and second-order differential equations. Under the precise conditions established by the authors, the corresponding models exhibited surprisingly high scores across those tasks, with close to 100% accuracy for the different kinds of integration and around 80% and 40% for ODEs of orders 1 and 2, respectively. The latter could be brought up to around 94% and 73% when considering if at least 1 of the top 10 hypotheses output by the model was correct (and 97% and 81% for the top 50). These results outperformed those of computer algebraic systems such as Mathematica (under runtime constraints), Matlab, and Maple. Generalization outside the training distribution proved challenging. However, the authors showed that the reason was a systematic bias toward specific lengths for the expressions of problems and solutions in different datasets. Training on mixed datasets substantially increased the performance on problems outside the training distribution.

While neural frameworks had been shown to successfully process symbolic properties of mathematical propositions and numerical properties of mathematical objects – at least to some extent – the relevance of Lample and Charton's (2020) work lies in the capacity of bridging the gap between both orientations by operating on objects through their symbolic expression. The application of DNNs to *symbolic regression* provides another significant confirmation of this potential. Symbolic regression is the task of finding the symbolic expression of a function given some of its values. Before the emergence of current DNN models, symbolic regression methods relied on evolutionary algorithms. Early applications of deep neural techniques to this task include Martius and Lampert (2016); Udrescu and Tegmark (2019); Kim et al. (2019); and Petersen et al. (2021), where different kinds of DNNs are used to improve the exploration of the space of possible mathematical expressions, in search of the one that best fits the set of given values. Martius and Lampert (2016), for instance, propose a model where, activation functions correspond to symbolic operators acting on different dimensions of each layer, in such a way that, after training the model through backpropagation, the symbolic expression can be read out of the learned coefficients. Petersen et al. (2021) adopt a different

strategy, where syntactic properties of symbolic expressions are encoded as constraints within a generative recurrent architecture, which is then trained through reinforcement learning. In all these cases, training takes place within one set of values (i.e., corresponding to one function) so the learning procedure only concerns the selection of the symbolic expression for those values. More recently, an alternative approach has been proposed, more in line with current trends in AI: The model is trained on pairs of sets of values and corresponding symbolic expressions, learning to predict the latter from the former (Biggio et al. 2021; Valipour et al. 2021; d'Ascoli et al. 2022; Kamienny et al. 2022). From this perspective, *the symbolic expression appears as yet another property of a mathematical object* (e.g., a function) initially given through numerical expressions. Significantly, symbolic regression thus becomes a special case of language modeling, as Valipour et al. (2021) point out: "Symbolic mathematics behaves as a language in its own right, with well-formed mathematical expressions treated as valid 'sentences' in this language. It is natural, therefore, to consider using deep language models to address tasks involving symbolic mathematics" (p. 2).

A notable case of this recent orientation is that of d'Ascoli et al. (2022), where a neural model is trained to discover the recurrence relation holding between the first terms of a numerical sequence. The data is generated by sampling symbolic expressions of functions to produce a recurrence relation and compute the first 25 terms of a sequence. Symbolic expressions were encoded as sequentialized trees, following Lample and Charton (2020), while numbers (integers and floats) were encoded as in Charton (2021) (cf. above). A simple transformer architecture is then trained to predict a symbolic expression. Since many equivalent expressions could correspond to the initial function, a prediction is considered correct if the difference between the n following terms in the sequence computed with the predicted expression and those calculated with the initial expression is smaller than a given tolerance error. The authors showed their model could achieve an in-distribution accuracy of over 90% for integer sequences and over 70% for floats for the prediction of the following five terms (falling to 78% and 43%, respectively, for ten terms), outperforming numeric methods. The model was also evaluated on the Online Encyclopedia of Integer Sequences (OEIS, Sloane (2007)) for out-of-domain generalization, where, even if performance dropped substantially, it showed some significant results and outperformed existing methods for this task in Mathematica.

Overall, despite many limitations and somewhat modest results often concerning relatively simple and well-known mathematical objects, this line of work has shown that the generic approach provided by DNNs can meaningfully manipulate, to a significant extent, both numeric and symbolic properties of mathematical objects of different kinds (arithmetical, algebraic, analytic, etc.). Admittedly, the relative simplicity of the objects under study, whose properties are in most cases well-known, is such that these results cannot be expected to replace or even compete with existing methods and practices (at least not anytime soon). But as Charton (2021) makes it clear, this is not the purpose of this work either. Its significance lies elsewhere, namely, in showing that those well-known practices are not entirely orthogonal to the seemingly unrelated principles of neural nets. At any rate, this line of research

demonstrates that if DNNs can do something with mathematical proofs, the same can be said for mathematical objects.

3.3 Skill-Oriented

A third group of works applying neural methods to mathematical knowledge follows a different strategy than the previous two. The idea is not so much to solve specific tasks involving mathematical statements (e.g., premise selection) or mathematical objects (e.g., symbolic integration) but to endow the neural model with general mathematical *skills* so that, once trained, it becomes capable of spontaneously solving mathematically related problems.

Stated in those abstract terms, one could wonder how such a goal could have any chance of success. However, hope in this research orientation is motivated by the turning point operated by transformer architectures in the field of natural language processing (NLP) and AI more generally. The original transformer model introduced by Vaswani et al. (2017) was conceived to improve machine translation performance while significantly reducing the training time. Its superiority compared to previous models had been demonstrated in practically all NLP tasks. Moreover, when trained as a *language model*, that is, as a model producing probability distributions over words or tokens given other tokens, Saxton et al. (2019) showed they could significantly outperform the best recurrent models on a variety of mathematical tasks (including algebra, arithmetic, calculus, and probability, among others). To that end, transformers are trained on synthetic corpora where mathematical tasks are expressed in terms of natural language question answering, processed as a sequence of characters, e.g., (p. 2):

Question: `Let x(g) = 9 * g + 1. Let q(c) = 2 * c + 1. Let f(i) = 3 * i 39. Let w(j) = q(x(j)). Calculate f(w(a))`
Answer: `54 * a - 30`

However, the true power of transformers was revealed with the emergence of what came to be known as *Large Language Models* (LLMs). Taking advantage of the scaling capabilities of transformers, primarily due to their exploitation of parallelization, the distinctive feature of LLMs lies in their training strategy. Popularized thanks to the BERT model (Devlin et al. 2018), the idea is to train, or "*pre-train*" as it is referred to in the field, a transformer architecture in a highly generic self-supervised fashion, namely, by learning to predict random masked words in natural language sentences over a gigantic amount of natural language text (typically scraped from all over the Internet). Despite its apparent simplicity, working in an autoregressive generative fashion,[4] such training endows the model with the general capacity to perform a large variety of natural language tasks in the form of a text to

[4] That is, iteratively generating text one word at a time from a sequence of words initially given, recursively augmented by the words produced by the model.

Table 2 Example texts of the three tasks examined in Shen et al. (2021)

Task	Text
Knowledge component prediction	Simplify the expression: (z2)2 Put parentheses around the power if next to coefficient, for example, $3x2=3(x^2)$,$x5=x^5$
Question answer scoring	Q: Explain your answer on the box below A: Because it is the same shape, just larger, making it similar
Knowledge tracing correctness prediction	Q: What is $2.6 + (-10.9)$? A: -8.3

be generated as the continuation of a given prompt. Once pretrained, the model can be further trained or "*fine-tuned*" on a significantly smaller domain-specific dataset, to perform tasks expressible in natural language and normally requiring domain-specific skills. A variant of this "transfer learning" strategy was introduced by the celebrated GPT-3 model (Brown et al. 2020), aiming to circumvent the need for fine-tuning by drastically scaling up both the model size and the training data.[5] This change in scale afforded generative pretrained transformer models a unique feature, known as "in context-learning," namely, the capacity to perform domain-specific tasks without any further training, just by prompting the model with a few examples or a short description of the intended task. Thus, once the model has been pretrained, one can input a short text with a few examples of, say, pairs of grammatically incorrect/correct English sentences before yet another incorrect sentence, after which the model will continue the text autoregressively, outputting the correct counterpart. Surprisingly, this highly unexpected behavior holds with reasonable accuracy for a remarkably wide range of tasks.[6]

Within this line of research, it became natural to prompt the generic model with tasks requiring mathematical skills, either after fine-tuning the model on mathematical data or in an in-context fashion. Shen et al. (2021), for instance, evaluate the performance of the original pretrained BERT model when fine-tuned for three general tasks associated with mathematics education: large-scale knowledge component prediction, open-ended question answer scoring or autograding, and knowledge tracing correctness prediction (see Table 2 for examples). Their work shows that such a model achieves a performance of almost 92%, 89%, and 87% for the three tasks, respectively (the first and third measured as accuracy, the second as area under the curve). Moreover, the authors could improve this baseline by adding a pretraining step on a dedicated mathematical corpus between the original pretraining and the fine-tuning. The mathematical corpus included paper abstracts from arXiv,

[5] Whereas BERT contained a maximum of 340 million parameters in its original version, and at the time of GPT-3's release, the largest model contained 17 billion, GPT-3 was composed of more than ten times this number: 175 billion parameters. As for the data, GPT-3 was trained on 500 billion words (tokens) compared to 3300 million for BERT.

[6] For a comprehensive account of LLMs, see Bommasani et al. (2021). For a precise yet accessible presentation of the mechanisms behind GPT, see, for instance, Wolfram (2023).

college MOOC syllabus, and textbooks and curricula ranging from prekindergarten to college graduate level.

As for in-context learning, the original GPT-3 paper by Brown et al. (2020) already showed the arithmetical skills of such a generic model, in particular, by trying to perform up to five-digit elementary arithmetic. The model is thus prompted with a few cases (typically a dozen) of arithmetic operations expressed in natural language, such as:

```
Q: What is 24 times 42?
A: 1008
```

ending with a line that stops at `A:`, after which the model is expected to fill in the subsequent text, hopefully with the correct answer to the last question. GPT-3 exhibited an accuracy of almost 100% on two-digit addition and subtraction and over 94% and 80% for three-digit subtraction and addition, respectively. These results are certainly surprising considering the model was never trained to perform such a task, but only to predict the next token given previous tokens. In addition, the authors provided evidence that only a negligible number of arithmetic problems tested could be found in the training data, suggesting that the model's performance cannot be attributed to simple data memorization. Nogueira et al. (2020) later showed that the capacity of LLMs to perform elementary arithmetical operations was highly dependent on the textual representation of numbers. By fine-tuning a pretrained T5 model (Raffel et al. 2019) on textual expressions like the ones used in GPT-3, the authors demonstrated that, while single-token decimal notation performed very poorly, representing numbers as digits separated by powers of 10, either using scientific notation or not (e.g., `3 10e1 2 10e0` or `3 10 2` for 32), could afford the model almost perfect accuracy for the addition of numbers up to 15 digits. Their study also confirmed the importance of pretraining on generic masked language modeling: A transformer trained from scratch needs ten times more training examples to reach the same performance on addition as a pretrained model.

A noteworthy characteristic of this approach is that *any mathematical problem is treated by expressing it in natural language* (see, for instance, Griffith and Kalita 2021; Meng and Rumshisky 2019; Cobbe et al. 2021). Yet, as surprising as those elementary results may be from a conceptual standpoint, tasks requiring mathematical skills have proved challenging for LLMs generically trained on natural language data when compared to other (mostly linguistic) tasks, such as machine translation, summarization, or sentiment analysis (see, for instance, Hendrycks et al. 2021a; Davis 2023). As suggested above, a possible strategy for the skill-oriented line of research to remain relevant despite this difficulty is to further train pretrained LLMs on large amounts of mathematical (or mathematically related) texts. This further training remains generic, in the sense of not being oriented toward any specific task: The training continues to take place as language modeling. A good example of this approach is provided by the Minerva model (Lewkowycz et al. 2022), where the PaLM LLM (Chowdhery et al. 2022) is further trained on a vast

mathematical and scientific dataset built from the arXiv preprint server and mathematical webpages scraped from the Internet. The objective is to be able to perform both numerical calculations and symbolic manipulation without training on specific tasks or relying on principles foreign to the neural architecture (like a calculator or a compiler). To improve the performance, the model also implements a series of techniques, including prompting methods that encourage detailed answers and majority voting on samples of possible results to a question. A distinctive feature of Minerva is that, unlike typical language modeling training, mathematical expressions are not normalized to keep only the natural language text but include the syntax of mathematical typesetting systems such as LaTeX and MathJax. Thus, an expression like $E = mc^2$ is represented as the string `$E=mc^2$` instead of `Emc2`. In this way, the model input and output can include a more precise representation of mathematical expressions. With all these tweaks and tricks, Minerva was able to significantly outperform previous models on different mathematical benchmarks, including high school, college, and graduate math problems and STEM-focused language understanding problems, with gains of up to 30% in some cases. Figure 3 shows an example from the MATH dataset correctly solved by Minerva. The model exhibited an accuracy of over 50% on this dataset, compared to less than 20% for existing neural models at the moment of publication.

The challenges faced by a skill-oriented approach are more than just technical. Among others, the notion of "mathematical skill" is difficult to characterize. The field has progressively converged to an implicit agreement assuming an institution-centered answer to this question: mathematical skills are the ones measured by formal education tests, exams, and competitions. This choice is reflected in the benchmarks established by the field, which are essentially composed of textbooks and curricula from different educational levels, national (which often means: American) admission tests, like the Scholastic Assessment Test (SAT), and institutional competitions such as the International Mathematics Olympiad (IMO). But such a choice is as uncritical as it is unreflective. Historical, geographical, and social biases

Question: A line parallel to $y = 4x + 6$ passes through $(5, 10)$. What is the y-coordinate of the point where this line crosses the y-axis?

Target: Since the line is parallel to $y = 4x + 6$, we know that its slope is 4. Thus the equation of the line is $y = 4x + b$ where b is the y-intercept. Plugging $(5, 10)$ into this gives $10 = 4 \cdot 5 + b \Rightarrow b = \boxed{-10}$ which is what we wanted.

Model output: The line is parallel to $y = 4x + 6$, which means that it has the same slope as $y = 4x + 6$. Since the line passes through $(5, 10)$, we can use the point-slope form of a line to find the equation of the line:

$$\begin{aligned} y - 10 &= 4(x - 5) \\ y - 10 &= 4x - 20 \\ y &= 4x - 10 \end{aligned}$$

Now that we have the equation of the line, we can find the y-coordinate of the point where the line crosses the y-axis by substituting $x = 0$ into the equation:

$$y = 4 \cdot 0 - 10 = \boxed{-10}.$$

Fig. 3 Example of a problem from the MATH dataset, correctly solved by Minerva (Lewkowycz et al. 2022). The solution is output in LaTeX, token by token

remain unassessed. Also, the field's research habits leave no room to address the consequences of reinforcing a school-like conception of mathematics, which is fundamentally focused on performance due in part to the uninterpretable characteristics of the proposed models. Let alone the encouragement of an anthropomorphic image of neural models resulting from the idea that they can acquire skills and "beat" humans on tests designed for specific human purposes. Yet, if we refrain from taking these models at their face value and admit from a critical stance that the notion of mathematical skill is not independent of how educational institutions measure it at a given space and time, the performance exhibited by these models appears no less surprising. Compared to the previous orientations, the lack of dedicated model architectures (like in proof-oriented cases) or datasets (as in object-oriented approaches), as well as the highly unstructured and underspecified goals and means characterizing the skill-oriented perspective, endows those results with enough room for philosophical inquiry.

3.4 Heuristic-Oriented

The fourth and final research orientation differs from the previous ones in that neural models are not conceived as generic devices performing multiple tasks in a way supposedly akin to human agents. Instead, the idea is to use neural nets as *tools* addressing specific aspects of advanced open problems in mathematics in order to guide the intuition of mathematicians at work. From a philosophical perspective, this approach can be deemed one of the most sensible and fruitful, if only because it grants a central place to working mathematicians, whose judgment is arguably indispensable when mathematical practice is not reduced to reproducing existing mathematical knowledge. However, this orientation is the most underdeveloped in the current state of the field. Many factors may explain this situation, starting with ideological reasons, related to a widespread belief in the field viewing AI systems as agents imitating, competing with, and ultimately replacing human agents. Such a position is not independent of economic conditions in a field that is largely driven by corporations. In its current state, AI research is not separable from profit-seeking through the development of market products, and the market size of advanced mathematical research is negligible compared to the one defined by the automation of known mathematical tasks in business and education. Finally, from a sociological perspective, the development of this orientation requires the close collaboration of two very different research communities, namely, professional mathematicians (often in the area of pure mathematics) and applied computer scientists and engineers, which represents an additional challenge to the success of the attempts in this direction.

As a result of their heuristic nature, works belonging to this category are less characterized by unified methods, frameworks, and goals than by the individual adaptation of new machine learning tools to specific and highly technical mathematical problems. Accordingly, this orientation might be better presented through concrete cases rather than general trends. We will present two such cases, which

represent two important aspects of mathematical heuristics: forming *conjectures* and searching for *counterexamples*.

Early uses of deep learning for conjecture generation can be found in Hughes (2020); Carifio et al. (2017); Levitt et al. (2019); Jejjala et al. (2019); and Heal et al. (2020). Most of this work concerned combinatorial properties in knot theory. But it was undoubtedly the paper by Davies et al. (2021) that attracted renewed attention to this line of research by presenting the use of machine learning in mathematical practice as a general framework for working mathematicians. The authors present deep neural methods as a tool "to guide [the mathematicians'] intuitions concerning complex mathematical objects, verifying their hypotheses about the existence of relationships and helping them understand those relationships" (p. 70). In their view, deep learning techniques can contribute to understanding the relation between two objects $X(z)$ and $Y(z)$ by helping to identify a function $\hat{f}$ such that $\hat{f}(X(z)) \approx Y(z)$ for some well-specified domain of z. Such a function can then be used to verify the existence of structural features in objects and understand them through attribution techniques. Accordingly, the proposed method consists in starting with a hypothesis for a function f, generating synthetic data, training a supervised model, and finding patterns through attribution techniques. The whole procedure can help mathematicians formulate conjecture candidates, informing, in turn, new hypotheses and data generation in an iterative and interactive process.

The authors of this work demonstrated the fruitfulness of their approach in the case of two open problems concerning the structure of knots and the combinatorial invariance of symmetric groups in representation theory. To better understand how neural techniques interact with mathematical objects and practices in this case, it may be helpful to examine certain aspects of this work. The result obtained concerns the connection between two distinct branches of knot theory, which could be addressed by finding relations between their respective characteristic invariants. In particular, the goal was to obtain a connection between the *signature* of a knot (a classical algebraic invariant in gauge theory), and typical geometric invariants in hyperbolic knot theory. Such a connection was not yet known in the field. The strategy was, then, to predict the signature of a knot from its hyperbolic invariants.

Using existing libraries in knot theory, the authors produced a sample set of 2,7 million knots with their corresponding invariants, divided into a train and a test set, which served to train a very classic fully connected feed-forward neural net to predict the signature from a list of hyperbolic invariants. Somewhat surprisingly, the model was able to predict the signature over the test set with very high accuracy, indicating that a connection does indeed exist between both kinds of properties of the knots. However, the high predictive accuracy can only do so much as verify that a structural connection exists. To obtain further insight into the structure of such a connection, the authors made use of gradient-based attribution techniques (saliency maps). The main idea is to measure the sensitivity of the prediction accuracy to small changes in the different dimensions of the input, which correspond, in this case, to the different hyperbolic invariants. In this way, the authors could determine that the prediction of the signature relied primarily on 3 out of 12 invariants, all 3 related to a geometric structural feature of the knot called "cusp." This observation allowed them to define

a new quantity attached to a knot, which they termed "natural slope" and used to state the conjecture that the signature was equal to the natural slope multiplied by some error. Interestingly, this first conjecture turned out to be false. Further investigation resulted in a refined version of the conjecture, including a fourth invariant also identified by the attribution analysis. This eventually led the authors to prove two theorems connecting the two kinds of invariants.

Our first example shows that the success of neural nets in targeting functions predicting some mathematical features from others can contribute to verifying and understanding structural properties of mathematical objects and formulating and refining conjectures. On the other hand, their capacity to find arguments in the domain for which the prediction fails can provide a powerful tool to discover counterexamples of existing conjectures. Such a practice can be no less informative of the objects' structure. Wagner (2021) provides a notable example of this approach. This paper uses neural models to find counterexamples and constructions to open problems in combinatorics and graph theory. The strategy adopted shows some originality compared to previous approaches. The idea is to train a neural model to find a mathematical object through *reinforcement learning*, following the success of this technique in training automatic systems to play games, such as Go. From this perspective, mathematics appears as a game where counterexample candidates are akin to possible winning moves. The model is highly general in that it requires only minimal prior knowledge, namely, the encoding of legitimate moves (i.e., the sort of mathematical construction expected) and a reward function providing a score to be used as feedback for how close the construction is from a winning move. Thus, by trying to maximize the score without knowing anything about the structure of the problem, the learning algorithm will improve the model's performance toward a winning move determined by the conjecture to be refuted.

Wagner (2021) applied this general strategy to a series of conjectures in extremal combinatorics, an area of mathematics interested in the maximum or minimum sizes of collections of finite objects (e.g., graphs) subject to some restriction. Significantly, within this framework, the general way to encode mathematic constructions is as *words*, in such a way that the combinatorial problem is translated into "a problem about generating a word of certain length from a finite alphabet" (p. 3). For instance, n-vertex graphs are represented as 0–1 sequences of length $\frac{n(n-1)}{2}$, which, as we have seen, can be generated by neural language models in an autoregressive fashion. The reward function computes a score once the entire word has been produced. The function typically includes a positive and a negative component reflecting desired and undesired properties given the problem studied. Using the reward function, the model is trained through gradient descent to minimize the cross-entropy loss.

A concrete example among the ones addressed by Wagner (2021) is a conjecture in spectral graph theory stating that for a connected graph of $n \geq 3$ vertices, with largest eigenvalue λ_1 and matching number μ, we have $\lambda_1 + \mu \geq \sqrt{n-1} + 1$. Although a counterexample to this conjecture already existed for $n = 600$, the method proposed was able to find a counterexample for $n = 19$ by setting n as a hyperparameter and training a neural model to minimize $\lambda_1 + \mu$ (set as the reward or loss function) hoping the loss would drop below $\sqrt{n-1} + 1$ for some construction

after enough training. Figure 4 shows a sample of the best construction at different moments of the training process. The last graph constitutes a counterexample for the targeted conjecture. Interestingly, while the trained model remains uninterpretable, the analysis of the evolution of the model's output during training provides some insight into the structural properties of the problem. We can see that the model first restricts the outputs to sparse graphs and then converges to a "balanced double star" structure. As exemplified in the paper, this kind of structural insight can be obtained even in the case where the model fails to provide an explicit counterexample for the problem in question, either indicating how a counterexample could be eventually obtained or providing hints as to why the conjecture holds and how it could be proved.

The relevance of neural methods for heuristic purposes in mathematical research has been called into question, for instance, by Davis' (2021) critique of Davies et al. (2021), for it is not clear what the specific value added by DNNs is compared to more classic statistical methods already in use by mathematicians in their own practice. Moreover, as pointed out by Wagner (2021), methods like the ones used in this paper are not adequate to address all problems and, as the author acknowledges, they "did not succeed in refuting any of the most famous conjectures in the field" (p. 1). Nevertheless, the renewed strengths and capabilities of statistical methods brought about by deep learning techniques are such that their most elementary application to the treatment of open problems in advanced mathematics is far from trivial, neither regarding the means deployed nor the kind of results obtained. In particular, neural methods seem to force the development of statistical approaches in fields classically considered foreign to them. Even though a heuristic use of DNNs relies more on a case-by-case approach rather than a solid general framework, it seems that the near future will show a fruitful development of works like the ones presented in this

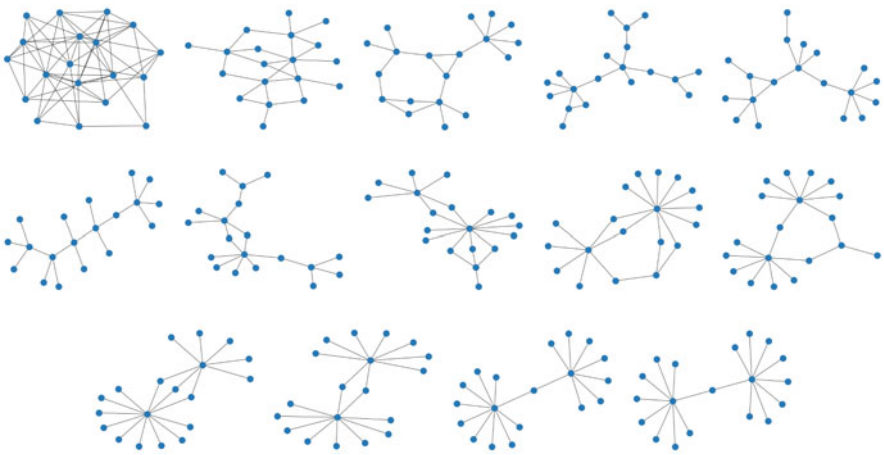


Fig. 4 Sample of the best construction over the training progress (from left to right, and from top to bottom) in Wagner (2021). The last graph (bottom right) constitutes a counterexample for the targeted conjecture

section. Consequently, this line of work could pave the way to more fundamental research on the statistical aspects of various fields in pure mathematics.

4 Philosophical Significance

Given the state of the field at the moment of writing these lines, the survey presented in the previous section is certainly fragmentary, partial, and guaranteed to be outdated by the time the reader encounters these pages. The four categories proposed to understand existing research orientations should be taken as temporary markers on moving soil rather than well-defined boundaries on a stable domain. Many of the papers presented lie at the crossroad of multiple research lines, and one can expect numerous blends of the four orientations in the near evolution of the field. It is nonetheless revealing that, faced with the renewed question of how to apply novel machine learning techniques to mathematics, instead of resorting to well-established distinctions, the research orientations tend to be spontaneously organized according to the AI researchers' implicit assumptions as to *what it is that we do when we do mathematics*. Indeed, as far as lines of research can be identified in this rapidly evolving field, they resist any traditional disciplinary (e.g., algebra, calculus, etc.), epistemological (e.g., pure, applied), or foundational (e.g., logic, set theory, category theory, etc.) categorization. Instead, they consistently explore specific *practices* assumed to characterize what mathematics is about: proving theorems, manipulating objects, acquiring skills, or addressing open problems.

The results in each identified direction exhibit countless limitations, many of which are already revealed in the literature, others yet to discover. We have referred to some of them in our account, including brittleness, poor generalization, restriction to elementary cases, lack of interpretability, absence of conceptual and theoretical foundations, and uncritical attitude toward mathematical corpora (see also Davis 2019; Welleck et al. 2022). All these limitations are real and could not be stressed enough if neural methods are to become a significant component of future mathematical practices. However, the simple fact that those mathematical practices can be addressed, if only with relative success, from the radically empirical perspective assumed by current machine learning approaches should be enough to arouse the interest of philosophers and historians of mathematics.

There are many reasons for the philosophy of mathematics, and the history and philosophy of mathematical practices in particular, to turn its attention to the recent mathematical applications of machine learning, even if, or precisely because, those reasons do not coincide with the ones advanced by the researchers in the field. Starting with the fact that, were these approaches to succeed to the point of becoming widely adopted as mathematical tools, the image of mathematics as we know it would certainly change in ways that are far from trivial, neither from a philosophical nor from a historical standpoint. Such a modification of mathematical practices would significantly impact the many dimensions philosophers and historians have become used to identifying in them, such as the role of proofs (Chemla 2012), the meaning of or even the need for foundations (Wagner 2019), the nature of

explanations (Mancosu 2008), or the role of experimental procedures (Borwein and Bailey 2003; Avigad 2008).

However, in the remaining pages, we would like to focus on a general feature underlying what, in our view, constitutes the philosophical novelty of neural models across their different applications to mathematical practices: *the renewed role of language in our understanding of mathematics*.

4.1 The Return of Language

If there is something our survey of neural applications to mathematics brings to light, it is at least the following: The success – actual or promised – of those methods in all their areas of application is inseparable from a whole new perspective on mathematical language, which goes far beyond the traditional understanding of mathematical expressions as simple notations for preexisting mathematical contents or a more or less arbitrary syntax for an independently determined semantics. The reason is that language, and more precisely, *writing*, including the wide variety of *textual forms* (statements, expressions, symbols, characters, etc.), *is all these models can rely on* to perform the mathematical tasks they are required to address. Whatever happens within the black box of a neural net happens on the basis of the processing of syntax and *syntax alone*. Granted, DNNs are not, as such, purely syntactic entities but the implementation of a relatively well-specified semantic procedure. However, such a procedure is highly generic, to the point that it can be successfully applied to very different kinds of data (images, language, sound, etc.) to perform significantly different kinds of tasks. The rapid convergence in recent years toward a single architecture (i.e., transformers) for all tasks confirms the resolve to rely on an unspecific analytic procedure. Given their generic character, it follows that if neural models can deal with mathematical content to any extent, such content is nowhere else to be found than in the corpus of (eventually supervised) textual expressions fed as inputs.

It is essential to understand that *this is not a bug but a feature* of this approach. Indeed, multiple ways exist to enhance neural models with hand-made rules and mathematical principles to increase their performance over specific tasks. As soon as accuracy becomes the primary concern, researchers do not hesitate to use all sorts of tweaks and tricks, introducing semantic features by other means. The resulting models progressively become, then, the object of an engineering practice, thus losing much of the philosophical interest concerning their relation to language. Nevertheless, the connectionist credo behind DNNs exhorts to develop learning models as generic concerning tasks as sparing regarding assumptions about the data. Consequently, whether openly acknowledged or unwittingly, current DNN applications to mathematical knowledge grant a critical place to mathematical language and textuality, providing the conditions for an original reflection on its role in the production and circulation of mathematical content.

Two primary aspects stand out when assessing the originality of the role played by language in this setting. The first is the effect of the fact that bare syntactic objects are

all these models can resort to for processing mathematical knowledge. Without dedicated axiom systems, deductive rules, logical operations, symbolic reasoning, material manipulation, or access to any other kind of context than textual, most of the conceptual focus within this line of research is put on mathematical *representations*. Be it logical statements, formulas, equations, expressions, or symbols,[7] researchers in machine learning have been led to raise the question of mathematical representations in ways that can communicate with current research interests in the philosophy of mathematics. Take, for instance, the proof-oriented approach and its reliance on multiple proof assistants' syntax to represent logical statements. The numerous strategies to process that syntax to improve the neural models' performance on inferential tasks (e.g., character- vs. token-based processing, explicit definition replacement, variable renaming, etc.) could provide novel perspectives for the design of mathematical languages, as addressed, for instance, by Avigad (2015) and "Mathematics as a Semiformal Language" (in this section). Furthermore, the various methods to encode mathematical objects in the object-oriented approaches (sequentialized binary trees for formulas, lists of numbers for matrices, different positional encodings for numbers, etc.) provide the elements for a fruitful dialogue with current investigations on the role of representations in the philosophy and history of mathematical practices, such as Schlimm (2018); Waszek (2018); and Kohlhase et al. (2018). In any case, practically no study across the different orientations fails to underline the critical character of the corpus and its representation for the success or failure of neural methods in mathematics. Yet, this centrality of representational aspects is not without novelty compared to existing philosophical perspectives on this problem. For the goal here is not to find the best representation for a given content (a suitable "notation") but the best representation for given texts so that tasks assumed to rely on their unknown content can be correctly performed by a statistical learning model such as neural nets. The careful investigation of this original question may bring unprecedented perspectives and insights into the classical problem of the relation between syntax and semantics, in mathematics as well as in language in general.

The second remarkable consequence concerning mathematical language is related to the connection between mathematics and *natural language*. Indeed, addressed as an artificial language from a modern logical perspective, most philosophical traditions have kept mathematical language apart from natural language for almost one and a half centuries. Moving away from this standard position, neural machine learning models are led to operate novel articulations between both, which are far from philosophically trivial. On a superficial level, one can find that, since many of the corpora used are composed of existing published papers (especially since the emergence of LLMs), the natural language necessarily present in those papers is usually leveraged to contribute to determining the content of mathematical expressions. A conceptual investigation on this topic is likely to provide compelling insight into the interaction between both

[7] To the best of our knowledge, there has been no work within this area of research involving the content of mathematical diagrams yet. From a different perspective, Sørensen and Johansen (2020) have used neural techniques to identify diagrams in mathematical texts.

natural and mathematical expressions in different corpora, be it from a historical perspective, in line with works such as those of Netz (1999) or Herreman (2000), or from a more cognitive viewpoint (Giaquinto 2008; Toffoli and Giardino 2013; Kohlhase et al. 2018). However, deep learning approaches bring natural language to the forefront of mathematical practice and knowledge in a more fundamental fashion. For the vast majority of the learning models and techniques used for processing mathematical expressions are none other than those specifically developed for the processing of natural language. The recent convergence toward transformer architectures also confirms this circumstance. This means that, within this setting, *mathematical expressions are treated as being themselves a sort of natural language*. This idea is not only implicit in all uses of neural language modeling architectures in mathematical applications but, as we have seen, is also repeatedly advanced by the researchers themselves across all research orientations. At odds with a view making a clear-cut distinction between natural and artificial languages or resolved to reduce the former to the latter through logical means, the success of neural methods in mathematics is inextricably tied to an essential yet unexplored link between mathematical contents and the mechanisms governing the ordinary practice of natural language. It could be, for instance, that mathematical and natural language share specific properties, or that the treatment of natural language constitutes a general framework for studying all kinds of languages, or yet that deep neural models constitute a general framework for the analysis of any type of data, including both natural language and mathematics. Whatever the case, the feats of neural models' applications to mathematics mark the return of natural language to the center of the reflection about what it means to do mathematics, in a philosophically original and compelling way.

4.2 New Challenges for a Linguistic Approach to Mathematical Practice

Despite the assumed divide between artificial and natural languages in modern philosophies of science, linguistic approaches to mathematics have not been altogether absent from the historical and philosophical scene, particularly since the emergence of the philosophy of mathematical practices. The chapters included in the current section of this Handbook provide an excellent account of much of that work. However, the many uses of neural models call for an original linguistic approach to mathematics, presenting new challenges to a language-driven philosophy of mathematics.

As already advanced, the originality lies in that *mathematical content is derived from expressions alone, semantics from syntax, meaning from pure text*. Yet, in modern mathematics, access to content is supposed to be provided through explicit and rigorous symbolic mechanisms, such as definitions within a formal language, logical operations, deductive rules, systems of axioms, or formal models. Those symbolic principles are assumed not only to determine what those expressions mean but also to control how expressions interact and what results from those interactions. Without direct access to explicit symbolic means – let alone cognitive faculties or

social conditions – it is difficult to know how the semantics or the content of expressions can be determined.

The difficulty is even greater if one considers the mathematical content DNNs manipulate is not limited to *referential* aspects of expressions. Indeed, the solution of mathematical tasks requires much more than associating, for instance, the expression `406` to a particular number or quantity, or to the same entity as the expression `four hundred and six`; it goes beyond referring the expression `A∧B` to a logical conjunction or $y''-y=0$ to a differential equation of second order. The tasks in question require, in addition, that enough *operational* content is involved when determining that `406` added to `326` equals `732`, that `A∧B` is likely or unlikely to be a premise in the proof of some given logical statement, or that $y(x)=c_1e^{x}+c_2e^{-x}$ is the solution to that differential equation.

Resorting to how content is dealt with in natural language, as most models do, might sound promising because the range of expressions does not depend, in this case, on explicit rules and other artificial procedures – typically absent in neural models – but tends to rely on more "natural" mechanisms. The latter could be invoked as being also at work in the relation between mathematical expressions and their content to explain the surprising capabilities of neural nets. However, as natural as it may seem, if we restrict ourselves to pure expressions, the access to content in natural language is far from simple and direct. The meaning of words is usually given through other words, risking an infinite regression, and ostensive definitions are doomed to inescapable ambiguities unless considerable segments of content are already known, as the radical critiques of Wittgenstein (2009) and Quine (2013), among others, have sufficiently shown.

To make things worse, the recourse to natural language imposes significant constraints on the treatment of expressions. Indeed, DNNs can only achieve their results by addressing the problem as a predictive task. Thus, given the expressions `406` and `326` as inputs, the model is required to predict the expression `732`, after being trained over a massive quantity of similar cases (although not this specific one). Or more precisely, it is expected to predict the expression `732` from any expression containing `406`, `326` and some expression of the additive operation, as in `406 + 326=` or `what is 406 plus 326?` (or some sequential encoding thereof). Now, by approaching the task as if we were dealing with natural language, we force models to perform those predictions by the same means by which they can predict, say, `you` from `she asked` when generating the expression `she asked you`, or `Paris` from `The capital of France is`, however differently structured those expressions and contents may be.

Finally, analyzing those underlying means presents innumerable obstacles. As it is widely acknowledged, DNN models are practically uninterpretable (see, for instance, Lipton 2018). In particular, we have no access to formal representations of a model other than its internal state, which can go from hundreds of billions up to over a trillion parameters in the current state of the art. Several methods have been developed in recent years, providing tools for the analysis and principles of interpretability of NLP neural models (see Belinkov and Glass (2019) and Madsen et al. (2021) for a survey). Among them, the method known as *probing* (Conneau et al. 2018) enjoys some success. The idea is to use the encoded vector representations of a model to train a classifier over a

specific linguistic property considered relevant for performing linguistic tasks (e.g., grammatical dependencies). If the classifier achieves good performance, the properties in question can be regarded as somewhat encoded in the internal state of the original neural model. However, while several mechanisms have been proposed to show that the information corresponding to the probed property is legitimately attributable to and effectively employed by the model (Hewitt and Liang 2019; Ravfogel et al. 2021), a fundamental gap remains between neural models and the symbolic feature being assessed. In other words, high – and ultimately superhuman – performance on such predictive tasks, which can indeed be conceived to grasp significant aspects of the content of expressions, might not reach that accuracy through the same means as humans. Neural models might be able to write sensible natural language or solve natural language tasks without ever following grammatical rules for any acceptable sense of grammatical rule-following. Likewise, those models might be able to solve mathematical tasks with reasonable and even surprising accuracy without ever manipulating any of the symbolic procedures by which we have historically controlled our mathematical practices. Converging results do not imply identical theories, representations, knowledge, or capabilities, which, in turn, does not mean that content is not being grasped. Interpretability attempts like the one of Charton (2022) presented above are certainly remarkable and promising but still limited and, all in all, largely insufficient given the scope of the problem.

5 Conclusion: But How?

In the face of all these obstacles, which are by no means unknown to researchers in the field, why should we place any hope at all in contemporary machine learning approaches to mathematical knowledge? Why should we rely on the orientations that a critical assessment could derive from them for a linguistically driven philosophy of mathematical practice? Is not the entire enterprise simply ill-conceived?

If the proof of the pudding is in the eating, it seems fair to acknowledge that, while the results exhibited so far are certainly limited from a mathematical point of view, they are far from insignificant from a philosophical one. Given the scale of the challenges, the slightest evidence that a seemingly impossible task can be achieved deserves philosophical consideration. Moreover, the models in question have shown surprising capacities in the treatment of natural language, where the richness of content makes the task otherwise challenging. Indeed, despite these models' duly identified limitations, the results exhibited by the application of DNNs to natural language processing in the last decade have been enough to catalyze, if not altogether generate, a revolution in the field (cf. Manning 2015).

Due to the close relation between mathematical texts and natural language within this framework, it seems legitimate to assume that the hope placed in neural methods for treating mathematical language is grounded on the mechanisms underlying the success of neural language models, whatever they may be. In this sense, the symbolic character of mathematical practices cannot be an objection *a priori* to this approach. On the one hand, it is not excluded that neural models are implicitly manipulating some higher-order representation akin to symbolic properties or structures (cf., for instance, Manning

et al. 2020). On the other, symbolic practices in mathematics as we know them are a relatively recent phenomenon. Although mathematical practices across different historical periods and cultures have constantly used expressive means that are irreducible to those of natural language, the recourse to a foundational role of symbolic practices such as axiomatic methods, formal systems of inference, or model-theoretical semantics has only become standard in the course of the past century. Mathematical knowledge has very well been produced and evolved for millennia across different cultures without such foundations (cf. Wagner 2019) and would arguably continue to do so if alternative principles and practices came to fulfill better the tasks expected from it in existing cultures, even if that requires to reinspect our existing notion of semantics, content, or even understanding, both in natural language and in mathematics.

Accordingly, a philosophical investigation of this phenomenon should focus on the elementary mechanisms responsible for the success of generic learning algorithms in manipulating natural language content when applied to a practically raw corpus of texts, and on the reasons and conditions for those same mechanisms to apply to the treatment of mathematical content. To avoid resorting to dubious, or at any rate unverifiable, analogies between artificial neural models and human learning capacities, let alone biological characteristics of the human brain, it is imperative to turn our attention to what those models *actually do*. As we have seen, DNNs are statistical models resulting from a statistical learning procedure in the context of predictive tasks. If such models can grasp linguistic content when operating with pure text, it is natural to ask what the source of that content can be. Since the expressions fed into the model are nothing but a sequence of empty identifiers, and models are highly generic before training, the source of all content, be it linguistic or mathematical, can lie nowhere else but in the training corpus. Which is another way of saying: *in the series of collective and historically determined linguistic and mathematical practices, as they are recorded in the many forms textuality can take*.

That significant segments of such content can be drawn from the expression of those practices is something the limited yet startling results of mathematical applications of neural language models are displaying before our eyes. The task remains to understand *how* that is even possible and invest this feat with the philosophy it deserves. With it, a whole new program opens up before a philosophy of mathematical practices informed by a renewed theory of language and signs.

Acknowledgments The author wishes to thank Deniz Sarikaya, Bharath Sriraman, David Waszek, Roy Wagner, and John Terilla for their patient support, precious feedback, and constant encouragement.

Funding Information This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 839730.

References

Alemi AA, Chollet F, Een N, Irving G, Szegedy C, Urban J (2016) DeepMath – Deep sequence models for premise selection. In: Proceedings of the 30th international conference on neural information processing systems, NIPS'16. Curran Associates, Red Hook, pp 2243–2251

Avigad J (2008) Computers in mathematical inquiry. Oxford University Press, New York, Chap 11, pp 134–150

Avigad J (2015) Mathematics and language. In: Davis, E., Davis, P. (eds) Mathematics, Substance and Surmise. Springer, Cham

Bansal K, Loos SM, Rabe MN, Szegedy C, Wilcox S (2019) HOList: An environment for machine learning of higher-order theorem proving (extended version). CoRR abs/1904.03241, http://arxiv.org/abs/1904.03241

Belinkov Y, Glass J (2019) Analysis methods in neural language processing: a survey. Trans Assoc Comput Ling 7:49–72. https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00254/1923061/tacl_a_00254.pdf

Biggio L, Bendinelli T, Neitz A, Lucchi A, Parascandolo G (2021) Neural symbolic regression that scales. In: Meila M, Zhang T (eds) Proceedings of the 38th International Conference on Machine Learning, PMLR, Proceedings of Machine Learning Research, vol 139, pp 936–945. https://proceedings.mlr.press/v139/biggio21a.html

Blechschmidt J, Ernst OG (2021) Three ways to solve partial differential equations with neural networks – a review. GAMM-Mitteilungen 44(2):e202100,006. https://doi.org/10.1002/gamm.202100006

Bommasani R, Hudson DA, Adeli E, Altman R, Arora S, von Arx S, Bernstein MS, Bohg J, Bosselut A, Brunskill E, Brynjolfsson E, Buch S, Card D, Castellon R, Chatterji NS, Chen AS, Creel K, Davis JQ, Demszky D, Donahue C, Doumbouya M, Durmus E, Ermon S, Etchemendy J, Ethayarajh K, Fei-Fei L, Finn C, Gale T, Gillespie L, Goel K, Goodman ND, Grossman S, Guha N, Hashimoto T, Henderson P, Hewitt J, Ho DE, Hong J, Hsu K, Huang J, Icard T, Jain S, Jurafsky D, Kalluri P, Karamcheti S, Keeling G, Khani F, Khattab O, Koh PW, Krass MS, Krishna R, Kuditipudi R, et al (2021) On the opportunities and risks of foundation models. CoRR abs/2108.07258. https://arxiv.org/abs/2108.07258, 2108.07258

Borwein JM, Bailey DH (2003) Mathematics by experiment – plausible reasoning in the 21st century. A K Peters, New York

Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A, Agarwal S, Herbert-Voss A, Krueger G, Henighan T, Child R, Ramesh A, Ziegler D, Wu J, Winter C, Hesse C, Chen M, Sigler E, Litwin M, Gray S, Chess B, Clark J, Berner C, McCandlish S, Radford A, Sutskever I, Amodei D (2020) Language models are few-shot learners. In: Larochelle H, Ranzato M, Hadsell R, Balcan M, Lin H (eds) Advances in Neural Information Processing Systems, Curran Associates, Inc., vol 33, pp 1877–1901. https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

Brunton SL, Kutz JN (2022) Data-driven science and engineering: machine learning, dynamical systems, and control, 2nd edn. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781009089517

Carifio J, Halverson J, Krioukov D, Nelson BD (2017) Machine learning in the string landscape. J High Energy Phys 2017(9):157. https://doi.org/10.1007/JHEP09(2017)157

Charton F (2021) Linear algebra with transformers. CoRR abs/2112.01898, https://arxiv.org/abs/2112.01898, 2112.01898

Charton F (2022) What is my math transformer doing? – three results on interpretability and generalization. 2211.00170

Chemla K (2012) The history of mathematical proof in ancient traditions. Cambridge University Press, Cambridge

Chowdhery A, Narang S, Devlin J, Bosma M, Mishra G, Roberts A, Barham P, Chung HW, Sutton C, Gehrmann S, Schuh P, Shi K, Tsvyashchenko S, Maynez J, Rao A, Barnes P, Tay Y, Shazeer N, Prabhakaran V, Reif E, Du N, Hutchinson B, Pope R, Bradbury J, Austin J, Isard M, Gur-Ari G, Yin P, Duke T, Levskaya A, Ghemawat S, Dev S, Michalewski H, Garcia X, Misra V, Robinson K, Fedus L, Zhou D, Ippolito D, Luan D, Lim H, Zoph B, Spiridonov A, Sepassi R, Dohan D, Agrawal S, Omernick M, Dai AM, Pillai TS, Pellat M, Lewkowycz A, Moreira E, Child R, Polozov O, Lee K, Zhou Z, Wang X, Saeta B, Diaz M, Firat O, Catasta M, Wei J, Meier-Hellstern K, Eck D, Dean J, Petrov S, Fiedel N (2022) Palm: scaling language modeling with pathways. 2204.02311

Cobbe K, Kosaraju V, Bavarian M, Chen M, Jun H, Kaiser L, Plappert M, Tworek J, Hilton J, Nakano R, Hesse C, Schulman J (2021) Training verifiers to solve math word problems. CoRR abs/2110.14168, https://arxiv.org/abs/2110.14168, 2110.14168

Conneau A, Kruszewski G, Lample G, Barrault L, Baroni M (2018) What you can cram into a single $&!#* vector: probing sentence embeddings for linguistic properties. In: Proceedings of the 56th annual meeting of the Association for Computational Linguistics (volume 1: long papers). Association for Computational Linguistics, Melbourne, pp 2126–2136. https://doi.org/10.18653/v1/P18-1198. https://aclanthology.org/P18-1198

d'Ascoli S, Kamienny P, Lample G, Charton F (2022) Deep symbolic regression for recurrent sequences. CoRR abs/2201.04600, 2201.04600

Davies A, Veličković P, Buesing L, Blackwell S, Zheng D, Tomašev N, Tanburn R, Battaglia P, Blundell C, Juhász A, Lackenby M, Williamson G, Hassabis D, Kohli P (2021) Advancing mathematics by guiding human intuition with AI. Nature 600(7887):70–74. https://doi.org/10.1038/s41586-021-04086-x

Davis E (2019) The use of deep learning for symbolic integration: a review of (lample and charton, 2019). 1912.05752

Davis E (2021) Deep learning and mathematical intuition: a review of (davies et al. 2021). CoRR abs/2112.04324, 2112.04324

Davis E (2023) Mathematics, word problems, common sense, and artificial intelligence. 2301.09723

Devlin J, Chang MW, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. CoRR abs/1810.04805, 1810.04805

Ferreira D, Freitas A (2020) Premise selection in natural language mathematical texts. In: Proceedings of the 58th annual meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, pp 7365–7374. https://doi.org/10.18653/v1/2020.acl-main.657. https://aclanthology.org/2020.acl-main.657

Ferreira D, Freitas A (2021) STAR: cross-modal [STA]tement [R]epresentation for selecting relevant mathematical premises. In: Proceedings of the 16th conference of the European chapter of the Association for Computational Linguistics: main volume. Association for Computational Linguistics, Online, pp 3234–3243. https://doi.org/10.18653/v1/2021.eacl-main.282. https://aclanthology.org/2021.eacl-main.282

Freivalds K, Liepins R (2017) Improving the neural GPU architecture for algorithm learning. CoRR abs/1702.08727, http://arxiv.org/abs/1702.08727, 1702.08727

Ghahramani Z (2023) Introducing PaLM 2. https://blog.google/technology/ai/google-palm-2-ai-large-language-model/

Giaquinto M (2008) Cognition of structure. Oxford University Press, New York, Chap 2, pp 43–64

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, Cambridge MA/London UK

Griffith K, Kalita J (2021) Solving arithmetic word problems with transformers and preprocessing of problem text. CoRR abs/2106.00893, https://arxiv.org/abs/2106.00893, 2106.00893

Heal K, Kulkarni A, Sertöz EC (2020) Deep learning gauss-manin connections. CoRR abs/2007.13786, https://arxiv.org/abs/2007.13786, 2007.13786

Hendrycks D, Burns C, Basart S, Zou A, Mazeika M, Song D, Steinhardt J (2021a) Measuring massive multitask language understanding. 2009.03300

Hendrycks D, Burns C, Kadavath S, Arora A, Basart S, Tang E, Song D, Steinhardt J (2021b) Measuring mathematical problem solving with the MATH dataset. In: Thirty-fifth conference on neural information processing systems datasets and benchmarks track (Round 2), https://openreview.net/forum?id=7Bywt2mQsCe

Herreman A (2000) La topologie et ses signes: Éléments pour une histoire sémiotique des mathématiques. L'Harmattan, Paris

Hewitt J, Liang P (2019) Designing and interpreting probes with control tasks. 1909.03368

Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2(5):359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Hughes MC (2020) A neural network approach to predicting and computing knot invariants. J Knot Theory Ramif 29(03):2050,005. https://doi.org/10.1142/S0218216520500054.

Jejjala V, Kar A, Parrikar O (2019) Deep learning the hyperbolic volume of a knot. Phys Lett B 799: 135033. https://doi.org/10.1016/j.physletb.2019.135033. https://www.sciencedirect.com/science/article/pii/S0370269319307555

Jiang AQ, Li W, Tworkowski S, Czechowski K, Odrzygóźdź T, Miłoś P, Wu Y, Jamnik M (2022) Thor: wielding hammers to integrate language models and automated theorem provers. In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A (eds) Advances in neural information processing systems, vol 35. Curran Associates, pp 8360–8373. https://proceedings.neurips.cc/paper_files/paper/2022/file/377c25312668e48f2e531e2f2c422483-Paper-Conference.pdf

Jiang AQ, Welleck S, Zhou JP, Lacroix T, Liu J, Li W, Jamnik M, Lample G, Wu Y (2023) Draft, sketch, and prove: guiding formal theorem provers with informal proofs. In: The eleventh international conference on learning representations. https://openreview.net/forum?id=SMa9EAovKMC

Kaiser Ł, Sutskever I (2016) Neural GPUs learn algorithms. 1511.08228

Kaliszyk C, Chollet F, Szegedy C (2017) HolStep: a machine learning dataset for higher-order logic theorem proving. CoRR abs/1703.00426, http://arxiv.org/abs/1703.00426, 1703.00426

Kamienny PA, d'Ascoli S, Lample G, Charton F (2022) End-to-end symbolic regression with transformers. 2204.10532

Kim S, Lu PY, Mukherjee S, Gilbert M, Jing L, Ceperic V, Soljacic M (2019) Integration of neural network-based symbolic regression in deep learning for scientific discovery. IEEE Trans Neural Netw Learn Syst 32:4166–4177

Kohlhase A, Kohlhase M, Ouypornkochagorn T (2018) Discourse phenomena in mathematical documents. In: Rabe F, Farmer WM, Passmore GO, Youssef A (eds) Intelligent Computer Mathematics. Springer International Publishing, Cham, pp 147–163

Lample G, Charton F (2020) Deep learning for symbolic mathematics. In: International conference on learning representations. https://openreview.net/forum?id=S1eZYeHFDS

Lample G, Lacroix T, Lachaux MA, Rodriguez A, Hayat A, Lavril T, Ebner G, Martinet X (2022) Hypertree proof search for neural theorem proving. In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A (eds) Advances in neural information processing systems, vol 35. Curran Associates, pp 26,337–26,349. https://proceedings.neurips.cc/paper_files/paper/2022/file/a8901c5e85fb8e1823bbf0f755053672-Paper-Conference.pdf

Lee D, Szegedy C, Rabe M, Loos S, Bansal K (2020) Mathematical reasoning in latent space. In: International conference on learning representations

Levitt JSF, Hajij M, Sazdanovic R (2019) Big data approaches to knot theory: understanding the structure of the jones polynomial. 1912.10086

Lewkowycz A, Andreassen A, Dohan D, Dyer E, Michalewski H, Ramasesh V, Slone A, Anil C, Schlag I, Gutman-Solo T, Wu Y, Neyshabur B, Gur-Ari G, Misra V (2022) Solving quantitative reasoning problems with language models. In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A (eds) Advances in neural information processing systems, vol 35. Curran Associates, pp 3843–3857. https://proceedings.neurips.cc/paper_files/paper/2022/file/18abbeef8cfe9203fdf9053c9c4fe191-Paper-Conference.pdf

Lipton ZC (2018) The mythos of model interpretability: in machine learning, the concept of interpretability is both important and slippery. Queue 16(3):31–57. https://doi.org/10.1145/3236386.3241340

Loos SM, Irving G, Szegedy C, Kaliszyk C (2017) Deep network guided proof search. CoRR abs/1701.06972, http://arxiv.org/abs/1701.06972

Lu P, Qiu L, Yu W, Welleck S, Chang KW (2023) A survey of deep learning for mathematical reasoning. In: Rogers A, Boyd-Graber J, Okazaki N (eds) Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Toronto, Canada, pp 14,605–14,631. https://doi.org/10.18653/v1/2023.acl-long.817. https://aclanthology.org/2023.acl-long.817

Madsen A, Reddy S, Chandar S (2021) Post-hoc interpretability for neural nlp: a survey. https://doi.org/10.48550/ARXIV.2108.04840., https://arxiv.org/abs/2108.04840

Mancosu P (2008) Mathematical Explanation: Why it Matters. Oxford University Press, New York, Chap 5, pp 134–150

Manning CD (2015) Computational linguistics and deep learning. Comput Linguist 41(4):701–707. https://doi.org/10.1162/COLI_a_00239. https://direct.mit.edu/coli/article-pdf/41/4/701/1807114/coli_a_00239.pdf

Manning CD, Clark K, Hewitt J, Khandelwal U, Levy O (2020) Emergent linguistic structure in artificial neural networks trained by self-supervision. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.1907367117

Martius G, Lampert CH (2016) Extrapolation and learning equations. CoRR abs/1610.02995, http://arxiv.org/abs/1610.02995, 1610.02995

McCarthy J, Minsky ML, Rochester N, Shannon CE (2006) A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. AI Mag 27(4):12

Meng Y, Rumshisky A (2019) Solving math word problems with double-decoder transformer. CoRR abs/1908.10924, http://arxiv.org/abs/1908.10924, 1908.10924

Netz R (1999) The shaping of deduction in Greek mathematics. Cambridge University Press, Cambridge

Newell A, Simon HA (1956) Plans for the Dartmouth summer research project on artificial intelligence. Typescript, Supplement to McCarthy, et al. (2006)

Nogueira R, Jiang Z, Lin J (2020) Document ranking with a pretrained sequence-to-sequence model. CoRR abs/2003.06713

Paliwal AS, Loos SM, Rabe MN, Bansal K, Szegedy C (2019) Graph representations for higher-order logic and theorem proving. In: AAAI conference on artificial intelligence

Petersen BK, Landajuela M, Mundhenk TN, Santiago CP, Kim SK, Kim JT (2021) Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients. 1912.04871

Polu S, Sutskever I (2020) Generative language modeling for automated theorem proving. CoRR abs/2009.03393, https://arxiv.org/abs/2009.03393, 2009.03393

Quine WVO (2013) Word and object, new edition, paperback edn. The MIT Press, London

Rabe MN, Szegedy C (2021) Towards the automatic mathematician. In: Platzer A, Sutcliffe G (eds) Automated deduction – CADE 28. Springer International Publishing, Cham, pp 25–37

Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, Zhou Y, Li W, Liu PJ (2019) Exploring the limits of transfer learning with a unified text-to-text transformer. CoRR abs/1910.10683, http://arxiv.org/abs/1910.10683, 1910.10683

Ravfogel S, Prasad G, Linzen T, Goldberg Y (2021) Counterfactual interventions reveal the causal effect of relative clause representations on agreement prediction. In: Proceedings of the 25th conference on computational natural language learning. Association for Computational Linguistics, Online, pp 194–209. https://doi.org/10.18653/v1/2021.conll-1.15. https://aclanthology.org/2021.conll-1.15

Saxton D, Grefenstette E, Hill F, Kohli P (2019) Analysing mathematical reasoning abilities of neural models. 1904.01557

Schlimm D (2018) Numbers through numerals. The constitutive role of external representations. In: Bangu S (ed) Naturalizing logico-mathematical knowledge: approaches from psychology and cognitive science. Routledge, New York. pp 195–217

Shen JT, Yamashita M, Prihar E, Heffernan NT, Wu X, Lee D (2021) Mathbert: A pre-trained language model for general NLP tasks in mathematics education. CoRR abs/2106.07340, 2106.07340

Sloane NJA (2007) The on-line encyclopedia of integer sequences. In: Kauers M, Kerber M, Miner R, Windsteiger W (eds) Towards mechanized mathematical assistants. Springer Berlin Heidelberg, Berlin/Heidelberg, pp 130–130

Sørensen HK, Johansen MW (2020) Counting mathematical diagrams with machine learning. In: Pietarinen AV, Chapman P, Bosveld-de Smet L, Giardino V, Corter J, Linker S (eds) Diagrammatic representation and inference. Springer International Publishing, Cham, pp 26–33

Szegedy C (2020) A promising path towards autoformalization and general artificial intelligence. In: Benzmüller C, Miller B (eds) Intelligent computer mathematics. Springer International Publishing, Cham, pp 3–20

Toffoli SD, Giardino V (2013) Forms and roles of diagrams in knot theory. Erkenntnis 79(4): 829–842. https://doi.org/10.1007/s10670-013-9568-7

Trask A, Hill F, Reed SE, Rae J, Dyer C, Blunsom P (2018) Neural arithmetic logic units. In: Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R (eds) Advances in neural information processing systems, vol 31. Curran Associates. https://proceedings.neurips.cc/paper_files/paper/2018/file/0e64a7b00c83e3d22ce6b3acf2c582b6-Paper.pdf

Turing A (1948/2004) Intelligent machinery (1948). In: The essential Turing. Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780198250791.003.0016

Udrescu SM, Tegmark M (2019) AI Feynman: a physics-inspired method for symbolic regression. 1905.11481

Valipour M, You B, Panju M, Ghodsi A (2021) SymbolicGPT: a generative transformer model for symbolic regression. ArXiv abs/2106.14131

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Lu, Polosukhin I (2017) Attention is all you need. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (eds) Advances in Neural Information Processing Systems, Curran Associates, Inc., vol 30. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Wagner R (2019) Does mathematics need foundations? Springer International Publishing, Cham, pp 381–396. https://doi.org/10.1007/978-3-030-15655-8_17

Wagner AZ (2021) Constructions in combinatorics via neural networks. 2104.14516

Wang M, Deng J (2020) Learning to prove theorems by learning to generate theorems. In: Larochelle H, Ranzato M, Hadsell R, Balcan M, Lin H (eds) Advances in neural information processing systems, vol 33. Curran Associates, pp 18,146–18,157. https://proceedings.neurips.cc/paper_files/paper/2020/file/d2a27e83d429f0dcae6b937cf440aeb1-Paper.pdf

Waszek D (2018) Les représentations en mathématiques. PhD thesis, http://www.theses.fr/2018PA01H231, thèse de doctorat dirigée par Panza, Marco Philosophie Paris 1 2018

Welleck S, Liu J, Bras RL, Hajishirzi H, Choi Y, Cho K (2021) Naturalproofs: Mathematical theorem proving in natural language. CoRR abs/2104.01112, https://arxiv.org/abs/2104.01112, 2104.01112

Welleck S, West P, Cao J, Choi Y (2022) Symbolic brittleness in sequence models: on systematic generalization in symbolic mathematics. 2109.13986

Williamson G (2023) Is deep learning a useful tool for the pure mathematician? 2304.12602

Wittgenstein L (2009) Philosophical investigations, 4th edn. Wiley-Blackwell, Chichester

Wolfram S (2023) What is ChatGPT doing. . . and why does it work? https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/. Accessed 09 July 2023

Wu Y, Jiang AQ, Li W, Rabe MN, Staats CE, Jamnik M, Szegedy C (2022) Autoformalization with large language models. In: Oh AH, Agarwal A, Belgrave D, Cho K (eds) Advances in neural information processing systems. https://openreview.net/forum?id=IUikebJ1Bf0

Zheng K, Han JM, Polu S (2022) MiniF2F: a cross-system benchmark for formal Olympiad-level mathematics. In: international conference on learning representations. https://openreview.net/forum?id=9ZPegFuFTFv

Index

A
Abelian group, 1868, 3180
Abelian integrals, 1636, 1658, 1673, 1674
Abel-Ruffini theorem, 763, 764
Abel's addition theorem, 1634, 1635, 1666, 1667, 1673
  elliptic and hyperelliptic integrals, 1651–1673
  general form, 1661–1673
Abel's approach, 719, 763
Abel's impossibility theorem, 763
Absolute arithmetic continuum, 1680
  and Archimedean arithmetic continuum, 1703–1705
  s-hierarchical, 1705–1707
Absolute continuity axiom, 1710
Absolute density, 1701
Absolute geometry, 725
Absolute linear continuum
  absolutely dense ordered class, 1702
  definition, 1701
Absolutism, 1238
Abstract, 2311–2318, 2320–2332
Abstracta, 2229
Abstract algebra, 70, 74, 3179
Abstraction principle, 159–166, 2384, 2395–2397
  characteristics, 2387
  definitions in mathematics, 2387
  Horwich's uniqueness problem, 2389
Abstract values
  beauty, 2105, 2107
  explanation, 2088, 2089, 2105, 2106
  fit, 2103, 2105–2107
  motivation, 2103, 2106, 2107
Abū Kāmil, 201, 208, 210, 222
Acceptable derivations, 2199
Accounting system, 1330
Ackermann function, 1941
Acquisition of system of computation, 3062
Active scientific realism, 1425
Act of communication, 3049
Actors' knowledge, 3066
Actors' practices and concepts, 3057, 3063
Actual number, 1604, 1611, 1614, 1616, 1618, 1620, 1622, 1624
Actual quantity, 1606, 1609, 1626
  for inverse trigonometric functions, 1640
Adequacy regions, 1444
Ad hoc polymorphism, 3174
Adhyardhā, 1049
Adversarial/non-adversarial arguments, 2669
Aemianalytic space, 1874
Affordance, 2062
Agency
  and mathematical activity, 2909–2911
  and mathematical artefacts, 2911–2914
  and mathematical proofs, 2914–2916
  and mathematical texts, 2916–2918
Agent-based approach, 2795
Agnis, 1033
Agreement, 2188
"Aha"-epiphany, 2187
Airlike/firelike sensibles, 564, 586, 589
Akhmatova, Anna, 849–852
Aleksandrov, A.D., 1896
Alexandria, 1728, 1740
  Degree-Minute-Second System (DMSS), 1158–1160
  rational expression arithmetic, 1160–1162
Alexandrov, P., 871, 875, 876, 878
Algebra, 64–66, 68–72, 79, 2140
  abstract, 70, 74
  algebraic equations, 75
  algebraic structures, 71, 72, 74
  algebraic topology, 71, 74
  cossic, 201–203, 207, 208, 210, 231, 232
  Hilbert space, 81, 82

© Springer Nature Switzerland AG 2024

B. Sriraman (ed.), *Handbook of the History and Philosophy of Mathematical Practice*,
https://doi.org/10.1007/978-3-031-40846-5

Algebra (*cont.*)
irrationals, 1627
origins of, 71
modern, 217, 232
number(s), 1596, 1612, 1627
premodern, 231, 232
spatial, 72–74
sum/difference, 1603, 1614
Algebraic construction, 3179
algebraic hierarchies, 3180–3182
implicit structure, 3179–3180
structures as objects, 3182–3183
Algebraic curves, 2021
Algebraic geometry, 735–745
Algebraic number theory, 1918
Algebraico-tree-theoretic condition, 1691
Algebraic proof, 2158, 2159
Algebraic structure of field, 2160
Algebra of line segments, 239–241
Algorithm, 2138, 2141, 2145
Algorithmic derivation, 432, 433
Algorithmic device, 2200
Algorithmic inferential system, 2209
Algorithmic norm, 433, 438–441
Algorithmic operation, 2191
Algorithms, 3184
Algorithm text, 2144
Al-Harawī's, 1748, 1749, 1757, 1759, 1765, 1771, 1776–1778, 1780, 1781
Alien arithmetic holomorphic structures, 2077
Al-Karajī, 200, 207, 209
Al-Khayyām, 230
Al-Khwārazmī, 203–205, 230
Al-Khwarizmi, 2141
Allele, 1386
Almagest, 1730–1731, 1741, 1760
Al-Māhānī, 1776–1778, 1780
Al Qiftī, 1776
Amalgamation, set models, 1712
Ambiguities, 3052
American Mathematical Society (AMS), 2436
Anabelian geometry studies, 1276
Anachronism, 1139, 1141
Analysis, 603, 629, 630
Analytical work, 3059
Analytic geometry, 73, 223, 226, 232, 263, 3020, 3025, 3026
Anaximander, 23
Ancient actors, 3050
Ancient China, 3064, 3067
Ancient Greece, 3051
Andersen, Line Edslev, 2040
Anderson, Alexander, 214, 216
Angle between two planes, 1750
Angler, 474, 475, 497–501, 503, 506, 917, 928–931, 935, 936, 951, 967, 979
division and collection of, 931
division of the, 928
logos criterion of, 929
tridentry to, 929
Angle trisection problem, 2342
Anhomeomers, 409
Anhupotheton, 603, 624–630
Anonymi Commentarius in Platonis Theaetetum (ACIPT), 943
Anselm, 103
AntConc software, 2938
Antecedents, 2623, 2631, 2632
Ante Rem Structuralism (ARS), 2323, 2327
Anthyphairesis, 467, 472–478, 481–483, 485, 487, 489, 493, 494, 502, 504, 506, 507, 510, 512–516, 602, 604–614, 617–619, 629, 632, 634, 636, 639
finite, 921
geometric, 921
infinite, 921
Logos Criterion for periodic, 922, 923
periodic, 604–608, 610, 613, 622, 624, 630, 634, 922, 923
philosophical, 925, 928
Anthyphairetic division, 524, 530, 531, 533, 535–537, 588, 594
Anthyphairetic infinity, 522, 523, 539, 560, 561
Anthyphairetic interpretation, 530, 533, 538
of Logos, 474, 498
of True Opinion, 474, 486
Anthyphairetic periodicity, 530, 552, 591, 592
Anticipating, 2140
Anti-Darwinian, 1394
Anti-de-Sitter spaces, 722
Anti-evolutionist argument, 1402
Antinomy, 805, 806, 833–835
Antipathy, 2459
Āpastamba, 911, 1047, 1048, 1055, 1065, 1070
Āpastamba Śulbasūtra, 1044
Apollonius of Perga, 212, 218, 223, 2172
Aposteriori science, 2435
Apotome line, 918, 919, 953–955, 958, 959, 962, 966, 981, 982, 988–991, 995, 998, 1001, 1003, 1004
anthyphairesis, 955–958
assumed line of, 962
definition, 953
Pell-related classification, 988–990
simple, 954
Appellation, 1596, 1599, 1603, 1610, 1618

Applied mathematical practices, 1237
Applied mathematics
  applications, 1241, 1242
  neutral technical knowledge, 1241
  social/ethical responsibility, 1241, 1242
Applied theoretical mathematical research practice, 1227
Apprehension, 3051
Appropriation, 3053
Approximation theorem for continuous functions, 1482
A priori knowledge, 2467, 2468, 2470–2472
Arabic, 2142, 3066
Arabic mathematics, 1725, 1748, 1778–1780, 1783
Arabic translation, 1776
Arbitrary exemplification, 2428, 2482, 2484
Arbitrary proposition/convention, 1613
Archimedean arithmetic continuum, 1703–1705
Archimedean axiom, 237, 1495, 1531, 1704, 1709
The Archimedean Oath, 1235
Archimedean ordered field, 1693
Archimedean property, 1470, 1486
Archimedes, 103
Archimedes' principle, 1476, 1480
Archives, 3064
Arch length of an ellipse, 1636–1637
Area, 3063
Areas as measure of cardboard weight, 2574
Argument, 2702
Argumentation
  adversarial/non-adversarial arguments, 2669
  argument-that/argument-about arguments, 2667
  A-schemes, 2675
  B-schemes, 2675
  centralized/distributed arguments, 2668
  C-schemes, 2677
  dialogue types, 2670, 2671
  hard/soft arguments, 2671
  mathematical communities, 2681–2684
  mathematical reasoning, 2678
  monologue/dialogue/polylogue arguments, 2667
  persuasive/directive/polemic arguments, 2669
  practices, 3051
  problem choice, 2679
  process/product arguments, 2667
  reasoning about refutations, 2680
  reasoning presentation, 2681
  rhetoric, 2678
  schemes, 2591, 2674, 2675
  sequential/parallel arguments, 2668
  small scale/large scale arguments, 2667
  static/dynamic arguments, 2668
  theory, 2073, 2591
  Toulmin layouts, 2672
Argument deductive mathematical proof, 2073
Argument-that/argument-about, 2667
Aristotelian biology, 400
Aristotelian physics, 404, 408
Aristotle, 31, 42, 103, 204, 213, 347, 348, 354, 355, 375, 467, 602, 603, 605, 608, 612, 613, 618, 620, 621, 623, 632–642, 1501, 1502
  metaphysics, 602, 618, 620–624
  physics, 603, 612, 613, 632, 633, 635–639
Aristotle's theory of hylemorphism, 405–406
Arithmetic, 2140, 3158, 3161
Arithmetica, 9
Arithmetical operation, 3050
Arithmetical content, 3139
Arithmetical hierarchy, 1911, 1912, 1917
Arithmetical induction, 448
Arithmetical operations, 2147–2149, 3050
Arithmetical point, 1493
Arithmetical relations, 1952–1954
Arithmetical rewriting, 3138
Arithmetical transfinite recursion, 1983
Arithmetic comprehension axiom, 1915, 1981
Arithmetic continuum, 1679
Arithmetic elliptic curve, 1919, 2019–2020
Arithmetic-geometric construction, 3083
Arithmetic of line segments, 242–244
Arithmetico-set theoretic continuum, 1679
Arithmetic Transfinite Recursion, 458
Arithmetization, 1466, 1482, 1486, 2763
  of analysis, 1969–1971
  of computation, 1975–1976
Arithmology, 833
Arrow's Impossibility Theorem, 2378
Arrow's theorem, 2368
Articles, 3049
Artificial agent, 3199
Artificial ("formal") languages, 2180
Artificial Intelligence (AI), 103, 2593, 2967, 3193
Artificial languages, 2967
Art of dividing 17 camels, 2540

ArXiv, 2436, 2446, 2448, 2453, 2454, 2934
Aryabhatiya, 3052
Ascending/descending sequence principle (ADS), 1984
A-schemes, 2675
Asian style options, 1351
Assertive, 3049
Assessment of this degree of generality, 3060
Assumed line, 962
Astral sciences, 3052
Astrology, 1733
Astronomy, 1780–1782
Aśvamedha yajña, 1067
Atoms, 95
Attempted enumeration, 1613
Aura, 3066, 3068
Aurel, Marco, 206, 207
Authorities, 2139
Autoformalization system, 3200, 3201
Autolycus, 1746–1758, 1782
Automated deduction, 2114, 2120
Automorphism, 1689, 1694, 1711, 1908, 1909, 1917, 2424, 3167
  group of structure, 1930
  nontrivial, 1929
Axiomata, 1802
Axiomatic approach, 1791
Axiomatic method, 1793, 2975
Axiomatico-deductive order, 3049
Axiomatico-deductive reasoning, 3051
Axiomatic pluralism, 2252
Axiomatics, 97–105, 1791, 2789, 2849, 2850, 2852, 2863, 2869–2872
Axiomatic set theory, 1224, 3157, 3178
Axiomatic system, 2293, 2297, 2298
  definition, 2397
  semantical universes associated with, 2394
  of Zermelo-Fraenkel set, 2392
Axiomatisation of the ZF system, 2380
Axiom of choice, 1541, 1542, 1914, 1916, 1917, 1966–1969, 2367
Axiom of Extensionality, 2370, 2380
Axiom of Foundation, 2381
Axiom of Infinity, 2371–2374, 2381
Axiom of Pairing, 2370, 2380
Axiom of reducibility, 2410
Axioms, 1722–1724, 1794, 1795, 2141, 2183
Axiom-schema, 2324
Axiom Schemas of Comprehension and of Replacement, 2374
Axiom systems, 2199
Axiom V Newton, 1182

B
Babbage, C., 102
Babylon, 1155–1158
Babylonian mathematics, 2802
Bach, J-S., 15
Bachelier's thesis, 1358
Back-of-the-envelope calculation, 1245
Bacon, Francis, 162, 170
Balancing point outside concave quadrilateral, 2573
Balls and Vase Problem, 2539
Bar induction, 1532
Barlow, William, 1837, 1841, 1845, 1848
Barton, Bill, 3004
Barwise-Kreisel compactness theorem, 1959
Basel problem, 2544, 2546
Base-2 number system, 177
Basic Argument from Improbability (BAI), 1402
Basic arithmetic operations, 169, 178, 179
Basic cognition, 2803, 2805
Baudhāyana, 911, 1055, 1059, 1066, 1073, 1078, 1081
Bayesian analysis, 1421
Bayesian approaches, 2591–2592
Bayesian epistemic probability, 1417
Bayesian inference, 1380
Bayesian inverse probability logic, 1421
Bayesianism, 1416
Bayesian perspective, 2713, 2716, 2718, 2721, 2723, 2724
Bayesian probabilities, 1448
Bayesian reasoning, 1413, 1449
Bayesian statistics, 1380, 1448
Bayes' Rule, 1379
Bayes's theorem, 1415
Beauty
  Cellucci's account, 2093, 2100
  Rota's account, 2092
Begriffsschrift, 2127, 2132, 2990
Behavior, 682
Being mode of existence, 652, 663, 692
Bella Muchnik's People's University, 876
Beltrami's model, 1965, 1966, 3027
Belyi, Andrei, 807, 835, 845
Benacerraf's dilemma, 2887
Benedetto da Firenze, 209
Beneficence, 1232
Bergman, I., 15
Berkeley, 163, 165
Bernard Bolzano Gesamtausgabe, 1597
Bernoulli, Daniel, 299

Bernoulli, Johann, 184, 297
Bernoulli numbers, 2485
BERT model, 3208
Bézout's Lemma, 1192, 1206, 1212, 1213, 1216
Bézout's theorem, 3160
Bhāskara I, 3052
Bhaskara II, 920, 966, 999, 1005, 1008–1011, 1015
Bhaskara II cakravala
  description, 1010
  Weil's proof, 1012
  with kuttaka and its relation to reconstructed Theaetetus' method, 1010–1012
Bhavana, 919, 920, 999–1002, 1004–1006, 1008, 1009, 1013, 1015
Biagio, Maestro, 209
Bibliographies, 3065
Bieberbach, Ludwid, 1849
Bieberbach-Frobenius' theorem, 1850
Binary clock, 195
Binary notation, 170, 173, 175, 176, 180, 184
Binary number system, 168, 171, 187, 188
  addition, 179
  calculating machine, 169, 179
  development of, 177–181
  early writings on, 172–177
  fractions and expansions, 181–184, 186
  medal, 186
  and philosophy, 184–187
  wax seals, 186
  and Yijing, 187–188
Binary numeration, 174
Binomial line, 918, 920, 953, 958, 960, 962, 964, 998
  assumed line of, 962
  conjugacy with apotome line, 959
  connection of Bombelli with complex number, 960
  definition of, 958
  role of, 959
  simple, 958
Binomial theorem, 1191, 1602, 1623, 3166
Bioethics, 413
Biogeography, 1388, 1391
Biometricians, 1385
Biometry *vs*. Mendelism, 1391
Birkhoff's equation, 1311
Bisector theorem, 1889
Blackboard, 3067
Black-Scholes formula, 1346, 1350, 1351
Blocks, 2149, 3068
Blog-posts, 2737
Bloor, David, 2603–2607
  strong program, 2605
  symmetry principle, 2607
Bodies of knowledge, 3069
Bohr, Harald, 2622
Bohr's 1913 atomic theory, 80
Bohr's correspondence principle, 2359
Bolzano's *Größenlehre*, 1597, 1606, 1608, 1625
Bolzano's insights and uncertainties, 1597, 1625
Bolzano's manuscripts, 1596, 1597, 1611, 1623
Bolzano's mathematical notebooks, 1597, 1600, 1602, 1605, 1607, 1608, 1614, 1616, 1625
Bolzano's *meßbare Zahlen*
  corrections and amendments, 1621–1623
  definition, 1615, 1623
  equality relation, 1619–1621, 1624–1625
  expression, 1616, 1620, 1622
  infinitely great, 1618, 1619, 1624
  infinitely small, 1617–1621
  notion, 1597, 1606, 1621, 1627–1629
  principle of trichotomy, 1608, 1624
  properties of, 1624
  rational number, 1617
  sequence of rational numbers, 1622, 1623
  theory, 1596, 1597, 1608, 1622
  unmeasurable expressions, 1616, 1622
  unmeasurable numbers, 1618, 1624
Bolzano's RZ, 1597, 1606, 1607, 1610–1614, 1618–1626
Bolzano's uncertainties, 1598, 1599
Bolzano-Weierstrass theorem, 1982
Bombelli's approach, 2548
Bombelli, Rafael, 200, 203, 205, 210
Bond (desmos), 507, 508, 510, 512, 513
Book X. of Euclid's *Elements*, 918, 919, 939, 950, 960, 979–981, 988, 991–993, 996, 998
Boolean algebra, 2473
Boolean connectives, 1924
Boolean exclusive-or operator, 181
Boolean logic, 2956
Borel's theory, 424
Born's rule, 77
Borwein, Jonathan, 2433, 2436, 2438–2441, 2445, 2453, 2455
Borweinian sense, 4
Bos, Henk, 227, 229
Boundary between drafts and publications, 3058

Bounded value theorem, 1977
Bounding quantities, 1601
Bourbaki, Nicolas, 24, 397
Bourbaki's sense, 768
Bourbaki-style axiomatic techniques, 2346
Bouvet, J, 169, 186, 187
Brahmagupta, 919, 920, 999–1002, 1004–1006, 1009, 1011, 1013, 1015
Braille, 194
Brain asymmetry, 26
Brasser, Jacob R., 210
Bravais, Auguste, 1836, 1838, 1846
Brazilian and black feminisms, 1317
Brentano, Franz, 1501–1506
Bresson, R., 15
Brian Rotman's semiotic approach, 2703
Brick-tiling plans, 1035
Brouncker, W., 1017
Brouwer, L.E.J., 30, 46–48, 57, 468, 1512–1515
  Compositio Mathematica, 654
  fixed point theorem, 658, 1980
  geometric topology, 659
  intuitionism, 468
  intuitionistic logic, 647
  language, 665–673
  lectures of Mannoury, 653
  life of, 652–661
  philosophy, 646–652
  as professor, 653
  view on mathematics, 675–689
Brouwer's intuitionism, 468
Brown, Tony, 3005
Brownian motion, 1333, 1334, 1336, 1399
Brute-force, 2833
B-schemes, 2675
Bugaev, Nikolai Vasilievich, 805, 832
Bulgakov, Mikhail, 16, 831
Bulgakov, Sergei, 23, 806, 812
Busemann, Herbert, 1886–1900
Busemann's geometry, 1886
Byrne, Oliver, 1807, 1811, 1813, 1817, 1821
  adult life, 1809
  books and contributions, 1808
  diagram fallacies, 1821
  experimentation, 1815
  geometrical entities/constructions and logical inferences of proofs, 1813
  geometrical proofs and constructions, new presentation of, 1808
  graphical representations, 1808
  imperfect diagrams, 1817
  innovative approach, 1809
  mathematics and geometry, practical applications of, 1809
  presentation of proposition, 1817, 1821
  visual aspect of diagrams, 1815
Byzantine lexicon *Suda*, 1730

C

Cakravala, 920, 921, 966, 999, 1000, 1002, 1005–1008, 1010, 1012, 1015, 1016
  Bhaskara II, 1008–1010
  cyclic meaning, 1015
  initial step of, 1007–1008
  origin of, 1005–1006
  proposition, 1006–1007
  similarity with Theaetetus' conjugacy method, 1015–1016
Calculating rods, 3048, 3064, 3065, 3067
Calculation, 2139, 2141, 2147, 3175
Calculus, 176, 1313, 1314, 1317, 2562, 2569, 2583
  early, trigonometric functions absence in, 324–329
  fundamental theorem of, 334
  modern, 328
  of variations, 3059
Calendar, terminology, 1098
Cancellation principle, 1537, 1539, 1540, 1543, 1555
Canonical comparison, 523, 551, 554, 555
Canonical full binary tree, 1683
Canonical polyhedra, 550, 554, 576
Cantor's continuum hypothesis, 2840
Cantor, Georg, 28, 75, 76, 82, 83, 86, 99, 105–107, 160, 162–165, 1484–1496, 2768
Cantor-Bendixson theorem, 1983
Cantor-Bernstein theorem, 2990
Cantor-Dedekind theory, 1679
Cantor normal form, 447, 448, 456, 1698
Cantor pairing function, 1971
Cantor's hypothesis, 1967
Cantor ternary set, 1494
Capital asset pricing model (CAPM), 1341
Cardano, Girolamo, 201, 205, 209, 210, 232
Cardinalities, 94
Careers, 2621, 2622, 2627, 2628, 2630
Carnap's principle of tolerance, 651
Cartan, Henri, 390
Cartesian Dualism, 92
Cartesian plane, 264
Cartesian space, 1713
Cartesian text, 3091

Cartesian theory of the equation, 3094
Cartwright, Nancy, 3021
Cases, 3060
Cashflows, 1333
CasualConc software, 2938
Catastrophe theory, 394, 404
C*ategoremata*, 345
Category-theoretic FOM, 2351–2353
Category theory, 703, 740–742, 756, 790, 1224
Cauchy completeness, 1586
Cauchy's convergence criterion, 1625, 1982
Cauchy sequences, 1526, 3164
Cauchy-Weierstrass conception, 1466
Cause for motion, 562
Caustics, 393
Cavaille, Jean, 424
Cavaillès's thematization, 2788
Cavalieri's geometry, 3075
Cavalieri's method, 367, 371
Cayley graph, 2530–2532, 2537, 2839
Celebrity, 2627, 2628, 2633
Celestial spheres (*orbium*), 1793
Centralized/distributed arguments, 2668
Centre of gravity, 1183, 2571
Centroid of a cardboard quadrilateral, 2572
Certainty, 2146, 2477
Ceva point, 2575
Chalk, 3048
Chance, 1389
Change of notation, 3059
Change of sign, 2561
Chapman-Kolmogorov equation, 1349
Characteristic field, 1710
Checking, 2140–2142, 2144, 2146, 2560
Chekhov, A., 15
Chemla, Karine, 2042
Chevalley-Tarski theorem, 1925
Chicago Board of Options Exchange (CBOE), 1348
China, 3051
Chinese language, 2142, 3049, 3054, 3067, 3069
Chinese mathematics, 2042
Chinese or Indian mathematics, 2787
Chinese Remainder theorem, 2559, 2560
Chladni figures, 847, 848
Choice function, 1542
Chomsky's grammar, 1937
Chomsky's hierarchy of languages, 1937
  Chomsky's grammar, 1937
  phrase structure languages, 1937
  production, 1937
  regular languages, 1937
Church-Turing's thesis, 1910, 1943, 1947, 1951
Cicero, 100
Circle, 183
Circular/revolutionary meaning of bond, 510
Circular process of modelling, 2576
Circulation, 3068
Circumcentre, 2577, 2578
Civic education, 904
Civic morality, 904
Civilisations, 23
Classe de mathématiques spéciales, 390
Classe de mathématiques supérieures, 390
Class field theory, 1919, 2016
Classical analysis, 1224
Classical geometries, 1888, 1891
Classical population genetics, 1398
Classical probability theory, 77
Classical semiological theories, 2961
Claude Chevalley's, 3058
Claudius Ptolemy's mathematical astronomy
  astronomical tables, 1742
  legacy, Pappus/Theon, 1732
  mathematical practice, 1729
  Museum, 1728
  philosophy, 1732, 1733
  production methods, 1729
  socio-historical and material aspects, 1738–1740
  sphericity, 1730, 1731
  tables, 1732
  technical fields, 1729
  texts/languages/genres, 1734–1738
Clavius, Christoph, 200–202, 205, 210, 231
Clement XI, Pope, 188
Clues, 3048, 3060, 3063, 3065, 3067, 3068
Coalescent tree, 1398
Cobordism, 391–392
Coda, 10–11
Codifications, 2331, 2333
Cognition, 2140, 2988–2989
Cognitive abilities, 2332
Cognitive apparatus, 3042
Cognitive history, 2050
Cognitive-metaphor analysis, 2704
Cognitive technologies, 3069
Cognitive unity, 2583
Cognitive work, 3063
Cohesive power, 1917, 1918, 2006
Cohesive sets, 1917, 2003–2007
Cohomological algebra, 740, 788
Cohomology theory, 740
Cole, J., 2861–2863
Colebrooke, 999–1001, 1005

Collective practices, 3057
Collinearity result, 2564
Collingwood, R.G., 104
Color symmetry groups, 1855–1857
Commandino, Federico, 212–214, 218
Commentaries, 2149, 2151, 2159, 3050, 3052, 3055
Commentariolus, 1790, 1793, 1798, 1800–1802
Commentary, 2154, 3053
Commentator, 2148, 2150, 2158
Common Core State Standards for Mathematics (CCSSM), 1150
Common fraction, 1152
Common sense model, 1237
Communicating knowledge, 3049, 3058, 3068
Communication, 3049, 3053, 3055, 3069
Communities, 2682, 2684
Commutation tests, 2963, 3007, 3136–3137
Commutative diagram, 2056
Compactness theorem, 1959, 1996–1997
Comparison between triangles, 1762
Compatibility logic, 1419–1423, 1449, 1452, 1455
Complete elementary Euclidean geometry, 1707
Completeness theorem, 117, 118, 2220
Complete ordered fields, 1486
Complete quadrilateral, 1769
Complete set, 1953
Complexity theory, 1906, 1907, 1909–1912
Complex proofs, 2076
Compound measurement, 1159
Comprehension, 2374
Comprehension axiom scheme, 2381
Compulsion, 2140, 2142
Computability theory, 1906, 1973–1975, 1997–2003
  Ackermann function, 1941, 1942
  arithmetical hierarchy, 1936, 1951, 1952, 1956
  arithmetical sets and relations, 1952–1954
  Church-Turing's thesis, 1935, 1943, 1947, 1951
  complete set, 1953
  composition, 1940
  computable enumeration, 1943–1946, 1953
  computable functions, 1935, 1939, 1951
  computable linear ordering, 1958, 1960
  computable pairing function, 1945
  computable relations, 1939
  computable set, 1939, 1944, 1953, 1958, 1959
  computable structure, 1958–1960
  computably enumerable sets, 1944, 1946, 1947, 1949
  definable relations, 1953, 1960
  fixed point theorem (recursion theorem), 1949
  halting problem, 1935, 1947
  halting set, 1935, 1946, 1947
  Hilbert's Tenth Problem, 1934, 1948, 1949
  incomparable Turing degrees, 1936, 1956
  index sets, 1950, 1952, 1953
  Kleene normal form theorem, 1945
  Kleene's predicate, 1945, 1951
  least non-constructive ordinal, 1957
  least Turing degree, 1955
  many-one reduction, 1953
  minimalization, 1942
  oracle Turing machines, 1954
  partial computable functions, 1939, 1941–1943, 1945
  partial recursive functions, 1939, 1942
  primitive recursive functions, 1935, 1940–1942, 1945
  productions, 1937
  recursive functions, 1935, 1936, 1942, 1943, 1951
  Rice's theorem, 1950, 1951
  standard model of arithmetic, 1959
  substitution, 1940
  supremum of Turing degrees, 1955
  Turing degrees, 1936, 1954, 1955, 1957
  Turing jumps, 1955, 1957
  Turing machine, 1935, 1938, 1939
  Turing reducible, 1936, 1954
  undecidable problems, 1934, 1935, 1949–1951
  universal Turing machine, 1938
Computable function, 1910–1913, 1915, 1917, 1918
Computable sets, 1998
Computable structure theory, 1912
Computably enumerable set, 1911, 1913, 1915, 1917
Computation, 2138, 2141–2143, 2152, 2158–2160, 3058
Computational opacity, 2434
Computational propaganda, 1247
Computational tools, 2422
Computer
  agency, 2791
  algebra system, 3205
  history, 20
  mathematical experiments, 2517

programming, 2563
proof, 2121, 2122
simulation, 2471
verification, 2474
Computer-assisted proof (CAP), 2460–2464, 2466, 2470–2476
Comte, Auguste
and positive philosophy, 2601, 2602
and sociology, 2598
Conceivable quantity, 1603, 1604
Concepts, 709, 758, 760, 2623, 2625, 2626, 2630, 3057, 3067
from a communal perspective, 144–150
and definitions, 137–138
and division of mathematics, 1607
explanatoriness, 153–154
form, 763, 764, 766, 767, 772, 773, 784, 785, 787, 791, 792, 794
Fregean view, 141
fruitfulness, 152–153
Lakatosian view, 141
mathematical, 712, 713
naturalness, 151–152
philosophical, 702, 703, 706, 707, 711–714
problems, 713–716
refining proofs, 435
and terms, 141–144
Conceptual analysis, 1467
Conceptual approach, 3058
Conceptual change, 2536, 2537
Conceptual engineering, 143, 2205
Conceptual innovations, 3061
Conceptualization of perpendicularity, 1045
Conceptual mathematics, 2646, 2975
Conceptual or metaphysical discursive analysis, 1465
Conceptual or logical analysis, 1467
Conceptual structuralism, 2856–2858
Conceptual thinking, 759
Concordance of sūtras, 1093
Conditions of validity, 2145
Confidence, 2620, 2630
Confidence distributions, 1422
Confidence intervals, 1419, 1420, 1423, 1428, 1439, 1444
Configuration, 2191
Conformal disk model, 1965
Confucian classics, 2149
Congruence formula, 1108, 1110, 1117, 1122
Congruence notation, 3144–3148
Congruence theorem, 1762
Conica, 2142
Conjectural density, 2451
Conjugacy propositions of Book X. of the *Elements*, 961
Heiberg's failure to realize importance of, 960
importance of, 960–962
Weil's view, 961–962
Conjugation of quadratic expressions, 960
Connection, 2153
Consensus, 2627
Consolidation, 527, 532, 539, 543, 544, 546, 548, 569, 570, 573, 578, 585, 587–590, 592, 594
Constancy axiom, 1546–1548
Constancy principle, 1543–1544
Constitutive power, 3114, 3115
Constructing the reals, 4
Construction, 2142
Construction of angles, 1760, 1761
Constructive analysis, 1525, 1528, 1530–1532, 1547
Constructive least upper bound principle, 1527
Constructive mathematics, 1514, 1521–1523
Constructive ordinal, 1957
Constructive proofs, 435
Constructive reals
algebraic operations on, 1525–1526
axiomatizing, 1529–1530
convergence of sequences and completeness, 1526–1527
functions on, 1527–1529
Constructivism, 1426, 2601, 2603, 2604, 2606, 2608
Constructivist school, 753
Constructivist sociologies, 2601
Contemporary analysis, 2423
Contemporary biology, 401
Contemporary mathematics, 2625, 2634, 2979–2980
Content, 2188
Content & Receptacle, 526, 539, 547, 548, 573, 584, 585, 592
Context of communication, 3051
Contexts, 3051
Contexts for mathematical activity, 3063
Contexts of Communication, 3051
Contextualization, 3053
Contiguity, 3064, 3065
Continental analytic methods, 1810
Continuant polynomials, 1190, 1192, 1197
Continued fraction, 524, 527, 542, 597
Continuity, 3064
axioms, 1704
definition of, 1582

Continuity (*cont.*)
in history, 21, 23
of polynomial function, 1583
principle, 1532
and sign preserving property, 1582–1583
of totally ordered set, 1584–1585
Continuous function, 1577
and continuity, 1577
intermediate value theorem, 1578–1579
preface version, 1582
Continuous quantities, 1596, 1607
Continuum hypothesis, 40, 75, 76, 82, 83, 86, 1490, 2374
Contradiction (philosophical notion), 806, 833, 838, 844, 850, 853–856
Conventionalism, 883–884, 890
Conventional option, 2229
Convexity theory, 1894–1897
Convolutional neural nets (CNNs), 3196
Conway, John Horton, 1841, 1842, 1851
Conway names, 1680
Coordinatization mapping, 1709
Copernican theory, 1728, 1790
Copernicus, Nicolaus, 1724
astronomy, 1790
axiomatic approach, 1791–1793, 1798, 1799, 1801, 1802
discovery, 1800, 1801
geometric models, 1791
internal logic, 1797, 1798
interpretations, 1803
interpretive framework, 1799–1801
mathematical creativity, 1801
mathematical practice, 1803
petitio, 1796
questions, 1793, 1794
Corpus linguistics, 2791, 2935–2937
Correctness, 2142, 2143, 2145, 2154, 2155, 2158, 2159, 2188
of an algorithm, 2142, 2146, 2153, 2160
of the procedure, 2148, 2150, 2158, 3051
of this operation, 2157
of this transformation, 2156
Cosmic Galois group, 721, 722
Cosmism, 23
Cosmological model, 22
Cosmos, 23
Countable nonstandard model, 1997, 2005
of arithmetic, 2000
Countable set, 1966
Countably categorical theory, 2003
Counterfactuals, 2265
Countermathematicals
axiomatic, 2267–2268
counterarithmetical, 2269–2271
Lewis's definition, 2262
metamathematical, 2267
stipulational, 2268
Counter-modernism, 651
Counterpossibles, 2262
countermathematical for class of, 2262
in mathematical practice, 2264
vacuist responses, 2273–2279
Covariance matrix, 1438
Covid Alert Level (CAL), 1245
Coxeter–Dynkin diagrams, 1200
Criterion of elimination, 139
Critical internalism, 1135, 1138, 1139
Critical mind, 2139
Crystallography, 1826, 1828, 1830, 1832–1834, 1836–1838, 1845
Crystal systems, 1833
C-schemes, 2677
Cubic parabola, 264
Cubic polynomials, 2017–2018
Cuneiform, 2142, 2146, 2148, 3050, 3054
Curriculum, 3062
Curve, 1919
Cyclic graph, 2528, 2547
Cycloid, 326

D
Dakṣiṇāgni, 1063
Dangicourt, P, 181
Dante (Divine Comedy), 821, 823–825, 827, 828, 835, 850
Dante's allegory, 821
Dārśapaurṇamāsika vedi, 1066, 1086
Darwin's Black Box, 1404
Darwin, Charles, 1384
Darwin's theory of evolution, 93, 1385–1388
Das Kontinuum, 1516, 1517, 1519
Data analysis, 1412
Datta's proof, 1085
Davidsonian approaches, 2222, 3082
Davies, B., 104
da Vinci, L., 108
Davis, Chandler, 4
Davis, P., 103
Davydov curriculum, 1171
de Billy, Jacques, 203, 210
Decidable structure, 1959

Decimal number, 175, 178, 189, 192
Decimal place-value notation of number, 2159, 3069
Decimal place-value numeration system, 2149
Decimal place-value system, 3066
Dedekind, Richard, 1482–1488, 1494, 3137–3138
Dedekind abstraction, 3162–3164
Dedekind continuity axiom, 1708
Dedekind cut, 542, 1969, 1970
Dedekind device, 1466, 1467
De dignitate et augmentis scientiarum (On the Dignity and Advancement of the Sciences), 170
Deduction, 2553–2555, 2557, 2583, 2584
Deductive mathematical proofs, 2070
Deductive reasoning, 2583, 2584
Deep neural networks (DNNs), 2967, 3192, 3194, 3195
  heuristic-oriented, 3211–3215
  object-oriented, 3202–3207
  proof-oriented, 3197–3202
  skill-oriented, 3207–3211
Defeasible Modus Ponens, 2674
Definite property, 1467
Definitional equality, 3165
Degree-Minute-Second System (DMSS), 1158–1160
Degree of undecidability, 117
Degrees of generality, 3061
Dehn's word problem, 1906
de l'Hôpital rule, 2495
Deliciae mathematicae, 192
Dembski's formalism, 1406
Dembski, William, 1403
Democracy, 2377
Dempster-Shafer belief functions, 1425
Dentawyanjana, 1105
Dependency graph, 2449
de re and de dicto, 2308
Derivation, 2183
de Saussure, F., 20
Descartean plane, 263–265
Descartes, René, 66, 68–70, 73, 200, 202, 225, 231
Descartes' geometry, 3076, 3086–3098
Descartes' operations, 242–244
Descartes's analytic geometry, 419
Descending order, 449
Descriptive geometry, 3025
Descriptive set theory, 836, 2773
Designation/symbol, 1603, 1611, 1613
Detachability, 1406
Deterritorialization, 754
Deuteronomic texts, 1732
Diachronic, 3068
Diagonal argument, 1951
Diagonal argument example, 1951
Diagram, 3049, 3058
Diagram-based formal systems, 2209–2214
Diagram-based geometric reasoning, 2908
Diagram-chasing technique, 2055
Diagrammatic devices, 3057
Diagrammatic formal systems, 2983–2985
Diagrammatic proofs, 438, 2039, 2538
  from contemporary mathematics, 2053–2057
  Greek diagram, 2050–2053
Diagrammatic reasoning, 1808, 1812, 1815–1817, 1821, 2982
Diagrams, 1808, 1810–1813, 1815–1817, 1819, 1821, 3068
Diagrams in the plane, 2149
Diakena, 525, 526, 563–568, 573, 574, 576–579, 586–588, 590, 595, 596
Dialectic, 474, 478, 479, 486, 487, 489, 506, 518
Diameter, 474, 475, 477, 478, 484, 487–489, 492, 494, 507, 509, 513–516, 518
Dichotomy, 1385
Differential equations, 2560, 2561, 2576, 3060
Diffusion theory, 2770
Digital Object Identifiers (DOIs), 2443
Digital philosophy of mathematical practice, 2442, 2454
Dimensional homogeneity, 331–333
Dimitrov's theorem, 1917
Dina mulya, 1117
Dinapitu, 1109
Diophantine arithmetic, 9
Diophantine equation, 1008, 1011, 1906, 1911
Diophantine functions, 1948
Diophantine relations, 1948
Diophantine sets, 1911
Diophantus' Arithmetica, 3039
Diophantus of Alexandria, 201–203, 217, 219, 232
Directed graphs, 2054
Directed weighted networks, 2751
Directive, 3049
Dirichlet L-function, 3174
Dirichlet's principle, 2556
Discontinuity (philosophical notion), 804, 805, 807, 830, 832–834, 854
Discrete quantities, 1607
Discursive exposition, 3060

Dispersion, 527, 532, 539, 543–544, 546, 548, 568–570, 573, 578, 585–590, 592–594
Distribution theory, 2621, 2625, 2629, 2630
Divided Line, 532, 536
"Dividing in combination" (*lianchu*), 2158
Division algorithm, 1208, 1213, 1214
Division of Noble Sophistry, 475, 498, 502
Division of the Angler, 499, 503
Division of the Sophist, 503
Division of triangles, 1766, 1767
Documents, 3048, 3058, 3065
Dogmatism, 761
Domain of structure, 1923
The Domestication of the Savage Mind, 3068
Dominique Dumont, 3058
Dostoevsky, F., 15
Double anthyphairetic inequalities, 574
Double barrelled sense, 1238
Double-entry accounting, 1228
Double geometric progression, 171
Double inequalities, 522–524, 540, 542, 543, 546, 574
Double mathematical inequality, 568
Double tight inequality, 539, 569, 570, 573
Doubling strategy, 1153
Dow-Jones industrial average, 1343
Downward-Skolem theorems, 2375
Drafts, 3056–3058, 3063, 3064, 3068
"Drag-test" of dynamic geometry, 2569
Dualgruppe, 3057
Duality, 3058
Dual Remainders Games (DRG), 1205, 1209
Du Bois-Reymond, Paul, 1496–1501
Dunameis, 939–942, 944
  distributive plural interpretation of, 941–942
  support for distributive reading of, 942–944
Dunamis, 939, 942
Duodecimal, 168, 172, 190, 192
Dürer, 108
Dust, 3064
Dynamic arguments, 2668
Dynamic geometry, 2553, 2563, 2566, 2568, 2569, 2573–2575, 2578
Dynamic Geometry Systems (DGS), 2577
Dynamic inscriptions, 3067
Dynamic mathematics, 2625
Dynamics of rigid bodies, 307–314

E

Earthlike sensible body, 526, 564, 569, 570, 585, 587, 588
Eastern mind, 663, 685
Eccentric models, 1735, 1736, 1738
E-centered syntactic-transition rules, 2200
Ecliptical length, 2582
École Normale Supérieure, 390
Edges, 2530
Edinburgh School of the sociology of science and technology, 2603
Effective recognition procedures, 2182–2184
Egorov, Dmitri Fedorovich, 806, 807, 832, 834, 836, 840, 842, 844
Egyptian mathematics, 36
Eidos, 408
Eikos logos, 535
Eilenberg, Samuel, 391
Einstein, A., 76, 79–81, 95
Electronic calculators, 2561
Elegant theorem, 986
Elementarily equivalent, 1907, 1908, 1917, 1918
Elementary arithmetic continua, 1700
Elementary continuity scheme, 1707
Elementary diagram, 1918
Elementary (disturbing) forces, 310
Elementary expression, 1611–1612
Elementary function arithmetic (EFA), 2030
Elementary school, 898
Elementary topos, 1551, 2352
Elements, 66, 523–525, 528, 529, 531, 536, 538, 541, 547, 549–551, 553–556, 560, 566, 575, 576, 580, 582, 587, 588, 593, 595–597, 1809–1812, 1964, 1969, 3049–3051, 3057
Eliminators, 449
Elliptic curve, 1919
  arithmetic, 2019–2020
  FLT, 2023–2025
  real, 2019
Elliptic functions, 1634, 1635, 1672, 1673
Elliptic geometry, 3027
Elliptic integrals, 1634–1636, 1643, 1666, 1669, 1671, 1673
  arc length of lemniskate, 1638–1639
  development in 18th century, 1641–1642
  Euler results, 1641–1646
  Fagnano's result, 1641–1643
  Legendre results, 1642–1651
  multiplication, 1662–1672
  oscillation of simple pendulum, 1637–1638
  review on Abel, 1669–1674
Embodiment, 3063
Embryology, 393, 395
Empirio-criticism, 885
Endomorphism, 2424
Enhanced manipulative imagination, 2063
Enumeration of cases, 3060
Enumerative induction, 2403, 2406, 2407

Environment, 3068
Eo-polynomial, 1196
Eo-sequence, 1193
Ephemera, 3064–3067, 3069
Ephemerides, 1780
Epicycles, 22
Epicyclic models, 1735–1737
Epistemic, 3058
Epistemic access, 2327–2331
Epistemic actions, 2912
Epistemic goodness, 1242
Epistemic idealisation, 1450–1451
Epistemic-probability-as-model, 1450, 1453
  epistemic idealisation, 1450–1451
  flexible handling of priors, 1451–1453
Epistemological value, 3058
Epistemological *vs.* social models, 1225, 1226
Epistemology, 2927, 2929, 2949
  application, 2886
  feminist, 2891
  ideal rationality, 2887–2888
  justification, 2899–2901
  mathematics, 2880, 2884, 2897
  non-ideal rationalit, 2888
  social, 2889, 2890
  subjects, 2892–2894
  traditional, 2892
  true propositions, 2899
  without a subject, 2885, 2886
Epstein, David, 2435–2438, 2440, 2442, 2444, 2452, 2454
Equalization, 2152
Equilateral triangle, 524, 549–551, 554, 576
Eridania Lake, 2328
Eristic, 474, 478, 479, 487, 489
Escher, Maurits Cornelis, 1857
Esenin-Volpin, Aleksandr, 876
Étale cohomology, 741
Ethical neutrality, 1250
Ethical rules, 1229
Ethical scrutiny, 1249
Ethical theory
  ethics in mathematics, 1236
  ethics-in-science, 1233, 1234
  medical ethics, 1232–1233
  social responsibility of mathematics, 1235, 1236
  social responsibility of science, 1234, 1235
Ethico-onto-epistemology, 1312
Ethics in mathematics, 900, 913
  common sense model, 1237
  content/subject matter ethics, 1229
  ethical and professional issues, 1237
  ethical awareness and action, 1237
  ethical common-sense, 1238
  ethical questions, 1236
  hidden ethical outcomes, 1230
  institution ethics, 1229
  personal ethics, 1229
  professional issues, 1236
  purism, 1241
  for working mathematician, 1236
Ethics-in-science approach, 1233, 1234
Ethnomathematics, 1099, 2601, 2608–2610, 2788, 2796, 2803, 2806, 2809, 2968
Euclid's approach, 93, 97, 99, 103, 201, 204, 212, 220, 1724, 3049, 3057
Euclidean algorithm, 1192, 1197, 1202–1204, 1207, 1208, 1212–1214, 1216
Euclidean arithmetic, 258
Euclidean diagrams, 1813, 2224–2236
Euclidean geometry, 4, 10, 64, 72, 73, 123, 160, 388–389, 445, 731, 883, 1054, 1707–1712, 1722–1724, 1747, 1751, 1756, 1759, 1760, 1771, 1774, 1808–1810, 1817, 2051, 2292, 2298–2299, 2538, 2547
  angle excess and area of triangles, 1892
  classical notions, 1891
  convexity theory, 1894
  sign of the curvature, 1892
Euclidean manifolds, 733
Euclidean metric spaces, 2531
Euclidean plane, 73, 74, 1746, 1748
Euclidean program, 2408–2410
Euclidean proportions, 236–239
Euclidean Remainders Game (ERG), 1205, 1215
Euclidean rigid motions, 1713
Euclidean solid geometry, 1747, 1774
Euclidean symmetry, 1827, 1854, 1858
Euclidean theory, 765
Euclid's elements, 1746, 1748–1750, 1759, 1760, 1763, 1782, 1794, 2142, 2342, 2345
Euclid's geometry, 2976–2979, 2982
Euclid's magnitudes, 3162
Eudemus, 613, 636–638
Eudoxian theory, 922, 1164
Eudoxus of Cnidus, 1463
Eudoxus's theory, 1476
Euler, Leonhard, 102, 104, 1017, 1779, 2014, 3058, 3059, 3068
  axioms, 298
  concept of observer, 305
  concepts of time and space, 289–291
  diagrams, 2989
  dynamics of rigid bodies, 307–314

Euler, Leonhard (*cont.*)
early life, 270
education, 270
equation of continuity, 1555
equations of motion, 302
first proposition, 305
force of inertia, 298
formula, 1784, 2680
hydrostatics and hydrodynamics, 276
interest in mathematics and natural philosophy, 271
law of generality of algebra, 3086
Letters, 277–279
and mass definition, 300
mathematics, 2482, 2484, 2493
matter and properties, 279–285
in *Mechanica*, 300
mechanics and mathematics, 291–293
method, 1351, 2545, 2563
naïve conjecture, 2695
notations, 3059
origin of forces, 285–288
philosophy of nature, 272–273
physical phenomena, 276
polyhedron formula, 2555
potential energy, 275
pragmatism, 304
and principle of inertia, 297
principles of internal motion, 297
proof, 2542, 2547, 2548
theorem, 2605
Theoria, 293, 303
working method, 2039
zeta function, 2557
Eurocentrism, 1240
European Mathematical Society (EMS), 2443
Evolutionary biology, 1395
Evolutionary perspective of history, 1142
Evolutionary trees, 1399
Ewens, Warren, 1399
Exact results, 2145, 2159
Examples, 3056
Excentric hypothesis, 2580, 2581
Exception-barring, 143
Exclusion idealizations, 2182
Exclusive imperatives, 2930
Execute operations, 3066
Execution, 2145, 2146, 2149, 2159, 2160
Existence theorem, 411
Expansions, 181–184
Expectation, 2625, 2628
Experimental algorithms, 2433
Experimental mathematics, 4, 2420–2423, 2427, 2428, 2469, 2552, 2553, 2788, 2789, 2822
computation, 2826–2829
cross-field enterprise, 2445, 2447
crucial experiments, 2440, 2441
empirical studies of, 2442, 2443
first-order scheme, 2823–2826
journal and its authors, 2444, 2445
linguistic analyses, 2448–2452
in mathematics teaching, 2430
quantitative analyses, 2447, 2448
real-world, 2832–2837
research field, 2829–2832
systematic experimentation, 2440
thought experiments, 2837–2841
unlikely circumstances, 2439
Experimental theorems, 2452
Experimental warrant, 2433, 2434
Experiment as hypothesis testing, 2420, 2426–2427
Experiment as verification, 2420, 2424, 2428
Experimentation in mathematics, 2552–2557, 2560, 2563, 2567, 2576, 2582–2584
Experiment indicator, 2448–2450
Experimenting with hypotheses, 2555, 2556, 2582
Expertise, 2622, 2629, 2630
Expert use, 3054
Explanation, 432, 437, 438, 440, 2088
of binary arithmetic, 169, 188
Kitcher's account, 2090
Lange's account, 2090
Steiner's account, 2089
Explanatory power, 3113, 3114
Explanatory proofs, 435, 437
Explorations, 3062
Exploring and Communicating Knowledge, 3059–3061
Expression, 3136
arbitrary, 1624
finite, 1618
imaginary, 1602–1603
irrational, 1605
measurable, 1616, 1622
of numbers, 1610, 1611, 1614–1617, 1619, 1620, 1622, 1623
of quantities, 1605, 1606, 1609, 1614, 1622
symbolic, 1599
unmeasurable/inestimable, 1616, 1622
Expressive power, 3109, 3110
Extended agency, 2791
Extended Turing machines, 2193

Extensional abstraction, 160
External representations, 2423
Extreme value theorem, 1979
Extrinsicism, 767

F
Faith, in external reality, 107–108
Faltings, Gerd, 2022
Fanaticism, 903
Feeling of Compulsion, 2142
Feferman, S., 454, 1801, 1802, 2856–2858
Feferman-Scott-Tennenbaum model, 2005–2007
Feit–Thompson theorem, 3168
Feminism, 1321, 1322
Feminist epistemology, 2890, 2891
Fermat, Pierre de, 100, 101, 223, 225, 1017
Fermat equation, 2014
Fermat point, 2575
Fermat point of a scalene triangle, 2575
Fermat principle, 393
Fermat's last theorem (FLT), 432, 1918–1920, 2012–2014, 2031, 2402
  elliptic curves, 2023–2025
  factorization, 2015–2017
  geometric result on, 2020–2022
  Grothendieck universes, 2028
  Peano arithmetic, 2028–2030
  Ribet's proof, 2026–2027
  Wiles's proof, 2027–2028
Fermat's little theorem, 173, 191
Ferreirós, José, 2858–2860, 2881, 2886, 2892, 2898, 2899
Ferreiros' approach, 2787
Feynman diagrams, 2979
Feynman-Kac theorem, 1348–1350
Fibonacci array, 1193, 1194, 1197, 1207, 1216
Fibonacci-like structure, 1196
Fibonacci polynomials, 1200
Fibonacci sequence, 1941, 2329, 2490
Fibre bundles, 1321, 1322
Fictional imagination for veridical mathematical reasoning, 2543
Fictionalism, 57
Fiction-style thought experiments, 2539
Field, H., 50, 51
Field axioms, 1530
Fields Medals, 2620, 2622, 2624, 2627
Final transformation, 2158
Financial mathematics, 911
Finding a proof, 2563
Finger operations, 3067
Finite order arithmetic, 2028
Finite simple groups, 2474
Finitism, 446, 452, 459
  Hilbert-Gentzen standpoint, 453–454
  PRA, 453
Fiodorov, N., 23
First-order continuum, 1510
First-order formulas classification, 1953
First-order Peano arithmetic, 1914–1916
First-order scheme, 2823–2826
Fisher, Ronald, 1384, 1394
Fisherian hypothesis testing, 1406
Five-colored maps, 2460
Fixed layouts, 3063
Fixed point theorem, 658, 684, 1949
The Flanders Panel, 1301
Flatland, 2540
Floating sexagesimal place-value notation, 3062
Florensky, Pavel, 24, 27, 469, 802–807, 809, 811–836, 838, 839, 841, 842, 844–862
Folk platonism, 2825
Fontaine's constructions, 2632
Footnotes, 3065
Foreign exchange, 1331
Formal and informal proofs, 2038
Formal atoms, 1464
Formal epistemology, 2592, 2593
Formalised mathematics, 2933–2935
Formalism, 2297–2298
Formalizability or explanation of proofs, 2432
Formalization, 2060, 2333–2334
  of group theory, 3168
  mathematician's proof and, 2114, 2126
  process of, 2114
  rigorous and gap free without, 2125
  with Leibnizian language, 2127
Formal languages, 3153
Formal logical calculus, 657
Formal pattern, 2152
Formal proofs, 434–437, 441, 2061
  in mathematical practice, 2122
  of mathematical theorems, 2114
  nature of mathematical reasoning, 2128
  philosophical problem of, 2132
  problem of, 2115, 2118
Formal strategies, 2152, 2160
Formula-based mathematical thought experiments, 2525, 2548
Formulaic language, 2963
Fosbury flop in high jumping, 2573
Foundationalism, 2408
Foundation Axiom, 2371

Foundations of Constructive Analysis, 2085
Foundations of Geometry, 70
Foundations of mathematics (FOM), 902, 2340
  Bourbaki's project, 2348–2351
  category theory as foundation, 2351–2353
  definition, 2357
  epistemological analysis of, 2343
  Hilbert-style FOM, 2343–2344, 2347, 2354
  metamathematics and received concept of, 2341–2343
  practical foundations, 2341
  standard semantic approach in, 2347
  tolerance and pluralism, 2360–2362
  univalent foundations, 2353–2357
Four-colour conjecture, 2435
Four-color theorem (4CT), 2422, 2428, 2435, 2460–2465, 2467, 2469, 2470, 2472–2474, 2476, 2477
Fourier series, 2557–2559
Fourier transform, 1304
Four of Hearts, 1368
Four primary bodies, 522, 524, 538, 549, 552, 557, 562, 564, 566, 579, 581, 582, 584, 596
Fractal images, 1318, 1320
Fractal thinking, 1319
Fraction(s), 181–184, 192, 1600–1602, 1605, 1608–1612, 1614–1617, 1619, 1621, 1623
Frankenheim, Moritz Ludwig, 1833, 1834, 1837, 1845
Frechet filter, 1993
Free-floating CM, 2271
Freeze frames, 2538
Frege, Gottlob, 43, 94, 160–164
Frege's Basic Law V, 2388
Frege's invention, 2185
Frege's system, 1973
Frequentism-as-model, 1428–1430, 1435–1439, 1453
  and falsificationist Bayes, 1448–1450
  i.i.d. models and subject-matter knowledge, 1432–1435
  as interpretation of probability, 1430–1432
  for statistical inference, 1439–1440
Frequentist inference, 1428, 1439, 1444, 1445
Freud, S., 22
Frey-Hellegouarch curve, 1919, 2023–2024
Frobenius, Ferdinand Georg, 1850
Fromm, E., 652, 663–665
Frontloading argument, 2407
Fruitful frameworks, 2055, 2062
Fuchs, Dmitry, 878
Full infinite anthyphairesis, 474, 487, 489, 494
Function concept, 2813
Function of bounded variation, 1302
Function theory, 836
Functoriality, 743, 785, 791, 793, 794
Functors, 743
Fundamental operations, 3050
Fundamental quantities, 3062
Fuxi sequence, 187
Fyodorov (Fedorov), Yevgraf Stepanovich, 1847–1849

G
Gadolin, Axel, 1838, 1841
Galilean physics, 3074
Galileo, 65, 66, 68–70, 97
Galley method, 179
Galois' group theory, 3080
Galois representations, 2027
Galois theory, 716–722, 761, 771
Game-theoretic definitions of semantics, 2369
Game theory, 2368, 2379
Gārhapatya agni, 1039, 1070, 1072, 1087
Gauss, C.F., 102
Gaussian binomial coefficient, 1191
Gaussian curve, 2775
Gaussian distribution, 1437, 1438, 1444
Gaussian elimination, 2563
Gavrilets, Sergey, 1407
Gelfand-Fuchs cohomology, 878
Gene pool, 1386
General algorithm, 2144
General cases, 3060
Generality, 2150, 2153, 2219–2224, 2982, 3054, 3057, 3059–3061
General laws, 3057
General patterns of computation, 2160
Generating functions, 2987
The Genetical Theory of Natural Selection, 1394
Genetic drift, 1386
Genetic polymorphism, 1386
Genetics, 901
Genotype space, 1402
Gentzen/natural deduction rules, 2201
Gentzen's proof theory, 446–447
  consistency of pure number theory, 448–449
  ordinal notations, 447–448
GeoGebra, 2517
Geography, 900
Geometrical configuration, 2148

Geometric algebra, 997
Geometrical interpretations of quantities, 2149
Geometrical model, 2776
Geometrical transformations, 2148
Geometrical vocabulary, 3083
Geometric anthyphairesis, 483, 486, 502, 527, 528, 539, 544, 548
  definition, 921
  logos criterion, 922–925
  Theaetetean theory of ratios of magnitudes, 922
Geometric calculation, 1887
Geometric configuration, 3051
Geometric constructions, 2509
Geometric conventionalism, 891
Geometric infinite anthyphairesis, 543
Geometric intuition, 73, 1482
Geometric language, 3121–3123
Geometric magnitudes, 201, 202, 212–225, 228, 230, 232
Geometric models, 1791
Geometric operation, 2147, 3050
Geometric point, 602, 620–624, 629
Geometric-topological (GT) diagrams, 2063
Geometric topology, 659
Geometry, 64–75, 78, 80–83, 679, 704, 2140, 2141, 2517
  algebraic, 704, 735–745
  analytic, 73
  Hilbert space, 82
  hyperbolic, 723, 727
  modernist, 72
  non-Euclidean, 707, 712, 715, 725
  projective, 73
  of triangles, 1763
Geontype, 1386
Gerard of Cremona, 1776, 1782
Gettier, Edmund L., 2882, 2884
Ghetaldi, Marino, 221
Ghost theorems, 2073
Giardino, Valeria, 2039
Gillespie, John, 1395
Gini-index, 2741
Global reading, 3050, 3057, 3059
GNU Octave, 1117, 1122
Gödel, K., 37, 39–41, 44–46, 58, 67, 99, 102, 106, 853, 854, 1812
  arithmetization, 1975
  incompleteness theorem, 67, 76, 83, 85, 1991, 2267, 2367
  procedure, 2058
  programme, 117, 118
  theorems, 708
Goldbach's conjecture (GC), 2402–2406, 2826, 2827
Goldberg, Sanford, 2890
Golden Age of Russian mathematics, 469, 845
Golius, Jacob, 226
Google, 1246
Gosselin, Guillaume, 206, 232
Goursat, Édouard Jean-Baptiste, 1843, 1849
Grain measures, 1153
Graph-algebras, 2423–2425
Graph diameter, 2529
Graph logic, 2473
Graph minor theorem, 1984
Great unread, 2442
Greco, Daniel, 2888
Greek astronomy, 1782
Greek axiomatics, 1792
Greek culture, 3022, 3023
Greek geometry, 352, 374, 379, 760, 2050, 2064, 2908
Greek mathematicians, 1462
Greek mathematics, 64, 69, 82, 911, 2039
  cognitive exercise, 3037
  history, 3034
  manipulation of proportion, 3036
  mathematical diagram, 3037, 3040
  schematic diagrams, 3041
  semiology, 3034
Greek numerical practice, 258–263
Greek philosophers, 899
Gregorian calendar, 1102–1104
Grosholz, Emily R., 3004
Grothendieck, A., 395, 1725, 1726
  algebraic geometry, 1870–1873, 2632
  Banach and Hilbert spaces, 1261
  biography, 1867–1869
  Classes de faisceaux et théorème de Riemann Roch (1955–1957), 1264–1266
  cohomology theory of abstract algebraic varieties, 1266–1268
  Éléments de Géométrie Algébrique (EGA), 1269–1271
  Esquisse d'un programme, 1280–1282
  Étalé and étale, 1290–1291
  geometry of forms, 1875–1879
  Hilbert spaces inside Banach spaces, 1260
  La clef des songes, 1286–1288
  La longue marche à travers la théorie de Galois, 1276–1278
  Les dérivateurs, 1282–1284
  localization, 1867

Grothendieck, A. (*cont.*)
mathematical heart and categorical reason, 1288–1290
mathematical techniques, 1259
mathematics, physics and philosophy, 1880–1882
methodological strategy, 1261
multilayered strategy, 1259
notion of space, 1867, 1880–1882
Ph.D. thesis, 1259
pursuing stacks, 1278–1280
Récoltes et semailles, 1284–1286
relativity and universality, 1291–1292
Résumé de la théorie métrique des produits tensoriels topologiques, 1260
Riemann's heritage, 1879–1880
schemes, 1870–1871
second period, 1266
Séminaire de Géométrie Algébrique (SGA), 1272–1274
standard conjectures on algebraic cycles, 1274–1276
Sur quelques points d'algèbre homologique, 1262–1264
synthetic account of, 1288–1294
tame topology, 1873–1875
techniques de construction en géométrie analytique, 1268–1269
third period, 1276
topos, 742, 1871–1873
types and archetypes, 1292–1294
Group theory, 39
G-space, 1888, 1889, 1891, 1893, 1894, 1896
Guala's social theory, 2790
Guessing, 2560, 2561, 2583
Gutkind, E., 656

H
Haab calendar, 1101
Haar measure, 3170
Hacking, 2140, 2144, 2145
Haṃsamukhī, 1049
Hahn decomposition theorem, 886
Haldane, J. B. S., 1384, 1394
Halting problem, 1935, 1947
Halting set, 1911, 1912, 1947
Hamiltonians, 101
Hamkins, Joel David, 121, 128
Hanacaraka, 1105, 1106
Handling of the text, 3050
Handy Tables, 1738, 1740–1742
Hard/Soft arguments, 2671
Hardy-Weinberg law, 1393
Harriot, Thomas, 168, 227, 230
Harris, Michael, 2620, 2621, 2623, 2624, 2626–2629, 2631, 2633, 2634
Hausdorff maximality principle, 2206
Haüy, René Just, 1826, 1831–1833
Heath, T., 66
Hebrew calendar, 1102
Hedden, Brian, 2888
Hedrai, 563–565, 576–579, 590, 595, 596
Heesch, Heinrich, 1854
Heine-Borel theorem, 1915, 1979
Heisenberg method, 79, 80
Heliocentric theory, 1790, 1797
Heliocentrism, 1797, 1798, 1803
Helmholtz–Lie problem, 1889
Henri Poincaré's, 3060
Heraclitus, 18
Hermann, Carl Heinrich, 1852
Hermann, J, 184
Hermeneutic vacuism, 2283
Heron's formula, 260–263
Hero's method, 2561
Hersh, Reuben, 3, 4, 93
Hertz, 104
Hessel, Johann Friedrich Christian, 1834–1838
Hessel's theorem, 1835–1837
Heterogeneous reasoning, 2186, 2214
Heterozygous, 1386
High-school
biological sciences, 901
philosophy, 901
physics, 901
Hijri, 1102
Hilbert's Paradox of the Grand Hotel, 2539
Hilbert, David, 45, 46, 98, 162, 1849, 2369
Hilbert-Gentzen standpoint, 453–454
Hilbertian axiomatics, 3078
Hilbert Problem IV, 1891
Hilbert's arithmetic, 240
Hilbert's axiomatization, 445
Hilbert's formalism, 445
Hilbert's foundational program, 2988
Hilbert space theory, 39, 64, 72, 74, 76, 80–82
Hilbert's plane model, 241
Hilbert's program, 445–446
Hilbert's tenth problem, 1906, 1911
composition, 1940
diophantine definition, 1949
diophantine equations, 1934, 1947, 1948
diophantine functions, 1948
diophantine relations, 1948
formulation of Hilbert's Tenth Problem, 1934, 1947
history of Hilbert's Tenth Problem, 1934
negative answer to Hilbert's Tenth Problem, 1949

Hilbert's theorem, 3172
Hilbert transform, 1303
Hill, T., 188, 190
Hindu-Arabic numeral system, 1331
Hindu Mathematicians, 919, 966, 992, 999, 1002, 1005, 1007, 1015
Hipparchos' excentric hypothesis, 2581, 2582
Hipparchus of Bithynia, 212
Hippasus of Metapontum, 1462
Hippocratic Oath, 1236
Hiraṇyakeśi śulba, 1026
Histography, 2634, 2635
Historical anthropology, 3053
Historiography, 2310, 2590
  history of relation between history and mathematics, 1128–1135
  reconstruction as historical method, 1141–1145
  role of mathematics in, 1135–1141
History of mathematics, 8–10, 898, 904, 2331, 2803, 2806, 2815
Hobbes, 94
Homeomers, 409
Homogeneous potential, 143
Homogeneous-universal, 1704
Homogenizations, 2152
Homogenizing and equalizing, 2151–2153
Homomorphism, 706
Homothetic paths, 2513
Homotopy type theory (HoTT), 775, 776, 2353, 3169
Horizons, 2625, 2634
Hössjer, Ola, 1404
Høyrup, Jens, 2146, 2147
Human cognition, 2140
Human computation, 2194
Human idiosyncrasy, 2593
Humanistic moral tradition, 1243
Human surveyability, 2465
Hume, D., 100
Hume, James, 227, 230
Hume's Principle (HP), 2389
Husserl, E., 52, 53
Husserl's analysis of origin of geometry, 711
Husserl's interpretation, 3078
Hyperarithmetical hierarchy, 1957, 1960
  Barwise-Kreisel compactness theorem, 1959
  computable infinitary formulas, 1959, 1960
  computable ordinals, 1956
  constructive ordinals, 1957
  hyperarithmetical hierarchy of sets, 1956
  hyperarithmetical sets and relations, 1956, 1957, 1960
  Kleene's ordinal notations, 1957, 1958
  least non-constructive ordinal, 1956
Hyperarithmetical set, 1958, 1960
Hyperbolic geometry, 727, 883, 3026
Hyperbolic group, 2532, 2542
Hyperbolic trigonometry, 1773
Hyperelliptic integrals, 1636, 1664, 1666, 1667, 1673
Hypernatural numbers, 1535
Hyperreal line, 1534, 1553
Hypnosis, 16, 17, 20
Hypotheses, 603, 623–630
Hypothesis tests, 1419
Hypothetico-deductive thinking, 2583

I

Ibn Abī Jarrāda, 1758, 1765, 1776, 1777
Ibn al-Bannā', 204, 207
Ibn al-Haytham, al-Ḥasan, 1747, 1756, 1781
Ibn Hūd, 1776
Ibn 'Irāq, 1758, 1777–1780
Ibn Tibbon, Jacob ben Makhir, 1776
Ideal, definition of, 3139
  congruence, 3144
  divisor, 3139
  power, 3141
  prime, 3141
  simple, 3141
Ideal divisors, 2016
Idealism, 95
Idealized decimal measure number, 1164
Ideal rationality, 2886, 2887
Ideals of rings, 2017
Ideational mathematics (IM), 2646
Identity, 2141
Ideology-neutral accounts, 2252
Idiosyncratic idealism, 646
Ignoramibusstreit, 2342
Image account, 535, 536, 553, 557
Imaginary, 821, 823, 827, 830, 831
  expressions, 1602–1603
  numbers, 1628
  quantities, 1599, 1603, 1605, 1606, 1608–1610, 1612
Imperatives
  arXiv, 2934
  controlled natural language, 2935
  diversity, 2945–2946
  exclusive, 2930
  frequency, 2943–2945
  inclusive, 2930

Imperatives (*cont.*)
in mathematical proofs, 2918
mood, 2926
uses of, 2931
Implicit definition, 2384–2388, 2390, 2394–2397
Impossible elements in thought experiments, 2541
Incentre, 2569
Incircle, 2569
Inclusive imperatives, 2930
Incommensurabilty, 1462, 2141
Incommensurable dyad, 525, 557, 561, 563, 564, 583
Incubation period, 2557
Independence of irrelevant alternatives, 2377
Independent from the writing, 3069
Independently and identically distributed (i.i.d.), 1432–1436
Independent random variables, 1373
Independent unknowns, 210–211
Indian mathematics, 911
Individual random variables, 1373
Indivisible line, 467, 602, 603, 607, 608, 610–613, 619–624, 640–641
in Aristotle, 603, 620–623
in Plato, 602, 607–611, 620–624, 640–641
in Proclus, 611–613, 619
pseudo-Aristotelian, 603, 640–641
in Simplicius, 602, 611–613, 619, 620
in Xenocrates, 602, 607, 608, 612, 613, 619
in Zeno, 603, 640–642
Indonesia, 1100, 1101, 1103–1105, 1109, 1121
Induction, 458, 1966, 1972, 2403, 2406, 2407
Inductive generalization, 2482, 2484, 2486, 2488
Inductive types, 3164
Inequality, 524, 526–527, 539–543, 546–548, 550, 558–563, 566, 568–570, 572
Inertia, 298
Inference packages, 2980
phenomenological faithfulness, 2241–2246
and Shin on perceptual inference, 2246–2251
Inferential actions, 2915
Inferential algorithmic functions, 2208
Inferential scope of expression, 2205
Inferring warrants, 2653
Infilateral polygons, 1464
Infimum, 1526
Infinita terminata, 368, 371
Infinite anthyphairesis, 524, 527, 529, 531, 533, 537, 540, 541, 547
Infinite hypernatural numbers, 1535
Infinite kosmoi, 524–526, 536, 539, 544, 548, 589–590, 594
Infinite polynomials, 2545
Infinite Ramsey theorem, 1923, 1927, 1928
Infinite sets, 1966, 2760, 2768, 2769
Infinitesimal linearity principle, 1537, 1552
Infinitesimals, 363, 1479–1481, 1495, 1496, 1500, 1508, 1533, 1535–1538, 1540, 1543, 1544, 1546, 1547, 1549–1553, 1557, 1558
Infinitesimal straightness principle, 1537
Informal mathematical pluralism, 2252
Informal mathematics, 2690, 2692, 2694–2697
Informal proof, 432, 434, 436, 441, 2059, 2060
mathematicians' standards of rigor, 2121
worked-out version of, 2126
Informal reasoning, 2215–2219
Informal techniques, 2188
Information technologies, 2733, 3049
Initial segment of the anthyphairesis, 474, 481, 485, 487
Ink blot problem, 2539
Inquiry-based learning (IBL), 2553, 2555
Inscriptions, 3048, 3065, 3069
Institutionalization of mathematics, 2861–2863
Institutions, 3064
Instructions, 2928, 2937, 2945
Instrumentalism, 95
Instrument of thinking, 3056
Integers and mathematical computation, 2274
Integral calculus, 1464
Integrative power, 3112
Intellectual satisfaction, 2038
Intelligible being, 467, 523–539, 547–548, 555–558, 561–564, 579–585, 589–597, 602, 603, 608–620, 622–629, 631, 634–642, 917, 919, 925, 928, 936–939, 946, 948, 949, 951, 967, 979
acquisition of knowledge of, 939
division and collection, 951
knowledge of, 939
Intelligible equalization, 591, 592
Intelligible infinity, 524
Intention, 3049
Intercept theorem, 2509
Interdisciplinarity, 2594
Interference, 706
Interior angle, 2568
Intermediate value theorem, 437, 1529, 1578–1579, 1700, 1969
Internalism, 2634
International Congress of Mathematicians, 1970
International cooperation, 2379

International Ethics in Mathematics conferences, 1237
Internet, 20, 21
Interplay of practices, 2787
Interpret, 3063, 3066
Interpretation, 2148, 2149, 2183, 3050, 3053, 3056, 3068
Interpretations of probability
  aleatory probability, 1413–1415
  controversy, 1417–1419
  epistemic probability, 1415–1417
Interpretative equivalence, 1430, 1439–1441, 1445
Interpreting, 2153
Intersectionality, 1321
Intersection of circles, 1757
Inter-universal Teichmüller theory, 2077
"Intradisciplinary" approach, 2792
Intuition, 1364, 2180, 2300
Intuitionism, 30, 468, 646
  critics of, 680
  first act of intuitionism, 665
  mathematical, 647
  second act of, 675
Intuitionistic analysis, 657
Intuitionistic continuum, 1531–1533
Intuitionistic field, 1545
Intuitionistic logic, 647
Intuitionistic set theory (IST), 1551
Intuitively-effective recognition procedure, 2193, 2200
Inverse probability logic, 1421–1423, 1430, 1448, 1449, 1454
Inverse process, 2292
Inverting rule, 1154
Irrational coefficients, 203, 209–210
Irrational number, 1476, 1479, 1548
Irreducibility, 152
Ishāq ibn, 1776
Isometries, 1965
Isomorphism, 706, 1908
Isosceles right triangle, 548, 549
Isosceles triangle, 1052
Iteration, 2561, 2569
Iterative algorithm, 2562
Iterative numerical methods, 2561
Itō's formula, 1338

J
Java island, 1104
Javanese calendar
  background, 1122
  characteristics, 1105–1119
  computer implementation, 1117–1119
  main cycles of, 1106
  view of time, 1119–1120
Javanese people, 1105
Jean-Baptiste Biot, 2138
John Austin, 3049
Jordan, Camille, 1827, 1838, 1847
Journal of Theoretical Biology, 1404
Jurassic Park, 1301
Jury Theorem, 2374
Justification, 2402–2404, 2406, 2408–2413, 2415, 2880, 2900
Justificatory proofs, 435, 436
Just Six Numbers, 1560

K
Kalpasūtras, 1025
Kalupahana, D., 668
Kāmya citis, 1034, 1088
Kāmya yajñas, 1025
Kandinsky, V., 23
Kant, I., 161, 1722, 1723, 2140
Kantian apriorism, 470
Kantian teaching, 673
Kartini, Raden Adjeng, 1103, 1104, 1109
Kātyāyana, 911, 1059, 1060
Kātyāyana Śulbasūtras, 1055
Kauffman's reformulation, 2473
Kawolu, 1121
Keldysh, Lyudmila, 871–873
Kepler, Johannes, 212
Kimura, Motoo, 1399
King Wen sequence, 188
Kitcher, Philip, 53–55, 2881
Kleene normal form theorem, 1945
Kleene's ordinal notation, 1957
Klein, Jacob, 201, 204, 221, 222
Knot diagram, 2055
Knot theory, 2839
Knowledge and Social Imagery, 2605
Knowledge as recollection, 478, 514–516
Knowledge (episteme), 487, 498, 507, 514, 517
Knowledge of the diameter, 472, 475, 478, 484, 489, 507, 509, 514
Kock-Lawvere axiom, 1537, 1552
Kolmogorov, A., 18, 875, 876, 878
Kolmogorov complexity, 21, 22
Kolomogorovs quantification, 99
Kosmesis, 580, 581, 590
Kosmos, 537, 584
Kripke's scheme, 1532, 1548
Kronos era, 969
Kruskal's tree theorem, 1984

Kuhn, Thomas, 2881
Kuhn's language, 703
Kurup, 1115–1117
Kuttaka, 920, 999, 1006, 1008, 1010–1013, 1015

L
Lagrange, J.L., 1017
  approach, 1646–1648
  four-squares theorem, 1948
  notation, 2985
  theorem, 2349
Lagrangian equation, 777
Lakatos, I.
  concept stretching, 2704
  critiques of formalist model of mathematics, 2691–2694
  practice of argument, 2703
  principle of fallibility, 2697–2698
  Proofs and Refutations, 2690
  translative rhetorical force, 2705–2707
Lakatos, Imre, 2900
Lakoff's and Núñez's approach, 2704
Lambang, 1115
Lambert, Johann Heinrich, 1893
Lamina centroid, 2574, 2575
Landen's theorem, 1642
Langlands program, 1919, 2012
Language, 17, 19, 31, 665–673, 2602, 2603, 2608, 2956, 2959–2961
Language-based algorithmic systems, 2199–2202
The language of proofs, 2949
Large cardinal axioms, 116, 117
Large cardinal hierarchy, 116
Large Hadron Collider, 2828
Large language models (LLMs), 3199, 3207
Lasker-Noether theorem, 780
Latin translation, 1776
Lattice, 1946
Lautman, Albert, 424
Law of double negation, 1524
Law of excluded middle (LEM), 1541, 1542, 1546
Law of generality of algebra, 3074
Law of the balanced lever, 2570
Lawvere's theory of elementary topos, 2353
Lawvere-Tierney axioms, 1551
Layouts, 3057, 3063, 3069
Leadership, 2628
Learning environments, 3064
Learning Mathematics with Texts, 3062
Learning of the lists, 3062
Least upper bound principle, 1531, 1585, 1625, 1626
Leaving mathematics for philosophy, 423
Lebesgue, Henri., 836, 837, 840
Lefschetz, Salomon, 394
Legendre, Adrien-Marie, 1830
Legendre's Conjecture, 2404–2406
Legendre symbol, 152
Leibniz, G.W.F., 102, 168, 194, 1464, 2141, 3056, 3057
  binary arithmetic, 9
  conception of proof, 2141
  definition of plane, 1887–1888
  intellectual workshop, 3055
  manuscripts, 3067
  monadology, 806
  notation, 2985
  principle, 380, 3055, 3164
  thinking, 342
  triangle, 1316
Leibnizian algebra, 336
Leibnizian calculus, 328
Leibnizian universal language, 2116, 2122
Leibniz's trajectory
  extension-shape-motion, 352
  generation of physical bodies, 353
  phase III in mathematics, 361
  phase III in physics, 360
  phase II in physics, 354–357, 359
  phase IV in mathematics, 367–369, 372
  phase IV in physics, 363–367
  phase V, 374–380
Length, 3063
Lenin Prize, 15
Leonardo's theorem, 1830
Les intuitions topologiques primordiales de l'aristotélisme, 407
Les mathématiques et l'intelligible, 409
Lettered diagrams, 1816, 3054
Lettering of the diagram, 3049
"Letter of the 99", 876
Letters, 3049
Lévi-Strauss, Claude, 17, 3008
Lexell, Anders, 1746
Lexicographically ordered full binary tree, 1679, 1682–1684, 1688
Li Chunfeng, 2143, 3050
Light bulb, 2092
Light proof, 2546
Limit, 1600–1602, 1614, 1617
Limit points, 1494
Linearity, 1815

Linear order, 1967
Lines of argumentation, 2154
Linguistic analysis, 3052
Linguistic approach, 3218–3220
Linguistics, 17, 19, 20, 413
The Little Commentary, 1732
Liu Hui, 2143, 2145, 2154–2158, 3050
Lloyd, Geoffrey, 3051
Lobachevsky's thinking, 722–727
Location arithmetic, 170
Locke, John, 163, 165
Locus, 1386
Logarithmic functions, 330–331
Logic, 4, 1906–1908, 1912, 1916, 2305
Logical empiricism, 1427
Logical formalism, 2349
Logical objects, 44–45
Logical positivism, 882, 885, 893
Logical power, 3106–3109
Logical principles, 667
Logical probability, 2716
Logical reasoning, 100
Logic (or formalization), 2333, 2334
Logismos, 474, 475, 507–509, 518
Logistice numerosa, 217–219
Logistice speciosa, 201, 212–215, 218–222, 227, 230, 232
Logos, 474–476, 478, 479, 487, 489, 493–504, 506, 507, 518, 602–608, 610–614, 616, 617, 619, 620, 624–635, 637–639
Logos and Catastrophe Theory, 392
Logos Criterion, 474–476, 478, 487, 489, 493, 494, 498, 499, 501–504, 510, 512–516, 922, 924, 925, 929, 931, 932, 935, 966, 967, 969, 971, 972, 976, 978, 983
  of the Angler, 499
  abbreviated representation, 924
  example, 924–925
  periodicity anthyphairesis, 922
Logos et théorie des catastrophes, 394
Longino, Helen, 2891, 2892
Long-term historical approach, 3067
Lorentz transformation, 76
Łoś's Theorem, 1917
Löwenheim-Skolem theorem, 437, 2375
Lower Dedekind cut, 1969
Lucas array, 1215
Lucasian oereo polynomials, 1215
Lunisolar calendars, 1102
Luzin, Nikolai Nikolaevich, 807, 836–843
Luzin's crisis, 839
Luzitaniya, 841

M
Macbeth, Danielle, 2041
Mach, E., 95, 97, 104
Machine reasoning, 2122
MacIntyre mathematics, 1243
MacIntyre's theory, 1242–1244
MacIntyre's virtue ethics, 1244
Maclaurin, Colin, 296
Macrocosmos, 23
Macrostates, 1390
Maṇḍala parikṣepa, 1067
Maghetti, Benedetto, 210
Magma, 1990
Magnetic groups, 1857
Magnitudes, 237–238, 256–258, 1568–1569, 1596, 1604
  real numbers, 1568–1569
Mahāvedi, 1033, 1046, 1064, 1065
Mahoney, Michael, 222, 226
Major mathematics, 750–754, 756, 757, 760, 762, 765, 767, 773–776, 779, 781, 785–788, 795
Malevic, C., 23
Mallarmé, S., 29
Manava, 911
Mānava's construction, 1073
Mānava Śrautasūtras, 1028
Mānava Śulbasūtra, 1056, 1078, 1080
Mānava version, 1072
Mancosu, P., 1812, 1813
Manders, K., 1816, 1817, 1819, 1821
Manin, Yuri, 4, 14–24, 26–31, 92, 872
Manipulations of signs, 2141, 2159
Manipulative imagination, 2912
Marketisation, 1249
Markov Chain Convergence Theorem, 1375, 1376
Markov chain Monte Carlo (MCMC) algorithm, 1372, 1376, 1377
Markov chains, 1372, 1374–1376, 1380
Markov models, 1401
Martian mathematics, 2461
Martin-Löf's type theory (MLTT), 2353
Marxism, 886, 2607, 2608
Maśaka śulba, 1026
Mataram, Sultanate, 1105, 1116
Material, 3064
Material and social environments, 3064, 3065
Material environment, 3066, 3067
Material instrument, 682
Material objects, 3068
Mathematica, 2561, 2763

Mathematical activity, 2309, 2310, 2321, 2331, 2332, 2909–2911
Mathematical argument, 2691, 2701, 2702
Mathematical artefacts, 2911–2914
Mathematical-astronomical theory, 1791
Mathematical astronomy, 1797, 1801
Mathematical cognition, 2796, 2802–2804, 2809
Mathematical communication, 2610
Mathematical community, 2679, 2680, 2682, 2684
Mathematical concepts, 2541, 2542, 2548
Mathematical conjecture, 2716, 2720, 2724
Mathematical crystallography, 1826, 1849
Mathematical culture, 3054, 3055
Mathematical data and models, 1251
Mathematical debates in 1990s, 2432, 2452
Mathematical definitions, 3170
  from a communal perspective, 144–150
  and concepts, 137–138
  explanatoriness, 153–154
  fruitfulness, 152–153
  naturalness, 151–152
  nature of, 138–140
Mathematical equation, 396
The Mathematical Experience, 2294
Mathematical experiments, 2504–2507
  research projects, 2507–2508
  with diagrams, 2508–2509
Mathematical foundations, 2367
Mathematical heuristics, 3212
Mathematical induction, 437, 2492
Mathematical innovations, 3061
Mathematical interpretation, 940–944, 2366
Mathematical intuition, 38–41, 647, 2535, 2536, 2539, 2540, 2542
Mathematical justification, 2402, 2403, 2406, 2408, 2410, 2411, 2790
Mathematical knowledge, 1221, 1223, 1224, 2470, 2601–2604, 2607–2611, 2613, 2963, 3053
Mathematical Knowledge Without Proof, 2403
Mathematical language, 2929–2935, 2937, 2958, 2959, 2962, 2965–2967
  abstraction management, 3167–3169
  abstract objects, 3161–3169
  algebraic construction, 3179–3183
  algorithms, 3183
  Dedekind abstraction, 3162–3164
  definedness and partiality, 3159–3161
  diagrams, 3183
  equality, 3164–3165
  intuitions and heuristics, 3185–3186
  methodological orientation, 3155–3156
  natural language, 3153
  numbers, 3161–3162
  object identification, 3165–3166
  operations, 3158–3159
  philosophical orientation, 3154–3155
  plans, 3184–3185
  predicates and relations, 3159
  set-theoretic abstraction (*see* Set-theoretic abstraction)
  sorts, 3156–3158
  structure identification, 3166–3167
Mathematical logic, 910
Mathematical meaning, 3050
Mathematical modelling, 1426, 2368
Mathematical models in theology, 2774
Mathematical modernism, 65, 68–71, 74–83
Mathematical objects
  as abstractions, 42–43
  alternative conception, 55–56
  as fictions, 50–52
  heuristic conception of, 56–58
  as idealizations of operations, 53–55
  as idealizations of sensible things, 52–53
  as independently existing entities, 37–41
  as logical objects, 44–45
  as mental constructions, 46–48
  as simplifications, 45–46
  as structures, 48–50
Mathematical papyrus, 2145
Mathematical pluralism, phenomenon of, 2366
Mathematical practice, 910, 2341, 2432, 2526, 2529, 2541, 2548, 2848–2851, 2853, 2855, 2858, 2859, 2861, 2864–2867, 2869–2873, 3048, 3054, 3058, 3061, 3068
  activities, 2639
  aesthetic judgements, 2659
  agendas, 2641
  claims, 2642
  cognitive mechanism, 2640
  communities, 2638, 2642, 2643, 2659
  computer-assisted proofs, 2642, 2643
  corroborate/challenge claims, 2641
  discourse, 2642
  DNA evidence, 2642
  education research, 2643
  educators, 2638, 2639
  expert, 2659
  fine-grained cognitive processes, 2641
  goal-directed activities, 2639
  goals, 2640
  groups of scholars, 2640

judgments, 2643
laboratory settings, 2659
mathematical language, 2659
mathematics journals, 2638
mechanism, 2641
methods, 2643
precision, 2640
problem solving process, 2640
psychologist, 2641
tasks of computers and experiments in, 2436–2438
teaching/research, 2641
truth of assertions, 2643
types of evidence, 2642
types of implications, 2640
Mathematical Practice as Philosophy, 468
Abel and Galois case study, 716–722
architecture of concepts and, 706–716
Lobachevsky and Riemann case study, 722–735
Weil and Grothendieck case study, 735–745
Mathematical practices, 9–11, 1227, 1228, 2810, 2956–2965, 2967, 2968, 3066
anachronism, 2172
basic, low-level practices, 2814–2816
cognitive aspects, 2796–2801
fictionalism, 2169
function concept, 2808–2816
Hamkins' multiverse view, 121
Kitcher's definition of practice, 2170
limits of, 2802–2806
mathematical and scientific practices, 2810–2814
mathematics education, 2167
ontological commitments, 2175
Platonic tradition, 2166
Platonism, 2168
Quine's ontological commitment, 2169
set-theoretic pluralism debate, 111
sociocultural aspects, 2801–2802
web-of-practices approach, 2806–2808
Woodin's Ultimate-L argument, 115
Mathematical probability theory, 1372, 1412
Mathematical problem, 2138, 3065
Mathematical procedures, 3050
Mathematical proof, 2039, 2914–2916
algorithmic derivation, 432
concept-refining proofs, 435, 440, 441
definition, 2124
diagrammatic proofs, 2039
explanatory proofs, 435, 437, 439
formal proofs, 2041
functions, 432
history of, 2042
justificatory proofs, 435–437
kinds of, 434
machine-executable, 2120
norms, 433
pluralist view, 441
problem of formalizing, 2121
social epistemology of, 2040
unification-inducing proofs, 435, 439, 440
values of, 2040
Mathematical reasoning, 2429, 2590, 2591
Mathematical revolution, 69
Mathematical signs, 2956, 2962–2966, 2968
Mathematical statement of Theaetetus' theorem, 945
Mathematical tablets, 2148
Mathematical texts, 2916–2918, 3055, 3069
Mathematical thinking, 2564, 2582, 2583
Mathematical, thought and experiment (M-T-E), 2525
knot theory, 2526
unknot diagram, 2526
visual symmetry in graphs, 2527
Mathematical thought experiments, 2837
Mathematical transformations, 754
Mathematical truth, 1220
Mathematical typesetting systems, 3210
Mathematical understanding
growth of, 2644
implications, 2649, 2650
open ended interviews, 2645–2647
task-based interviews, 2647–2649
teaching, 2644
universal agreement, 2644
university level, 2644
Mathematical work, 3060, 3061, 3066, 3067
Mathematical writings, 3064
Mathematicians, 870–873, 878, 1240
Mathematicians' proof reading
arguments, 2656, 2657
components, 2658
correctness, 2653–2656, 2658, 2659
eye-tracking studies, 2658, 2659
hypothesis generation *vs*. hypothesis testing, 2650, 2651
lectures, 2650
mathematical community, 2650
mathematical practice, 2658
proof validation, 2651–2653
substantial training, 2650
textbooks, 2650
Mathematicians reflections on practice, 4
Mathematician's work, 3060

Mathematics, 870–874, 877, 878, 2290
as a cultural practice, 911
apps and computer applications, 1228
central paradigm, 1225
city model, 1226
contradiction and ambiguity, 2304–2305
creativity and insight, 2301
creativity as reification, 2301–2303
definition, 1607
disciplines, 1607
division, 1611
education, 1132, 2564, 2565, 2702, 2787, 2792, 2930–2933
elementary mathematics, 1598
Euclidean geometry, 2298–2299
formalism as process, 2297–2298
historiography of, 912
history of, 1310–1312
identity and difference, 1313, 1314
infinite, 1598
intuition/fast thinking, 2300
knowledge fields, 1225
linear and as fractal, 1318
logic, algorithms and human mind, 2305
machines, 3193
mathematician(s), 1596, 1603, 1612, 1629
metric and topological, 1320
negation and indeterminacy, 1317
notion/concept/understanding, 1596–1598, 1604, 1607, 1627
objectivism and absolutism, 1238
parallel and intersectional, 1321
performativity, 1247–1250
perspective of process, 2292–2294
practices, 1225
presentation, 1599
problematic applications, 1246, 1247
process of reason, 2296–2297
public communications, 1244, 1245
pure, 1608
and purity, 1238–1241, 1250
rationality, 2299
reason/revelation, 2300
shared concepts, 1225
and social order, 2598–2599
social practice, 1228
sociologies (*see* Sociologies of mathematics)
soul, 1626
specialists, 1250
strong and weak objectivity, 2303
teaching, 2556, 2557
in twentieth century, 850
understanding and learning, 2295–2296
virtue theory, 1242–1244
work(s), 1596–1599, 1603, 1604, 1606, 1626
Mathesis biceps, 170, 171
Mathoverflow, 2683
Math worlds, 2608, 2611
Matière, forme et catastrophes, 407
Maupertuis, P.L., 101
Maximality thesis, 2195
Maximum likelihood (ML) estimators, 1438
Maxwell, James Clerk, 1390
Mayan calendar system, 1101
Meaning/intention *yi* 意, 2148, 2155–2157
Meaning of the algorithm, 2152
Meaning of the computations, 3051
Measurable expressions, 1615
The Measurable Figures, 1778
Measurable numbers
addition, 1622
BBGA, 1601
Bolzano's definition, 1623
Bolzano's project, 1606
multitude, 1621
non-rational, 1618, 1619
notions, 1627
not-yet-modern conception, 1597
number continuum, 1597
real numbers, 1596
RZ, 1607
Measurable quantities, 1601, 1628, 1629
Measure theory, 2369
Mechanical procedures, 2192
Mediaeval developments, 2764
Mediation, 3008, 3013, 3018
Medical ethics
beneficence, 1232
clear and exemplary offers, 1233
Hippocrates, 1232
judgements, 1233
justice, 1232
non-maleficence, 1232
principles, 1233
respect for autonomy, 1232
Mediveal scholasticism, 103
Mekhmat, 870, 872–877
Memorized, 3062
Mendel, Gregor, 1393
Mendelejev-effect, 2822
Mendelian populations, 1394
Mendelians, 1385
Menelaus of Alexandria, 1724
Menelaus' theorem, 1724, 1746–1785

Menger, K., 888–892
Meno, Socrates, 2149
Meno, 466, 473–475, 477–490, 492–496, 498, 507–509, 514–517, 2140–2142
Mentality, 3051
Mereotopology, 1320, 1321
Mersenne, Marin, 171, 1746
Mersenne number, 176
Metalanguage elements, 2222
Metamathematics, 2341–2343
Metaphor, 18, 21, 23, 25
Metaphorical model, 795
Metaphysical realism, 663
Metaphysics, 2459
Métaphysique extrême, 398
Methodological power, 3111
Metric geometry, 1726, 1886, 1891, 1897
Metric spaces, 1887, 1888, 1890, 1892, 1893, 1896
Metrological lists, 3062
Metrological table for surfaces, 3063
Metrological tables, 3062, 3064
Metropolis algorithm, 1377–1380
Microcosmos, 23, 834
Microstates, 1390
Midpoint path construction
  experimental research project, 2518–2519
  genesis and explanation, 2510–2513
  geometric experiments on paper, 2513–2515
  geometric experiment with computers, 2516–2518
Midpoint series, 2511
Mill, J.S., 94
Miller–Rabin primality test, 2071
Minimalization, 1942
(Mini-)Polymath, 2683
Minkowski spaces, 1890
Minkowski's theorem, 2534
Minority mathematics, 750–755, 757, 758, 760, 762, 765–769, 773–775, 781, 782, 785, 786, 791, 795
Minor mathematics, 1311, 1313
Miscellanea Mathematica, 1597, 1625
Mistake, 3060
Mizar proof language, 3173
Modal logic of forcing, 110, 124, 125, 128
Modal notions, 2883
Modèles mathématiques de la morphogenèse, 386, 395
Modelling stages, 2577
Modelling the real world, 2576–2582
Model theory, 1907, 1908, 1990, 1992, 2369
Modern algebra, 766, 767, 775, 779–781, 784, 3179
Modern coalescent theory, 1399
Modern European mathematics, 3040
Modernism, 69, 70, 754, 1310, 1311
Modernist geometry, 72
Modernity, 69, 75
Modern mathematical practice, 3066
Modern mathematics, 750, 757, 766, 768, 769, 772–775, 779, 791
Modern physics, 64
Modern rewriting of the proof, 2159
Modern science, 2139
Modern set theory, 2375
Modern vulgate, 407
Modularity, 1919, 2024–2025
Modular representations, 2027
Moment, 3063
Monadological concept, 743
Monologue/dialogue/polylogue arguments, 2667
Monomials, 201–209, 221
Monotonicity, 1768–1769
Monster-barring, 143
Monte Carlo simulation, 1350, 1399
Monte Carlo technique, 1352
Montucla, Jean-Étienne, 219
Moral philosophy, 900
Mordell's conjecture, 2022
Morgenstern, O., 105
Morphism, 784–785
Morphogenesis, 394–396
Morphology, 414, 1388
Morris, Rebecca Lea, 2041
Morse theory, 412
Motchane, Léon, 396
Motivation, 2150
  Morris's account, 2099
  Pólya's account, 2097–2099
Motive, 722
Multiplication, 179, 1154, 2150
Multiplicative type, 3062
Multi-spoof perfect number, 2274
Multitude
  arbitrary, 1604
  determinable, 1615
  finite, 1601, 1603, 1611, 1615, 1619
  of ideas, 1613
  indeterminable, 1615
  infinite, 1611–1615, 1626
  innumerable, 1627
  of measurable numbers, 1619, 1621
  of numbers, 1601, 1612, 1614

Multitude (*cont.*)
of operations, 1611, 1614
of points, 1613
of quantities, 1615
of rational number, 1613
of things, 1598, 1626
Museion (museum), 1728, 1740
Mutilated chessboard problem, 2535, 2537, 2539
Mysticism, 689

N
Nameable set, 837, 857
Napier, J, 170
Naproche-SAD, 2935
Narodnaya Volya, 808
National Mathematics Advisory Panel (NMAP) Report, 1166
Natural language processing (NLP), 2182, 2185, 2448, 2957, 2962, 2967, 3207, 3217
Natural laws, 104
Natural numbers, 1475, 1970, 1971, 2322, 2323, 2326, 2329, 2332, 2333
Natural philosophy, 395
Natural products of ordinals, 1700
Natural science, 470
Natural selection, 93, 1386, 1387
Natural sum, 447
Nature of the signs, 2146
Negropontis, S., 67, 466
Neointuitionism, 1513
Neo-platonism, 11
Networked-numbers Game, 1197–1202, 1204
Netz, Reviel, 3004, 3022
Neural nets, 3196
Neuralyzer, 1367, 1368
Neutral geometry, 725, 1913, 1964, 1965
New Math movement, 1169, 2350, 2555
New perspectives, 4
Newton, Isaac, 1180
Newtonian mechanics, 101
Newtonian quantity of matter, 300
Newton's advancements, 1180
Newton's color wheel, 1183
Newton's Opticks, 913
Newton's physics, 3074
New type of mathematical statement, 3061
New types of statements, 3061
Neyman-Pearson logic, 1442
Nicomachus, 212
Nietzsche, F., 68, 70
Niggli, Paul, 1851
The Nine Chapters on Mathematical Procedures, 2143, 2145, 2146, 2150, 2151, 2153, 2154, 2157–2159, 3049, 3050, 3055, 3057, 3066
Nippur, 3062, 3064
Nirūḍhapaśubandha vedi, 1046
Noble Sophistry, 474, 475, 497, 498, 502
Noether's theorems, 779
Nominalism, 2311–2313, 2330
Nominalist, 2311–2315, 2317, 2319
Non-anthyphairetic approach to incommensurabilities, 996
Non-Archimedean ordered field, 1703
Non-commutative algebra, 2366
Non-congruence theorems, 1763
Nondecimal number systems, 169, 171, 172, 177, 191–193
Non-deductive evidence, 2402–2407, 2410, 2412–2415
Non-deductive knowledge of mathematics, 2412, 2415
Nondegeneracy of Δ, 1537, 1546–1548
Nondictatorship, 2377
Non-Euclidean geometry, 96, 123, 648, 707, 712, 715, 723, 725, 731, 883, 1512, 1723, 1725, 1799, 1812, 1908, 1966, 1969, 2293, 2366
Non-Euclidean metric space, 2531
Non-Identity (NI) principle, 537
Nonlogical vocabulary, 2221
Non-maleficence, 1233
Non-Ptolemaic tables, 1741
Nonstandard analysis (NSA), 1224, 1470, 1549–1551
and hyrerreal line, 1533–1536
and SIA, 1552–1553
Non-standard models, 1906–1909, 1916, 1922
Nonstandardness, 1922
Nonstandard sets, 1534
Nonstandard universe, 1469
Nontrivial automorphism, 1923
Non-vacuism, 2284–2286
Non-verbal signs
Euclid's geometry, 2976–2979
norms in contemporary mathematics, 2979–2980
philosophical ramifications, 2980–2982
Non-Western mathematics, 2806
Norms, 2070, 2072, 3058, 3067
Notations, 3057, 3059, 3174
Novikov, Sergey, 872, 873
Number line, 246

Number theory, 448–449, 1223, 1224, 2764
Numeration system, 2144
Numerical evidence
  evaluation, 2716
  overwhelming, 2723
  in pure mathematics, 2714
Numerical methods, 2561, 2563
Numerical observation, 2433, 2434, 2450
Numerical values and arithmetical operations, 2148
Numerical variable, 2563
Nyañchana method, 1043, 1064
Nystrom, J, 169, 188, 190

O
Oberwolfach, 1302
Objective Bayesianism, 1416, 2716
Objectivity, 2848–2851, 2856, 2857, 2860, 2861, 2863–2868, 2870, 2872, 2873
Objectless, 1605, 1606, 1608, 1610–1612
Observation, 2433–2435, 2438, 2441–2443, 2447, 2448, 2450, 2453, 2454, 3061
Observation indicator, 2448, 2450
Observer-independent reality, 1425, 1426
Observers, 3050
Occam's razor, 887
Oe-polynomial, 1196
Oereo polynomials, 1196
Oereo sequences, 1193
Oe-sequence, 1192
O'Halloran, Kay L., 3004
Old-Babylonian, 3064
Old-Babylonian Mesopotamia, 3054
Old-Babylonian period, 3062
o-*minimal* (order-minimal), 1909
1-Eliminator, 450, 451, 457
One kosmos model, 536
Online mathematics interactions, 2611
Ontology of mathematics, 3018–3020, 3023, 3028
Open exploration, 3056
Open interval topology, 1549
Operational calculus, 2986
Operationalism, 1427
Operations, 2144, 2159, 2160
Operations qua operations, 2153
Opticks, 1181, 1186
Oracle Turing machines, 1954
Oral performances, 3064
Order axioms, 1530
Ordered field of line segments, 246
Order of λ_α, 2425, 2426
Order of presentation, 3143–3144
Order-preserving permutation, 1926
Order relation, 1524
Ordinal notations, 447–448
Ordinal numbers, 1956
Orientals and computation, 2141
Origin, 1385
Original proof, 3060
Origin of forces, 303
The Origin of Species, 1384, 1387, 1390
Origins of mathematics, 10, 11
Orismos, 408
Ottoman empire, 902
Oughtred, William, 227, 230
Ouranos, 537
Overbeck, T, 181
Oxymoron, 400

P
Pacioli, Luca, 206
Padding Lemma, 1945
Padinan, 1109
Painting, 26
Paleontology, 1386, 1388, 1391
Palindromically periodic theorem, 552
Palindromicity, 918–920, 939, 961, 962, 966–968, 970–976, 980, 982–984, 988, 990, 993, 997, 998, 1004
Palindromic periodicity
  myth for division and collection of the Statesman, 969–972
  myth with Kronos era and Zeus era, 969
Pancawara cycle, 1105, 1106, 1109
Pangeometry, 723, 726
Paper tools, 3056
Pappus, 630, 1746, 1748, 1773–1776
Pappus of Alexandria, 231, 232, 1732, 1734, 1737–1741, 1773–1775
Pappus problem, 251–253
Paradigm shifts, 22, 23, 2144
Paradoxes and foundations, 4
Paradoxes of the Infinite (PU), 1605, 1606, 1615, 1626
Parallel axiom, 1913, 1964–1966
Parallelism, 1891
Parallelogram law, 72
Paramagnetism, 16
Parameter estimation, 1424
Parametric polymorphism, 3174
Pareto efficiency, 2377
Parimaṇḍala kūrmaciti, 1039
Paris-Harrington theorem, 1982

Parmenides, 473, 485, 492, 493, 499, 512, 513, 516, 517, 523–524, 527–528, 534–535, 537, 548, 561, 568, 591, 596, 925–928, 950, 987
Parmenides dyad
  equalization of parts, 927
  infinite anthyphairesis, 926
  logos criterion, 926
  self-similiar property of one and many, 927
  true intelligible Being, 928
Partial computable functions, 1938, 1939
Partial Differential Equation (PDE), 2557, 2558
Partial fraction decomposition, 1652
Partial recursive functions, 1939, 1942, 1955
Pasaran day, 1108, 1109
Pascal, Blaise, 102, 804
  discourse, 3083
  triangle, 1194
Pasternak, Boris, 808, 848
Pauli, JW, 175
Pax romana (Roman peace), 1729
Peano arithmetic (PA), 446, 447, 454, 1908, 1914–1916, 1920, 1990, 2013, 2028–2030, 2220
Peano axioms, 1972
Peano second-order axioms, 2333
Peano's formulation, 1223
Peano syntactic-transition rules, 2200
Pearson, Karl, 1393
Peirce, Charles Sanders, 1506–1509, 2956
Peletier, Jacques, 206, 208–210
Peletier's algebra, 3084
Pell equation, 984, 987–992, 997–1005, 1008–1010, 1013, 1015, 1016
  by Bhaskara II Cakravala, 1005–1008, 1010–1012, 1015, 1016
  by Brahmagupta, 1000, 1006
  Brahmagupta's bhavana/composition, 1000
  Hindus' solution to, in Europe, 999, 1000
  Pigeonhole principle, 953, 966, 979, 984
  by Pythagoreans, 984–987
  tools for proof, 988, 990
Perceptual inference, 2246
Perceptual invariants, 2536
Perceptual manipulations theory, 2913
Perceptual training, 2988
Perfect numbers, 172, 175, 177
Perfectoid theory, 2621, 2623, 2624, 2627, 2629, 2632
Perfect set theorem, 1983
Periodic anthyphairesis, 472, 502, 504, 506, 510, 512, 513, 516, 517, 524–528, 532, 534, 548, 552, 555, 562–564, 582, 584, 586, 591, 596
Perpendicular bisector, 2577
Perpetual calendar, 1098
Personal ethics of mathematicians
  ethical norms, 1231
  judgments, 1231
  knowledge construction, 1230
  mathematical practice, 1232
  mathematical results, 1231
  peer review and revised, 1231
  responsibilities, 1231
  rewards and advancement, 1231
  scientific results, 1230
Personal idiosyncrasies, 3057
Personal reality, 1425
Personal training, 3063
Perspectivist form, 3127
Persuasion tool, 2493–2500
Persuasive/directive/polemic arguments, 2669
Petimus, 1796
Petitiones, 1795, 1802
Petitiones quas axiomata vocant, 1795
Peucer, Caspar, 208
Phaedrus, 935, 937, 973
Phenomenological flavor, 2236
Phenomenological perspective of history, 1142
Phenomenology, 2186–2188, 2237
Phenotype, 1386
Philebean infinite, 560
Philebus, 466, 917, 919, 925–927, 935, 936, 938, 945–948, 950–953, 961, 979, 996
Philosophical analogue, 522, 527–531, 548, 591
Philosophical Essay on Probabilities, 1388
Philosophical ideas, 900
Philosophical Investigations, 2602, 2606
Philosophie de la singularité, 411, 412
Philosophy, 184–187
Philosophy of action, 2786, 2789
Philosophy of mathematical practice (PMP), 2638, 2795, 3005, 3020
Philosophy of mathematics, 14, 160, 161, 2308–2311, 2314, 2317, 2318, 2320–2323, 2760
  applications, 1304, 1305
  epistemological *vs.* social models, 1225, 1226
  human verification, 1222
  individual/joint work, 1302, 1303
  level of individual experience, 1306
  mathematician, 1221, 1222
  (M)athematics *vs.* (m)athematics, 1223, 1224
  numbers/arithmetic calculations, 1299
  number theory, 1223

open and democratic, 1222
periodic function, 1303
philosophical paradoxes, 1298
positions, 1222
traditional, 2794
unique and unified subject, 1224, 1225
well-composed theory, 1303
Philosophy of probability, 1369
PhiMSAMP, 2308
Phrase structure languages, 1937
Phylogenetics, 1386
Phylogeny, 24
Physical continuum, 1511
Physical reality, 68, 76, 79, 81, 84, 86
Physical theory, 76, 77, 79, 80
Piaget's reflecting abstraction, 2788
Pictorial form, 3119
Picture proofs, 2524, 2535, 2538, 2539
Pierce, C.S., 95
Pigeonhole principle, 918, 966, 980, 981, 983, 984, 995
Pisano, Leonardo, 1330, 1331
Planck scales, 1560
Plato, 9, 11, 30, 36, 37, 55, 56, 65, 67, 68, 83, 93, 106, 473–475, 478, 484–487, 489, 492–495, 497, 498, 503, 508–510, 512, 514–517, 602, 603, 606–615, 617–624, 627–641, 1463, 2140, 2142
dialogues, 466
indivisible line, 467
Meno, 30
methodos, 1795
Parmenides, 602, 603, 607, 612–615, 619–629, 637, 639
Phaedrus, 602, 603, 608–612, 620, 635–639
philosophy, 466, 467
problem, 2316–2319
Receptacle, 467
Republic, 602, 603, 611, 624–630
Sophist, 602, 607–614, 617, 620, 625, 627, 630, 635, 636, 639
Theaetetus, 633
Timaeus, 635, 636
Platonic idea, 944
Platonic ideal of mathematics, 1130
Platonic option, 2229
Platonic realism, 8
Platonism, 8, 11, 29, 30, 68, 83, 84, 86, 92, 93, 95, 764, 1223, 1224, 2291, 2295, 2308, 2311, 2312, 2320, 2322, 2323, 2330, 2334
Platonist, 1240, 2312, 2313, 2315, 2316, 2318–2321, 2326, 2331
Plausible reasoning, 2434
Playfair, John, 1809, 1810
Playing cards, 1369
Pluralism, 4, 1802, 2394–2397
in foundations, 2390–2394
in mathematics, 2384
signs of, 2390
Pluralist, 2252
Plurality of mind, 683
Poetry, 831, 847, 848, 850, 861
Poincaré, Henri, 1509–1512, 1722, 1723, 1725, 3060, 3061, 3068
Poincaré disc model, 2531
Pointed topological space, 3180
Point groups, 1828–1843
Point mass centroid, 2574, 2575
Polar triangle, 1779, 1780
Politeia, 918, 932, 935, 952, 974, 997
Politicus, 918, 919, 925, 935–939, 945, 949, 951, 952, 961, 962, 967, 969–971, 973, 975, 976, 979, 980, 983, 984, 998
double measurement, philosophical, 940
double measurement in, 949
geometric mean of dyad, 950
initial steps of, 967
intelligible generation, 950
Philebean finite, 950
Politicus division rules of the game, 972–973
Pólya, György (George), 1850–1851
Pólya's conjecture, 2827
Polyhedra formula, 143
Polyhedron, 2696
Polymath project, 2592, 2728–2731
affiliations and joint affiliations, 2743–2746
description of data, 2737–2740
features, 2733
global participation data, 2740–2743
perspectives on Polymath, 2731–2734
Polymath as e-research, 2734–2737
relative weight of participants, 2746–2750
specific project communities, 2750–2753
Polymath-wiki, 2737
Polynomials, 2016
cubic, 2017–2018
equations, 2562
modern, 205, 207, 221
premodern, 209
Poncelet's geometry, 2994
Pontriagin's theorem, 1565
Popper, Karl, 93
Population genetic models, 1397
Position-of-speech tagging (PoS), 2449
Positions, 3066
Positivism, 2602

Posterior probability, 1452
Post-formal proofs, 434
Postmodernity, 69
Post-structural semiotics, 3014, 3015, 3018
Power, 918, 919, 930, 938–941, 943–945, 954, 979, 983, 995
Power notation, 3141–3144
Power Set Axiom, 2371, 2381
Practical applied mathematics, 1228
Practical certainty, 2141, 2146, 2149, 2153, 2159, 2160
Practical values
  discovery, 2085, 2086
  mathematical communication, 2086
  power of different methodologies, 2084
  systematization, 2085, 2086
  transferability, 2084
Practice, 2144, 3049, 3053, 3055, 3056, 3066–3068
  proof, 2148
  of sharing written productions, 3063
  of teaching, 3064
Practicing mathematician, 3057
Pragmatism, 2795
Pragmatist theory of science, 401
Pranatamangsa, 1120, 1121
Prauga citi, 1089
Predicate logic, 1973
Predicative mathematics, 1916
Predicativism, 459
Prediction quality, 1423–1424
Pre-Euclidean dialectics, 2555
Pre-Eudoxian theory of magnitudes, 922
Pre-formal proofs, 434
Pre-Gregorian calendars, 1101–1102
Premodern numbers, 201, 203–208
Pre-Platonic dialectics, 2555
Prescriptions, 2138, 2148
Presentist internalism, 1135
Pre-Socratics, 8, 68, 83
Primary bodies, 524–526, 538, 549, 552–554, 557–559, 561–567, 569, 573, 574, 577, 579–586, 589–590, 593–596
Primary continuum, 1503
Primary earth (PE), 524, 527, 552–553, 555–557, 561, 566, 567
Primary function, 3069
Primary water (PW), 526, 552–553, 555–557, 561–562, 564, 566–568, 570, 575–579, 586–590, 592
Prime factorization app, 2565
Primeness, 152
Prime numbers, 137, 172
Prime number theorem, 2101, 2405, 3160
Primitive p-th root of unity, 2016
Primitive recursive arithmetic (PRA), 453
Principia Mathematica, 1973, 1975, 2607
Principle of composition, 1005
Principle of continuity, 1465
Principle of fallibility, 2697–2698
Prior probability assessments, 1415
Private mathematical practice, 3058
Prix Peccot, 2622, 2623, 2627
Prizes, 2623, 2627–2629
Probability models, 4, 1364–1369, 3061
  compatibility logic, 1419–1421
  epistemic-probability-as-model, 1450–1453
  frequentism-as-model, 1428–1440
  interpretative equivalence, stability and robustness, 1440–1441
  inverse probability logic, 1421–1422
  mathematical models and reality, 1425–1427
  model assumptions, 1445–1448
  prediction quality, 1423–1424
  reality, observations, operations, 1427–1428
  significance tests, frequentist inference, 1441–1445
Probability theory, 77, 79, 2767
Probability zero, 3061
Problem, 3059
Problem choice, 2679
Problem of induction in mathematics, 2722–2723
Problems, 1364–1366, 1369, 2150
Problem-solving procedures, 2038, 2732, 3083
Procedures (*shu* 術), 3054, 3065
Process/product arguments, 2667
Processes, 3067
Proclus, 477, 494, 506, 514, 515, 518, 602, 611–613, 619, 629, 630, 935, 947, 950, 984, 985, 1792, 1794, 1795
  Commentary to Timaeus, 611, 612
  in Euclidem, 606, 629, 630
Production of mathematical concepts, 3061
Product rule, 1539
Professional mathematicians, 1244
Program, 94, 102
Programmable graphing calculators, 2561
Programming languages, 3158
Projective form, 3124, 3127
Projective geometry, 73, 822, 823, 3025
Proof-generated definition, 149
Proofiness, 1245

Proofs, 4, 19, 25, 2139–2142, 2152, 2160, 2402–2406, 2408, 2409, 2411, 2412, 2414, 3049, 3058, 3171–3173
activity, 2914
based mathematics, 1239
and computation, 2139, 2141, 2159
and deduction, 2554, 2555
as dialogues, 2075
epistemic autonomy, 2071
as narratives, 2075
norm challenge, 2076
as recipes, 2075
and refutations, 2555, 2680
result of dialogues, 2076
sketches, 2074
validation, 2076
Propensity, 1414
Proportions
Euclidean, 236–239
geometrical analysis, 262
use of, 238–239
Propositional attitudes, 2880–2882
Propositional axiomatic approach, 2346
Propositional calculus, 2293
Propositional equality, 3165
Propositions, 3049
Protein space, 1402
Proto-mathematics, 2805
Prover/Skeptic dialogues, 2683
Pseudo-freegan set theory, 2326
Pseudoholomorphic curve, 739
PSLQ program, 2830, 2831
Psychoanalysis, 415–416
Psycholinguistics, 17, 19
Ptolemaic astronomy, 1739, 2355
Ptolemy, C., 22, 201, 212, 215, 217
Publications, 3058
Public practices, 3058
Published texts, 3068
Pure and applied practical work, 1228
Purely Analytic Proof, 1603
Pure mathematics, 1240
Pure practical mathematical practices, 1227
Pure Theory of Law, 682
Pure Theory of Numbers (RZ), 1596, 1606
Purism, 1238–1241
Purity
descriptive sense, 1238
disciplines and serves, 1241
ethical neutrality, 1239
ethical values, 1241
history of mathematics, 1239
ideologies, 1240
mathematics, 1238, 1239
ulterior motives, 1238
unmixed/undiluted, 1238
valuation, 1239
Puruṣa, 1058, 1059
Pushkin, A., 15
Pushkin's poetry, 845
Putnam, Hilary, 2896
p-values, 1419, 1420, 1422
Pythagoras principle, 8, 10, 11, 18, 247–249, 477, 488, 605, 611, 632–634, 731, 1036, 1037, 1041, 1046, 1052, 1054, 1055, 1057, 1964, 2057, 2082, 2546, 2762, 2832, 2886
Pythagorean mathematics, 64, 65, 83–87
discovery of, 64
geometry and algebra, modernity to modernism, 69–74
original Pythagorean attitude, 65
radical Pythagorean mathematics and mathematical modernism, 74–83
Pythagorean ordered field, 1710
Pythagorean proof of incommensurability, 475, 477, 488
Pythagorean triples, 1036, 1037, 1041, 1043, 1044, 1047, 1055, 1063–1065, 1093, 1918, 2013, 2014

Q

Qua algorithm, 2145
Quadratic incommensurabilities, 524, 551
Quadrilateral, 2563, 2566–2568, 2570–2574
Qualitative robustness, 1446
Qualitative science, 400
Quantifier-free formula, 1975
Quantitative science, 400
Quantities, 3062
absence of, 1599, 1606
in abstracto, 1609
actual/real, 1606, 1609, 1626
bounding, 1601
concept/notion, 1597, 1598, 1603, 1608, 1609, 1614, 1627, 1628
constant, 1601
continuous, 1607
discrete, 1607
division, 1596
expression/idea, 1606, 1609, 1614, 1621
general theory, 1606, 1625
imaginary, 1603, 1606, 1608, 1609
infinite, 1608, 1614, 1622
irrational, 1600–1603, 1605, 1609, 1614

Quantities (*cont.*)
measurable, 1601, 1628
nature of, 1627
negative, 1603
numerical, 1616, 1628
ω/Ω, 1603
rational, 1608
science of, 1596, 1598, 1607, 1627
theory, 1597, 1605, 1626
transcendental, 1599, 1601
unmeasurable, 1609
variable, 1601, 1603
Quantum and relativity theory, 1227
Quantum electrodynamics, 720
Quantum field theory (QFT), 79–81, 720
Quantum mechanics (QM), 69, 72, 75–82, 727, 738, 902
Quantum phenomena, 734
Quantum physics, 704, 712
Quantum theory, 66, 70, 74–76, 79–81, 84, 712, 722, 727, 901
Quasi-empirical theory, 764, 2435
Quaternary arithmetic, 191
Quaternary number system, 168, 171, 172, 191, 193

R

Race, 1316, 1317, 1321
Racial injustice, 1251
Radical consequences, 1469
Radical Pythagorean mathematics, 71, 74–83
Radical translation, 3081
Radon–Nikodym derivative, 3170
Railroad building, 2562
Rainbow Newton, 1184, 1185
R-algebra, 3169
Ramsey theory, 1982
Ramus, P, 171
Randomness, 1414
Random variables, 1373, 1374
Rathacakraciti, 1051, 1075, 1090, 1091
Ratio, 249, 252
Rationality, 2297, 2299, 2303–2305
Rational numbers, 1475–1478, 1548, 1971
Rational probabilistically/statistically predictive theory, 82
Rational representational theory, 82
Rational structure, 2916
Ratios of infinitesimals, 1541
Readers, 3048
Readership, 3049
Real analysis, 1466
Real closed field, 265
Real-closed ordered fields, 1700, 1706
Real elliptic curve, 1919, 2019
Realism, 8, 727
Realistic portfolios, 1341
Reality-without-realism (RWR), 76–80, 83, 84, 86
Real number, 1165, 1468
Brentano, Franz, 1501–1506
Brouwer, L.E.J., 1512–1515
Cantor-Dedekind theory and arithmetization of continuum, 1481–1496
constructive and intuitionistic mathematics, 1521–1533
decimal representation of, 1478
du Bois-Reymond, Paul, 1496–1501
nonstandard analysis and hyrerreal line, 1533–1536
origins of, 1474–1481
Peirce, Charles Sanders, 1506–1509
Poincaré, Henri, 1509–1512
smooth infinitesimal analysis, 1536–1560
Weyl, Hermann, 1515–1521
Receptacle and content, 526, 532, 543–547, 564, 573, 585, 592
Recipe model of proofs, 2060, 2061, 2927–2929
Recollection, 474, 475, 478–483, 485, 486, 489, 491, 493, 494, 507, 508, 514–518
Rectification, 1776–1778
Recurrent neural nets (RNNs), 3196
Recursion theorem, 1949
Recursive comprehension axiom, 1915
Recursive formulas, 966
Recursive functions, 1910
Recursive law, 2492
Reed, D., 72, 73
Referees, 3049
Reference potential, 143
Refutations, 2564–2570
Regularities, 3056
Regular languages, 1937
Re-idealization, 2541, 2542
Reification, 2291–2293, 2301, 2322–2327
Rein analytischer Beweis, 1579–1580
Lagrange on intermediate value theorem, 1581
two functions pattern, 1580–1581
Reinforcement learning, 3196, 3213
Relational data, 2739
Relational organization, 709
Relativisation, 2371
Relativity, 76

Reliabilism, 2883
Religion, 902
Religious inspirations, 2768
Religious metaphors, 2760
Renormalization, 720
Replacement, 2374
Replacement axiom scheme, 2381
Representational abstraction, 165
Representation of computation, 2139
Representations, 2422–2424, 3067
Republic, 67, 474, 485, 487, 489, 495, 497, 506, 510, 512, 513, 516, 517
Research process, 3056
Resemblance interpretations, 2222
Resources, 3059
Restivo, Sal, 2603, 2604, 2607, 2608, 2611
Restore the ephemera, 3067
Restoring of inscriptions, 3066
Restricted solution domain, 2579
Reverse mathematics, 26–28, 831, 859, 1471, 1907, 1912, 1913, 1915, 1916
  arithmetic comprehension, 1981
  arithmetization of analysis, 1969–1971
  arithmetization of computation, 1975–1976
  axiom of choice, 1966–1969
  axioms for numbers and sets of numbers, 1972
  bounded value theorem, 1977
  computability, 1973–1975
  parallel axiom, 1964–1966
  theorems and nontheorems of ACA0, 1982–1983
  theorems and nontheorems of RCA0, 1976–1977
  theorems and nontheorems of WKL0, 1979–1980
Reviel Netz, 3049
Revolutionary vacuism, 2283
Rewriting, 2156, 2159
Rhetoric and mathematics
  dialectical heuristic for informal mathematics, 2694–2697
  mathematical discourse, 2700–2705
  principle of fallibility, 2697–2698
  problem situations and historical contexts, 2699–2700
  problems with formalism, 2691–2694
  translative rhetorical force, 2705–2707
Rhind, 2145
Rhind Mathematical Papyrus, 1151
Rhombus, 1052
Ribet's theorem, 1919
Rice's theorem, 1950, 1951
Richard Dedekind, 3057
Richard's paradox, 1466
Riemann hypothesis, 2403, 2414, 2714, 2718–2720, 2723
Riemannian geometry, 64, 76, 81
Riemannian manifolds, 733
Riemann-Roch theorem, 1868
Riemann's hypothesis, 2557
Riemann's theory, 722–735
Riemann surface, 84, 85, 2021
Ries, A, 171
Right-triangular Fibonacci array, 1194
Rigor, 396, 2188, 3058
Rigorous proof, 2929, 2938
Rings of integers of algebraic number fields, 2015
Robinson, Abraham, 1533
Robinson's nonstandard approach, 1679
Robustness, 1424, 1430, 1441, 1446, 1450, 1454
Rodin, Andrei, 3026
Romé de l'Isle, Jean-Baptiste Louis, 1826, 1831
Rostropovich, M., 15
Rudolff, Christoff, 209, 210
Rules, 2138, 2140, 2142, 2147, 2159
Russell, Bertrand, 44, 93, 160, 2140
Russell's paradox, 1913
Russian mathematical community, 843
Russian poetry, 845
Russian revolution, 836
Russian Silver Age, 846
Russian Symbolist movement, 807, 845

S
Sacks, Oliver, 2294
Saka calendar, 1105
Sanskrit, 2142, 3052, 3066
S&P 500, 1343
Saptawara cycle, 1105, 1109
Śatapatha Brāhmaṇa, 1064, 1065
Satisfactory resolution, 3061
Saturation principle, 1534
Satyāṣaḍha śulba, 1026
Saussure, Ferdinand de, 2957–2959
Scalar data, 2740
Scales, 3050
Sceptical prognosis, 690
Schematic diagrams, 3041
Scheubel, Johann, 203, 206, 209
Schlimm, D., 1800
Schlimm-Feferman framework, 1801
Schoenflies, Arthur Moritz, 1845, 1847–1849

Scholastic assessment test (SAT), 3210
Scholia, 995
Scholze, Peter, 2620, 2623, 2624, 2626–2629, 2631–2633, 2635
School and teaching
  biological sciences, 901
  elementary school, 898
  geography, 900
  history teacher, 898
  maths teacher, 899
  philosophy, 901
  physics, 901
School documents, 3064
School geometry, 1240
School texts, 3068
Schwartz, Laurent, 2620–2624, 2626, 2627, 2631, 2632, 2634
Schwartzenberg, L., 16
Schwenter, D, 192
Scientific calculators, 2561
Scientific principles in mathematics
  areas as measure of cardboard weight, 2574
  balancing point outside concave quadrilateral, 2573
  centroid of a cardboard quadrilateral, 2572
  centroids/centres of gravity, 2570, 2571
  Fermat point, 2575
  Fosbury flop in high jumping, 2573
  lamina centroid, 2575
  law of the balanced lever, 2570
  point mass centroid, 2575
  Steiner tree network for a square, 2576
Scientific revolutions, 21
Scientific theories, 893
Scientific tradition of mathematical knowledge, 1150
Scriabin, A., 23
Scribal schools, 3062
Scribes, 3062, 3064
Secondary continuum, 1503
Secondary intuition, 1364, 1365
Second-order logic, 1957
Second order ZFC set theory, 2375
Second problem, 1364
Sector Figure, 1769, 1771–1772
Sedecimal, 168, 169, 188–191
Seidenberg-Tarski theorem, 1927
Self-contradictory, 1603, 1605, 1613
Self-supervised learning, 3196
Semantic field, 2153
Semantic-transition rules, 2184
Semantics, 2849–2851, 2861, 2863, 2864, 2867, 2868, 2872
Semicircle, 1063
Semiformal language, 3153
Semiological system, 3042
Semiology of mathematical practice, 4, 2958
Semiophysics, 416
Semiotic components, 2802
Semiotic duality, 3038
Semiotics, 2930, 2931, 2958, 3120
  commutation tests, 3007
  mathematical example, 3009–3014
  ontology, 3018–3021, 3024
  post-structural semiotics, 3015–3017
  structuralist, 3005–3007, 3009
Sensed pattern, 2547, 2548
Sensory imagination, 2524, 2543, 2544
1-sequence, 450–452, 457
Sequence of operations, 2142
Sequential/parallel arguments, 2668
Set content, 3139
Set-theoretic abstraction
  functions, 3175–3177
  quotients, 3178–3179
  sets, 3177–3178
Set-theoretic geology, 110, 111, 124, 125, 127, 128
Set-theoretic mind set, 2537
Set theory, 66, 70, 75, 76, 83, 86, 110, 111, 116, 122, 130, 439, 441, 805–807, 809, 836, 837, 839, 840, 855, 1681, 1907, 1908, 1912–1914, 1916, 1967, 2409–2412, 3160, 3164, 3165, 3178
  axioms of set theory, 2370
  axioms of ZF(C) and deductions, 2370
  epistemic pluralism in, 113
  foundational role in mathematics, 2369
  modelling power, 2369
  natural interpretation, 2369
  philosophical questions, 2369
  set-theoretic knowledge, 2376
  silent majority, 2376
  social ideas, 2369, 2370
  social concepts, 2370
  societal interpretation, 2369, 2374, 2375
Sexagesimal place-value system (SPVS), 1155, 1156
Shakespeare, W., 15
Shapiro, S., 48–51
Shelley, P.B., 82
Sher, Gila, 2899
s-hierarchical Archimedean arithmetic continuum, 1705–1707
s-hierarchical mapping, 1693
s-hierarchical morphisms, 1693

s-hierarchical ordered algebraic structure, 1681
s-hierarchical ordered fields, 1689–1692, 1695–1696, 1706
s-hierarchical structure, 1679
Shikhanovich, Yuri, 877
Shimura-Taniyama-Weil conjecture, 1919
Shrödinger, E., 21
Shubnikov, Alexei Vasilievich, 1855–1857
Sibṭ al-Māridīnī, 206
Side and diameter numbers, 956, 958, 966, 984–986, 1011
Sigma functions, 2279–2281
Significance tests, 1442
Signified, definition of, 3136–3137, 3148–3149
Sign of the curvature (Busemann), 1892
Signs in mathematics, 2146, 2972
  cognition, 2988–2989
  computational limitations, 2990–2992
  for finite agents, 2987–2995
  non-verbal signs, 2976–2980
  open-endedness and historical development, 2985–2987
  philosophical stakes, 2992–2995
Silver Age of Russian Mathematics, 469
Silver Age of Russian poetry, 849
Simil-Proof (SP), 2900
Simon Says Analogy, 2434
Simple apotome line, 918, 954, 955, 958, 988–990, 993, 998, 1001, 1003
Simple harmonic motion, 325
Simplicius, 602, 611–613, 619, 620, 636–639
Simson, Robert, 1810
Sine Rule, 1769–1771, 1779–1781, 1785
Single-component concept, 709
Single real number, 1971
Singular value decomposition (SVD), 3203
Situated proofs, 2583
Skolem's discovery of nonstandard models, 1922
Skolem's model, 2000–2003
Skolem's theorem, 439
Skyscraper metaphor, 1226
Small scale/large scale arguments, 2667
Smooth infinitesimal analysis (SIA), 1469, 1536, 1543, 1549–1551
  arithmetical axioms, 1545–1546
  cohesiveness of continuum, 1544–1545
  consistency and models of, 1551–1552
  constancy axiom, 1546–1548
  constancy principle, 1543–1544
  differential calculus in, 1538–1540
  infinitesimal linearity axiom, 1546
  infinitesimal straightness principle, 1537
  internal logic of, 1541–1542
  and NSA, 1552–1553
  and physics, 1553–1560
  smooth and Dedekind real lines in, 1548–1549
  smoothness of functions, 1540–1541
  smooth real line R, 1545
  strict order relation $<$ on R, 1545
Soap bubble demonstration, 2576
Social agency, 2791
Social constitution of mathematical knowledge
  axiomatics, 2869–2872
  conceptual structuralism, 2856–2858
  institutionalization of mathematics, 2861–2863
  interplay of practices, 2858–2860
  objectivity with/without objects, 2863–2867
  representational/non-representational semantics, 2867–2869
Social constructivism, 1224, 1425
Social contexts, 3054
Social environments, 3064
Social epistemic virtue, 2072
Social epistemology, 2786, 2789, 2889, 2890
  deductive mathematical proofs, 2070
  inductive reasoning, 2070
  mathematical community, 2071, 2072
  mathematical proofs, 2071
  permanence of mathematical results, 2072
  possibility of autonomy, 2072
  randomized proofs, 2071
Social ontology, 2786, 2789, 2790, 2850–2856, 2863, 2865, 2870, 2872, 2873
Social order, 2378, 2598–2599, 2605
Social physics, 2598
Social practice
  aspects, 1249
  definition, 1227
  external goods, 1242
  institutions, 1251
  internal goods, 1242, 1244
  mathematics, 1227, 1233, 1234, 1243, 1244
  participants, 1229, 1230
  science, 1235
  space and time, 1227
Social process
  argument, 2073
  complex proof, 2074
  mathematical proofs, 2073
  proofs as dialogues, 2076
  of proof validation, 2070
  role of mathematicians, 2074
  of verification, 2074

Social realism, 1223
Social reality, 1425
Social responsibility of mathematics, 1220, 1235, 1236, 1242, 1250
Social responsibility of science, 1234, 1235, 1250
Society mathematization, 1249
Sociologies, 2611
  of knowledges, 2599–2601
  of mathematics, 2598–2602, 2604–2614
Sociology of scientific knowledge (SSK), 2607, 2610
Socrates, 466, 478–496, 503, 507–509, 515, 516, 1463, 2140
Soft typing, 3158
Sohncke, Leonard, 1834, 1841, 1847, 1850
Solipsism, 95
Solovetsky Monastery, 816
Soloviev, V., 23
Solzhenitsyn, Aleksandr, 15, 24, 815
Solzhenitsyn's exile, 875
Sophist, 100, 474, 475, 487–489, 493, 495, 497–499, 502–507, 510–513, 516, 517
Sophists, 100
Sophist's' division in the *Sophistes*, 917, 932, 934, 935
  cyclical nature-periodicity, 935–937
  division and collection of, 934–935
  division for, 931
  fundamental analogy in divided line of, 932
  Logoi of division snd collection, 932
  logos criterion, 932
  Sophistes, 918, 925, 926, 928, 929, 931, 935, 939, 944, 967, 971, 979
  Theaetetus, 917–922, 924, 929, 933, 937–939, 943, 945, 946, 952, 953, 955, 960, 961, 966, 970, 979, 980, 984, 996, 997, 1013, 1015
Sosa, Ernest, 2894
Soul, 603, 608–612, 625, 627, 635–639
Souslin hypothesis, 2375
South Asia, 3052
Soviet Union, 469
Space groups, 1828, 1829, 1844–1854
Space-number, 1870
Space of a page, 3056
Spatial algebra, 72–74
Spatial and dynamic entities, 3067, 3068
Spatial devices, 3057
Spatial imaginary, 1308, 1309, 1312, 1314, 1316, 1321
Spatial layouts, 3056
Specialist mathematicians, 1250
Species (in Viète and Fermat), 211, 218–223, 225, 227, 231
Spectatorial conception, 2893
Speech acts, 3049, 3051
Speiser, Andreas, 1852, 1855
Sphere, 1746–1751, 1757, 1758, 1760, 1769, 1774–1775, 1781, 1793
Spherical geometry, 1746–1748, 1756, 1760, 1765, 1768, 1773, 1774, 1779, 1780
Spherical triangles, 1772
Spherical trigonometry, 1746, 1772, 1779, 1781, 1784, 1785, 3058, 3059
Spherics, 1724, 1725
Spoof perfect numbers
  definition, 2272
  mathematically acceptable, 2275–2277
  theory, 2274–2275
Square numbers, 180
Square root, 1062, 1080, 2561
Square-shaped tablets, 3062
Śrautasūtras, 1025
Stabilité structurelle et morphogénèse, 395
Stability, 1439, 1449, 1454
Stalnaker, Robert, 2895
Standard Bayesian analysis, 1421
Standard model of arithmetic, 1908, 1909, 1911, 1912, 1916, 1918, 1922
Standard realism, 1426
Standard sets, 1534
Static arguments, 2668
Static mathematics, 2625
Statistical falsification, 1420
Statistical hypothesis tests, 1441
Statistical inference, 1412, 1413, 1419, 1422, 1424, 1428, 1430, 1438, 1440
Statistical methods, 1393
Steiner tree network for a square, 2576
Step-by-step translation, 3060
Stevin, Simon, 200, 203, 210–211, 230
Stifel, Michael, 171, 205–211, 230
Stirling numbers, 1303
Stochastic calculus, 1333
Stochastic differential equation (SDE), 1337, 1338
Straightforward solution, 2220
Strict norms, 3066
Strict rules, 3066
String figures, 3067
String theory, 97
Strong Theaetetus' palindromic periodicity theorem, 979–980
Structural and foundational axioms, 1801
Structural embedding, 3128

Structuralism, 50, 772, 775, 779, 2394–2397
Structural isomorphism, 3127
Structuralist semiotics, 3005–3009, 3136, 3137, 3139, 3140
Structural mathematics, 2026–2027
Structure, 772
  computable structures, 1958, 1960
  decidable structures, 1959
  definable relations, 1953, 1960
  elementary diagram, 1959
  identification, 3166–3167
  open diagram, 1958
  of publication, 3060
  standard model of arithmetic, 1953, 1959
  of text, 3060
Structure theory, 118
Style, 2927, 2928, 2930, 2932, 2933, 2937, 2947
Subbotovskaia, B., 878
Subbotovskaya-Muchnik, Bella, 876
Subduple ratio, 171
Subjective idealism, 673–675
Subjective probabilities, 1415
Sub-scientific mathematical tradition, 1150
Substantive models, 1437
Subtractive abstraction, 160
Successive (ephexes), 481, 484–486
Successor function, 1972
Sui generis objects, 2330
Śulbasūtras, 911
  commentaries, 1028
  concordance of sūtras, 1093
  darśpaurṇamāsika vedi, 1087
  enhancement of figures, 1038
  proofs, 1039
  Rathacakraciti, 1090
  rectilinear constructions, 1035
  square-roots, 1084
  transformation of figures, 1037
  units of measurement, 1031
  Vedis and *citis*, 1033
Sultan Agung, 1105, 1113, 1114
Sumerian base-60 numeration system, 1155
Summulae Logicales, 345
Supervised learning, 3196
Support for his research, 3059
Supremum, 1526
Surakarta, Sultanate, 1104, 1116
Surface of the page, 3056, 3057
Surreal numbers, 1508, 1679
  construction of ordered, 1684–1687
  Conway names, 1696–1700
  ordered class, 1681–1682
  ordered field of, 1700–1702
  structure of ordered, 1687–1689
Surveyability of proofs, 2070, 2465, 2468, 2473
Surveyable, 2140
Suslin, Mikhail Yakovlevich, 837, 842
Sūtras, 1029, 3052, 3053
Swerdlow's criticism, 1794
Swerdlow's critique, 1793
Syllogisms, 2184
Symbolic algebra, 2965
Symbolic notation, 2149
Symbolic reasoning, 2913
Symbolic regression, 3205
Symbolism, 820, 835, 845, 846
Symmetrical parts of text, 3059
Symmetry, 1825–1827, 2607, 3058, 3061
Symposium, 484–486, 516, 517
Syncategoremata, 345
Synchronistic effects, 2204, 3068
Syntactic algorithmic functions, 2208
Syntactic-transition rules, 2200
Synthetic-centered approach, 1810
Synthetic differential geometry (SDG), 1552
Synthetic theory of evolution, 1397
Systematic experimentation, 2440, 2453
System of computation, 3063

T

Tables, 3056, 3069
Taittirīya Brāhmaṇa, 1086
Taittirīya Saṃhitā, 1086
Takeuti's proof, 449
  1-eliminator, 450
  well-ordering of positive numerals, 449–450
Takeuti-Veblen finitist standpoint, 454, 458–460
Tame space, 1874
Tame topology, 1873–1875
Tannery, J., 14
Tarkovsky, A., 14–18
  Andrei Rublev, 15
  Mirror, 17
  Sacrifice, 24
Tarskian-inspired model theories, 2222
Tarski's model theory, 2183, 3080
Tarski's system, 1707
Tartaglia, Niccolò, 204, 205, 209, 232
Task, 3065
Taun, 1113–1115

Taylor, Richard, 2629
Taylor series, 2493, 2497
Technologies of writing speech, 3069
Tr̥ṣṇa, 681
Temporality, 2621, 2625, 2634
Tennenbaum's theorem, 1999
Terms for operations, 3050
Terquem's problem, 1191, 1192, 1216
Testimonial knowledge, 2471
Tetrabiblos, 1734
Tetractys, 171
Texts of procedures, 3054, 3065
Textual act of communication, 3052
Textual artefacts, 3062
Textual cultures, 3052
Textual environment, 3065
Textual learning, 3062
Textual organization, 3050
Textual phenomenon, 3059
Textual resources, 3059
Thābet ibn Qurra, 1776, 1779
Thales' theorem, 2509
Theaetetean theory of ratios of magnitudes, 922
Theaetetus, 474–478, 485, 487, 494, 496–499, 502–504, 508–512, 514–518, 604–608, 611, 625, 631–634, 918, 919, 938, 940, 945, 952, 953, 962, 979, 980, 984, 988, 993, 996, 997, 999, 1006, 1016, 1463
Theaetetus' mathematical discovery by Plato, 939–940
Theaetetus' palindromic periodicity theorem
  reconstruction of, 980–983
  strong, 979–980
  weak, 979
Thebes, 1151–1155
Theodicy, 186
Theodorus, 633, 1463
Theodosius, 1746–1760, 1763, 1768, 1774, 1781–1783
Theological mathematics, 2762
Theon of Alexandria, 1734, 1737, 1740–1742
Theorem-credit economy in mathematics, 1231
Theoretical mathematical research practice, 1227
Theory of ratios, 3161
Theory of signs, 2956, 2961, 2963, 2965
Theory of statistical (probabilistic) metric space, 891
Theory of transformations, 1643, 1672, 1673
Thesaurus mathematicus (Mathematical Thesaurus), 172, 190
Thick concepts, 2542, 2548
Third Man Argument (TMA), 537
Thom, René
  admitted to École Normale, 390
  Aristotle's biology, 410
  choice of, 389
  classes, 390
  cobordism, 391
  discovery of Aristotle's work, 403–406
  at École Normale, 398
  Esquisse, 406, 409
  Euclidean geometry, 388–389
  in France, 390
  high-school years, 389–390
  interested in dynamics, 389
  at *Lycée Saint-Louis*, 390
  Mathematical Collected Works, 391
  model theory, 392
  morphogenesis, 395
  morphological scheme and linguistics, 413–415
  phenomenological equivalence relation, 410
  philosophical work, 407
  preference to Aristotelian logic, 400
  in Strasbourg, 394
  in Switzerland, 390
  theory of models, 413
  theory of morphogenesis, 401
Thorvaldsen, Steinar, 1404
Thought experiments, 2422, 2429, 2837–2841
  art of dividing 17 camels, 2540
  Basel problem, 2544, 2546
  fiction-style thought experiments, 2539
  Flatland, 2540
  graph diameter, 2529
  hyperbolic group, 2532, 2533
  impossible elements, 2541
  infinite polynomials, 2545
  light proof, 2546
  mutilated chessboard problem, 2535–2537, 2539
  thick concepts, 2542
  veridical fictional imagination, 2542, 2543
Three-body problem, 3061
3-Dimensional Euclidean space, 413
Thureau-Dangin's interpretation, 2147
Thurston, Bill, 2290, 2295, 2297, 2299
Timaeus, 467
Time Space Matter, 75, 76
Time value of money, 1332
Tokens, 3048
Tolstoi, L., 23
Topological continuum, 1566
  Bolzanian continuum, 1589–1590
  Cantor's theory, 1590–1591

Topology, 68, 71, 72, 74, 392, 1309, 1313, 1320, 1324, 1325
Topos, 2352
Total cost function, 2563
Totalitarian regime, 871
Toulmin layouts, 2591, 2672, 2674
Trace of his working process, 3060
Tractatus Logico-Philosophicus (TLP), 3119
Traditional
  definition of mathematics, 1607, 1610
  denomination, 1628
  non-traditional notion/theory, 1627
  notion(s) of quantity/number, 1596, 1598, 1600, 1604
Traditional calculus, 2562
Traditional Chinese calendar, 1102
Traditional classroom teaching, 2564
Traditional mathematics, 2465, 2467
Traditional statistical modelling, 1423
Transcendental numbers, 1487
Transcendental quantities, 183
Transfer principle, 1534
Transformation of signs, 2141, 2146
Transformations of the list of operations, 2156
Transformed into a number, 3063
Translation, 3059
Translations of Manelaus, 1775–1778
Transmission, 3052
Transparent, 2148, 2153, 2158
Treatises, 3052
Trial and error, 96, 98, 100, 102, 2560, 2562
Triangle condition, 2532
Triangle inequality, 1763, 1773, 1774
Triangles, 2563, 2564, 2566, 2569, 2570, 2572–2576, 2578
Trichotomy law, 1525, 1547
Tridentry, 929
Trigonometric functions by Euler, 329–330
Trinomial coefficients, 2490
Trobriand Islands, 3067
True Arithmetic (TA), 1922
True Opinion, 474, 475, 483–487, 490, 491, 494, 495, 507–509, 514–517
True propositions, 2895, 2896, 2899
Truncated pyramid with a circular base, 2154, 2157
Truncated pyramid with a square base, 2155
Truth, 25, 651
Turing-head configuration, 2191
Turing jump, 1955
Turing machines, 1910–1913, 1938, 1939, 1954, 2190–2192, 2196
20th and 21st Century Mathematical Practice, 4, 910
Two concepts of number, 3062
Two's complement, 179
Tymoczko, T., 2460–2462, 2464–2466, 2469, 2470, 2475–2477
Type checking, 3158
Type of manipulation, 2146
Type of reading, 2146
Type of statement, 3061
Types of inscriptions, 3068
Types of text, 3058, 3069
Type theory, 3178
Typology, 1228
Tzolkin, 1101

U
Ubhayataḥ prauga, 1048
Ultrafilters, 1916, 1992–1993
Ultrapower, 1916–1918
Ultraproduct construction, 1916
Ultraproduct of first-order structures, 1994
Uncountable set, 1966
Understanding, 2140, 2142, 2146, 2148–2150, 2152, 2153, 2160, 2626, 3051
Undirected weighted networks, 2750
Unification, 432, 435, 439, 440
Unification-inducing proofs, 435
Uniform continuity, 1482
Union set axiom, 2371, 2380
Uniqueness problem, 2220
Univalence, 2353, 2354, 2356, 3169
Univalence axiom, 2353, 2354, 2356
Univalent foundations, 2357
Universal commensurability, 1475, 1477
Universal democracy, 2377, 2378
Universal domain, 740
Universal Turing machine, 1938
Universe of sets, 122
Universe of set theory, 2375
Unrestricted domain, 2378
Unsurveyable computer-assisted proofs, 2434
Unsurveyable proofs, 2467
Uptake, 3050
The Uses of Argument, 2673
Uses of writing, 3069
USSR, 870

V
Validity, 2145
Valid transformations, 2154
Value of firm, 1331
Value of recognisability, 126

Values of mathematical proofs
abstract value, 2083, 2108
beautiful proofs, 2108
explanatory proof, 2099
practical, 2083
practical value, 2099, 2108
Value-system, 1233
van der Waerden on the history of Pell problem, 992, 997, 1000
Van Hiele model of learning, 2583
Varese Lake, 2327–2329
Variation of Cantor normal form (VCNF), 455, 456, 458
The Variation of Plants and Animals Under Domestication, 1389
Varignon parallelogram, 2574
Vaulezard, Jean-Louis sieur de, 211, 221, 227
Veblen hierarchy, 455
Vedāṅga, 1025, 1029
Vedic people, 1026
Venn diagrams, 1322, 2984
Verheideggert phenomenology, 649
Veridical fictional imagination, 2542
Vershik, Anatoly, 872
Victorian mathematics, 911
Victorian Payment, 1248
Viète, François, 200, 202, 211, 222
Views from other domains, 4
Vinogradov, Aleksandr, 877
Virtue theory, 1242, 1243, 2141, 2146, 2155
Virtuosity, 1243
Visualization, 2245, 3056
Visual symmetry in graphs, 2527
Visual synopsis, 2538
Viviani's theorem, 2567
Vlady, M., 16
Vocabulary types, 2182
Von Neumann, John, 95, 105, 160, 163–165
von Neumann ordinals, 1685
Vuillemin's analysis, 716

W
Wästlund's proof, 2547
Ways of reading, 3052
Weak König's lemma, 1915
Weak Theaetetus' periodicity theorem, 937
Philebean finite *vs.* Pythagorean finite, 952
in Philebus, 946–948
philosophical imitation of, 946–951
Plato, 939–945
in Politicus, 949–951
reconstruction of proof of, 953–967
Weber, Leonhard, 1852
Web of mathematical practices, 2806–2810
Weierstrass, Karl, 1482
Weierstrass's new foundation of theory of functions, 1144
Weigel, E, 168, 170, 171
Weil on the history of Pell problem, 918, 920, 960, 961, 980, 998–1000, 1005, 1007, 1011, 1012, 1015
Weiss, Christian Samuel, 1832
Weldon, Walter, 1391
Well-ordered sets, 1956
Well-ordering of positive numerals, 449–450
Well-ordering principle, 1913
Western society, 1330
Wetonan cycle, 1105, 1109, 1112, 1113
Weyl, Hermann, 75, 76, 1467, 1515–1521, 1553
Wilf-Zeilberger theory (WZ), 2439
William Whewell, 2139
Windu, 1115
Wisdom, 689
Wittgenstein, Ludwig, 2602, 2603, 3017, 3018
WolframAlpha app, 2565, 2566
Woodin, W. Hugh, 115, 128
WordPress, 2728, 2734
Work, 3061
Working practices, 3056, 3057
Working procedure, 3060, 3061
Wright, Sewall, 1384, 1394
Writing act, 3049
Writing of speech, 3069
Writing of utterances, 3069
Writing practices, 3069
Writings on Mathematical Procedures, 3055
Written evidence, 3058, 3063, 3067
Wuku, 1109
Wulan, 1113

X
Xenocrates, 467, 607, 608, 611–613, 619
Xylander, 217, 219, 232

Y
Yajamāna, 1031
Yijing, 187–188
Fuxi sequence, 187
King Wen sequence, 188
Yogyakarta, Sultanate, 1104, 1111, 1113, 1116, 1120

Z
Zagzebiski, Linda, 2900
Zamorzaev, Alexander Mihailovich, 1856
zbMATH, 2443, 2446
Zeeman, Christopher, 392
Zeilberger, Doron, 2436, 2438, 2439, 2441, 2445, 2452
Zeno, 603, 613, 629, 630, 635–641
Zeno's paradox, 106
Zermelo-Fraenkel set theory, 437, 1923, 2367
Zermelo's axioms, 1968
0-thin triangles, 2532
Zeus era, 969
Zones of the same page, 3067